EVOLUTIONARY ECOLOGY OF RIBOSOMAL RNA COPY NUMBER: STUDIES IN SOIL PROKARYOTE AND NEMATODE COMMUNITIES

by

Mark Reid Williamson

Bachelor of Science, University of North Dakota, 2014

A Dissertation

Submitted to the Graduate Faculty

of the

University of North Dakota

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Grand Forks, North Dakota

December 2019

This dissertation, submitted by Mark Williamson in partial fulfillment of the requirements for the Degree of Doctor of Philosophy from the University of North Dakota, has been read by the Faculty Advisory Committee under whom the work has been done and it hereby approved.

Dr. Brian Darby, Chairperson

Dr. Rebecca Simmons, Committee Member

Dr. Peter Meberg, Committee Member

Dr. Vasyl Tkach, Committee Member

Dr. Chris Nelson, Committee Member

This dissertation is being submitted by the appointed advisory committee as having met all of the requirements of the School of Graduate Studies at the University of North Dakota and is hereby approved.

Chris Nelson, Associate Dean
School of Graduate Studies

12/2/19
Date

PERMISSION

Title Evolutionary Ecology of Ribosomal RNA Copy Number: Studies in Soil Prokaryote and Nematode Communities

Department Biology

Degree Doctor of Philosophy

In presenting this dissertation in partial fulfillment of the requirements for a graduate degree from the University of North Dakota, I agree that the library of this University shall make it freely available for inspection. I further agree that permission for extensive copying for scholarly purposes may be granted by the professor who supervised my dissertation work or, in her/his absence, by the Chairperson of the department or the Dean of the Graduate School. It is understood that any copying or publication or other use of this dissertation or part thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of North Dakota in any scholarly use which may be made of any material in my dissertation.

Mark R. Williamson
December 5th, 2019

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

It's been quite the journey. I'd like to first thank my committee members. To Dr. Vasyl, who taught me that North Dakota is a great place to pursue a Ph.D. during winter, because there is nothing to do but stay inside doing research. To Dr. Simmons, who was an endless source of encouragement and delicious baked goods. To Dr. Meberg, who advised me since I was a newly minted freshman. To Dr. Larson, who opened the doors to a wider graduate school experience. Most of all, to Dr. Darby who was there every step of the way with generous support, feedback, nematodes, more nematodes, and a great model for approaching science in a problem-based, rather than method-based manner.

I'd like to also thank many members of the biology department. My fellow graduate students, especially my office mates. We were the quiet office, and it probably saved me a semester of lost productivity. Thanks also to Bob Sheppard, the office staff, Dr. Newman, and the great instructors I'd been given the chance to work with: Dr. Sally Pyle, the queen of no-nonsense, Dr. Ralph, the hardest working instructor I ever saw, Dr. Manu, with his tireless efforts to get students to not sit in class passively, and Dr. Brett Goodwin, who taught me to love statistics.

Next, I'd like to thank my family. Thank you for all the love and support Mom and Dad. I know 10.5 years in school was a long time, I'm so grateful for you both seeing me through to it. I love you both. Thanks to my siblings, Sonia, Sarah, Zane, Grant, and Gretchen.

An absolutely essential part of this whole journey has been my beautiful and incomparable wife, Sarah Marie. Thank you for keeping my science human. You were the reason I was able to not only make my Ph.D. work reflect truth, but also goodness and beauty. Thanks also to my son, David Athanasius. Thanks for reminding me during those long nights that there are more important things than dissertations. A special thanks goes out to you, dear reader, who took the time to read the acknowledgements; go get yourself a churro, you deserve it.

Finally, thanks to God, who endowed me will all the gifts I possess and who grounds all in Himself, the Truth behind the truth we as scientists seek.

-Mark

ABSTRACT

Enrichment of ecosystems with excess nutrients often results in microbial communities with an overabundance of just one or a few species, such as algae in eutrophic waters, or micro-invertebrates in organically amended soils. To explain this, the Growth Rate Hypothesis postulates that rapid growth rates are constrained by cellular ribosome concentrations and that nitrogen and phosphorous act as limiting nutrients. However, the genetic mechanism to the Growth Rate Hypothesis has not been clearly identified. The objective of this work is to test the hypothesis that rRNA copy number variance is the causative genetic mechanism of the Growth Rate Hypothesis because it codes for vital and phosphorus-rich ribosomal RNA and appears in repeats that vary in copy number between organisms. In the first chapter, I suggest that there are three pieces of evidence needed to confirm this hypothesis: 1) that rRNA copy numbers are phylogenetically autocorrelated, 2) that nutrient enrichment disproportionately benefits high-copy number species, and 3) that rRNA copy number is positively correlated with rapid intrinsic population growth rate and generation time. The second chapter tested the hypothesis that rRNA operon copy number is a heritable trait with variation, as evidenced by phylogenetic auto-correlation in bacteria, archaea, and nematodes, using rRNA copy numbers obtained from publicly available databases as well as from experimental findings. Prokaryote copy number was phylogenetically autocorrelated as a whole. Copy number also clustered in most phyla, which indicates that copy number is under selection and therefore a viable mechanism for the Growth Rate Hypothesis. The third and fourth chapters tested the hypothesis that nematodes and bacteria with high copy number benefit from nutrient enrichment more so than species with low copy

number. I examined the abundance and distribution of nematode and bacteria from Illumina sequencing for long term enrichment plot samples, short term soil core mesocosm experiments from the spring, and short term mesocosms experiments from the fall. High copy number prokaryotes and fast-growing nematodes had significant responses to nutrient enrichment that differed across season and nutrient type, which indicates that copy number operated as the genetic mechanism for the Growth Rate Hypothesis. The fifth chapter tested the hypothesis that rRNA gene copy number is positively correlated with growth metrics in enrichment type soil nematodes. I measured life-history traits and sequenced copy number and found a significant positive correlation. The correlation indicates that fast-growing nematodes have high copy number, linking copy number to nutrient enrichment from the soil core experiments.

Chapter 1: Ecology, Evolution, and Genomics in the Soil Environment

Soil Health Monitoring

Healthy soil depends on monitoring and understanding soil biota. Growth rate is a consistent factor in ecological interactions within the soil. There are two contrasting growth rate strategies; one is the utilization of nutrient-rich environments, the other is the utilization of nutrient-poor environments. This corresponds to high versus low growth rates. These two strategies are evident across multiple soil taxa. Examples are copiotrophic— adapted to locating and responding quickly to nutrient spikes—versus oligotrophic—adapted to maximizing resources in low-nutrient environments—bacteria (Fierer, Bradford, & Jackson, 2007) and enrichment-type bacterivore versus basal bacterivore nematodes.

Nutrient availability drives the community composition of growth rate strategists. Ecosystem enrichment with excess nutrients often results in communities with an overabundance of fast-growing species, such as algae in eutrophic waters, or microinvertebrates in fertilized soils. This leads to an overall decrease in community diversity (Coolon et al., 2013; Ceulemans, et al., 2014; Kearns, et al., 2016; Wang et al., 2016). Nutrient-poor environments, in contrast, tend to have a more even distribution of slower growing species (Furman, McCallum, & Davis, 1993).

Species with different growth rates display different life-history traits. The r-k selection continuum describes a broader contrast between reproductive strategies (MacArthur and Wilson 1967). Organisms tend to either produce many offspring with low parental care (r-type) or few offspring with high parental care (k-type). There is wide variation between the two extremes among actual species, as well as exceptions. Despite the shift to a more nuanced life-history

paradigm (Reznick et al., 2002), the r-k selection paradigm is still useful in ecological studies, including the categorization of soil bacteria (Fierer, Bradford, & Jackson, 2007) and nematodes.

Given their ubiquity, diversity, and importance in the soil, prokaryotes are a good target taxon for monitoring soil health. They play an important role in the carbon, nitrogen, and phosphorus cycles. Bacteria uptake organic carbon through decomposition and liberate atmospheric carbon (in the form of CO_2) with photosynthesis and chemosynthesis. Nitrogen-fixing bacteria, symbiotic or free-living, take atmospheric nitrogen (N_2) and convert it into ammonia and other nitrogen compounds. Anaerobic bacteria perform denitrification, which releases nitrates and nitrites back into the atmosphere. Bacteria also solubilize phosphate found in mineral form. More broadly, soil bacteria and other members of the soil food web help decomposition, nutrient/mineral cycling, nutrient/mineral redistribution, serve as nutrient reservoirs, sequester carbon, help detoxification of the soil, physically structure the soil, and aid in biological control (Ferris & Matute, 2003).

Human and prokaryotic soil activity is interconnected. Many bacteria are beneficial and supply nutrients to crops, stimulate plant growth, control or inhibit plant pathogens, improve soil structure, bio-accumulate and leach inorganics, and bio-remediate polluted soils (Hayat, Ali, Amara, Khalid, & Ahmed, 2010). More generally, they aid in decomposition and nutrient recycling.

Interactions between prokaryotes and plants include both positive and negative feedback loops (Bardgett, 2005). Furthermore, in addition to plant-pathogens, there are a variety of soil bacteria species that are potentially pathogenic to humans, including *Clostridium, Corynebacterium, Escherichia, Legionella*, *Listeria*, and *Staphylococcus*. Pathogenic bacteria can enter the soil through infected hosts and human activities; for example, potentially

pathogenic gamma-Proteobacteria are increased in wastewater irrigated soil compared to rainwater control (Broszat et al, 2014).

Nematodes are another important taxon for use in soil health monitoring. Nematodes have a broad potential and are a growing target for use as bio-indicators, both with classical and new molecular approaches (Wilson & Kakouli-Duarte, 2009). Different species are used as indicators for soil conditions such as enrichment, drought, heavy metal toxicity, or altered land use (Neher 2001). Over 200 recent papers have covered the use of nematode indices to evaluate soil status (Ekshmitt & Korthals, 2009).

Of the major threats to soil health, nematodes are best for assessing contamination and loss of biodiversity (Ekshmitt & Korthals, 2009). Studies of heavy metal pollution on nematodes include Zinc, Copper, Nickel, Cadmium, Lead, Chromium, and Selenium (Nagy, 2009). In utilizing nematode communities, Neher and Darby (2009) showed how both single indexes that are independent of taxonomic identity (like Shannon's diversity) and multivariate analyses (including cluster analysis and ordination) were useful to reflect past events or predict future ones.

Together, prokaryotes and nematodes represent a wide range of dominant, important, and functionally diverse taxa. They are also linked trophically, as bacterivore nematodes are the second most dominant feeding group in the soil. Both groups are important for soil health and their community composition can be used to detect changes in the soil and nutrient availability.

Ecological Stoichiometry

Ecological stoichiometry is the study of the ratio and balance of chemical elements in ecological interactions (Elser et al. 2000b). It examines how nutrients affect the elemental ratios of organisms and how that shapes their ecology. In principle, the elemental output organisms put

into growth and reproduction must be balanced by the same elemental input. Elements not used in outputs like biomass production, respiration, or reproduction become waste products. Nutrient balance affects many physiological traits such as organismal growth rate, nutrient recycling rate, mineralization rate of nutrients (Elser et al., 2000c), organism size, life history (Elser et al, 1996), food webs (Elser et al. 2000a), and population dynamics (Anderson et al. 2004). Stoichiometry cannot be neglected when studying the effect of specific nutrients on growth rate.

While organic matter is predominately composed of carbon, hydrogen, oxygen, and nitrogen, other elements such as phosphorus, sulfur, and potassium are still essential for living organisms. Carbon is the backbone for most macromolecules as well as the basis for chemical energy. It forms the dominant molecular structure for base pairs in DNA and amino acids in proteins. Nitrogen is present in the ring structure of the base pairs and forms a functional group in all amino acids. Phosphorus is a component of ATP, nucleic acids, and many complex organic compounds, as well as bone in vertebrates.

Elemental composition is commonly expressed in ratios of C:N:P, C:N, C:P, or N:P. The C:N:P ratio is roughly 100:15:1 for deep ocean plankton biomass (Redfield 1934). For terrestrial plants, the N:P ratio can shift greatly and availability of each can limit growth (Güsewell 2004). P:C ratios in the algae *Rhodomonas* increased when subject to *Daphnia* predation when P was limited; showing the influence of grazing and nutrient recycling (Vrede, 1998).

The C:N:P ratio is variable within the macromolecules. Carbohydrates—with a few exceptions such as chitin, peptidoglycan, and glycosaminoglycans—are composed of C and little to no N or P. Lipids are also are usually C-rich with no N or P; the important exception being phospholipids, which have a fatty acid tail and phosphoric acid head. Proteins contain no phosphorus structurally and only gain it by post-translational phosphorylation. However, their

C:N ratio can range from 9:1 down to 3:2. Nucleotides contain all three, with C:N:P being 10:5:1 for the bases guanine and adenine, and 9:3:1 for cytosine, thymine, and uracil. Thus, even though there is only one phosphorous atom per nucleotide, phosphorous is stoichiometrically more abundant in nucleic acids than many other biomolecules. Therefore, DNA and RNA can be considered phosphorus-rich and potentially a limiting factor in genome size and RNA transcription (Elser et al. 2000c).

Growth Rate Hypothesis

The addition of nitrogen and phosphorus to soil and aquatic environments tend to stimulate the growth of organisms with disproportionately rapid growth rates, such as copiotrophic bacteria (Fierer et al. 2007) or enrichment-type nematodes (Ferris and Bongers 2006). This results in relatively few species dominating counts of abundance in microbial community abundance counts. The species that benefit from nutrient enrichment must have some cellular or physiological limitation to growth rate under normal conditions that are overcome by excess nutrients.

The Growth Rate Hypothesis (GRH) posits that the number of ribosomes (comprised mostly of ribosomal RNA) in the cytoplasm constrains an organism's intrinsic growth rate because ribosomes are the cell's primary site and rate-limiting factor of protein translation (Shah et al. 2013). Ribosomal RNA makes up over 80% of the cell's total RNA, and because RNA is one of the most disproportionately nitrogen and phosphorus-rich biomolecules—nucleic acids have a high P:C and N:C ratio relative to most other biomolecules (Elser et al. 2003)—it is possible that intrinsic growth rate is in part limited by dietary nitrogen and phosphorous (Elser et al. 1996).

For example, in marine environments, opportunistic gamma-proteobacteria respond quickly to nutrient enrichment (Eilers et al. 2000; Beardsley et al. 2003). Conversely, members of the alpha-proteobacteria SAR11 cluster, while the dominant group in ocean surface communities, do not respond quickly (Morris et al. 2002). Members of the SAR11 clade have low copy number (n=1), while Gamma-Proteobacteria have a higher than average ribosomal copy number (n<6). The growth rate of bacterial groups are affected by nutrient enrichment, but other factors such as mortality can be important in environments such as estuaries (Yokokawa et al. 2004).

In metazoans, a major percentage of total phosphorus (P) is found in nucleic acids in *Daphnia* species, as opposed to other P sinks like phospholipids (Vred et al. 1999). Additionally, *Daphnia* grow faster in P-rich as opposed to P-poor environments (Acharya et al 2004). More broadly, growth rates in crustacean zooplankton are positively correlated with percent P (Main et al. 1997). The stoichiometry between herbivores and autotrophs is not balanced and this imbalance is exaggerated in terrestrial environments (Elser et al. 2000a). In a diverse study, growth rate, RNA content, and %P were all correlated strongly in multiple taxa that included microbes, insects, and crustaceans (Elser et al. 1996).

Genetic Hypotheses for the Growth Rate Hypothesis

Under the GRH, it is expected that the growth rate of consumers with abundant cytoplasmic rRNA will be most limited by prey nutrient content and that organisms with abundant cytoplasmic rRNA should be competitively advantaged in nutrient-enriched environments (Watts et al. 2006; Kay et al. 2005). However, the underlying genetic basis for this phenomenon has not been clarified. Three main hypotheses have been proposed: genome

streamlining (Hessen et al. 2010a), intergenic spacer region length variation (Gorokhova et al. 2002), and rRNA gene copy number variation (Klappenbach et al. 2000).

Genome streamlining postulates that under selection for rapid growth, P is shifted from DNA to RNA and the genome is reduced (Hessen et al. 2010a). Hessen et al. (2010b) suggested that phosphorus limitation might lead to genome streamlining because fewer genes and gene transcripts would require less phosphorous-containing mRNA. However, mRNA is a small component of total RNA, and it might rather be the case that genomic rRNA operon copy number is a more direct limitation on the number of ribosomes, their nutrient demand, and an organism's intrinsic growth rate. Vieira-Silva (et al. 2010) found no evidence of streamlining in prokaryotes. Hessen (et al. 2010b) clarified that it was possible in eukaryotes as opposed to prokaryotes.

Streamlining may occur from a variety of mechanisms (Martnez-Cano et al., 2015) and follows two major paths: parasites, symbionts, and commensals on one side and free-living organisms on the other (Giovannoni, Thrash, & Temperton, 2014). Small genome size has been linked to adaptations to high-temperature environments (Sabath, Ferrada, Barve, & Wagner, 2013), symbiotic lifestyles (Gao et al., 2014), and free-living prokaryotic species (Giovannoni et al., 2014; Grote, Thrash, & Huggett, 2012, Swan et al., 2013). Low rRNA copy number was found in conjunction with a small genome (Strehl, Holtzendorff, Partensky, & Hess, 1999; Tripp et al., 2010). Most studies of small genomes are focused on marine or non-free-living prokaryotes.

Genomic streamlining studies help establish the minimum genome size for free-living organisms (Maniloff, 1996), as well as providing possibilities for biotechnology applications (Leprince, van Passel, & dos Santos, 2012). There is growing evidence that genome reduction is

the dominant form of evolution (Csuros & Miklos, 2009; Wolf & Koonin, 2013; Cuypers & Hogeweg, 2012; Mira, Ochman, & Moran, 2001). Long periods of time go by with a general trend towards reduction, punctuated by short bursts of innovation. But gene expansion is still an evolutionary mechanism. In Cyanobacteria, larger genome size was related to gene family expansion, which increased adaptability but also added to non-coding DNA accumulation (Larsson, Nylander, and Bergman 2011).

Genome streamlining is most evident in species that are parasitic or symbiotic. The symbiotic marine bacterium *Candidatus Endolissoclinum faulkneri L2* has a reduced genome size, resulting in a loss of most vital functions other than producing toxic metabolites as a chemical defense for its tunicate host (Kwan et al 2012). For free-living species, small/streamlined genomes are expected to be found in specialists but is not always the case as evident from the SAR11 Clade (Grote et al, 2012). Exaggerated streamlining, also called erosion, has occurred in symbiont strains of the already small genome of *Polynucleobacter necessaries* (Boscaro et al, 2013).

Another possible GRH mechanism is the reduction of the intergenic spacer region (ISR). This region is found within the rRNA operon but does not code for rRNA. It is cut out during the processing of raw rRNA towards the finalized forms (Srivastava & Schlessinger, 1990; Cooper, 2000). Each time the 16S, 23S, and 5S rRNA are synthesized from the operon, the unused spacer region is also synthesized. A high rate of rRNA synthesis results in inefficient P usage in spacer regions. Reducing those regions would increase P efficiency in rRNA synthesis (Gorokhova, Dowling, Weider, Crease, & Elser, 2002)**.**

The ISR is less conserved than rRNA and subject to more rapid duplication/deletion (Matyášek et al., 2012), but its phylogeny is consistent with rRNA trees (Lan, Pérez Luz,

Reeves, & Rodríguez-Valera, 1998). Sequence comparison of the ISR can be used to distinguish species that are all but identical in their 16S sequence (Chun, Huq, & Colwell, 1999; Osorio et al, 2004; Graham et al, 2009; Gürtler & Stanisich, 1996). Its evolution appears driven by recombination, at least in prokaryotes (Janezic, Indra, Rattei, Weinmaier, & Rupnik, 2014). Few studies have examined the reduction of the region related to P amount. One study found the reverse expectation; longer intergenic spacer variants were positively correlated with juvenile growth rate and body RNA and P (Gorokhova, Dowling, Weider, Crease, & Elser, 2002).

While there is merit to both, genome streamlining and ISR reduction as a possible mechanism of overcoming environmental nutrients constraints, neither appears to satisfactorily explain the variation in growth rate between species. We believe that the third mechanism, variation of ribosomal RNA copy number, is a better explanation. RNA is differentially expressed resulting in variable amounts in each cell, compared to the relatively static amount of cellular DNA; eighty percent of cellular RNA is represented by rRNA, indicating that phosphorus is a limiting factor in organismal growth. Reduction of CN for species adapted to low P environment may allow for more efficient ribosome synthesis, similar to how reduction of genome size is linked to oligotrophy (Hessen, Ventura, & Elser, 2008; Grote, Thrash, & Huggett, 2012; Giovannoni, Thrash, & Temperton, 2014; Martinez-Cano et al., 2015). Increase of CN for species that exploit high P fluxes would allow for a higher maximum growth rate, but less P efficient synthesis in times of low P (Nemergut et al., 2016; Niederdorfer et al., 2017; Ortiz-Alverz et al., 2018)

The link between high CN and growth rate has been examined previously. High copy number correlates with rapid colony formation in complex mediums (Klappenbach et al 2000). Furthermore, high copy number species grow extremely fast (Dan Buckley, in a talk during the

2017 Argonne Soil Metagenomics Meeting) and tend to have strong promoters to assist that high growth rate (Aiyar, Gaal, and Gourse, 2002). However, this link is not universal; bacteria of the *Bacillus* genus are notable for having growth rate and copy number decoupled (Valdivia-Anistro et al, 2016).

We posit that copy number variation provides a more explanatory mechanism than genomic streamlining and intergenic spacer reduction. Because rRNA is rich in phosphorus and ribosomes are needed for protein production, the rRNA operon is a plausible genetic candidate. The mechanism is as follows; the higher the rRNA operon copy number count, the higher the maximum growth rate due to faster synthesis of rRNA from increased copies. However, unless phosphorus is abundant, maximum growth rate cannot be attained, as P acts as a limiting factor.

Ribosomal RNA

Ribosomal RNA is a vital component of the ribosome, the protein building organelle for all living organisms. This attribute makes the rRNA operon the dominant barcoding gene in modern prokaryotic phylogenetics. Ribosomes require ribosomal RNA (rRNA) for its primary structure along with ribosomal proteins. This rRNA is coded in the genome as rDNA and can come in varied copy number. Figure 1.1 shows the general structure across major groups and includes various spacer regions.

Because the 16S/18S portion of the operon has both variable and invariable regions, it is supremely useful for molecular work. The rRNA operons have been used as molecular chronometers because of easy sequencing, universality, and consistency (Woese, 1987). Their use in phylogenetic analysis predates the use of PCR (Lane et al., 1985), but was highly benefitted by PCR methods (Wilson, Blitchington, & Greene, 1990). With PCR, the variable

regions distinguish between taxa, while the invariable allow for primer attachment to gain amplicon sequencing.

Copy Number Variation

Ribosomal RNA copy number is widely variable across the divisions of life. Most eukaryotes, including yeast, have tandem arrays that can range from dozens to thousands of copies. Copy number can vary over a 100-fold between different eukaryotes (Long and Dawid 1980). Copy number estimates ranged from 56 to 323 in six, well-studied nematode species (Bik, Fournier, Sung, Bergeron, & Thomas, 2013). In fungus, the mean copy number is 133, but is highly variable and phylogenetically correlated; in contrast, fungal copy number is not related to trophic mode, ecological guild, or the size of the genome (Lofgren et al., 2019). For microorganisms, copy number is most variable in ciliates, ranging in the tens of thousands (Gong, Dong, Liu, & Massana, 2013).

In prokaryotes, ribosomal operon copies are spread across the genome and are constrained in number. Ribosomal RNA operon copy numbers vary from 1 to 15 between species of bacteria (Klappenbach et al. 2000), with the most recent update of the Ribosomal RNA Operon Copy Number Database (rrnDB) having a few isolates with 16, 17, and 21 copies. The mean prokaryotic copy number is 4.8 with a mode of 7. The model species E. coli has 7 copies. Archaea are even more constrained in copy number, with a range of only 1 to 4 copies (mean=1.7, mode=1). Prokaryotic ribosomal copies can differ in expression levels in response to different environmental factors (Condon & Philips, 1992). Copies are not always identical within an organism but are still more phylogenetically related to one another compared with the homologous copy in other species (Cilia, Lafay, & Christen, 1996). Copy number is almost always constant within species. Notably, however, some strains of *E. coli* vary in copy number,

ranging from 6-8. More complex animals have much higher copy numbers, but proportionately have similarly low levels of overall intergenic copy number variation.

The rRNA segments in the rRNA operon are co-transcribed; large numbers of tandem operons in a genome are thought to facilitate the simultaneous transcription of many rRNA subunits (Weider et al. 2005). This results in a greater number of complete ribosomes in the cytoplasm, increasing the cell's capacity for mRNA translation in a ribosome-limited system (Shah et al. 2013). Ribosomal RNA copy number is positively correlated with genome size (Prokopowich et al. 2003) and "colonizing" lifestyles (Klappenbach et al. 2000). Elser et al. (2000c) suggested that ribosomal RNA content may be linked to the stoichiometry of phosphorous limitations in rapidly growing species; this was demonstrated in *Daphnia* (Acharya et al. 2004). Deleting operons in *E. coli* leads to decreased growth rate and competitive ability in nutrient-rich environments (Gyorfy et al., 2015; Stevenson & Schmidt, 2004), reinforcing the links between ribosomes, phosphorus, and growth rates.

While the rRNA operon contains multiple pieces of rDNA that are expressed as one unit, it is not clear if the expression of these multiple copies is constitutive. Constitutive expression is more likely for eukaryotes, as they are in tandem arrays, but should be theoretically possible in prokaryotes with some sort of regulatory mechanism. If they are constitutively expressed, organisms with high copy numbers would be at a disadvantage in low nutrient environments. Expression would be "all or nothing" for organisms which would make these organisms more subject to limited phosphorous.

Prokaryotes

Bacteria: There are approximately 52 identifiable, major phyla of bacteria, though many do not have culturable representatives (Giovannoni and Rappe 2003). Many bacteria taxa, both

cultured and uncultured, are found in the soil. Community diversity in the soil varies but is strongly influenced by pH (Fierer & Jackson, 2006). Dominant soil bacteria phyla are Acidobacteria, Bacteroidetes, Firmicutes, Actinobacteria, alpha-Proteobacteria, and beta-Proteobacteria (Fierer et al 2007). The most important functional types are decomposers, nitrogen-fixers, Actinobacteria, and sulfur-oxidizers. The approximate stoichiometric elementary analysis of bacteria is shown in tables 1.1a and b. While the major elements are quite stable, the macromolecule composition can vary. Bacteria with cell walls contain peptidoglycan, which is similar to chitin and contains nitrogen. Some bacteria additionally have teichoic and teichuronic acids, which are unique structural polymers that contain both nitrogen and phosphorus.

Archaea: Archaebacteria were first described in the landmark paper by Woese and Fox (1977). They represent a lineage of life as distinct from bacteria as from Eukaryota. Woese et al. (1990) later proposed a three-domain tree of life (Figure 1.2). The four main groups of archaea are methanogens, extreme halophiles, sulfate-reducing archaea, and extreme thermophiles. While all archaea were initially thought to be extremophiles, phylogenetic studies indicate that archaeans have an array of habits and are found in diverse environments (Delong 1992; Furhman, McCallum, and Davis 1993; Barnes et al. 1994; Munson, Nedwell, and Embley 1997; Preston et al. 1996). Using the SSU rRNA locus, Bintrim (et al. 1997) found diverse archaea in soil communities and Buckely, Graber, and Schmidt (1998) analyzed non-thermophilic Crenarcheaota in soils. Both of these studies describe new species with uncertain phylogenetic affinities. That uncertainty is typical for environmental metagenomics; new species discovered often have no close relatives that have been previously described.

Nematodes

Traditional nematode identification has relied on morphology but is time-consuming and difficult even for expert nematologists (Neilson et al, 2009). Molecular based methods are becoming the dominant form of identification, starting with a phylum-wide study using the 16S rRNA operon (Holterman et al., 2006), and are often used in tandem with morphological methods for additional verification. Within metagenomics, some researchers saw interest in rRNA copy number and began working to correct for high throughput sequencing (Darby et al., 2013) and examine intra-genomic variation in different copies of the rRNA (Bik et al., 2013). Other recent work in nematode genetics includes the use of molecular operational taxonomic unit (MOTU) to access nematode diversity in soil samples (Floyd et al. 2002), parallel sequencing using multiple rRNA (Porazkinska et al. 2009), the discovery of a biogeographical pattern of distribution (Porazkinska et al. 2012), and copy number estimation and manipulation of *C. elegans* lines (Bik et al. 2013). Methodologically, there has been work in the development of a better pan-metazoa primer set, including nematodes (Capra et al., 2016) and multiple testing of next-generation sequencing methods for detecting nematode diversity (Peham, Steiner, Schlick-Steiner, & Arthofer, 2017; Sapkota & Nicolaisen, 2015; Seesao et al., 2016).

Maturity Index

The Maturity Index (MI) is a very useful tool for using nematodes as soil health indicators (Bongers, 1990). Nematodes resistant to disturbances, "colonizers", are given a lower score, while nematodes sensitive to disturbances, "persisters", are given a higher score. The Maturity Index is the weighted means of the CP (colonizer-persister) values. The same format followed for the Plant Parasitic Index developed shortly after. While a low score on the indexes is indicative of a disturbance, the index alone does not tell the whole story; it has to be followed

up with investigation. Thus the Maturity Index is a tool that can be used alongside other soil measurements to test for health by indicating the level of disturbance. Further developments have led to a whole family of Maturity Index metrics (Ferris & Bongers, 2009).

Multiple papers have used the Maturity Index in subsequent years. Ferris, Bongers, and Goede (2001) studied the abundance and distribution of nematode functional groups using a weighting system. There was a decrease in mean MI values for nematodes in high nitrogen-enriched soils but no difference for phosphorus-enriched soils (Ghani, Cox, Sarathchandra, Yeates, & Burch, 2002). Soil plots with a variety of nutrient enrichments besides carbon found increase in bacterivores and fungivores and over time suggested succession from enrichment to generalist bacterivores (Ferris & Matute, 2003). Combining feeding groups with the life history characteristics of the MI disclosed complementary information about the soil systems studied and overall biodiversity (Bongers & Bongers, 1998).

More recently, the MI was more similar between nematodes identified with morphology and nematodes identified with DNA sequencing from soil-isolated specimens compared to nematodes identified from the soil through direct soil sequencing; stressing the need for standardization (Griffiths, de Groot, Laros, Stone, & Geisen, 2018). In fields, grasslands, and woodland, the body size of nematodes better explained environmental variation (Sechi, De Goede, Rutgers, Brussaard, & Mulder, 2018) than a study that used the MI on the same dataset (Vonk, Breure, & Mulder, 2013). The MI was combined with feeding type, tail shape, body length, and body shape to form a biological traits analysis in Po River lagoons, which gave more ecological information, but did not change distribution patterns compared to single-trait analysis (Franzo & Del Negro, 2019). In subalpine forests, nitrogen deposition increased abundance but decreased richness, diversity, and overall MI, as microbial increases would benefit colonizer

nematodes (Shaw, Boot, Moore, Wall, & Baron, 2019) and P addition decreased the MI as well (Yang, Zhang, Huang, Bhusal, & Pang, 2019). In greenhouse plots, nematode abundance and community composition were not affected by P addition alone (Olatunji et al., 2019).

Another closely related idea to the Maturity Index is that of enrichment profile (such as basal and enrichment type). The papers of Ferris & Bongers (2006), Li et al. (2014), and Jones et al. (2006) worked on aspects of this profile. "Enrichment-type" bacterivores are opportunistic colonizers associated with localized patches of organic enrichment such as dung, compost, or decaying carcasses. They disperse microbes throughout the soil, regulate the microbial communities on which they feed, and regulate the rate and chemistry in which soil organic matter is transformed into inorganic and dissolved organic nitrogen and phosphorous (Ferris and Bongers 2006). Understanding how dietary phosphorus affects the evolution of bacteria and enrichment-type bacterivorous nematodes will also help us to understand how the species composition of bacteria and nematodes might affect the consequences of phosphorous enrichment.

Objectives

Regardless of the mechanism that maintains copy number (CN) concordance in the rRNA operon, copy number variation is interesting from an ecological, evolutionary, and genetic level. It is tied to a number of biological phenomena that are worth exploring. This approach is a broad one, in line with the growing field of evolutionary and ecological functional genomics, the focus to know more about the interaction of genes, evolution, and ecology (Feder and Mitchell-Olds, 2003).

The objective of this project is to identify the genetic basis of the link between growth rate, nutrient limitation, and rRNA copy number. It posits rRNA copy number variation as the

primary genetic mechanism of the Growth Rate Hypothesis. Each chapter focuses on accomplishing a piece of the project.

Chapter 2: Phylogenetic Autocorrelation

Hypothesis: Ribosomal RNA copy number is expected to be phylogenically conserved in the sense that copy number evolves incrementally under selection rather than in large or random steps. Therefore, closely related species are expected to have closely related copy number.

Data Requirements: All publically available 16S rDNA sequences and copy number estimates for prokaryotes, as well as taxonomic information.

Approach: Test the correlation of bacteria sequences differences to copy number difference, the phylogenetic signal of copy number within bacterial phyla, and the accuracy of methods used to estimate unknown copy number.

Chapter 3: Oakville Prairie Soil Core Studies

Hypothesis: Prokaryote species with high rRNA copy number and nematode species with high colonizer-persister (CP) values should benefit most from chronic nutrient enrichment, especially phosphorus

Data Requirements: 16S and 18S amplicon sequencing data from two seasons of soil cores at Oakville prairie.

Approach: Collect soil core mesocosms to run short term enrichment studies. Sequence prokaryotic 16S and nematode 18S from the soil. Determine if high copy number prokaryotes and colonizer nematodes are more abundant in nutrient-enriched soil.

Chapter 4: Konza Prairie Soil Studies

Hypothesis: Species with high rRNA copy number should benefit most from chronic nutrient enrichment, especially phosphorus

Data Requirements: 16S/18S amplicon sequencing data from soil plots from Konza prairie.

Approach: Collect soil samples from long term enrichment study. Sequence nematode and prokaryotic 16S/18S from the soil. Determine if high copy number prokaryotes and colonizer nematodes are more abundant in nutrient-enriched soil.

Chapter 5: Nematode Life History Traits and 18S rRNA Copy Number

Hypothesis: Ribosomal RNA copy number should be positively correlated with growth rate, as faster-growing species need high levels of rRNA for protein production

Data Requirements: Life history traits and copy number estimations from phylogenetically diverse bacterivore nematodes.

Approach: Grow bacterivore nematodes from local soil and from CGC on NGM plates. Obtain growth rate metrics through life history observations. Obtain copy number estimation by sequencing pooled 18S amplicons.

Figures and Tables

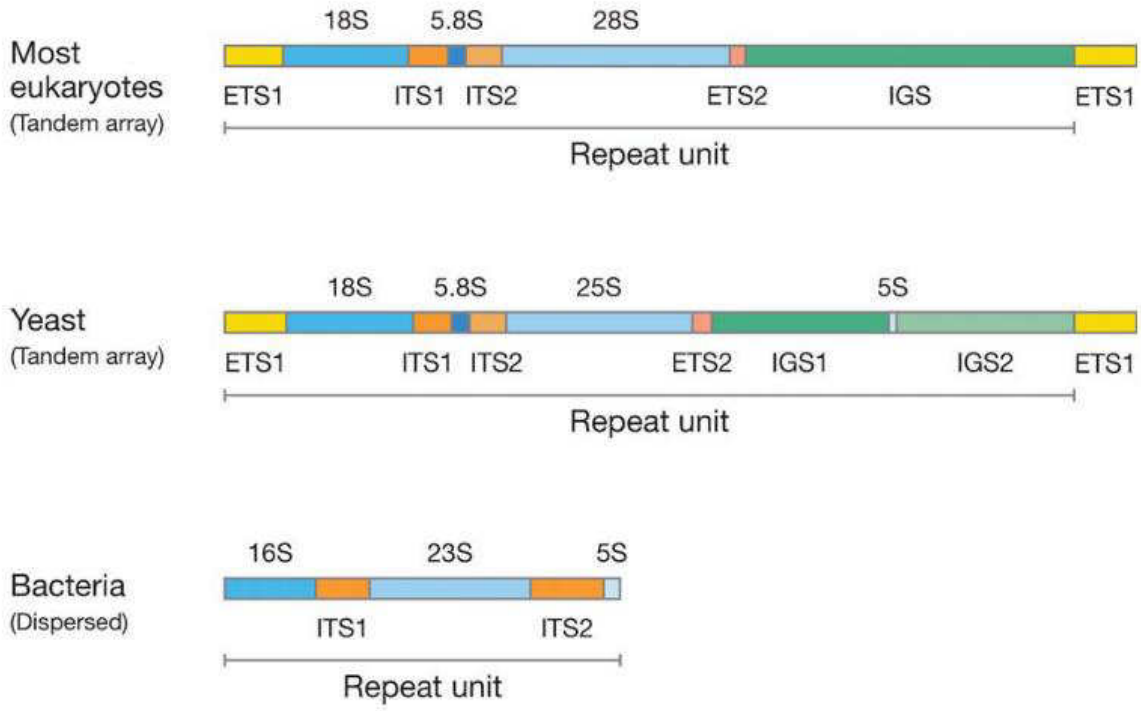

Figure 1.1: rRNA Structure across Organisms. From Nei and Rooney (2005).

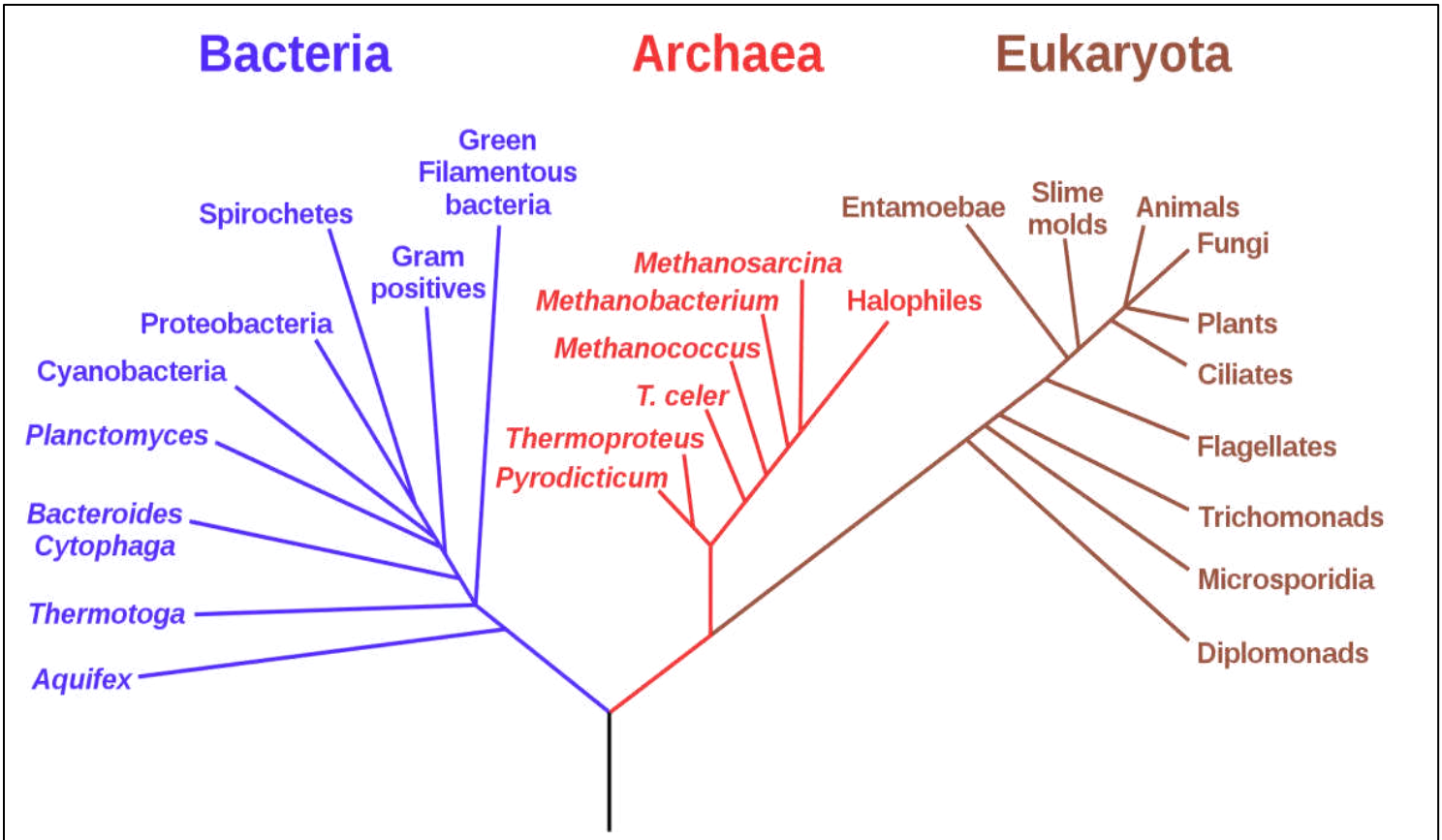

Figure 1.2: Tree of Life. Three domains of life, Bacteria, Archaea, and Eucaryota, are pictured in different colors. Major taxonomic groups are displayed for each domain. (http://commons.wikimedia.org/wiki/File:PhylogeneticTree.png)

Table 1.1a: Prokaryote elemental composition

Element	Percentage
Carbon	~50%
Nitrogen	~15%
Phosphorus	~1%

Table 1.1b: Prokaryote molecular composition

Macromolecule	Percentage
Protein	40-80%
Carbohydrate	1-30%
Lipid	1-40%
RNA	~20%
DNA	~3%

Chapter 2: Phylogenetic Autocorrelation

Abstract

This chapter tests the hypothesis that ribosomal RNA copy number variation is under positive, stepwise, natural selection and tests the accuracy of predicting copy number using phylogeny. Positive, stepwise selection is important for copy number to be a viable genetic mechanism for the Growth Rate Hypothesis and predicting copy number is important for quantitatively accurate metagenomic samples. We used the Ribosomal RNA Database (rrnDB) to test for selection by measuring the correlation between copy number and sequence identity using two different methods, phylogenetic autocorrelation and semi-variance. We next measured the phylogenetic signal of copy number across prokaryote phyla and the accuracy of copy number estimation by ancestral state reconstruction for each phylum. We found copy number to be phylogenetically autocorrelated and under selection across all prokaryotes, making it a viable mechanism for the Growth Rate Hypothesis. We were able to accurately estimate copy number for some phyla but not others, which means that metagenomic samples can be corrected for copy number on a phylum-level basis.

Introduction

The Growth Rate Hypothesis (GRH) postulates that high growth rate is linked to high concentration of ribosomal RNA (rRNA) and is limited by phosphorus (Elser, Sterner, et al. 2000). The genetic mechanism for the GRH is not known, but several studies (Gorokhova et al, 2002; Klappenbach et al, 2000, Prokopowich, Gregory, & Crease, 2003; Weider, Elser, et al, 2005) have suggested that rRNA copy number variation is a possible candidate. This is because ribosomal RNA copy number is highly variable.

The DNA that codes for rRNA are found in operons. The most utilized section is the 16S, which makes up part of the small-subunit of the ribosomes. Between prokaryotic species, rRNA operons range in copy number from 1 to 21, with an average of 4.8. Furthermore, copy number distribution is not consistent across phyla. The diverse and fast-growing Firmicutes have an average copy number of 6.9. Proteobacteria, a large phylum that includes *E. coli*, has a wider range but a lower mean of 4.9. The small and poorly characterized Verrucomicrobia has a range of 1-3 copies and a mean of 1.6. On a domain level, archaea are more constrained in copy number than bacteria. Archaea range from 1-4 copies and have a mean of 1.7.

For copy number variation to be a plausible mechanism for the GRH, it would need to evolve in a stepwise manner. That would indicate that the specific copy number a species has is not random but is related to their fitness. Preliminary work showed positive phylogenetic autocorrelation for a subsample of the rrnDB. Furthermore, Louca, Doebeli, and Parfrey (2018) found significant autocorrelation of the most recent version of the rrnDB.

Because ribosomal RNA operon copy number is variable, amplicon sequencing of the rRNA operon results in amplicon numbers that do not relate directly back to cell or genome count. This is problematic because metagenomics pipelines with an end goal of characterizing the abundance and distribution of organisms in environmental samples will not be quantitative. Truly quantitative cell counts in metagenomics studies require accurate copy number correction.

Current attempts to estimate copy number using phylogenetic methods have difficulty with highly disparate isolates, a common component of environmental samples. This is because phylogenetic autocorrelation tests how much a trait is correlated with evolutionary distance. If a species has no relatives with known copy number within a certain evolutionary distance, estimation has a low degree of accuracy. Ancestral state reconstruction is another method of

estimating copy number at higher points of dissimilarity. It estimates character traits of ancestral species to predict traits of unknown species. However, it suffers from the same issues with highly dissimilar sequences. Accurate copy number estimation is an ongoing problem.

Objectives

The objective of this chapter is to test the hypothesis that closely related bacterial species tend to have a similar number of rRNA operons in their genome. To accomplish this, we obtained copy number estimates and 16S rDNA sequences from the rrnDB (version 5.4), tested the strength of phylogenetic autocorrelation, determined the phylogenetic signal of copy number, accessed the success rate of copy number estimation by ancestral state reconstruction, and described a best-practices pipeline for copy number estimation.

Methods

Databases

The publicly available Ribosomal RNA Operon Copy Number Database (Stoddard et al. 2015) was used to obtain prokaryotic copy number estimates. Each entry in the database included scientific name, strain name, taxonomy, RefSeq accession, INSDC accession, and copy number. The rrnDB is a representative subsample of all available sequenced prokaryotes from GenBank, based on proportions of major phylogenetic groups (Figure 2.1). A 16S sequence file of all species from the most recent version (5.4) was used for alignment and phylogenetic tree work. Version 5.4 had 8223 Bacterial records (2913 species) and 262 Archaeal records (201 species). A custom Python script (Python Software Foundation. Python Language Reference, version 3.5), which included Biopython (Cock, et al, 2009) was used to connect entries from the

rrnDB to their respective sequences. Only isolates with unique sequence and copy numbers were retained, 5185 isolates representing 2598 species.

Alignment and distance calculations

The 16S sequences were aligned in CLC Genomics Workbench 10 (Qiagen Bioinformatics) using default fast settings. The resulting pairwise distance matrix was converted to tabbed columns. The same was done for copy number difference. Columns were binned by 0.1% intervals and any entries below 60% identity were removed. Finally, the percent identity values were converted to percent difference values.

Autocorrelation Graph

The pairwise sequence difference and copy number difference columns were aggregated in R (R Core Team 2017) by binning category and the average copy number difference was plotted as a function of the average sequence percent difference. Next, a Mantel test was run on the two pairwise matrixes (sequence percent difference and copy number difference) using the R-package ape (Paradis, Claude, and Strimmer 2004). Mantel tests are used to test for correlation between two matrices (Mantel 1967), so a significant test statistic would indicate the two-variable matrices are autocorrelated.

When visualizing the autocorrelation plot, a spike in copy number difference was observed at 10% sequence dissimilarity. It was examined and determined to be a true biological event from high copy number variation in the Phylum Firmicutes. For further confirmation, a Firmicutes phylogenetic tree was created, colored by copy number, and tested for phylogenetic signal (see phylogenetic signal section).

Semi-variogram

A semi-variogram was created using the package pylin in R (Tarroso, Velo-Anton, and Carvalho 2015) to complement the autocorrelation graph. Semi-variograms display spatial autocorrelation of measured sample points and can be adapted for phylogenetic distance instead of geographic distance. The difference matrix and the copy number difference matrix were modeled using an exponential model, a lag of 1, and a lag tolerance of 0.5. The sill, range, and nugget were calibrated to 1000, 5.5 and 650 respectively.

Phylogenetic Signal

To test if copy number was under selection within each phylum, the 16S sequences were separated by either phylum or another appropriate grouping. Gammaproteobacteria was aligned in CLC Genomics Workbench 10 using default fast setting, followed by a neighbor-joining tree, while all other groups were aligned in Mega7 (Kumar, Stecher, and Tamura 2015) using ClustalW on default settings, followed by a maximum likelihood tree.

Each isolate in the tree was colored by copy number using Fig Tree (v1.4 Andrew Rambaut) and a python recoloring script (Figtree-recolor.py, Terry Cojones). To supplement the archaeal data, 155 archaeal records from the Silva Database, a quality-checked and aligned ribosomal RNA sequence database (Quast et al 2013), were also used. Those sequences were aligned used to create a maximum likelihood tree in Mega7 that was manually colored for copy number. Phylogenetic signal of copy number was tested using the R package phytools (Revell 2012) on each nexus tree using the lambda method. All signal tests generated a test statistic and resulting p-value, except three trees that could not be resolved due to size or structure.

Ancestral State Reconstruction

To test ancestral state reconstructions by phyla, Verrucomicrobia was selected as a representative phylogenetic tree. This phylum has only a few described species and accounts for about 5% relative abundance in the soil (Delagado-Baquerizo et al, 2018). Ancestral states were estimated under an ER model using the R-package mnormt (Azzalini and Genz 2016), followed estimated marginal ancestral states, and a final simulated stochastic character map from 200 replications using an empirical Bayes method.

To access the accuracy of copy number prediction using ancestral state reconstruction, phylogenetic trees with copy number information were subject to two trials. For the first trial, copy number information was randomly excluded for a fourth of the species in each tree. Using the package castor (Louca and Doebeli 2017), the hidden copy number states were predicted using maximum parsimony. The percentage of incorrectly predicted copies were recorded for 1000 replications. For the second trial, each species on the tree was systematically deprived of copy number information, then had copy number estimated by ancestral state reconstruction. The estimated difference for the species was the absolute difference between true copy number and estimated copy number. The estimated difference was modeled as a function of Phylum with a negative binomial distribution.

Minimum Taxonomic Cutoffs

Each isolate's 16S sequence in the rrnDB was compared against every other isolate's sequence within each taxonomic level from species to domain. The lowest sequence difference value was retained. Isolates that matched a lower taxonomic level were excluded from higher levels. Then, all the lowest sequence differences values across all isolates were averaged

together for each taxonomic level. Those means were added as colored lines to the autocorrelation plot.

Cumulative Distribution Function

To test the generality of the rrnDB results, a cumulative distribution function was used to visualize how closely the Silva database blasted to the rrnDB because the function describes the probability that a variable will take on a value across a whole range. A global search was run for each entry in the SSU Ref 99 132 dataset—a Silva de-replicated amplicon file—using the usearch_global function of USEARCH (Edgar 2010) and the rrnDB 5.4 amplicons as the database. The resulting file had the percent identity of the nearest hit in the rrnDB, which was translated to percent difference. Those values were used to plot an empirical cumulative distribution function.

Results

Autocorrelation

Copy number was positively autocorrelated with sequence identity up to a threshold of 15-30% sequence difference, or a copy number difference of four (Figure 2.2). A Mantel test found copy number difference and sequence percent difference to be significantly positively correlated (nperm=999, z-stat=1791443203, p-value=0.001). The semi-variogram also reported positive autocorrelation (Figure 2.3). Comparisons became uncorrelated at a range of 5.5. The colored lines in Figure 2.2 are the mean minimum percent identity values to be included in the same taxonomic group, hereby called taxa-cutoff values. Taxa-cutoff values became increasingly inaccurate as sequence percent difference increased; for example, the phylum cutoff

was larger than the domain cutoff. The cutoff for order could not be plotted because no prokaryote order in the rrnDB contained more than one family. Thus, the mean minimum distance to be included in the same order was no different from the mean minimum distance to be included in the same family. The cumulative distribution function of Silva entries to the rrnDB showed that over 50% of Silva species blasted to the family level, and almost 90% blasted to the phylum level (Figure 2.4).

The sequences within the spike in the autocorrelation graph had 9-11% sequence difference and greater than 4 copy number difference. This included 72667 pairwise comparisons from 1248 isolates. Plotting the log number of pairwise comparisons within the spike against rank and a histogram of log hits revealed two statistical populations. The population centered around zero were isolates with higher copy number difference than expected that were causing the spike. The second population was species with typical copy number differences that had a pairwise comparison with the first population.

There were 9 genera in the top 100 isolates with the highest contributions to the spike. Four of them were from candidate phyla (*Candidatus Hoeglandella*, *Erwinia, Moranella,* and *Gullanella*) with a copy number of 1. The other 5 genera were from Firmicutes (*Bacillus, Lysinibacillus, Listeria, Staphylococcus,* and *Macrococcus*), a phylum known for high average copy number (9.72, 10.25, 6, 5.14, and 4.67 for each genus respectively). The majority of the next several hundred isolates were also Firmicutes. The Firmicutes tree (Figure 2.5) shows that copy number is highly variable across the whole tree. There was a significant phylogenetic signal for copy number (λ=0.982, p-value < 0.0001). Copy number is diverse across Firmicutes but is still under selection. Firmicutes was responsible for the spike in the autocorrelation plot.

Phylogenetic Signal

Positive phylogenetic signal occurs when species in a phylogenetic tree cluster by a certain trait; related species are more similar to a randomly selected species because the trait is under selection (Münkemüller et al., 2012). Copy number had a significant phylogenetic signal in 17 of the 22 phylogenetic trees, non-significance in 2, and no estimation in 3 (Table 2.2). Tenericutes and Deltaproteobacteria had issues with their tree structure while Gammaproteobacteria was too large to analyze. Figures 2.6 and 2.7 show representative phyla with significant and non-significant phylogenetic signal. Copy number is indicated by coloration of the tip label and is clustered in the significant tree and randomly distributed in the non-significant tree.

The archaeal tree from the Silva sequences showed copy number clustered by group and changing in a stepwise manner (Figure 2.8). Single copy lineages appear the most conserved (green). The highest copy numbers (purple) only appear in clades with species of 2-3 copy number. Strikingly, copy number variation only appears in one phylum, Euryarchaeota. The other three smaller phyla (Crenarchaeota, Candidatus Korarchaeum, and Thaumarchaeota) all have a single copy of the operon.

Ancestral State Reconstruction

Ancestral state reconstruction was partially successful for the representative phylum Verrucomicrobia tree (Figure 2.9). Each node has an ancestral copy number state probability in pie-chart format. Across phyla, the success rate of ancestral state reconstruction in trees with 25% dropped branch information ranged from 35-80% depending on the phyla (Figure 2.10).

When systematically dropping copy number from each branch, there was a significant difference in copy number estimation error ($F_{20,8131}$=22.90, p-value<0.0001, Person Chi-Square/DF=1.15) and again estimation was more successful for some phyla than others (Figure 2.11). Some taxa had a low estimated difference—Chlamydia, Tennericutes, Thermotagae, Chloroflexi, and Archaea—which makes estimation of copy number reliable. Others, such as Verrucomicrobia, Firmicutes, Bacteriodetes, and Planctomycetes, have much more variable differences. There was variability within the sub-phyla of Proteobacteria; Beta and Epsilon-proteobacteria both had high accuracy, while Delta-proteobacteria had low accuracy.

Discussion

The best resource for copy number information is the rrnDB. The current version has copy number information for 3484 bacterial species and 220 archaeal species (version 5.5). There are also several software tools to predict or correct for copy number. Notable ones are PICRUSt (Langille et al., 2013), CopyRighter (Angly et al., 2014), 16Stimator (Perisin, Vetter, Gilbert, & Bergelson, 2016), and UNBIAS (Edgar, 2017). PICRUSt (the Phylogenetic Investigation of Communities by Reconstruction of Unobserved States) predicts copy number with ancestral-state reconstruction using a specific weighting method. CopyRighter uses copy number estimates to correct microbial profiles. 16Stimator estimates copy number from draft genomes, which can be used to supplement a reference database. UNBIAS uses an algorithm which estimates copy number by finding the top match to a reference database and then divides the sequence count by that number.

Copy number can be determined by two major molecular methods. The first is genome annotation. If the whole genome is sequenced, copies can be found using 16Stimator or other gene prediction software. The second method is qPCR amplification with a reference gene. One approach is using a single copy reference gene within the same organism being examined (Zhang et al. 2009). Another approach is to use the copy number of a reference species, such as *E. coli* for bacteria, alongside the copy number of the organism being examined (Lee et al. 2008). Both use real-time PCR to amplify the rRNA gene fragments. From there, copy number can be determined by the ratio of amplification between the reference and 16S gene.

Adding copy number information has been shown to improve metagenomic estimates. Kembel et al (2012) used reference data and ancestral state reconstruction to estimate copy number and it improved the accuracy of both microbial community diversity and abundance. However, there has been limited success with correctly estimating copy number. Edgar found that the UNBIAS algorithm had poor correlations between true counts and the algorithm's estimation. Louca, Doebeli, and Parfrey's (2018) results—using PICRUSt, CopyRighter, and PAPRICA (Bowman and Duckow 2015, a pipeline that used PICRUSt to ultimately predict metabolism from phylogeny)—on test datasets found sequence identity was only able to predict copy number accurately for taxa closely related to known taxa. They also ran autocorrelation between phylogenetic distance and copy number difference and found that correlation was inaccurate after 15% divergence and random at 30%. One paper—on *Drosophila* gut microbiota—has already taken their advice and ignored copy number correction in their pipeline (Erkosar et al., 2018). Another paper—on pollen DNA metabarcoding—acknowledged the bias of copy number but did nothing to correct for it (Bell et al., 2018).

Our examination of copy number estimation did not use any software, but rather manually examined estimation using ancestral state reconstruction after removing copy number information in two different ways. Our results suggest that accurate estimation is possible for some prokaryote phyla but not others. This is more nuanced conclusion than the general summary by Edgar and Louca that copy number estimation is unreliable.

Despite the difficulty of accurate copy number prediction, correcting for copy number is still necessary for quantitative counts of metagenomics samples. Typical metagenomics analysis separates sequence reads into operational taxonomic units (OTUs) and estimates diversity. OTUs are then linked to known taxonomic categories. Those proportions are not accurate without copy number correction.

Using data from the rrnDB, copy number was found to be heritable and evolving in a stepwise manner. Copy number was visibly clustered on most copy-number-colored phylogenetic trees. Both the autocorrelation and the semi-variogram plots gave similar results, with a significant relationship between copy number difference and sequence identity.

The value at which correlation disappeared for the semi-variogram was larger than expected based on the autocorrelation plot plateau. But it matched 30% sequence dissimilarity, the point where Louca, Doebeli, and Parfrey found autocorrelation to disappear. Given that the semi-variogram nugget was much higher than zero, there was a nugget effect. This means there were sources of variation at smaller distances than the sampling interval. This may be an artifact of copy number being a discrete variable, while distance is continuous.

In the autocorrelation graph, there was a large and unexplained spike at 10% species percent difference. The spike is not an error; it is due to real variation found in nature. It

appeared to be driven by the phylum Firmicutes, which are known for fast growth rate and high copy number. Thus, the pairwise comparison between a Firmicutes isolate and an average bacterial isolate has a much higher copy number difference than expected by phylogenetic distance. The evolution of copy number is not the same across phyla and ought to be considered on an individual phylum basis.

Most phyla had significant phylogenetic signal, showing copy number was under selection. Non-significant phyla were small ($n<25$) and dispersed. Small sampling size may be the driving force behind the non-significance in Planctomycetes and high evolutionary rates in Fusiobacteria. While the Gammaproteobacteria, Deltaproteobacteria, and Tenericutes could not be analyzed, visual inspection of their colored trees suggests that copy number is under selection in Gammaproteobacteria and Tenicutes but not for Deltaproteobacteria. Overall, copy number is heritable and increases under stepwise selection for most bacteria phyla.

In both the autocorrelation and semi-variogram, copy number difference and sequence difference became decoupled at significant evolutionary distances. These significant distances are what make copy number estimation difficult for isolates that have no closely related cultured representatives. Using the taxa-cutoff values as a guide, autocorrelation becomes unreliable outside the class cutoff.

The taxa-cutoff values also give a more quantitative answer to define taxonomic levels in prokaryotes. It may be ontological impossible to distinguish some down to species levels in all cases, but finding the limits of detection are of pragmatic use.

Predicting copy number using ancestral state reconstruction at the phylum level was marginally successful. For some—such as Archaea, Spirochaetes, Chloroflexi, and

Thermotogae—the chance of getting an incorrect prediction had a reasonable error rate of 20%. Others—such as Planctomycetes and Verrucomicrobia—had only a 50% chance or worse to be assigned the correct copy number, which does not justify trying to correct for copy number rather than leaving the abundance count alone. Archaea despite being domain level, had one of the lowest mean error rates, so copy number prediction was very reliable. It is also not unexpected, as Archaeal copy number is known to be much more constrained and only species found in Phylum Euryarchaeota have copy number greater than one.

Interestingly, the Small category, which consisted of fifteen phyla with ten or fewer representatives, had about a 60% chance of correct copy number prediction, or on average a deviation from true copy number of 0.55. Ancestral state reconstruction was more likely than not to produce the right copy number precisely on the types of candidates that are likely to have few close relatives. More work is required to authenticate this, but it may be possible to use ancestral state reconstruction to get a better-than-random estimate of copy number for disparate isolates.

Since most Silva records were blasted to a member of the rrnDB at phylum or lower, the cumulative distribution function provides support for scalability and reusability. Similar copy number trends are expected in more general datasets. This is important, as inferences with other data depend on the rrnDB being representative.

Conclusion

We found that rRNA copy number is phylogenetically autocorrelated with sequence identity, falling off at 15-20% dissimilarity. Positive phylogenetic signal by phylum proved that most lineages have conserved copy number evolving under selection. Ancestral state

reconstruction showed that copy number is more accurately estimated for some phyla than others. This points out the need to approach copy number estimation in a taxonomic-specific way. This research aids in the continued need for accurate copy number estimation and a best practices pipeline.

We found taxa-cutoff values, which indicated how closely species' 16S sequence needed to be included in the same taxa. While doing so, we found an issue with single-family orders. An interesting future effort would be to determine how widespread the problem of single-family orders in prokaryotic taxonomy. More broadly, taxa-cutoff values are an important step for distinguishing taxonomic levels by sequence identity.

Taken together, the results show that 16S rRNA copy number variation is a plausible genetic mechanism for the GRH in prokaryotes. The next step is to determine if high copy number species respond more readily to nutrient enrichment than lower copy number ones.

Best Practices Pipeline

Correcting for copy number is necessary for making metagenomics quantitative. As more copy number information is generated for environmental samples, theoretically, copy number information should eventually become available for all isolates. But presently, we need a more pragmatic approach. Therefore, we propose a pipeline for copy-number-corrected amplicon-sequencing (Figure 2.12). This is in response to some general attitudes that copy number variation should be ignored due to the difficulty of accurate copy number estimation for environmental samples. We agree that it is difficult but believe a thoughtful effort that considers the nature of copy number across phyla can achieve useful copy number corrections for metagenomics samples.

Figures and Tables

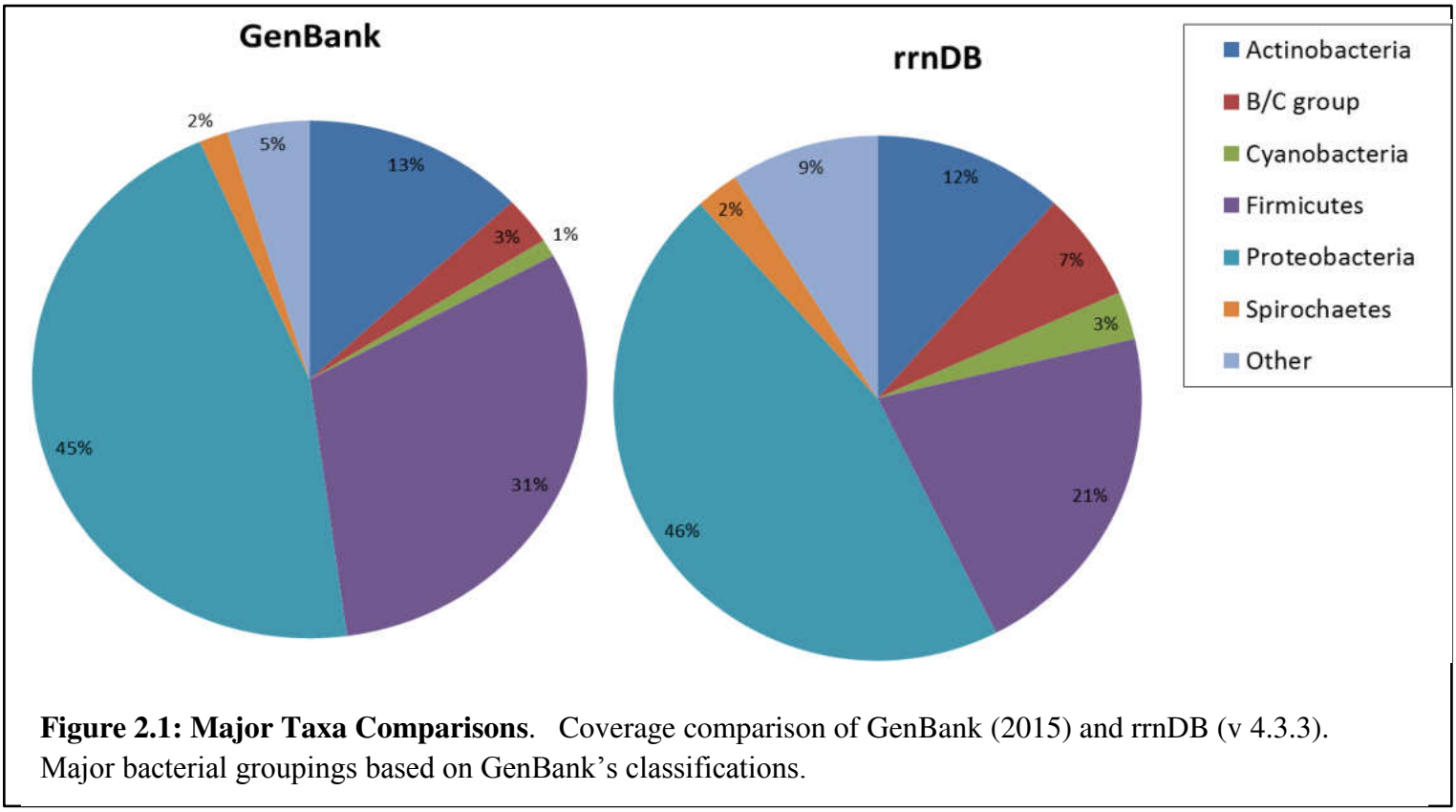

Figure 2.1: Major Taxa Comparisons. Coverage comparison of GenBank (2015) and rrnDB (v 4.3.3). Major bacterial groupings based on GenBank's classifications.

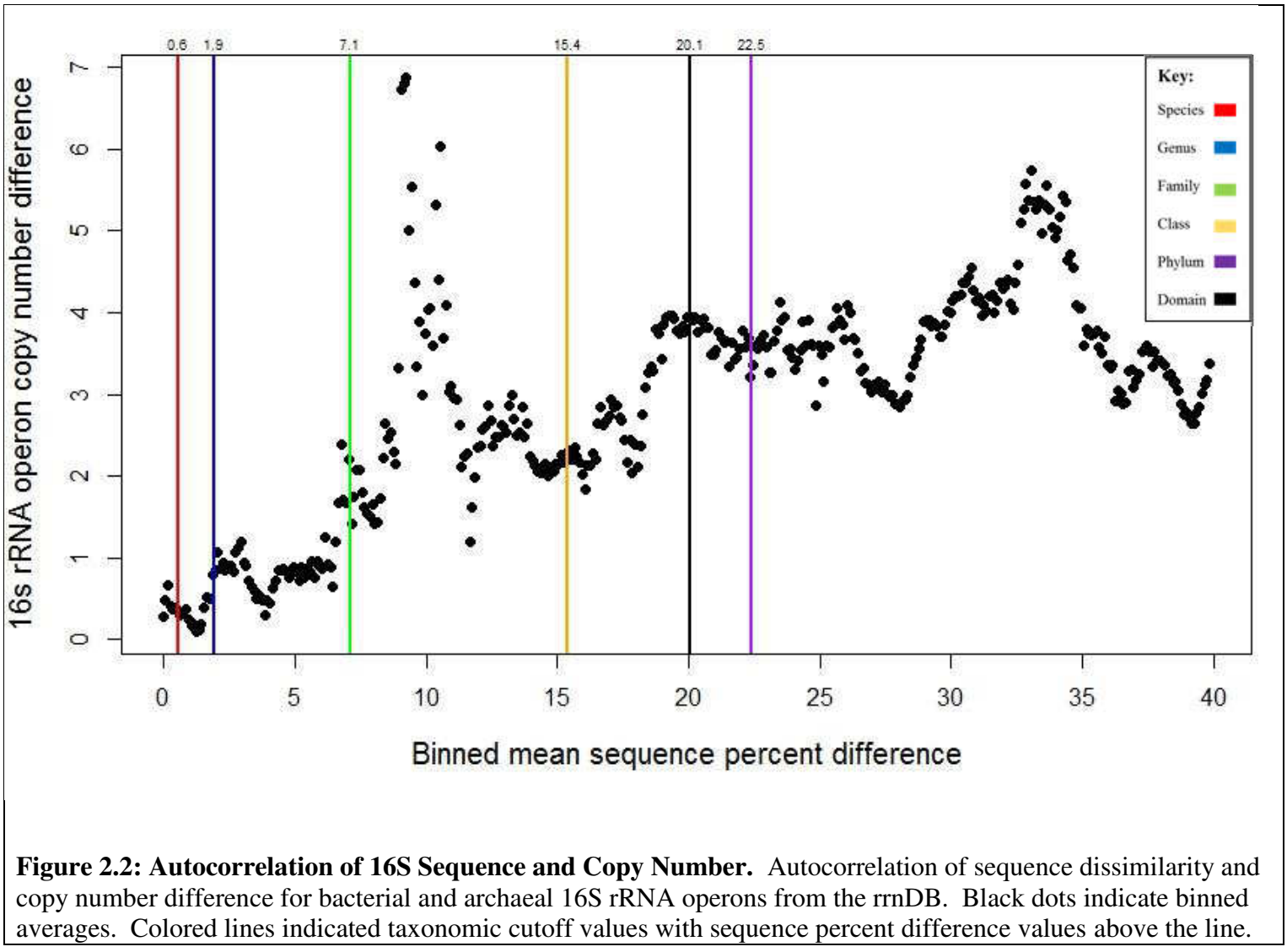

Figure 2.2: Autocorrelation of 16S Sequence and Copy Number. Autocorrelation of sequence dissimilarity and copy number difference for bacterial and archaeal 16S rRNA operons from the rrnDB. Black dots indicate binned averages. Colored lines indicated taxonomic cutoff values with sequence percent difference values above the line.

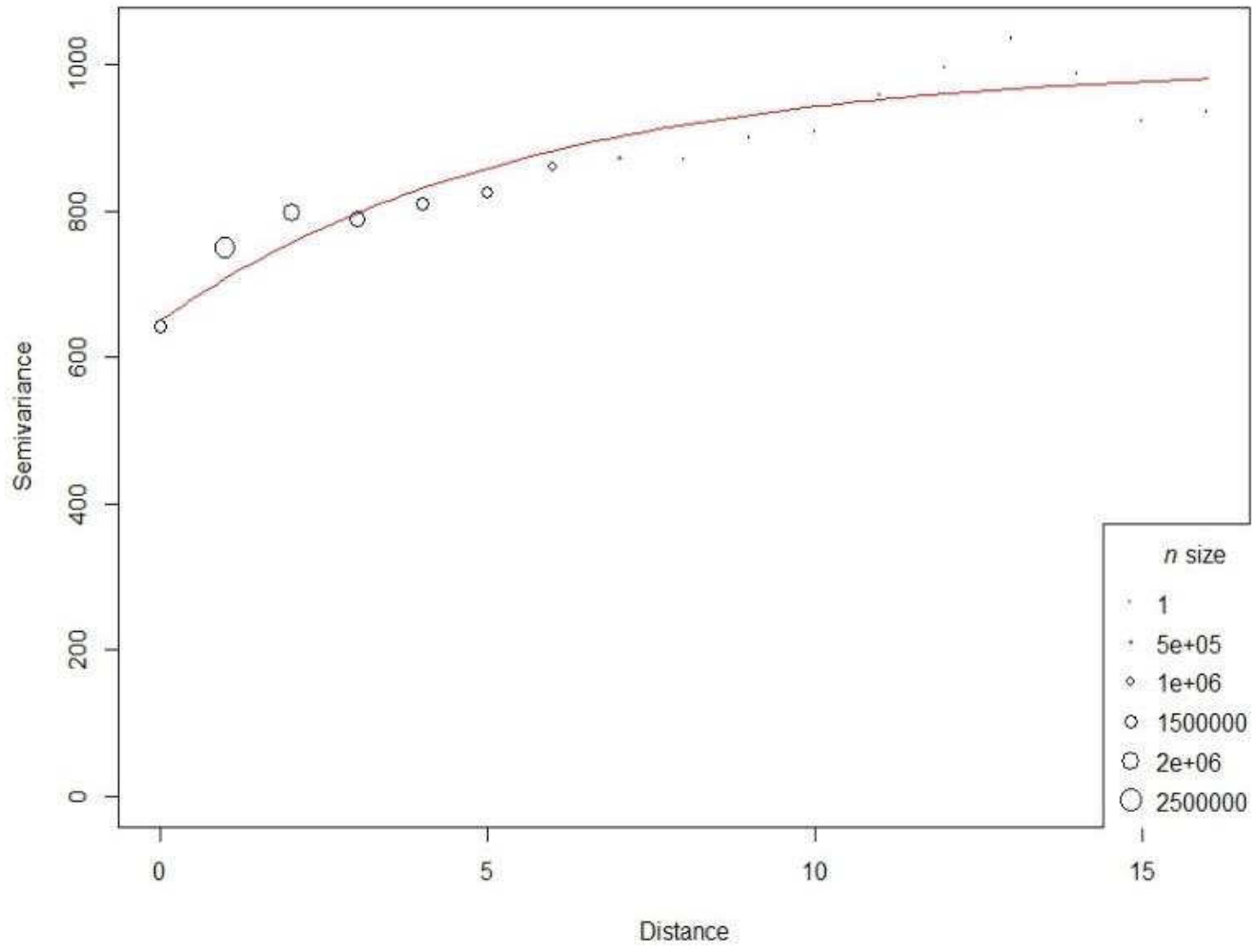

Figure 2.3: Semi-variance of 18S Sequence and Copy Number. Semi-variogram of pairwise identity difference and copy number difference (exponential model, mean=874.4, R squared=0.8591, lag=1). Sill was set at 1000, range at 5.5, and nugget at 650.

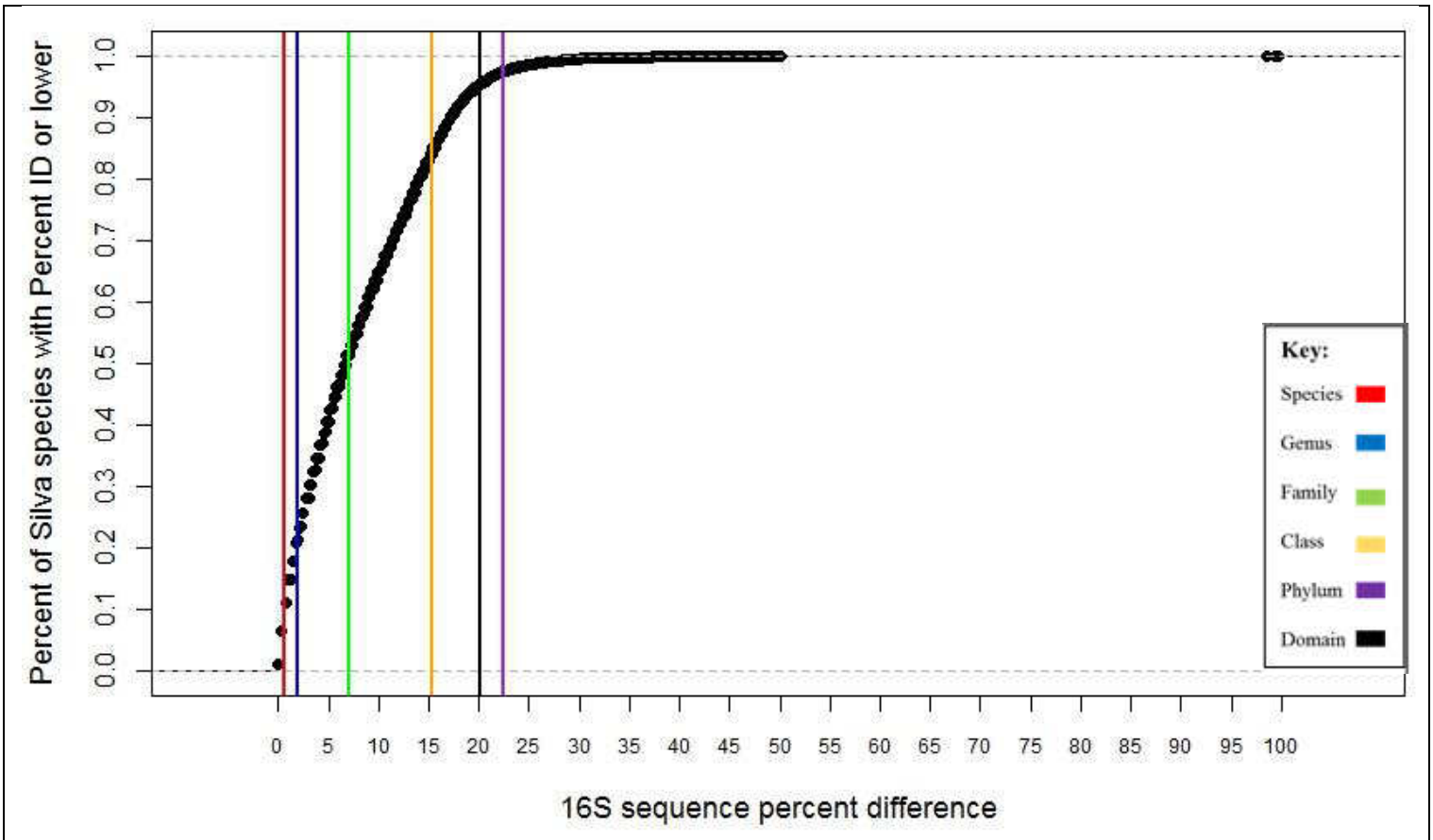

Figure 2.4: Cumulative Distribution Function. Black dots represent the sequence dissimilarity of Silva prokaryotic 16S sequences compared to sequences found in the rrnDB. Colored lines indicated taxonomic cutoff values.

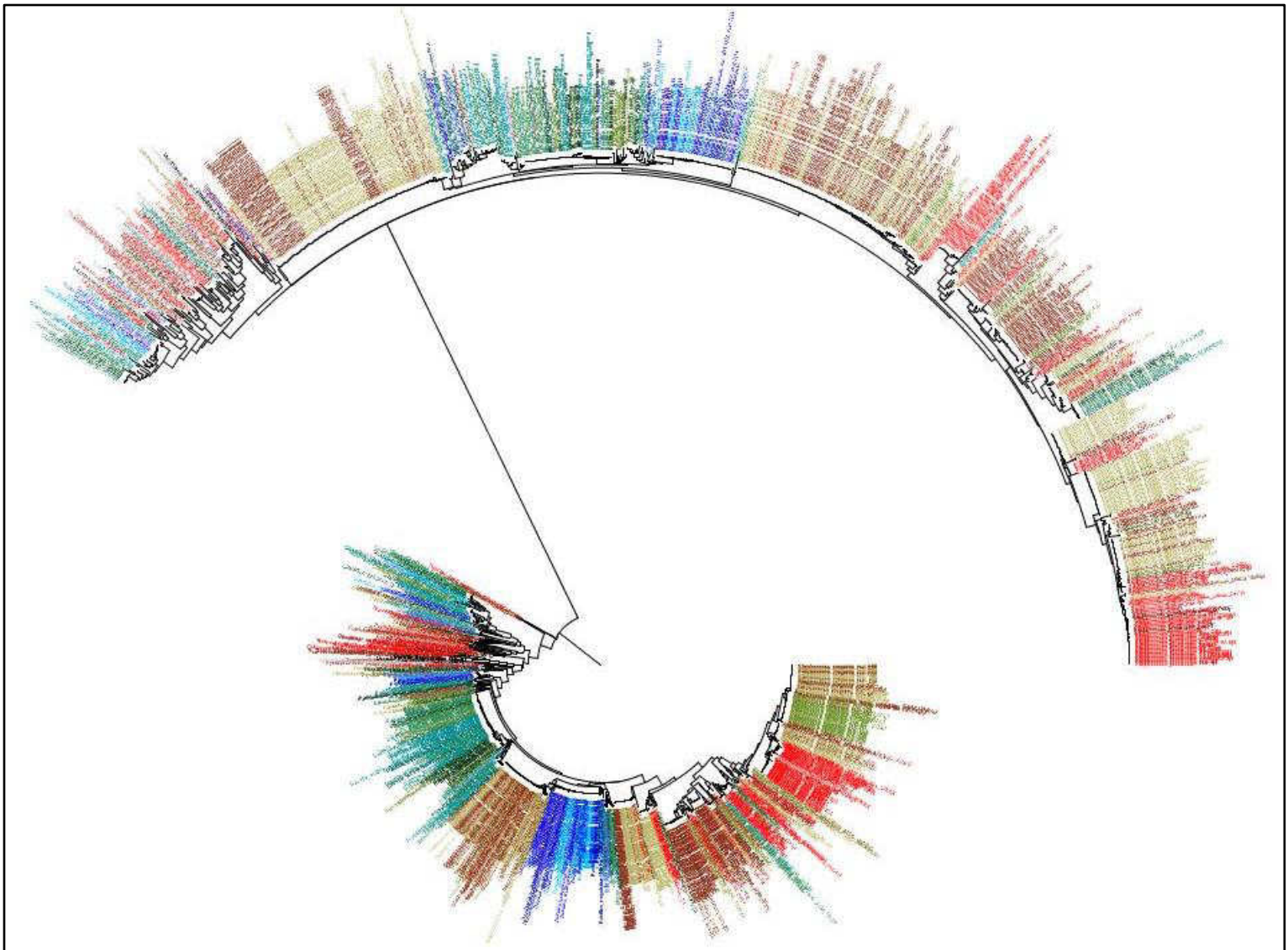

Figure 2.5: Firmicutes Copy Number Tree. Phylogenetic tree of Firmicutes colored by copy number. Phylogenetic signal had significant lambda (0.982, with p-value < 0.0001). Figure made in FigTree version 1.4.2.

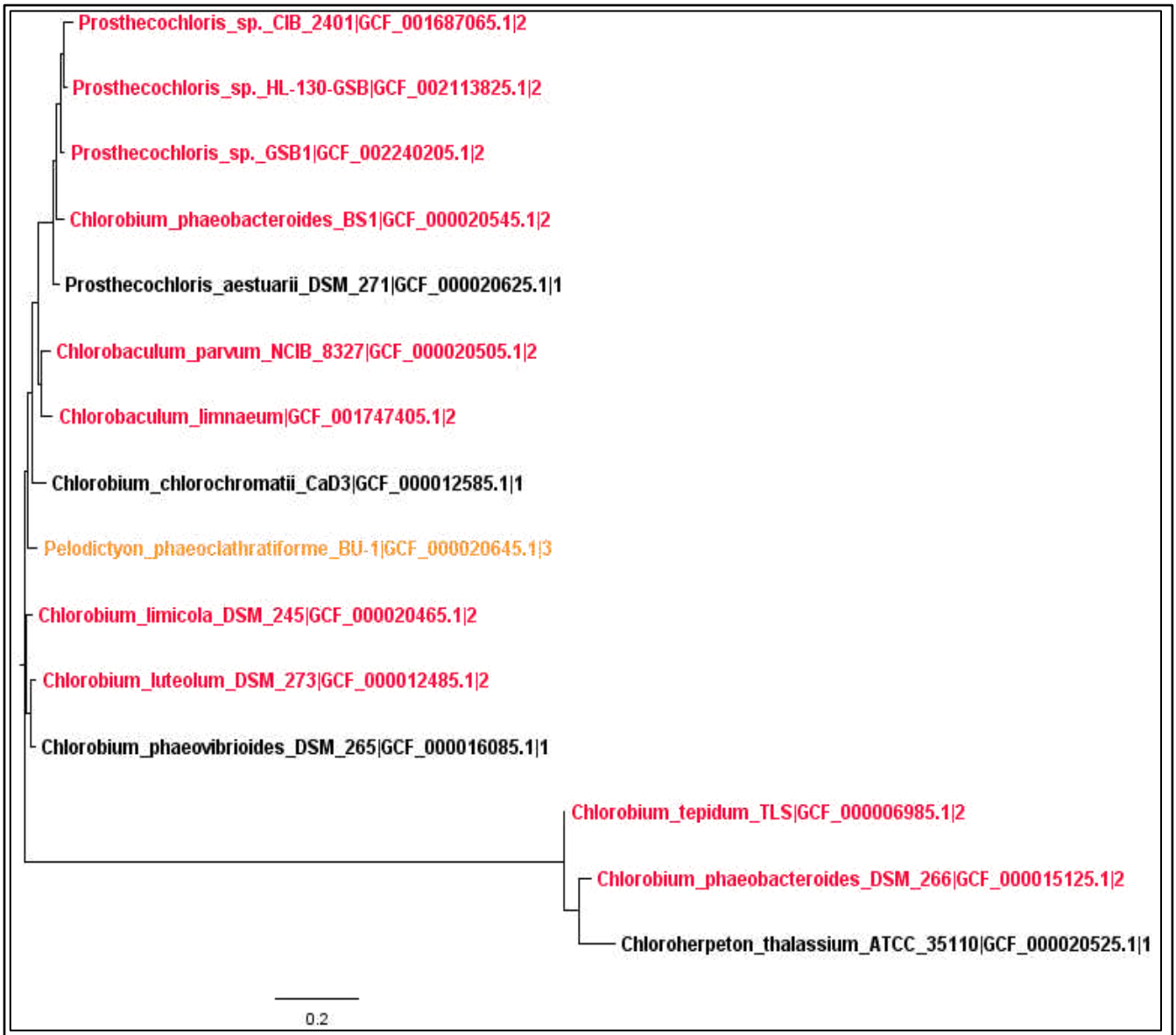

Figure 2.6: Example of Significant Phylogenetic Signal. Phylogenetic tree of Chloroflexi colored by copy number with a significant lambda value (0.9999, with p-value <0.0001). Figure made in FigTree version 1.4.2.

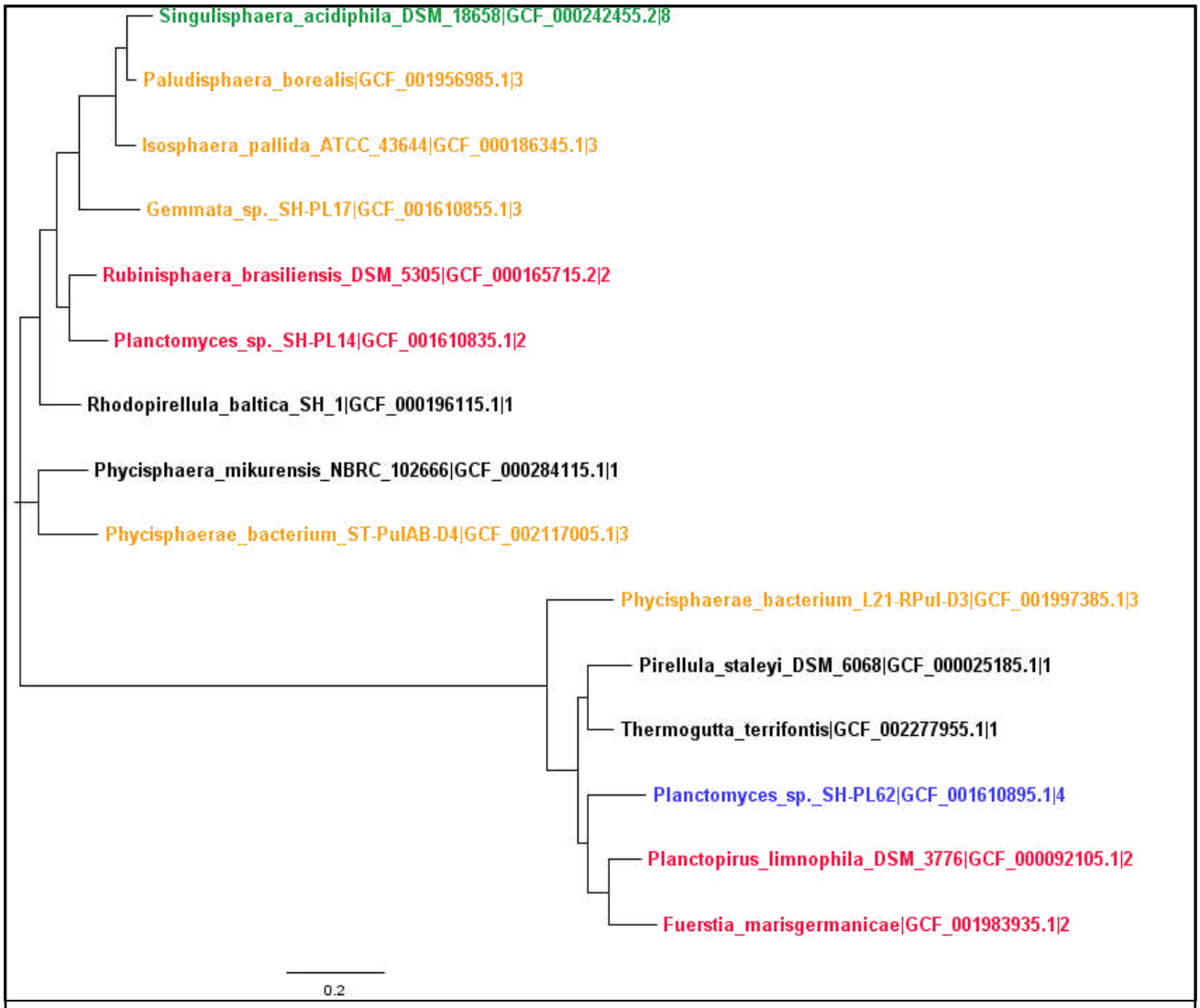

Figure 2.7: Example of Non-significant Phylogenetic Signal. Phylogenetic tree of Planctomycetes colored by copy number with a non-significant lambda value (0.3976, with p-value = 0.6346). Figure made in FigTree version 1.4.2

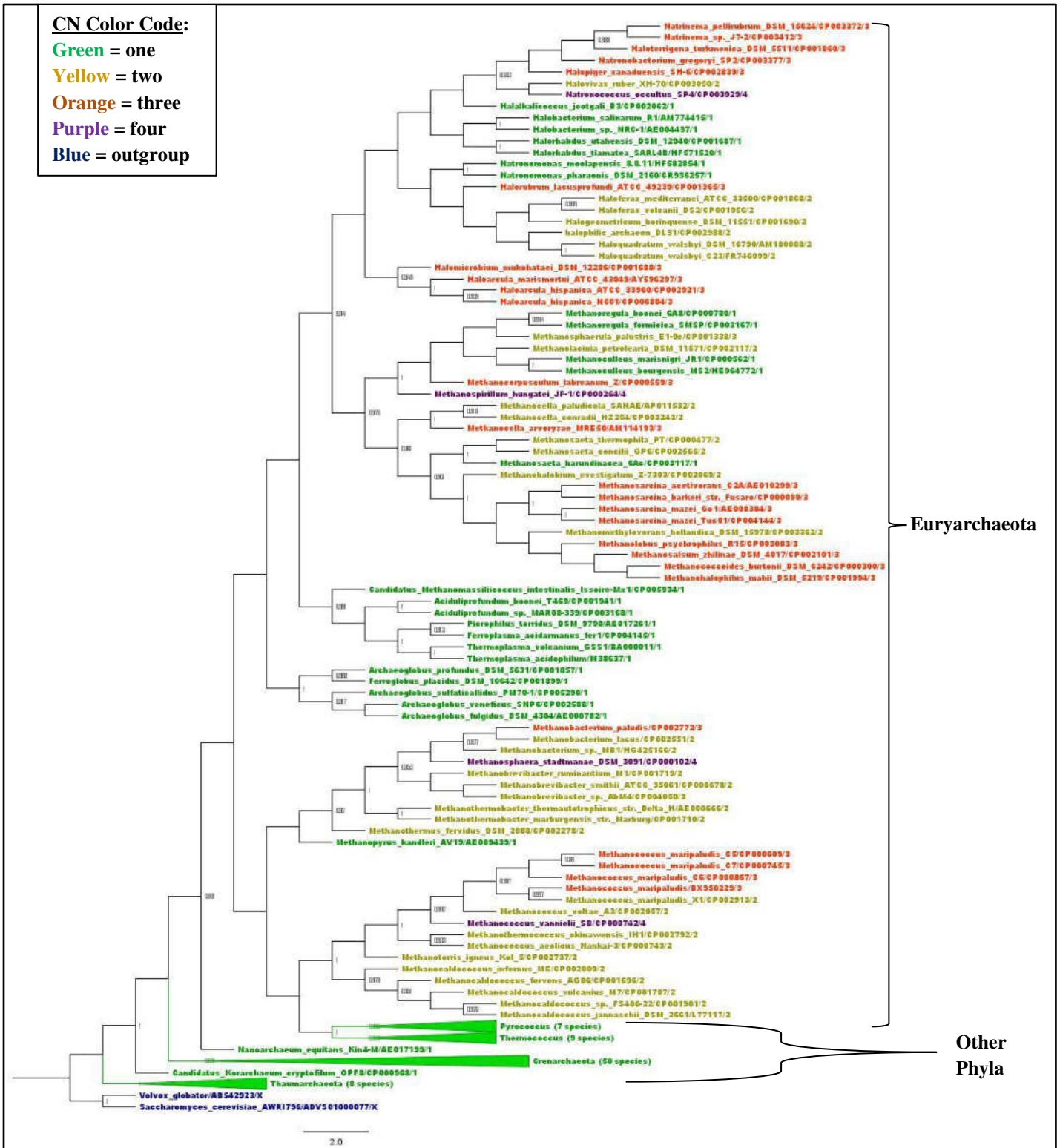

Figure 2.8: Silva Archaea Copy Number Tree. Phylogenetic tree of archaeal 16S sequences from the Silva database. Color code based on copy number. Copy number variation occurred only in the Euryarchaeota phylum. All other species in other phyla had a copy number of one. Green algae and yeast

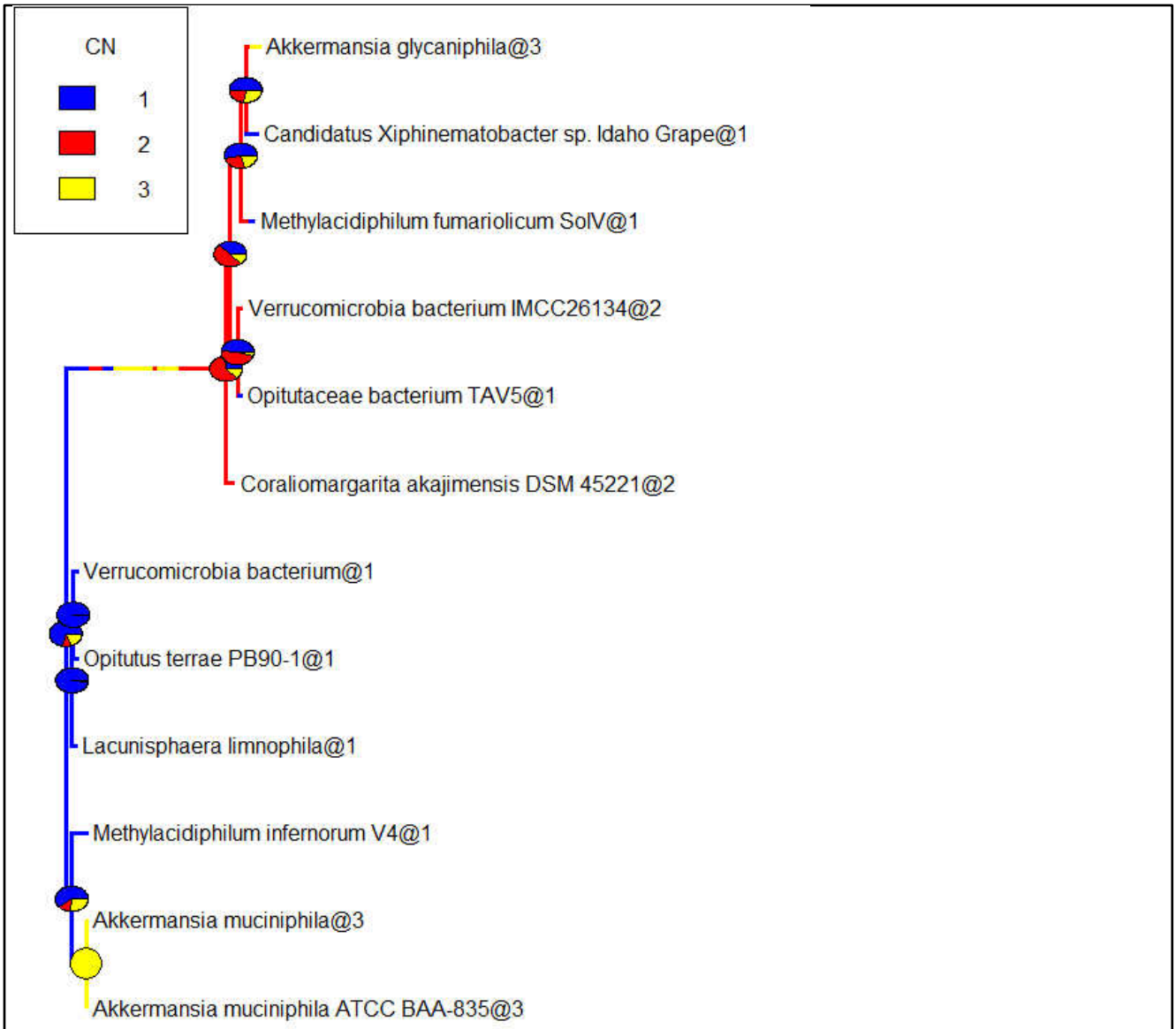

Figure 2.9: Ancestral State Reconstruction Example. Stochastic character map tree for the phylum Verrucomicrobia. State estimated using empirical Bayes method with 200 simulations under the ER model.

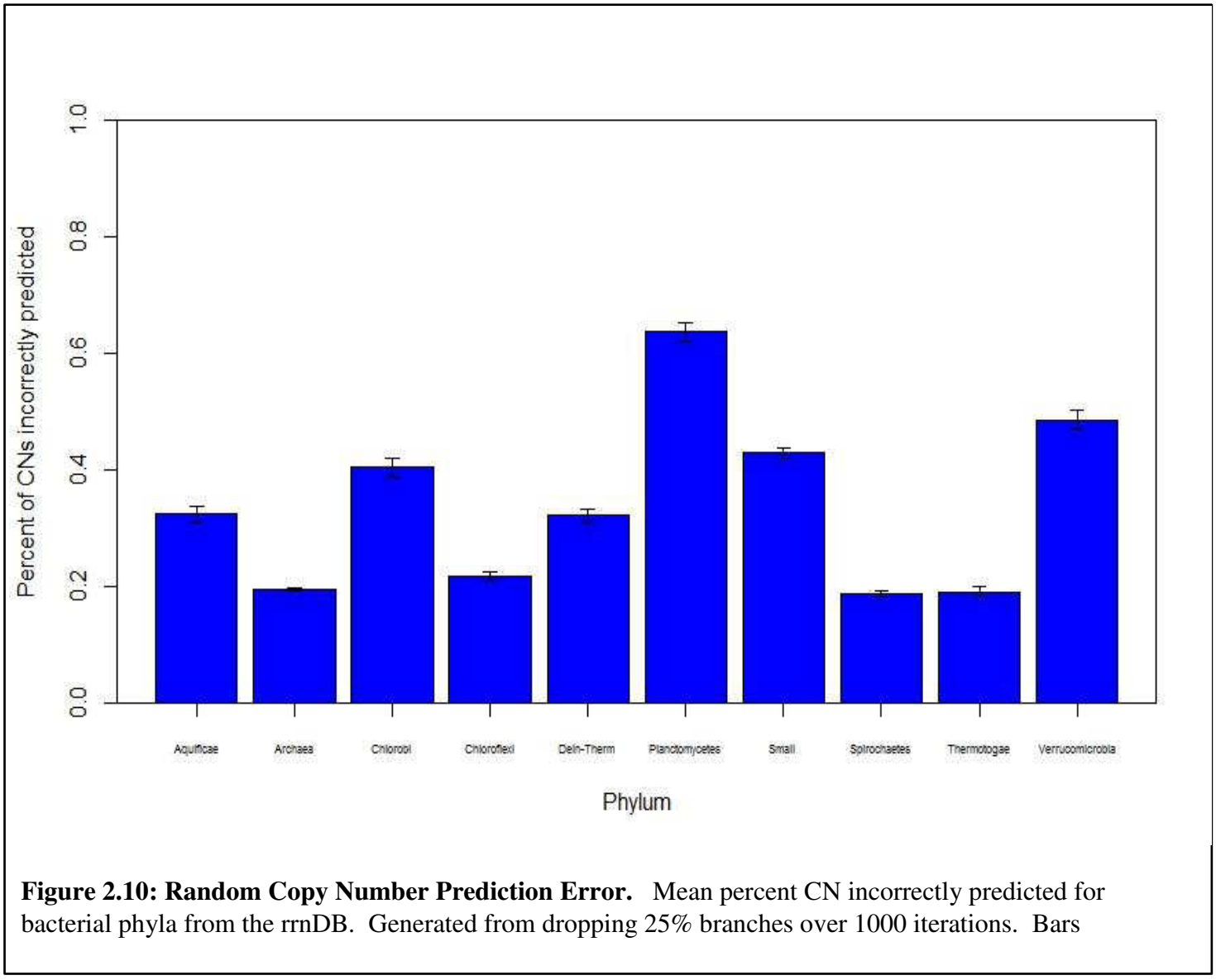

Figure 2.10: Random Copy Number Prediction Error. Mean percent CN incorrectly predicted for bacterial phyla from the rrnDB. Generated from dropping 25% branches over 1000 iterations. Bars

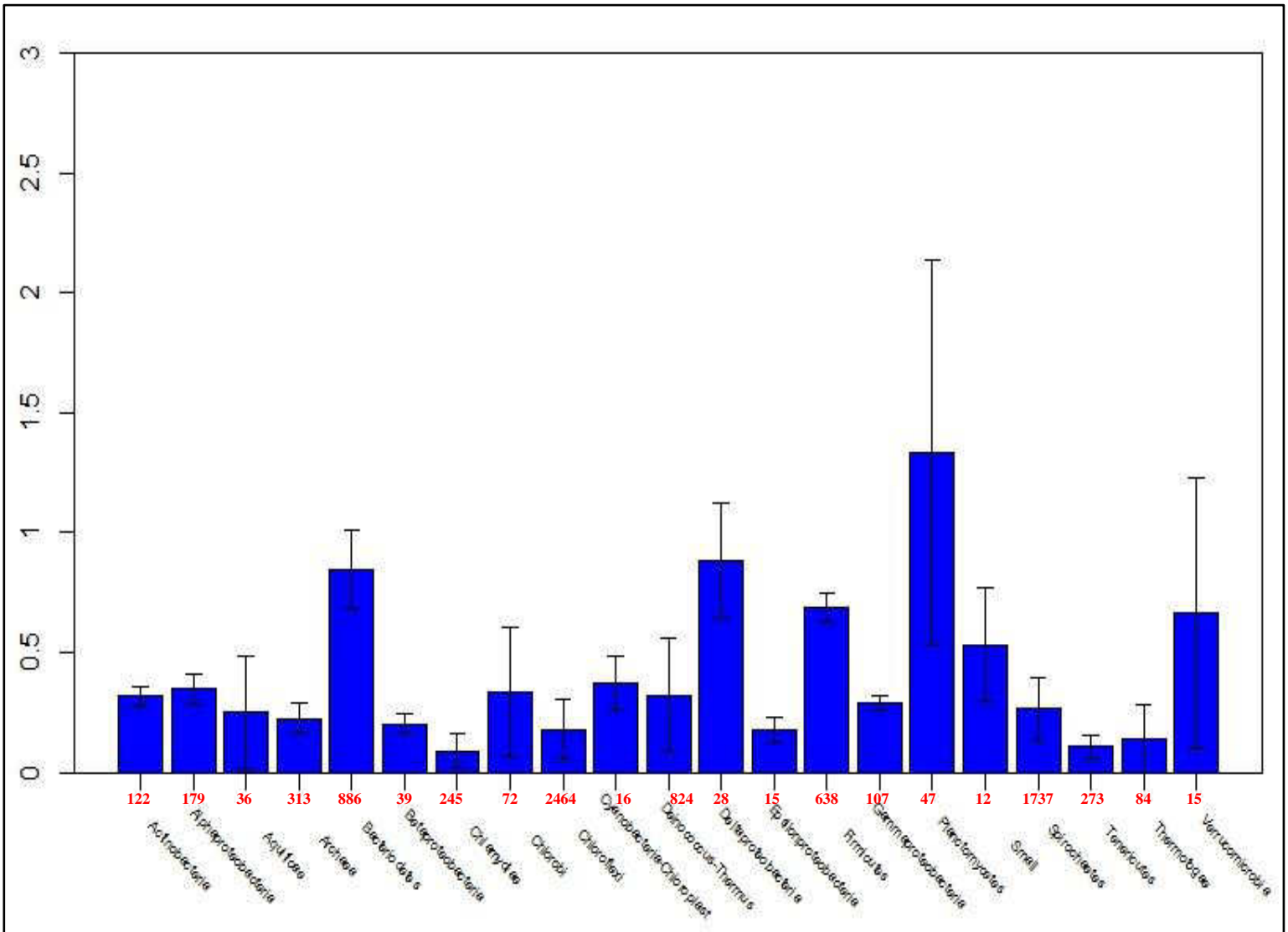

Figure 2.11: Systematic Copy Number Prediction Error. Differences between true copy number and ancestral state reconstruction estimated copy number by phylum. Generated by systematic dropping of copy number from each branch of the phylum tree. Diamond represents the mean, the bar represents the median, and open black circles are outliers. Red numbers are the number of species in each phylum.

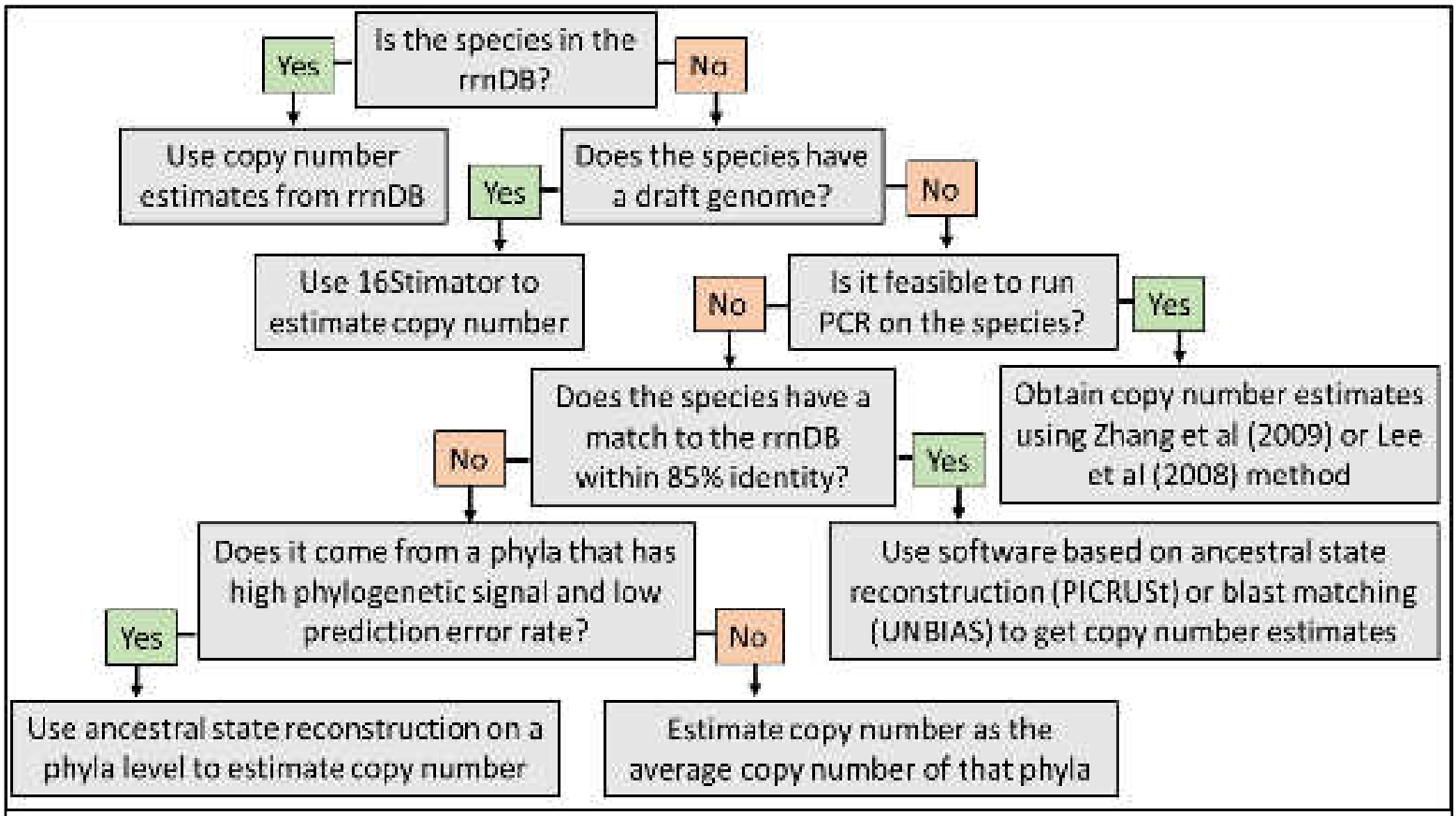

Figure 2.12: Copy number correction best practices pipeline.

Table 2.1: Minimum Taxonomic Cutoff Values

Taxa	Mean distance (% dissimilarity)
Species	**0.57**
Genus	**1.93**
Family	**7.12**
Class	**15.40**
Phylum	**22.38**
Domain	**20.08**

Table 2.2: Phylogenetic Signal Values for Prokaryote Phyla

Phyla	Isolate Number	Mean Copy number	Lambda Value	P-value
Actinobacteria	824	3.2	0.9948	>0.0001
Bacteroidetes	273	3.7	0.8919	>0.0001
Tenericutes	179	1.5	n/a	n/a
Chlamydiae	122	1.7	0.9279	>0.0001
Cyanobacteria_Chloroplast	107	2.6	0.9999	>0.0001
Spirochaetes	72	1.6	0.9999	>0.0001
Chloroflexi	39	1.3	0.9999	>0.0001
Thermotogae	36	1.8	0.9999	>0.0001
Deinococcus-Thermus	28	2.6	0.9999	>0.0001
Fusobacteria	24	4.7	0.7000	0.4574
Aquificae	16	2.1	0.9999	>0.0001
Chlorobi	15	1.8	0.9999	>0.0001
Planctomycetes	15	2.4	0.3976	0.6346
Verrucomicrobia	12	1.6	0.9999	0.0005
Other (phyla =15)	47	1.8	0.9999	>0.0001
Archaea	245	1.7	0.9999	>0.0001
Firmicutes	1737	6.9	0.982	>0.0001
Alpha Proteobacteria	638	2.6	0.998	>0.0001
Beta Proteobacteria	886	3.6	0.999	>0.0001
Gamma Proteobacteria	2464	6.1	n/a	n/a
Delta Proteobacteria	84	2.8	n/a	n/a
Epsilon Proteobateria	313	2.3	0.837	>0.0001

Chapter 3: Oakville Prairie Soil Core Studies

Abstract

This chapter tests the hypothesis that nutrient-enrichment benefits soil microbial and microinvertebrate species with high rRNA operon copy number more so than species with low copy number. This would support the overall theory that copy number variation is the genetic mechanism for the Growth Rate Hypothesis. Soil cores were collected from Oakville Prairie during two different years and enriched with nitrogen, phosphorus, both nitrogen and phosphorus, and a control. Nematodes and prokaryotes were enumerated by high-throughput sequencing of a portion of the rRNA operon. For prokaryotes, we found that nutrient enrichment increased the representation of high-copy number prokaryotes in the mesocosm experiment conducted after the fall of 2017, but not in the mesocosm experiment conducted after the spring of 2016. However, for nematodes, we found that phosphorus enrichment increased the representation of enrichment type microbivores in the mesocosm experiment conducted after the spring of 2016, while nitrogen increased them in the mesocosm experiment conducted after the fall of 2017. Nutrient enrichment affected soil communities but differed across seasons.

Introduction

Background

Nutrient enrichment affects soil microbial community diversity and abundance. Enrichment often results in the rapid growth of one or a few specific species, such as an algal bloom in a lake. These specific species are fast-growing to take advantage of the acute nutrient source. In the soil, fast-growing bacteria take advantage of short-term nutrients over the slower-growing fungal pathways, while enrichment-type bacterivore nematodes characterize enriched soils over slower-growing plant parasites and predators. Long-term addition of low rates of

organic carbon increases prokaryotic but not fungal biomass (Marschner, Kandeler, & Marschner, 2003). High nitrogen increases copeotrophic taxa and decreases oligotrophic taxa (Fierer et al., 2012), but there is evidence that community composition can resist changes in nutrient enrichment (Bowen et al., 2011). Fast-growing bacterivore nematodes are indicative of enrichment. In contrast, slow-growing and dauer-forming species are more robust against poor soil conditions. Nitrogen enrichment resulted in greater abundance but lower diversity (Shaw et al, 2019), while phosphorus enrichment resulted in lower richness and abundance of omnivores and predators (Yang et al., 2019) in subalpine forests. Nematode maturity index values were lower in nitrogen but not phosphorus-enriched pasture plots (Ghani et al, 2001). Phosphorus in conjunction with tree type and optimum water resulted in higher total nematode density but otherwise did not affect nematode communities (Olatunji et al., 2019).

The Growth Rate Hypothesis (GRH) suggests that there is a connection between the C:N:P stoichiometry of organisms and their growth rate (Elser, Sterner, et al. 2000). High growth rate is correlated with high ribosome number and limited by phosphorus (P) availability since RNA is P-rich and rRNA is used to transcribe mRNA into peptides (Elser, Obrian, et al, 2000). To date, the GRH has been examined almost exclusively in aquatic systems. Thirty-five to sixty-nine percent of total body P was associated with nucleic acids in three crustacean zooplankton species (Vred et al. 1998). Growth rates were positively correlated with percent P in five Daphnia species (Main et al. 1997), and again in two Daphnia species (Archarya et al. 2004). Elser et al. (2003) also showed a positive correlation between P-content and growth rate for several species of microbes, insects, and crustaceans.

The exact genetic mechanism of the GRH is not clear, but ribosomal RNA (rRNA) copy number variation is one possible mechanism because rRNA is rich in phosphorus and rRNA

levels fluctuate far more than DNA. Sequences are grouped together into operons that are found in tandem or dispersed repeats across the genome. Finally, the number of repeats is variable across species, so species may have higher or lower copy number depending on growth rate.

Ribosomal RNA copy number is related to growth rate. Overall, copy number predicts growth rate and growth efficiency in bacteria (Roller, Stoddard, and Schmidt, 2015; Kurm et al, 2017). Low copy number species are assumed to be oligotrophic—adapted to maximizing resources in low-nutrient environments—while higher copy number species are copiotrophic—adapted to locating nutrient spikes (Větrovský and Baldrian, 2013). In specific lineages, *Escherichia coli* needs all seven copies active for rapid adaptation to new conditions (Condon, Liveris, Squires, Schwartz, & Squires, 1995) and *Mycobacteria* grows slower with one natural copy than theoretically modeled *E. coli* with one or two operons (Cox, 2014). However, the relationship between copy number and growth is non-linear and asymptotic (Samad et al, 2017). Strong rRNA operon promotors are needed in addition to higher copy number for high growth rate (Aiyar, Gaal, & Gourse, 2002) while genome size does not appear significantly related to copy number (Espejo and Plaza 2018) but may be related to log copy number (Stoddard et al. 2016). Copy number appears conserved within species and even genera (Rastogi et al., 2009). Phylogenetic work in Chapter 2 confirmed this to the phylum level. Sequence diversity in copies increases with increasing copy number (Větrovský and Baldrian, 2013).

Copy number in nematodes is much higher and similarly variable. However, unlike prokaryotes that have the Ribosomal RNA Operon Copy Number Database (Stoddard et al., 2015), only a few nematode species have copy number estimations.

The few studies that have looked at nematode copy number found a high variability (Long & Dawid, 1980; Sulston & Brenner, 1974; Ellis, Sulston, & Coulson, 1986; Bik et al., 2013; Darby et al., 2013; Pereira & Baldwin 2016), up to a magnitude of difference.

Given the diversity of soil communities and the effect of nutrients on community composition, soil cores are ideal experimental units to test the Growth Rate Hypothesis. Intact soil-core mesocosm experiments leverage both the natural soil community environment and control for random effects. Soil cores are commonly used to explore soil communities, including the environmental impact of genetically modified organisms (Bentjen et al 1989), survival and dispersal of biocontrol agents (Longa and Pertot 2009), fates of microorganisms by fermentation class (Providenti et al 2004), phosphate transport by fungi (Jakobsen et al 2001), variability of chemical fluxes (Yao et al 2010), root microbe differences across sites (Gagliardi et al 2001), and a variety of fertilizer effects, including natural versus synthetic (Brock et al 2007; Liang et al 2013; Dodd et al 2014).

Understanding copy number is not only important for exploring the GRH. The 16S and 18S rRNA operons are used regularly for amplicon sequencing of environmental samples. As the goal of most metagenomic studies is to understand the abundance and distribution of the microbiome, accessing and correcting for copy number is vital for making metagenomics studies quantitative.

Objectives

The objective of this chapter is to test the hypothesis that nutrient-enrichment benefits prokaryote species with high rRNA operon copy number and enrichment-type bacterivore nematodes more so than species with low copy number or non-enrichment-type. To accomplish this, I took soil cores from Oakville Prairie, treated them with excess nitrogen, phosphorous, or both, enumerated prokaryotes and nematodes with high-throughput rRNA amplicon sequencing, and analyzed read counts to compare the rRNA copy number or colonizer-persister value associated with each read between the treatments.

Methods

Soil Core Experiment:

Study Area

The Oakville Prairie Field Station consists of over 900 acres of upland and lowland prairie and contains several types of remnant prairie communities. The southern portion of Oakville Prairie contains tall-grass prairie communities, which include grasses such as little and big bluestem. Oakville also includes two beach ridges, created from the recession of Glacial Lake Agassiz, along with eight saline seeps; another geological feature not common elsewhere.

Core Collection:

Soil cores (2 cm diameter by 4 cm length) were taken at five locations across Oakville Prairie in June 2016 and October 2017. Twelve cores were taken at each location in June 2016 (60 total), and eighteen cores were taken at each location (90 total) in October 2017 (Figure 3.1).

The locations were intended to roughly sample the diversity of soil types of Oakville Prairie. Cores were collected with a core sampler head on a slide hammer. A 2 cm by 4 cm plastic liner was placed inside the corer head and pounded down into the ground below the soil horizon. The ground was cleared of vegetation beforehand. The core was leveled off at the bottom and capped on both ends for transport back to the laboratory.

Core Treatment:

From each location and year, one core from each location was set aside for baseline measurements, while the rest were randomly assigned into treatment groups (half to be destructively harvested after one month, and the other half to be harvested after 2 months, all blocked by location).

Nutrient solutions were created by dissolving nitrogen in the form of ammonium nitrate (NO_3), phosphorus in the form of commercial fertilizer (superphosphate) for 2016, and potassium phosphate buffer in 2017 into water. Treatments in the 2016 cores were: control (water), low phosphorus (100ppm), high phosphorus (1000ppm), nitrogen (1000ppm) and both (100ppm superphosphate and 1000ppm NO_3 solution). Treatments were applied to each core in 20 mL amounts. Each week, cores were weighed, and the water loss was replaced with more solution. The treatments for the 2017 cores were: control (water), phosphorus (4000ppm), nitrogen (4000ppm), or both (4000ppm each). The solutions were added once a week in 5ml doses. For high phosphorus cores, treatment was such that it had the effect of quadrupling available phosphorus at both 1 and 2 months (Figure 3.2).

Cores were set on tables at room temperature (25°C). Throughout the experiment, the cores were covered to slow evaporation and limit plant growth. Cores were monitored each day,

weighed, and checked for plant growth. Any plant growth was trimmed. The initial set of cores were deconstructed upon return to the lab. At one month, the next set of cores was deconstructed and at two months, the final set.

Soil Core Processing:

Core Deconstruction:

Soil was removed from cores, broken down, and homogenized. Five grams of soil was added to 15ml ultra-pure deionized water and vortexed for 30 seconds to create a "soil slurry" that would be used for both Biolog EcoPlates and bacterial DNA extraction. The remaining slurry was placed on a shaker for 2 hours at 200 rpm. Slurry was then centrifuged at 3000 rpm for 10 minutes. The supernatant was transferred to microfuge tubes for chemical analysis.

Biolog EcoPlates:

Each Biolog EcoPlate well contained substrate, redox nutrients, and a tetrazolium violet dye that turns purple upon respiration of sole carbon sources. There were 31 substrates and one control with three replicates for a total of 96 wells. Measurement of the violet color with a spectrophotometer after reaction correlates with the amount of respiration taking place in the wells. The 45 µL of soil slurry from the processing step was mixed with 15 mL ultra-pure water and 100 µL was added to each well of the plate. Plates were placed in an 18° C fridge for several days. Each day, plates were analyzed at 590 nm on a BioTek Epoch microplate reader using the Gen5 2.01 (Biotek) software. Day 6 was used for the final measurements. The OD values for each substrate were averaged across the 3 technical replicates. Ecolog Plates were only used in Oakville 2016.

The following calculations are from Weber and Legge (2010). For calculating community metabolic diversity (CMD), each averaged substrate was classed as a 1 if it was significantly higher than the water blank or 0 if it was not. For each sample, the classes were summed across all 31 substrates to get the CMD of that sample. For average metabolic response (AMR), each of the water blank OD values were subtracted from the summed substrate OD value. All those values were then summed and divided by 31 (Equation 1). This gave the AMD, which explains the respiration of the bacterial community across the substrates. In SAS Studio, CMD was modeled as a function of treatment by Month 1 and 2 separately with a Poisson distribution. AMD was also modeled as a function of treatment for each month with a Beta distribution.

$$\text{Equation 1: AMD} = \sum(\text{O.D well} - \text{O.D neg})/31$$

Nematode Extraction:

Eighty grams of soil was combined with 300 mL of water in a plastic jar. Nematodes were isolated from the soil slurry by sieving (2 mM, 150 μM, and 45 μM). Nematodes from the Oakville 2016 samples were extracted using Bearmann trays (Carter & Gregorich, 2007) while 2017 samples were isolated with sucrose flotation (Jenkins 1964). Extracted nematodes in water were placed in 100 mL hoc bottles. The extraction method was changed to better capture the full range of nematode diversity, only active nematodes are captured by Bearmann trays.

Molecular Work:

DNA extraction:

Prokaryotic DNA was extracted from a soil slurry using a modified protocol of the PowerSoil® DNA Isolation Kit (Mo Bio Laboratories). Instead of adding 0.25 grams of dry soil,

200 µL of soil slurry was added to the tubes initially and initial vortexing was reduced to 8 minutes. Nematode DNA was extracted using the UltraClean® Tissue and Cells DNA Isolation Kit (Mo Bio). All extracted DNA was kept in a -20° freezer until PCR.

Mi_seq PCR:

The PCR pipeline (Figure 3.3) started at step 1b with adapter-loci amplification using Mi_seq primers with adapters. After optimization tests found the best annealing temperature, bacterial DNA was amplified by PCR using Q5 hot-start polymerase for 25-35 cycles. The 515f_miseq and 816r_miseq primers were used to amplify a portion of the 16S operon of prokaryotes, while the NF1_miseq and NR2_miseq primers were used to amplify a portion of the 18S operon in nematodes. The initial denaturation step was at 98°C for 30 seconds, followed by denaturing steps of 10 seconds. Annealing occurred at 55° for 20 seconds and extension at 72° for 20 seconds. The final extension was 2 minutes.

Indexing PCR:

ZR-96 DNA cleanup kits (Zymo, Irvine, CA) were used to purify the PCR product. Then, the product was re-amplified using indexing primers. Each bacterial sample was given a unique combination of forward and reverse indexing primers. Q5 Hotstart polymerase was used with 1 µL of product for 15 cycles. Second round products were cleaned and concentrated (Zymo DNA Cleanup and Concentrator), then quantified using a broad range assay for double-stranded DNA with the Qubit fluorimeter.

Cleanup and Submission:

The 2016 product had an excess of smaller DNA fragments. The product was run on a gel and cut out. The larger 16S amplicons were recovered with a gel extraction kit (Zymoclean

Gel DNA Recovery Kit). The 2016 samples were added to a master library tube along with other experiments' samples at a concentration of 10 nM for a total volume of 100 µL. The same was done for the 2017 samples, minus the gel extraction. Master libraries were submitted to the Genomics Core at the University of North Dakota School of Medicine and Health Sciences for Illumina Sequencing using a MiSeq Reagent Kit v2 (600 cycles).

Data Analysis:

Bioinformatics:

Using the USEARCH software (version 11, Edgar 2010) and custom Python scripts (Python Software Foundation), Illumina sequence files were merged (-fastq_mergepairs), demultiplexed to separate out nematodes from prokaryotes (-fastx_demux), trimmed to a length of 250 for prokaryotes and 325 for nematodes (-fastx_truncate), filtered (-fastq_filter), dereplicated to a minimum unique size of 2 (-fastx_uniques), and cleaned of chimera (unoise3). Resulting sequences were identified as zero-radius operational taxonomic units ("ZOTUs"). Abundance and sample sources were added to each ZOTU and all ZOTUs were pooled.

From there, the ZOTUs were matched to a dereplicated SSU Ref NR 132 dataset from Silva to create per-sample read frequency table of nematodes and prokaryotes that also had copy number estimates from the Ribosomal RNA Operon Copy Number Database (rrnDB) associated with prokaryotic sequences, and CP values (colonizer-persistor values from the Maturity Index of Bongers, 1990) for nematodes. Only prokaryotic sequences that matched to the rrnDB were retained so that copy number information would be included; reads that did not match were discarded (32.5%). Only nematode species that had Colonizer-Persister values were retained; species that did not have values were discarded (22%). Colonizer-Persister values, which consist

of incremental numbers from 1 to 5, represent a nematode species ecosystem function. Species with values of 1 are enrichment indicators, 2 are basal fauna, and 3-5 are indicators of structure. These values were used in lieu of copy number estimates, which do not exist for most nematode species. Accurate copy number estimation of eukaryotes remains a difficult technical problem.

Generalized Linear Mixed Models:

Sequencing reads were analyzed by multinomial cumulative ("proportional odds") logistic regression using PROC GLIMMIX (SAS, Statistical Analysis Software, Cary, NC). For both Spring 2016 and Fall 2017 data, ordered rRNA copy number of each prokaryote sequence and ordered CP-value of each nematode sequence, was modeled as the response variable as a function of treatment as the explanatory variable. Using the cumulative logit link, this multinomial logistic regression models the probability of "having a sequentially higher rRNA copy number" for prokaryotes, and the probability of "having a sequentially lower CP value" for nematodes. Thus, treatment means were computed as the probability of a read coming from a genome with rRNA copy number of 2 or greater for prokaryotes, and the probability of a read coming from a CP-value of 1 for nematodes.

Nematodes sequencing reads were also analyzed by a modeling counts as a function of treatment by species using a negative binomial distribution with an offset of the natural log of the total count in PROC GLIMMIX. Parameter estimates for the best model in each year were used in a second GLIMMIX. Estimates were modeled as a function of treatment, CP-value, and interaction with a normal distribution.

Additional Metagenomic Analyses:

Multivariate community composition was analyzed by Principle components analysis using the prcomp function in R (R Core Team). Rarefaction curves were constructed in R using package vegan (Oksanen et al 2018) using a step number of 20. Cluster analysis was done by sample using hierarchical clustering in R: Euclidean distance matrices were clustered using the Ward method (ward.D) and the resulting trees were converted to dendrograms and colored by treatment, block, and month using the R package dendextend (Tal Galili 2015).

Results

Prokaryote community composition:

By examining the proportions of uncorrected reads across bacterial phyla to proportions that corrected for copy number, there were noticeable shifts in some phylum abundances (Figure 3.4). In Oakville 2016, without copy number correction, Verrucomicrobia, Thaumarchaeota, and Chroflexi were underestimated; Bacteriodetes, Firmicutes, and Actinobacteria were overestimated; Proteobacteria, Planctomycetes, and low abundance phyla were unchanged. In Oakville 2017, Chroflexi was underestimated; Bacteriodetes and Actinobacteria were overestimated; Proteobacteria and low abundance phyla were unchanged.

For prokaryote phylum proportions across treatment (Figure 3.6), cores that were not part of the experiment (treament=none) were more different than cores that were treated, even the control, indicating the effect of isolating, watering, and incubating the cores. In 2016, low phosphorus was similar to the control except Firmicutes was lower. High phosphorus had an increase in Verrucomicrobia and Acidobacteria, but a decrease in Planctomycetes, Firmicutes, and Proteobacteria. Nitrogen had an increase in Thaumarchaeota, Bacteriodetes, and Candidatus

Sacchrobacteria, but a decrease in Planctomycetes, Firmicutes, and Actinobacteria. Both had an increase in Verrucomicrobia and Acidobacteria, but a decrease in Thaumarchaeota, Proteobacteria, Firmicutes, and Bacteroidetes. In 2017, the treatments were generally similar to the control and had a greater proportion of Thaumarchaeota relative to the non-incubated baseline samples.

For the proportions of the top ten species in each year, one-fifth of the sequences in the Oakville 2016 cores were from a single species of Verrucomicrobia. Five of the top ten species were Proteobacteria, and the rest were individual species from Thaumarchaeota, Chloroflexi, Actinobacteria, and Candidatus Saccharibacteria (Table 3.1a). Over half of the sequences in the Oakville 2017 cores were from a single species of Thaumarchaota. Five of the top ten species were Actinobacteria, three were Proteobacteria, and one was Verrucomicrobia (Table 3.1b). Five species were found in the top ten proportion for both 2016 and 2017.

Most species from Oakville 2016 were found in only a few samples and almost none were found in more than 30 (Figure 3.6a). Most species from Oakville 2017 were also found in a few samples, but a low, steady percentage was found up to complete coverage of samples (Figure 3.6b). Rarefaction curves indicated that most of the diversity present was captured by the amplicon sequencing, as most lines plateaued (Figure 3.7). Cluster analysis by treatment resulted in almost no clustering; samples by treatment were randomly scattered across the tree (Figure 3.8a, 3.8b). Clustering by month was random for Oakville 2016 (Figure 3.8c) but in Oakville 2017, Month 1 and 2 had several conserved clusters (Figure 3.8d). Clustering by block was similarly random for Oakville 2016 but had well-conserved clusters for Oakville 2017 (Figure 3.8e, 3.8f). This suggests that block, which represented location, was the most important

factor for Oakville 2017 prokaryotic community composition. PCA by block recovered distinct groups as well (Figure 3.9).

Nematode community composition:

Most nematode sequences were from plant parasites, followed by free living. There were very few fungus-feeding or animal parasites. A significant fraction of sequences were classified as unknown or from an environmental sample, more so in Oakville 2017. Free-living nematodes were more prevalent in all treatments but the control and nitrogen plots for Oakville 2016 and in all treatments in 2017. The disparity between the overall dominance of plant parasites with the dominance of free-living within plots may be due to the environmental cores (treatment=none).

For feeding groups, semi-endo plant parasites were the top feeding group in all treatments across Oakville 2016 and 2017. The exception was for high phosphorus plots in Oakville 2016, in which bacteria-feeding nematodes were the top group and semi-endo plant parasites were the second highest. In the control plots of Oakville 2016 and the phosphorus plots of Oakville 2017, the second-highest group was plant ectoparasites. For all other plots, the second-highest was bacterial feeding.

There was a total of 105 nematode species identified in Oakville 2016 and 85 species in 2017. By family, Oakville 2016 had a quarter Hopoliamidae, a sixth Rhabditidae, and smaller portions of Plectidae, Tylenchidae, and Telotylenchidae. Oakville 2017 had a quarter Tylenchulidae and Rhabditidae, and a small number of Tylenchidaea (Figure 3.10).

The top species across treatments were varied (Table 3.2a and 3.2b). In Oakville 2016, *Helicotylenchus* was in the top five species in all treatments except high phosphorus. *Rhabditus*

was the top species in high phosphorus and both treatments. In Oakville 2017, *Rhabditus* was present in all treatments, but the top in Nitrogen and Both.

Treatment effects on community composition:

Multinomial logistic regression of copy numbers showed that the effect of nutrient enrichment on the representation of rRNA copy number among prokaryotes differed between the spring 2016 experiment and the fall 2017 experiment. In the spring 2016 experiment, the treatment coefficients for prokaryotes were negative relative to the "control", indicating that reads from the nutrient-enriched cores had a lower probability of coming from a high rRNA copy species than reads from the non-amended cores (Figure 3.11a). However, in the fall 2017 experiment, the treatment coefficients were positive relative to the "control", indicating that reads from the nutrient-enriched cores had a high probability of coming from a high rRNA copy species than reads from the non-amended cores (Figure 3.11b).

This pattern was almost the opposite for nematode communities. In the spring 2016 experiment, the treatment coefficients for nematodes were positive relative to the "control", indicating that reads from the nutrient-enriched cores had a higher probability of coming from an enrichment-type (low CP-value) species than reads from the non-amended cores (Figure 3.11c). In the fall 2017 experiment, the treatment coefficients from "Nitrogen" and "Both nitrogen and phosphorous" treatments were positive relative to the "control", indicating that reads from the nitrogen-enriched cores had a higher probability of coming from an enrichment-type (low CP-value) species than reads from the non-amended or phosphorous enriched cores (Figure 3.11d).

The two-step GLMM found a significant effect of CP-value ($F_{4,360}$=3.9, p=0.0041) and the interaction between CP-value and treatment ($F_{16,360}$=1.84, p=0.0246), but not for treatment in

Oakville 2016. Mean fold abundance for CP 1 species in low phosphorus and both treatments were significantly higher than the control. For CP 2 species, nitrogen plots had lower counts than control and both had lower than low phosphorus. For CP 3, 4, and 5 species, there were no significant differences between treatments, though low phosphorus was significantly higher than zero and high phosphorus was significantly lower than zero in CP 3 (Figure 3.12). For Oakville 2017, there was a significant effect of treatment ($F3_{,192}$=3.52, p=0.0161) and CP-value ($F_{4,192}$=4.68, p=0.0013), but the interaction was not significant.

Discussion

The most significant finding of this study was that nutrient enrichment increased the representation of high-copy number prokaryotes in the mesocosm experiment conducted after the fall of 2017, but not in the mesocosm experiment conducted after the spring of 2016. Conversely, phosphorous enrichment (but not nitrogen) increased the representation of enrichment-type (low CP-value) nematodes in the mesocosm experiment conducted after the spring of 2016, while nitrogen (but not phosphorous) increased the representation of enrichment-type (low CP-value) nematodes in the mesocosm experiment conducted after the fall of 2017. The seasonal difference between nutrient enrichment effects on copy number representation was not initially predicted but would be consistent with the understanding that microbial communities in the spring are already likely to be relatively copiotrophic in nature due to the flush of newly available nutrients during and after the spring thaw. These largely copiotrophic species are expected to have higher rRNA copy numbers, and perhaps this is the reason we did not see nutrient enrichment having an additional benefit to these species which had already just flourished. Indeed the multi-copy

Verrucomicrobia and Proteobacteria were a larger portion of the baseline samples from spring 2016, while the single-copy Thaumarchaeota dominated the baseline samples of fall 2017.

The original hypothesis predicted that phosphorous would benefit high-copy number bacteria as well as nitrogen, but in our experiments, phosphorous appeared to have a less pronounced effect than nitrogen alone. There are several possible reasons why nutrient enrichment, and phosphorous specifically, might not benefit high-copy number bacteria in the spring 2016 experiment, besides the difference in community composition. First, copy number is not evenly distributed among species. It is highly skewed towards smaller copy number; most species have a copy number below 5. In the experiments, most counts come from species with a mean copy number of one. Second, the effect of nutrient amendment may be more complex than simple uptake. Microbes utilize nutrients in a particular stoichiometric ratio, and phosphorus may not be a limiting nutrient for many bacteria. There also may be a threshold where at a certain level of nutrient amendment become unusable or that forms of nitrogen and phosphorus may not have been reliably utilized by microbes or utilized at the same rates. Third, both experiments had single species that were dominant. Candidatus *Xiphinematobacter* contributed to 22% of the reads in 2016 while *Nitrosophaera viennesis* accounted for 67% in 2017.

An important possible explanation for why nutrient enrichment did not appear to increase the representation of high copy-number prokaryotes in the spring 2016 experiment is that the fastest-growing microbes were in fact preyed upon by fast-growing nematodes. Indeed, phosphorous enrichment did increase the representation of enrichment-type bacterivores in spring 2016, this shift in nematode community could be sufficient to offset, or even reduce, the abundance of their microbial prey. Furthermore, the results of the two-step GLMM showed that enrichment-type nematodes responded in a species-specific manner to phosphorus in 2016.

Prokaryote community composition:

In Oakville 2016, Firmicutes saw an unexpected drop from control to nutrient plots, as the phylum contains many fast-growing species and should respond to enrichment. Low and high phosphorus plots also unexpectedly increased in Candidatus Saccharibacteria, a lineage with highly reduced genomes (Nelson et al. 2014) and a single copy of the rRNA operon. High phosphorus plots also had higher levels of Verricomicrobia and Acidobacteria. Acidobacteria are oligotrophs (Fierer, Bradford, and Jackson 2007), so they are not expected to respond to enrichment. Nitrogen plots, compared to the control, had lower Actinobacteria, higher Chloroflexi, Thaumarchaeota, and Bacteriodetes, and much higher Candidatus Saccharibacteria.

For Oakville 2016, some general trends emerge. Proteobacteria and Bacteriodetes were highest and Verrucomicrobia was lowest in the environmental samples (treatment=none). Verrucomicribia responds positively to phosphorus inputs. Thaumarchaeota responds negatively to nutrient treatments, especially both treatments together. Proteobacteria was dominant whether nutrient enrichment was present or not. Candidatus Sacchrobacteria responds positively to nitrogen alone. Bacteriodetes responds negatively to high Phosphorus or Both treatments. Acidobacteria responds positively to high phosphorus.

Oakville 2017 had different distributions than that of Oakville 2016, suggesting an effect of both block and season. Only two of the five core collection locations were the same between the two years. Furthermore, Oakville 2016 was started in the early summer, while Oakville 2017 was started in the mid-fall. All cores had lower Actinobacteria compared to the environmental samples (treatment=none). Within treatments, there was no significant change other than

Candidatus Saccharibacteria increasing in nitrogen treatments. Overall, the Oakville 2017 soil community resisted nutrient enrichment change.

The cluster analysis failed to resolve treatments as meaningful groups of comparison. Instead, both month and block were better at grouping. In the case of Oakville 2017, there was a strong clustering by block. Principle components analysis confirmed this. This indicates that soil microbe communities across the soil landscape are dynamic, diverse, and do not respond unilaterally to nutrient enrichment. Proximity is a better measure of community similarity than treatment type. While not recorded, we observed different plant communities across different blocks, so the aboveground biotic composition will drive the belowground composition.

Nematode community composition:

Nematode families were much more constrained and less defined than bacteria phyla. In both Oakville experiments, there was a large proportion of reads that were listed as other. Apart from that, plant parasites (Hoploliamidae, Tylenchulidae, Telotylenchidaea, and Tylenchidae) made up a third of Oakville 2016 and 2017 reads. Bacterial feeders (Plectidae and Rhabditidae) made up the other third or quarter. There were strong shifts in the families present in Oakville 2016 versus 2017, despite the feeding group proportions remaining relatively constant.

Familywise, Hoplolaimidae and Telotylenchidae were replaced by Tylenchulidae from 2016 to 2017. This may be due to different plant compositions between the summer and fall, as different plant-parasites may target different plants. For bacterial-feeders, Rhabditidae was present in both, but Plectidae was only present in the summer samples of Oakville 2016. Plectidae may take advantage of nutrient enrichment in the summer and declines afterward. This

is backed up by the virtual dominance of Plectidae in the long term enrichment plots of Konza (data not shown).

On a species level, Oakville 2016 had bacterivores as the top species in plots that included phosphorus. Plots without phosphorus had plant parasites as the top species. This fits with the previous understanding about enrichment-type bacterivores but highlights the specific importance of phosphorus.

Oakville 2017 only had bacterivores as the top species in the plots that had nitrogen in them. This may be due to general enrichment promoting bacterivores with a low signal to noise ratio in the data, or a more specific mechanism of enrichment that depends on season. Phosphorus may be more limiting in the summer, while nitrogen may be more limiting in the fall. Therefore, the limiting nutrient by season will affect bacterivores the most. This is backed up by significantly more bacteria-feeding nematode reads present in phosphorus-containing samples in 2016.

The lower abundance of root-feeding nematodes in 2016 nitrogen plots is unclear, as is the absence of hyphal-feeding in nitrogen plots compared to phosphorus in 2017. While statistically significant, it does not appear to be a biological significant.

The cluster analysis failed to resolve treatments as meaningful groups of comparison. Instead, month was marginally better at grouping, as was block. This means nematode communities were diverse across soil cores, even those with similar treatments and locations.

Conclusion

Nutrient enrichment increased the representation of high-copy number prokaryotes in the mesocosm experiment conducted after the fall of 2017, but not in the mesocosm experiment conducted after the spring of 2016. Furthermore, phosphorous (but not nitrogen) enrichment increased the representation of enrichment-type (low CP-value) nematodes in spring 2016, while nitrogen (but not phosphorous) enrichment increased the representation of enrichment-type (low CP-value) nematodes in fall 2017. We interpret these results as a combination of the seasonal difference in community composition at the beginning of each experiment, along with a combination of top-down control by invertebrate consumers (such as nematodes) and the possibility that nutrient limitations in grassland soils differ between spring and fall. The Growth Rate Hypothesis was upheld for both prokaryotes and nematodes, though the effects of the season were non-negligible.

Future Work:

More investigation is needed to determine how different copy number organisms respond to nutrient enrichment. A next step would be a meta-analysis of nutrient enrichment using metagenomics data, if sufficient data exists, to examine if high copy number species respond disproportionately to nutrient enrichment in other systems. Another method would be the analysis of soil communities across different natural or applied chemical gradients, such as the NEON sites. While nematode growth rate was connected to phosphorus enrichment, we still need to confirm that fast-growing nematode species have higher copy numbers. Nematode copy number needs to be estimated and correlated with growth rate using sequencing and life-history trait cultures.

Figures and Tables

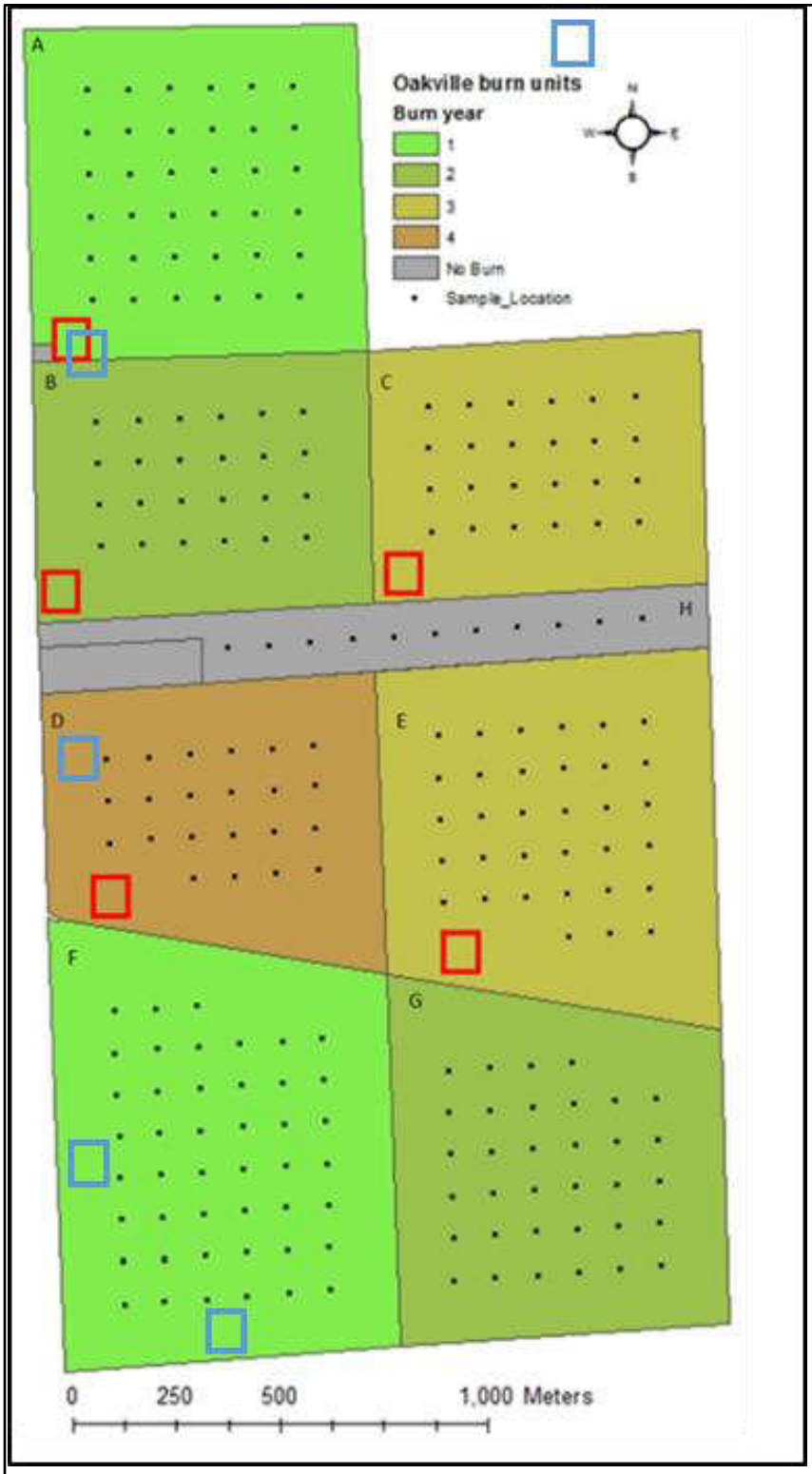

Figure 3.1: Oakville Prairie. Schematic map with soil core collection locations. Red indicates 2016 samples and blue indicates 2017.

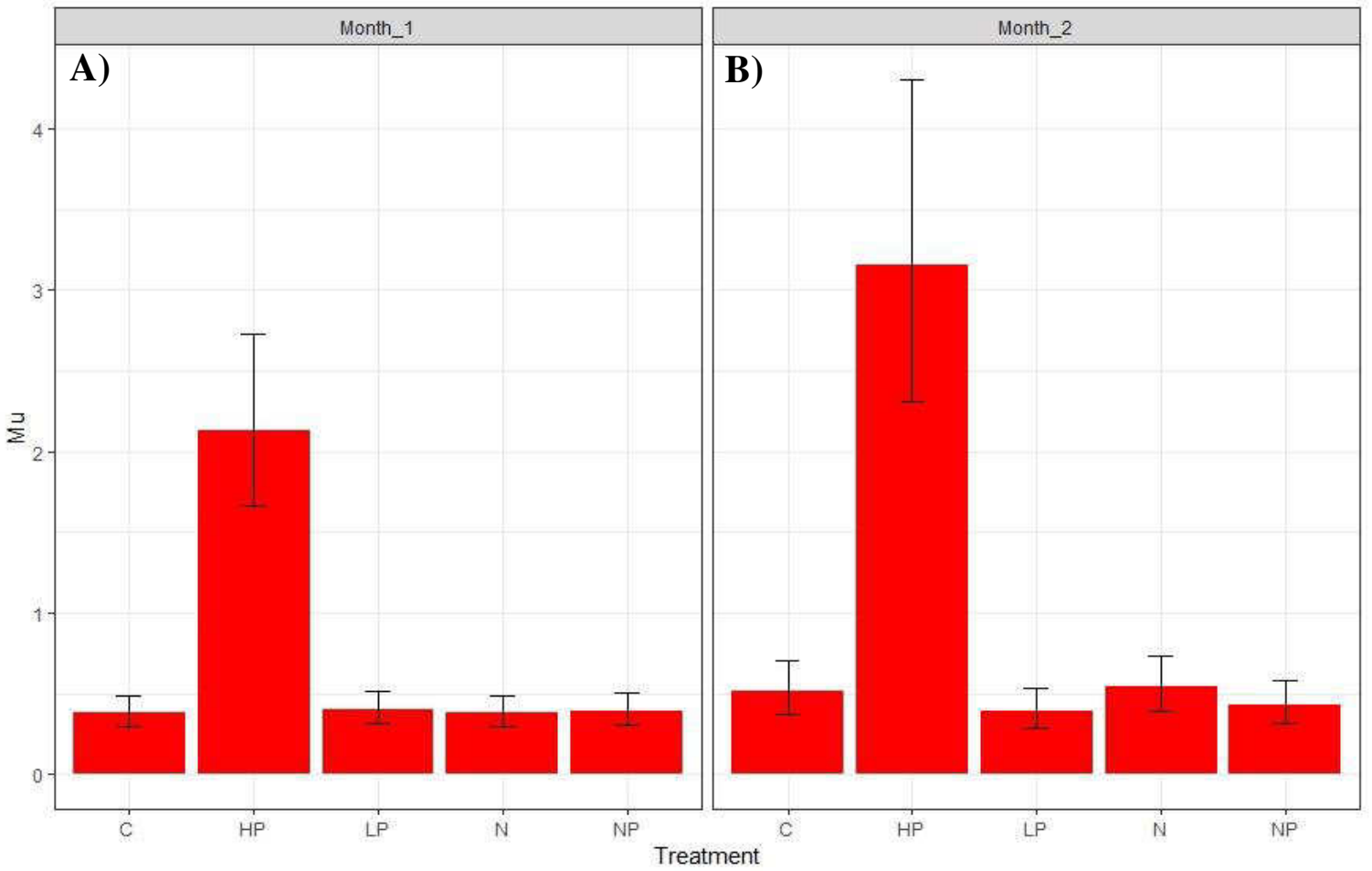

Figure 3.2: Phosphorus Chemical Concentrations. A) Month 1 Phosphorus, Oakville 2016; B) month 2 Phosphorus, Oakville 2016.

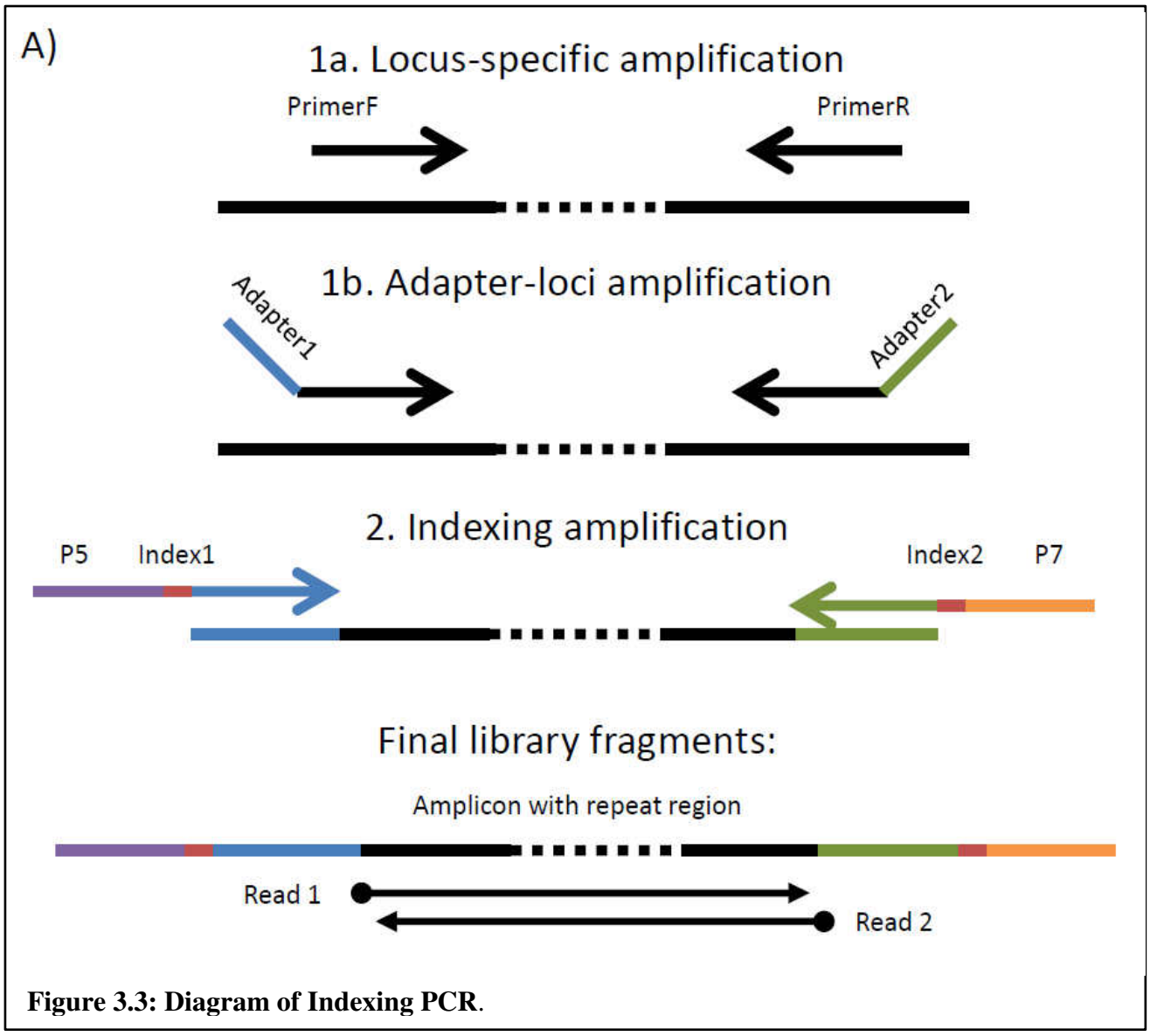

Figure 3.3: Diagram of Indexing PCR.

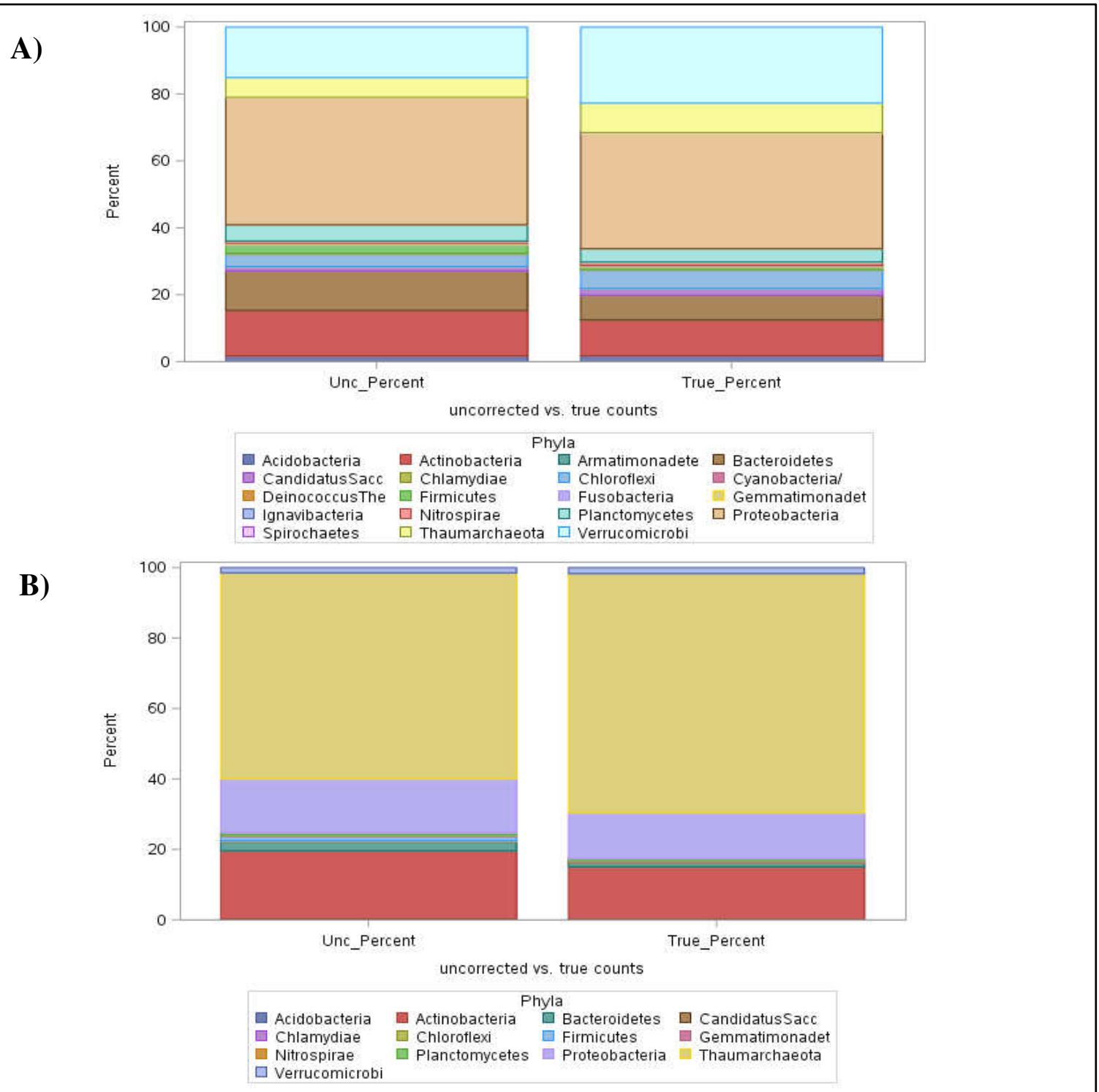

Figure 3.4: Bacterial Phyla Copy Number Corrections. Proportion bar plots of bacterial phyla for A) Oakville 2016 cores and B) Oakville 2017 cores.. The left bar plot shows the uncorrected proportions, while the right plot shows the copy number corrected proportions.

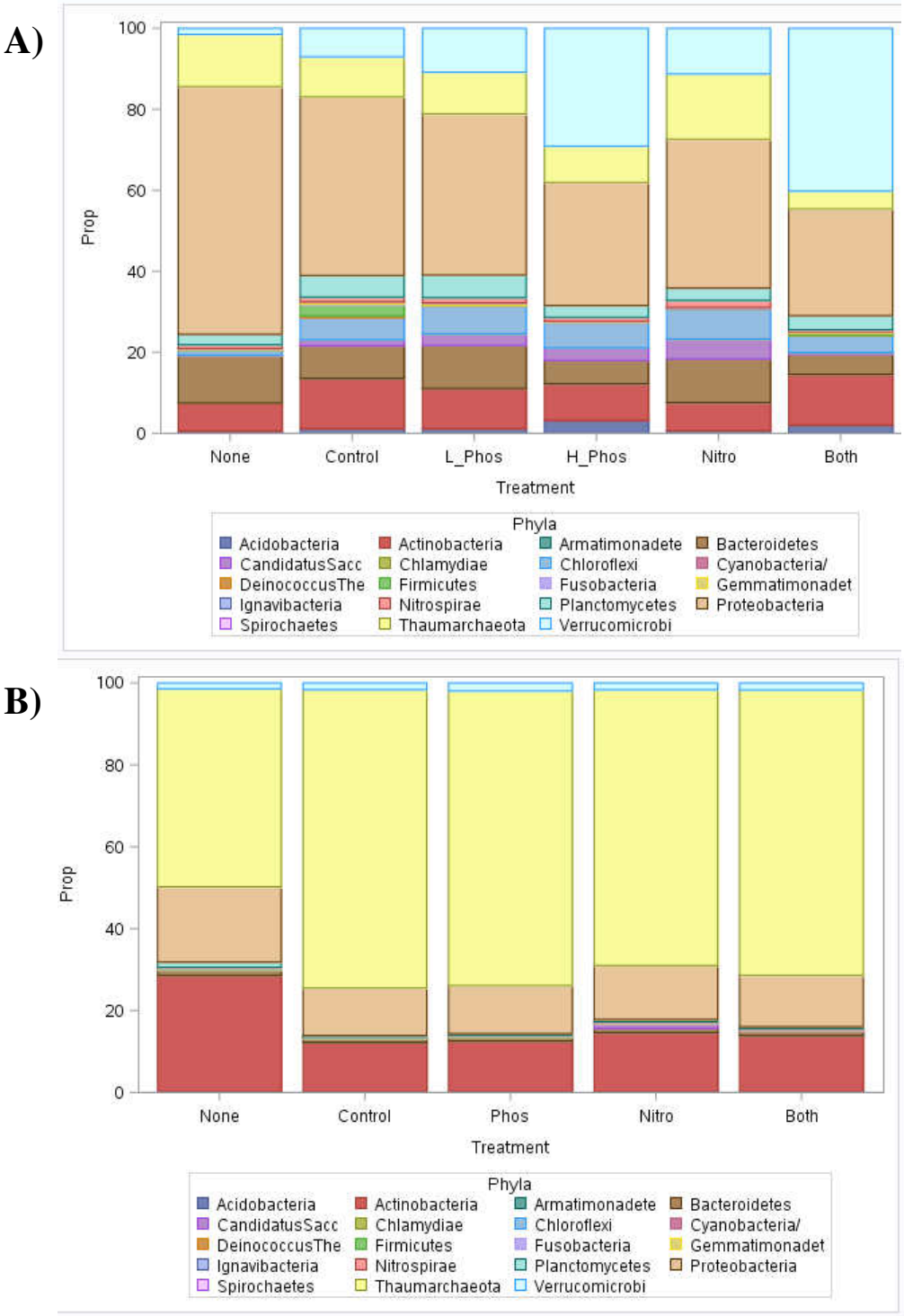

Figure 3.5: Bacterial Treatment Proportions. Proportion bar plots of bacterial phyla across Treatments. A) Oakville 2016 cores, B) Oakville 2017 cores.

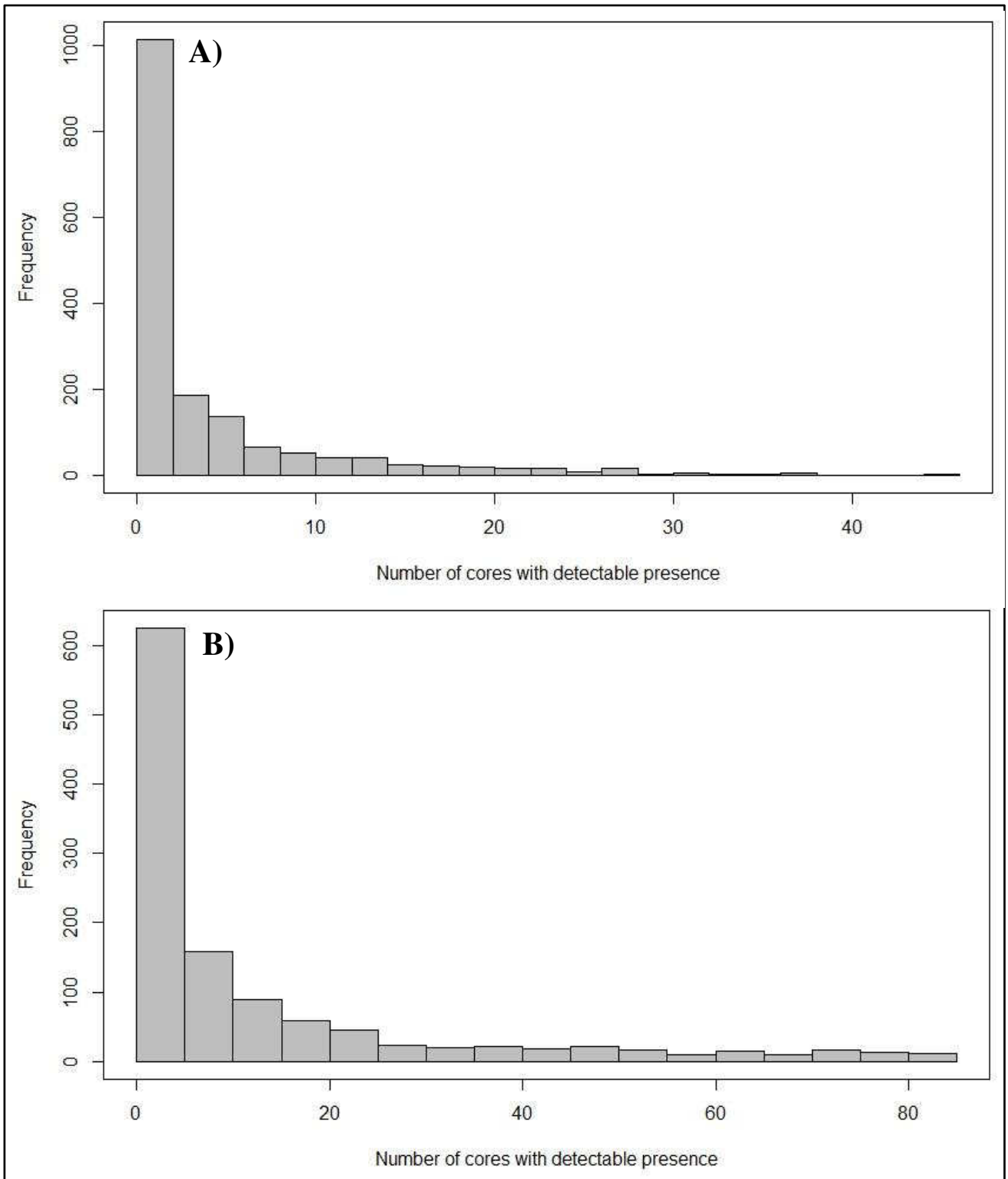

Figure 3.6: Proportion Histograms. Histograms of number of core samples each species is found in for A) Oakville 2016 and B) Oakville 2017.

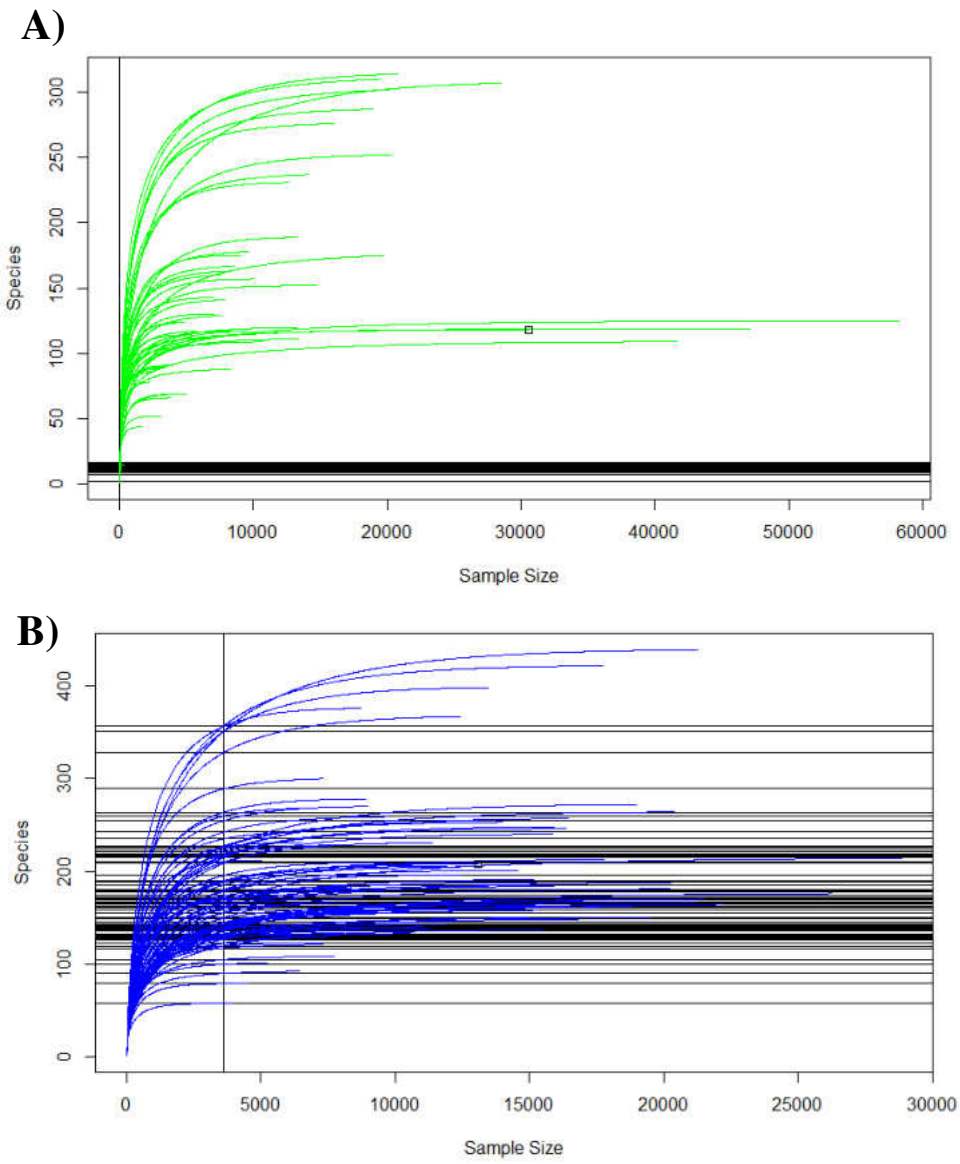

Figure 3.7: Rarefaction Curves. A) Oakville 2016 Bacteria, B) Oakville 2017 Bacteria. Each line indicates a different soil core sample.

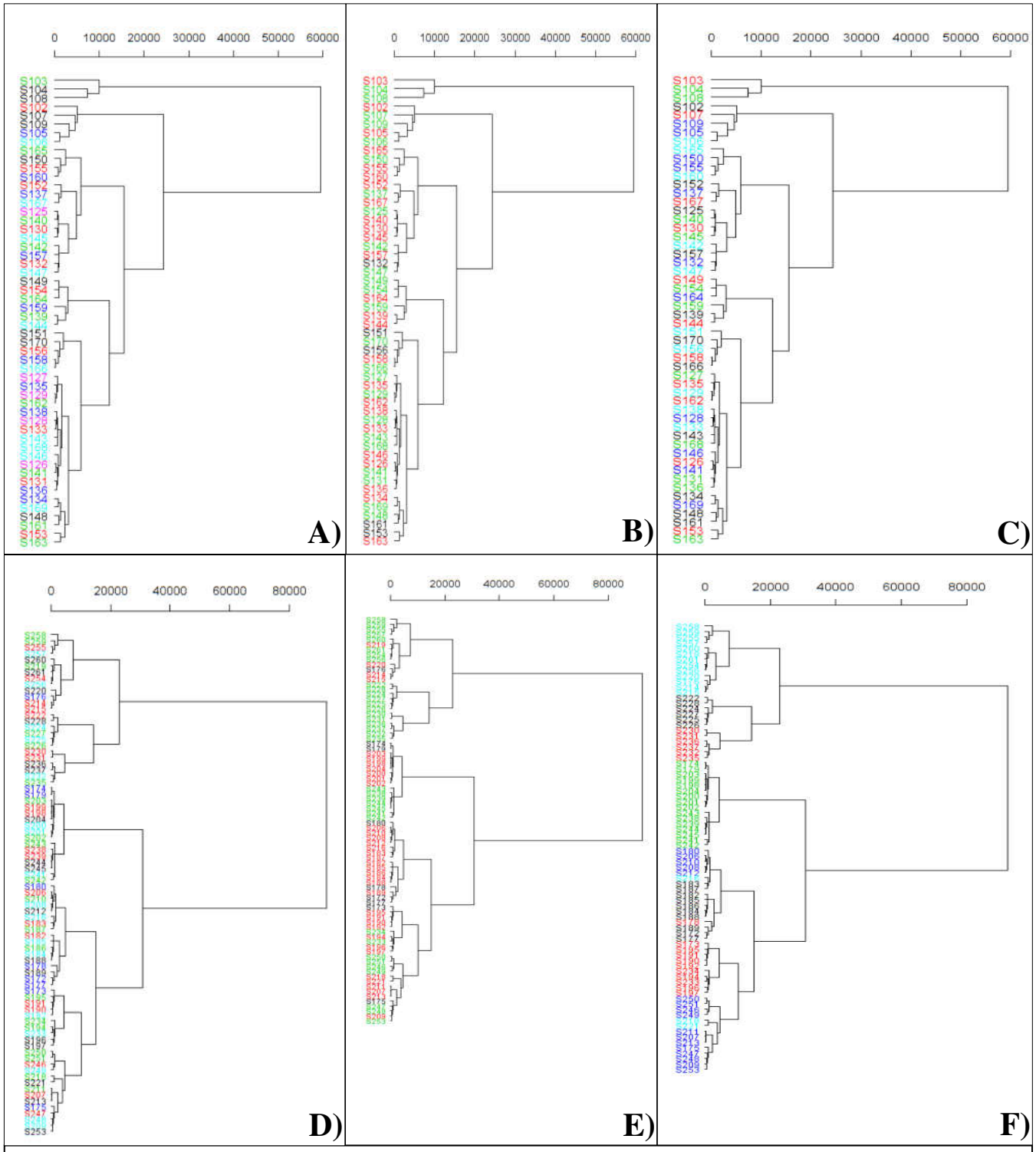

Figure 3.8: Cluster Analysis. Oakville Spring 2016 by A) Treatment, B) Month, and C) Block. Oakville Fall 2016 by D) Treatment, E) Month, and F) Block.

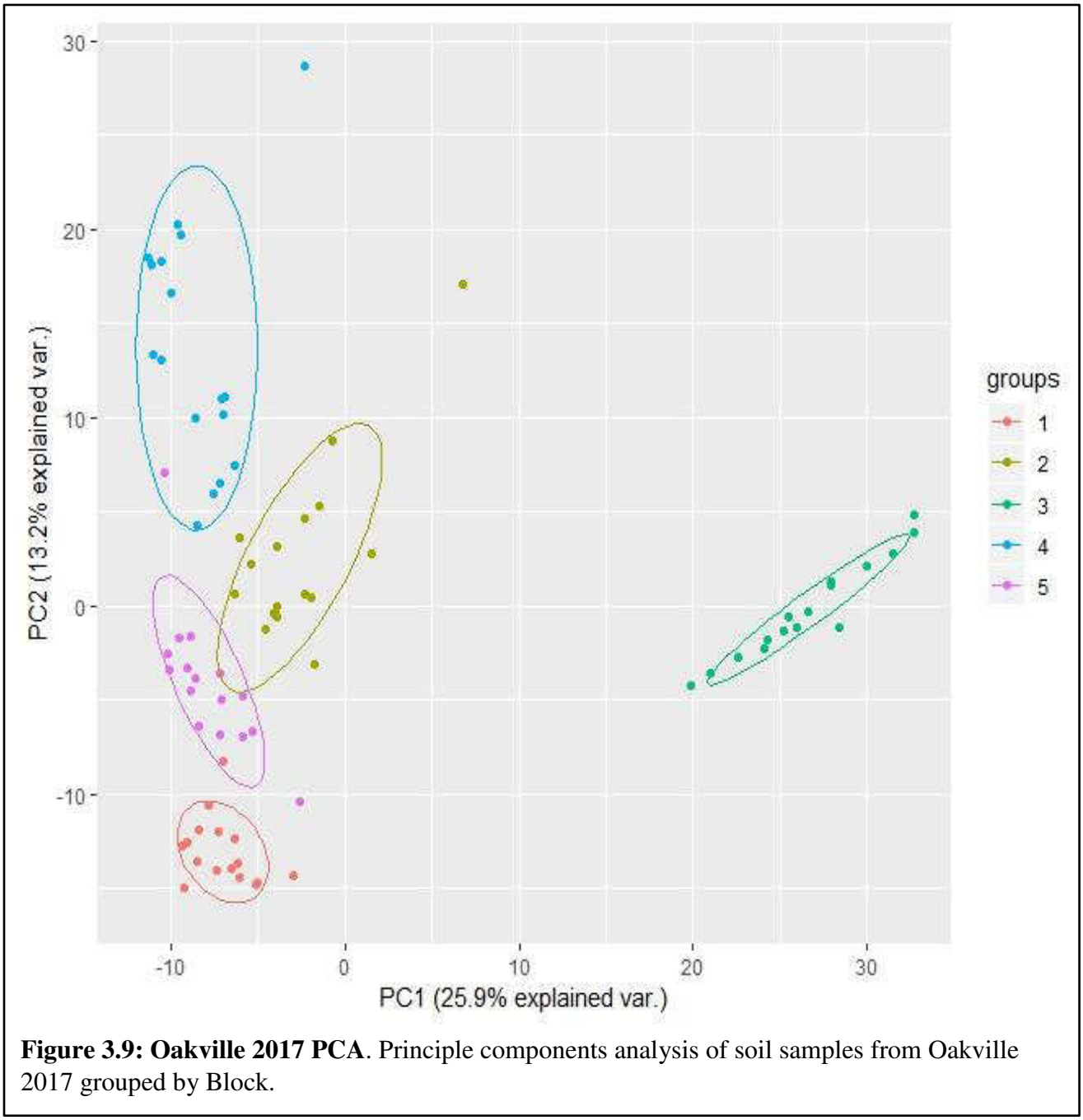

Figure 3.9: Oakville 2017 PCA. Principle components analysis of soil samples from Oakville 2017 grouped by Block.

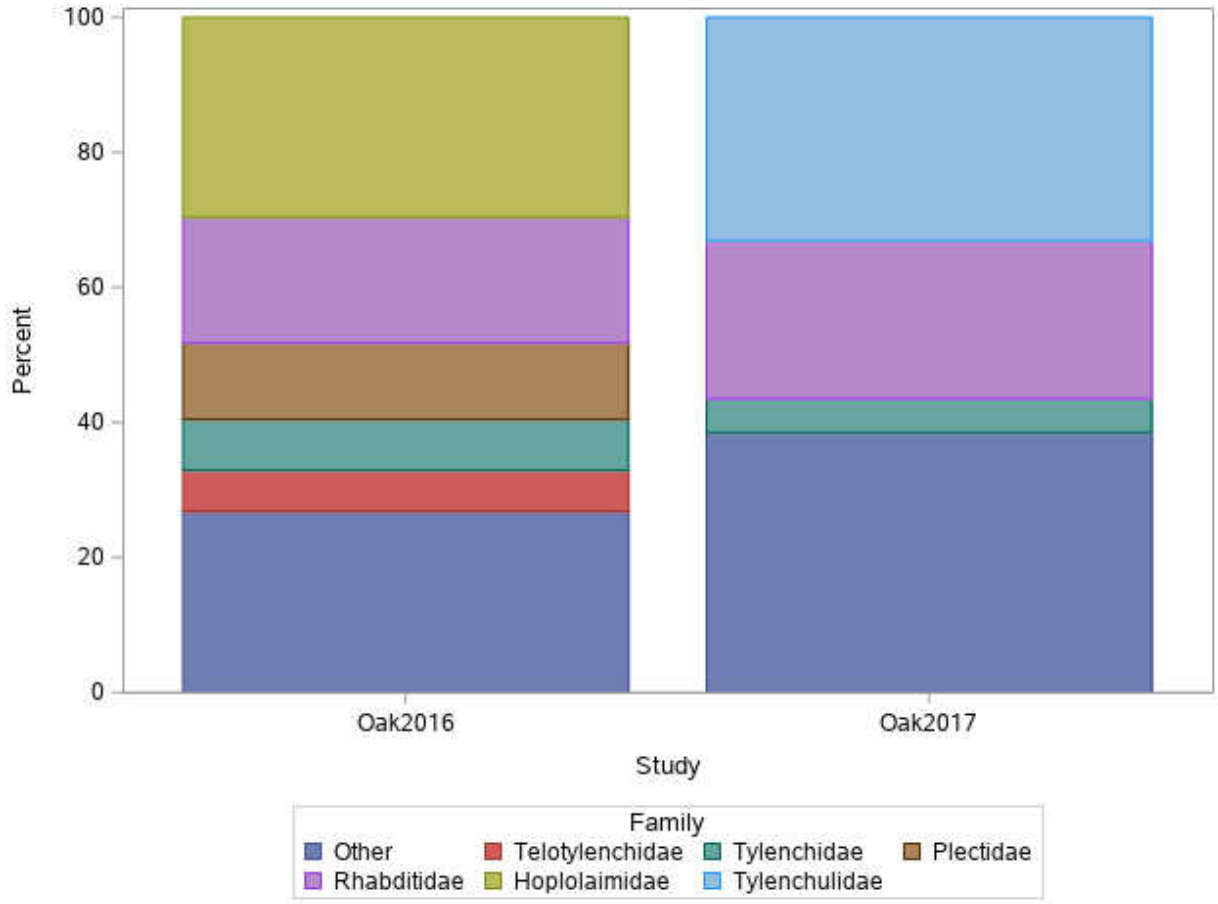

Figure 3.10: Nematode Proportions. Proportions of nematode families between Oakville experiments.

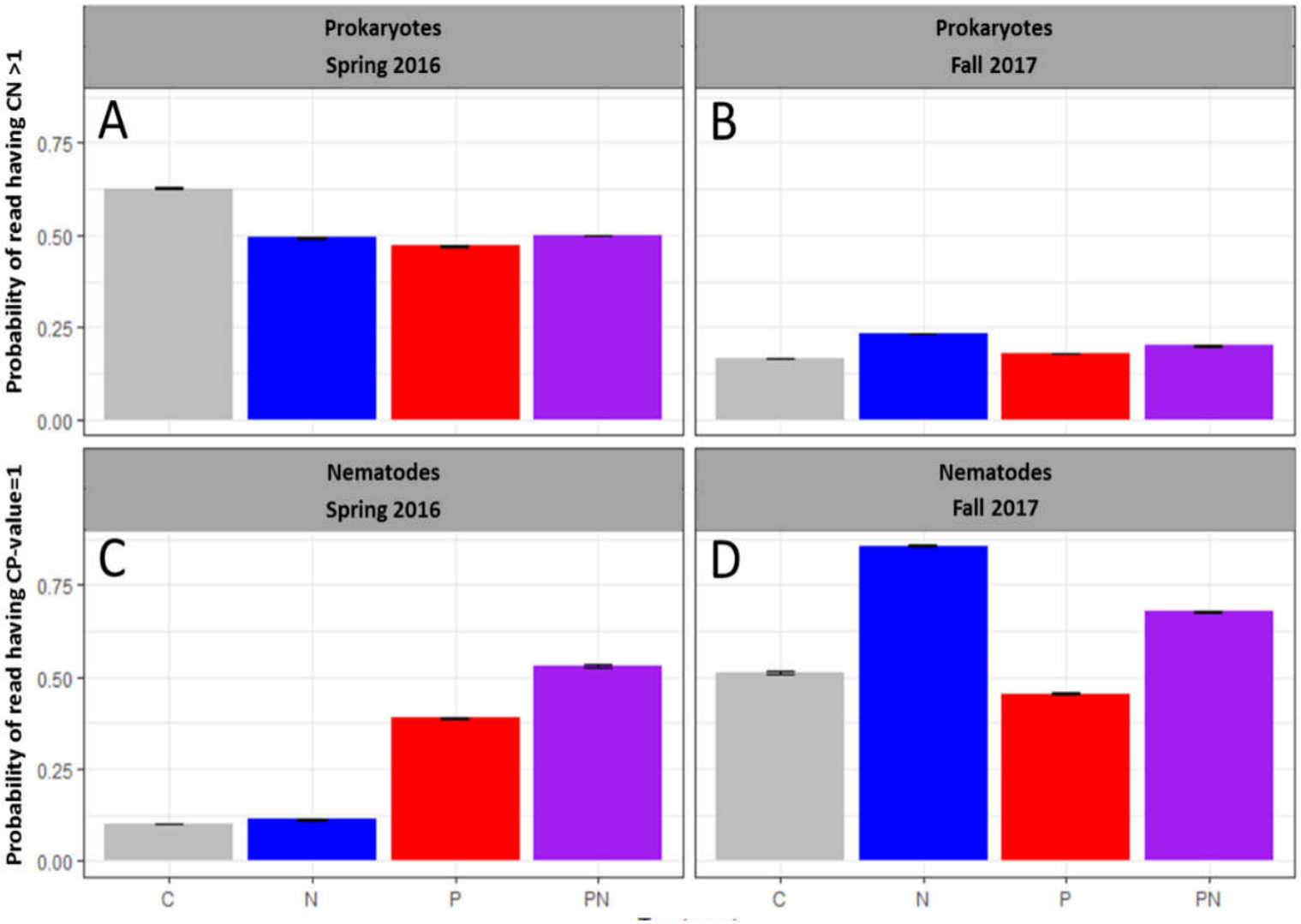

Figure 3.11: Treatment Effects of Prokaryote rRNA Copy Number and Nematode CP-values. Proportional odds of having copy number >1 across treatments for prokaryotes in A) spring 2016 samples, and B) fall 2017 samples. Proportional odds of having a CP-value=1 across treatments for nematodes in C) spring 2016 samples and D) fall 2017 samples.

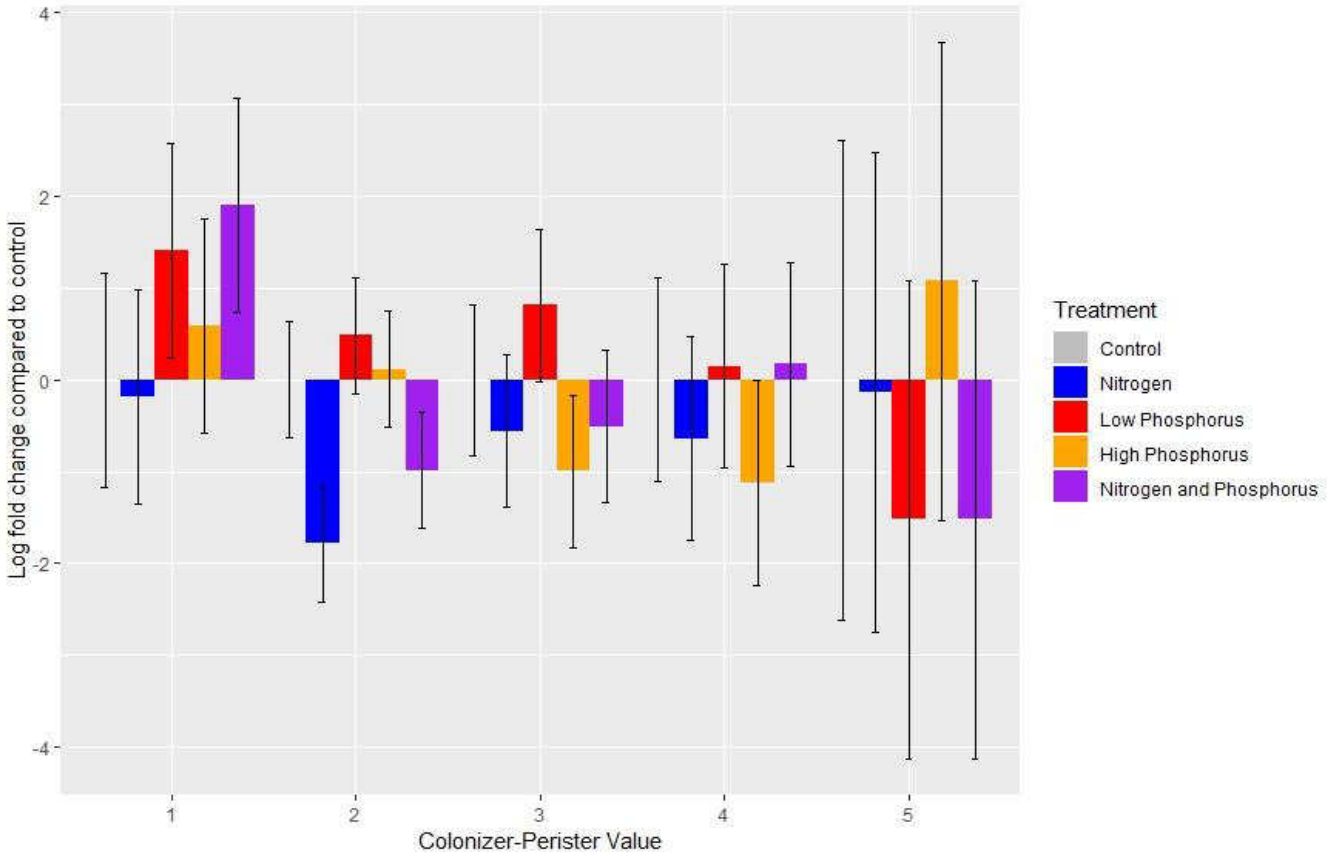

Figure 3.12: Nematode Abundance by Growth Rate across Treatment. Plots of Treatment*CN from parameter estimate GLIMMIX results for Oakville 2016 nematodes.

Table 3.1a: Most abundant prokaryote species in spring 2016.

Rank	rrnDB species ID	Phylum	Proportion
1	*Candidatus Xiphinematobacter sp. Idaho Grape*	Verrucomicrobia	0.22349
2	*Nitrososphaera_viennensis_EN76*	Thaumarchaeota	0.08811
3	*Steroidobacter_denitrificans*	Proteobacteria	0.05772
4	*Brevefilum_fermentans*	Chloroflexi	0.05431
5	*Rhodoplanes_sp._Z2-YC6860*	Proteobacteria	0.02644
6	*Conexibacter_woesei_DSM_14684*	Actinobacteria	0.02390
7	*Woeseia_oceani*	Proteobacteria	0.02064
8	*Candidatus_Saccharibacteria_oral_taxon_TM7x*	Candidatus Saccharibacteria	0.02024
9	*Ramlibacter_tataouinensis_TTB310*	Proteobacteria	0.01876
10	*Thioalkalivibrio_sulfidiphilus_HL-EbGr7*	Proteobacteria	0.01867

Table 3.1b: Most abundant prokaryote species in fall 2017

Rank	rrnDB species ID	Phylum	Proportion
1	*Nitrososphaera_viennensis_EN76*	Thaumarchaeota	0.67904
2	*Conexibacter_woesei_DSM_14684*	Actinobacteria	0.04602
3	*Ilumatobacter_coccineus_YM16-304*	Actinobacteria	0.02116
4	*Candidatus Xiphinematobacter sp. Idaho Grape*	Verrucomicrobia	0.01640
5	*Steroidobacter_denitrificans*	Proteobacteria	0.01597
6	*Rubrobacter_xylanophilus_DSM_9941*	Actinobacteria	0.01581
7	*Actinobacteria_bacterium_IMCC26256*	Actinobacteria	0.01547
8	*Pseudorhodoplanes_sinuspersici*	Proteobacteria	0.01526
9	*Microlunatus_phosphovorus_NM-1*	Actinobacteria	0.01387
10	*Rhodoplanes_sp._Z2-YC6860*	Proteobacteria	0.01362

Table 3.2a: Most abundant nematode species in spring 2016 by Treatment

Rank	None	Control	Low Phosphorus	High Phosphorus	Nitrogen	Both
1	*Helicotylenchus pseudorobus*	*Helicotylenchus pseudorobus*	*Plectus minimus*	*Rhabditis cf. terricola JH-*	*Helicotylenchus pseudorobus*	*Rhabditis cf. terricola JH-*
2	*Hoplolaimus columbus*	*Tylenchorhynchus claytoni*	*invertebrate environmental*	*Plectus minimus*	*Nematoda environ. sample*	*Helicotylenchus pseudorobus*
3	*Paratylenchus projectus*	*Hoplolaimus columbus*	*Helicotylenchus pseudorobus*	*Tylenchorhynchus claytoni*	*Plectus minimus*	*Plectus minimus*
4	*Basiria duplexa*	*Basiria duplexa*	*Hoplolaimus columbus*	*Nematoda environ. sample*	*Hoplolaimus columbus*	*Pristionchus americanus*
5	*Pristionchus americanus*	*Plectus minimus*	*Rhabditis cf. terricola JH-*	*Hoplolaimus columbus*	*Rhabditis cf. terricola JH-*	*Hoplolaimus columbus*

Table 3.2b: Most abundant nematode species in fall 2016 by Treatment

Rank	None	Control	Phosphorus	Nitrogen	Both
1	*Paratylenchus projectus*	*Nematoda environ. sample*	*Paratylenchus projectus*	*Rhabditis cf. terricola JH-*	*Rhabditis cf. terricola JH-*
2	*Rhabditis cf. terricola JH-*	*Paratylenchus projectus*	*Nematoda environ. sample*	*Paratylenchus projectus*	*Paratylenchus projectus*
3	*Xiphinema floridae*	*Rhabditis cf. terricola JH-*	*Rhabditis cf. terricola JH-*	*Nematoda environ. sample*	*Nematoda environ. sample*
4	*Tylenchorhynchus claytoni*	*metagenome*	*Ditylenchus dipsaci*	*Pristionchus americanus*	*Acrobeloides sp. KJA*
5	*Coslenchus cf. pastor 1 JH-*	*Basiria cf. aberrans 1 JH-2*	*Xiphinema floridae*	*Tylenchorhynchus claytoni*	*Mylonchulus brachyuris*

Chapter 4: Konza Prairie Soil Studies

Abstract

This chapter tests the hypothesis that nutrient-enrichment benefits prokaryote species with high rRNA operon copy number more so than species with low copy number. This would support the overall theory that copy number variation is the genetic mechanism for the Growth Rate Hypothesis. Soil samples were collected from the Konza Prairie Belowground Experiment plots. Bacteria and archaea were enumerated by high-throughput sequencing of a portion of the 16S rRNA operon. Abundance and distribution of prokaryote species were combined with treatment and copy number data, then copy number was modeled as a function of treatment using a generalized linear mixed model. As opposed to short-term experiments, long term enrichment resulted in no difference in copy number probability across treatments.

Introduction

Background

Chapter 3 tested the effect of nutrient enrichment on short term soil cores. It is similarly important to test the effect of nutrient enrichment on soil microbe communities by copy number on longer time scales. Fortunately, the Konza Praire Biological Station (Figure 4.1a) has a series of belowground plots (Figure 4.1b) that are a part of the Konza Prairie Long-Term Ecological research program. Known as the Belowground Plot Experiment, it was initiated in 1986 to assess above and belowground responses of native tallgrass prairie communities to contrasting fire regimes and nutrient addition treatments (Coolon et al, 2013). It consists of a series of 8 blocks of 8 plots each (total of 64) with eight treatments. The treatments are burned/non-burned, mowed/non-mowed, and nitrogen/phosphorus/ N+P/control.

As expressed in Chapter 3, nutrient enrichment disproportionately affects fast-growing species. This is the core of the Growth Rate Hypothesis (GRH); fast-growing species have higher ribosome concentrations but are limited by phosphorus. In a phosphorus-enriched environment, the limitation is removed. Ribosomal RNA copy number is a plausible genetic mechanism of the GRH because it highly variable, an integral component of ribosomes, and phosphorus-rich. If soil microbe species have higher mean copy number in phosphorus plots compared to the control, this would uphold the GRH in soils and provide evidence for copy number variation as the genetic mechanism.

Objectives

The objective of this chapter is to test the hypothesis that nutrient-enrichment benefits prokaryote species with high rRNA operon copy number more so than species with low copy number. To accomplish this, I took soil samples from Konza Prairie that had been treated with excess nitrogen, phosphorous, or both, enumerated prokaryotes with high-throughput rRNA amplicon sequencing, and analyzed read counts to compare the rRNA copy number associated with read between the treatments.

Methods

Study Area:

Konza Prairie is a native prairie owned and managed by a joint effort of Kansas State University and the Nature Conservatory. It is the largest unplowed tallgrass prairie in North America. The Konza Prairie Biological Station has been in operation for over 40 years with long term ecological projects revolving around three key processes: periodic fire, ungulate grazing, and variable continental climate.

Soil Collection:

Soil samples were collected in May 2016 from the Konza Prairie Belowground plots. A subset of 16 plots were sampled. All subsets were from burned, and non-mowed plots as they best represent a native prairie but differed by nutrient treatment. The nitrogen treatment was 10 g/m^2 and the phosphorus was 1 g/m^2. The combination treatment (both) was the same amounts of nitrogen and phosphorus but added together on the same plot (Table 4.1). At each plot, five core samples were taken and homogenized. Samples were placed on ice in a cooler and returned to Grand Forks, where they were placed into a 17°C fridge the following day.

Soil Preparation:

Soil was removed from bags, broken down, and homogenized. Five grams of soil were added to 15ml ultra-pure deionized water and vortexed for 30 seconds to create a "soil slurry" that would be used for both Biolog EcoPlates and bacterial DNA extraction.

Biolog Ecoplates:

Assays were run as described in Chapter 3. Briefly, soil slurry was diluted and added to Biolog Ecoplates. Plates were measured every day for respiration and day seven was used for the final analyses. Community metabolic diversity (CMD) was modeled using a Poisson distribution and average metabolic response (AMR) was modeled using a Beta distribution. Comparing across Konza and Oakville, CMD was modeled as a function of prairie, treatment, and interaction with a Poisson distribution. AMR was modeled as a function of prairie, treatment, and interaction with a Beta distribution.

Molecular Work:

DNA extraction:

Soil slurry was prepped for bacterial sequencing using a modified protocol with the PowerSoil® DNA Isolation Kit (Mo Bio Laboratories). Instead of adding 0.25 grams of dry soil, 200 µL of soil slurry was added to the tubes initially and initial vortexing was reduced to 8 minutes. All extracted DNA was kept in a -20° freezer until PCR.

Mi_seq PCR:

The PCR pipeline was adapter-loci amplification using mi_seq primers with adapters. After optimization tests found the best annealing temperature, bacterial DNA was amplified by PCR using Q5 hot-start polymerase for 25-35 cycles. The primers and 515f_miseq and 816r_miseq were used to amplify a portion of the 16S operon. The initial denaturation step was at 98°C for 30 seconds, followed by denaturing steps of 10 seconds. Annealing occurred at 55° for 20 seconds and extension at 72° for 20 seconds. The final extension was 2 minutes.

Indexing PCR:

ZR-96 DNA cleanup kits (Zymo, Irvine, CA) were used to purify the PCR product. Then, the product was re-amplified using indexing primers. Each bacterial sample was given a unique combination of forward and reverse indexing primers. Q5 hot-start polymerase was used with 1 µL of product for 15 cycles. Second round products were cleaned and concentrated (Zymo DNA Cleanup and Concentrator), then quantified using a broad range assay for double-stranded DNA on the Qubit platform.

The Konza samples were added to a master library tube along with other experiments' samples at a concentration of 10 nM for a total volume of 100 μL. Master libraries were submitted to the Genomics Core at the University of North Dakota School of Medicine and Health Sciences for Illumina Sequencing using a MiSeq Reagent Kit v2 (600 cycles). Nematodes were also extracted and sequenced. However, they were in low abundance, so they could not be used for analysis. The low abundances may have been due to wet soil conditions or improper extraction of nematodes from the soil.

Data Analysis:

Bioinformatics:

Using the USEARCH software (version 11, Edgar 2010) and custom Python scripts (Python Software Foundation), Illumina sequence files were merged (-fastq_mergepairs), demultiplexed with the 515f primer to separate out bacteria (-fastx_demux), trimmed to a length of 250 (-fastx_truncate), filtered (-fastq_filter), dereplicated to a minimum unique size of 2 (-fastx_uniques), and cleaned of chimera (unoise3). Resulting sequences were identified as zero-radius operational taxonomic units ("ZOTUs"). Abundance and sample sources were added to each ZOTU and all ZOTUs were pooled.

From there, the zotus were matched to a dereplicated SSU Ref NR 132 dataset from Silva to create a per-sample read frequency table of prokaryotes that also had copy number estimates from the Ribosomal RNA Operon Copy Number Database (rrnDB). Only sequences that matched to the rrnDB were retained so that copy number information would be included.

Generalized Linear Mixed Models:

Sequencing reads were analyzed by multinomial cumulative ("proportional odds") logistic regression using PROC GLIMMIX (SAS, Statistical Analysis Software, Cary, NC). For both Spring 2016 and Fall 2017 data, ordered rRNA copy number of each sequence was modeled as the response variable as a function of treatment as the explanatory variable. Using the cumulative logit link, this multinomial logistic regression models the probability of "having a sequentially higher rRNA copy number". Thus, treatment means were computed as the probability of a read coming from a genome with rRNA copy number of 2 or greater.

Additional Metagenomic Analyses:

Rarefaction curves were constructed in R using package vegan (Oksanen et al 2018) using a step number of 20. Cluster analysis was done by sample using hierarchical clustering in R. Distance matrices were made using an Euclidean method and then clustered using the Ward method (ward.D). Resulting trees were converted to dendrograms and colored by treatment using the R package dendextend (Tal Galili 2015). Venn diagrams were constructed for bacteria species presence across prairies and within-study treatments. Diagrams were constructed manually or in R using package VennDiagram (Chen 2018).

Results

Community Composition:

Copy number correction resulted in noticeable shifts in some phylum abundances (Figure 4.2). Without copy number correction, Verrucomicrobia and Thaumarchaeota were underestimated; Actinobacteria, Bacteriodetes, Firmicutes, and Planctomycetes were overestimated; Proteobacteria and low abundance phyla were unchanged.

Phylum proportions across treatment showed shifts in abundance (Figure 4.3). Compared to the control, Verrucomicrobia were more abundant in phosphorus plots and to a lesser degree in both plots. Thaumarachaeota were less abundant in phosphorus and both plots. Proteobacteria was unchanged across treatments. Planctomycetes were less abundant in phosphorus plots. Chloroflexi was more abundant in phosphorus plots, while less abundant in nitrogen and both plots. Bacteriodetes were less abundant in nitrogen and both nitrogen and phosphorus plots. Actinobacteria was less abundant in phosphorus plots, while more abundant in nitrogen and both nitrogen and phosphorus plots.

For the proportions of the top ten species, forty percent were from a single species of Verrucomicrobia. Four of the top ten species were Proteobacteria, three were Actinobacteria, and the other two were individual species from Thaumarchaeota and Chloroflexi (Table 4.2). Rarefaction curves indicated that most of the diversity present was captured by the amplicon sequencing, as most lines plateaued (Figure 4.4). Cluster analysis by treatment failed to resolve treatment (Figure 4.5). Even long-term enrichment was not enough for soil community composition to be dominated by treatment effects.

Treatment effects on rRNA copy number:

Multinomial logistic regression of copy numbers showed no effect of nutrient enrichment on the representation of rRNA copy number among prokaryotes. Treatment coefficients were not different from the control, so reads from the nutrient-enriched plots had the same probability of coming from a high rRNA copy species than from the non-amended plots.

Prairie Compositions:

CMD was not different across treatment, but AMR was ($F_{3,12}$=5.86, p=0.0106) (Figure 4.6). The combination treatment (N+P) was significantly higher than the control. Prairie was significant for both CMD ($F_{1,61}$=47.66, p<0.0001) and AMR ($F_{1,60}$=53.48, p<0.0001). CMD was lower in all Konza treatments compared to Oakville while AMR was higher in all Konza treatments compared to Oakville (Figure 4.7).

For the general distribution of species (Figure 4.8), Konza prairie had the most unique bacteria species, while the Oakville studies had similar numbers high species overlap. Across treatments (Figure 4.9), the bulk of bacterial species were found in phosphorus or both plots in Konza, while the bulk of Oakville 2016 and 2017 species were found in all plots. Most bacteria were found in all three months. Across block, there were several unique bacteria species in each block for both Oakville 2016 and 2017.

Discussion

Community Composition:

Copy number correction was important. Otherwise, low copy phyla were underrepresented while high copy species were overrepresented. Some groups, like Proteobacteria, were unaffected.

The distribution of bacterial phyla across treatments saw an increase in Chloroflexi, a diverse phylum, in phosphorus compared to the control, as was Verrucomocrobia, a suspected copiotroph. In contrast, Thaumarchaeota and Actinobacteria were decreased. Archaea have low copy number and Actinobacteria are neither copiotrophs nor oligotrophs. Nitrogen plots had a large increase in Actinobacteria and almost disappearance of Chloroflexi. There was also the appearance of Candidatus Saccrobacteria The plots with both treatments had a modest increase in Actinobacteria, decrease in Thaumarchaota, and increase in Verrucomicrobia. In short, the plots with both treatments had a combination of the shifts found in the phosphorus and nitrogen treatments alone. Different bacteria phyla respond to different nutrient inputs, while some—Proteobacteria and low abundance taxa—remain static. The cluster analysis failed to resolve treatments as meaningful groups of comparison.

Treatment effects on rRNA copy number:

We expected a significant effect of the interaction between copy number and treatment, but that did not occur. Because an effect was seen for the short term in Oakville, but not the long term, communities must shift over time to adapt to the nutrient changes.

Biolog Ecoplates:

Respiration values from the Biolog plates indicated that differences in average metabolic response occur only at long time scales. Because CMD was lower in most Konza treatments and AMR was higher compared to Oakville, this suggests that while microbial diversity has gone down in long-term enrichment plots, the productivity has gone up.

Venn Diagrams:

Oakville cores had a more similar species assemblage to one another than to Konza prairie. Still, there was considerable overlap because both were native tallgrass prairies. The major differences in treatment between Konza and Oakville suggest that long-term enrichment shifts species compositions across treatments, while short-term does not. The comparison by month further supported this, as unique species were found in each month. The comparison by block suggests a strong location effect on bacteria distribution. In contrast, most nematode species were found in all blocks. Nematodes were much more cosmopolitan in their environmental distribution.

Conclusion

For Konza, there was no response in abundance by copy number to the treatments. Nematodes were isolated in too small of numbers for analysis. The metagenomic analyses showed that there are some shifts in specific phyla proportions. Some of these shifts are prairie specific, but rather may reflect a more general trend. Overall, individual cores were more similar to other cores by distance (ae, being in the same block) rather than time or even treatment. Even at for long term enrichment, treatment did not erase the diversity of soil communities.

The growth rate hypothesis was not upheld for long-term soil enrichment. Furthermore, copy number did not appear to predict high growth rate of high copy number species in enriched plots, as opposed to Oakville. This suggests that short and long-term dynamics of community response to nutrient enrichment are different.

Future Work:

Because nematodes were not successfully analyzed from long term Konza plots and short term experiments from Oakville showed a link between phosphorus and fast-growing species, further sampling of Konza nematodes is needed to determine if this link persists long term.

Figures and Tables

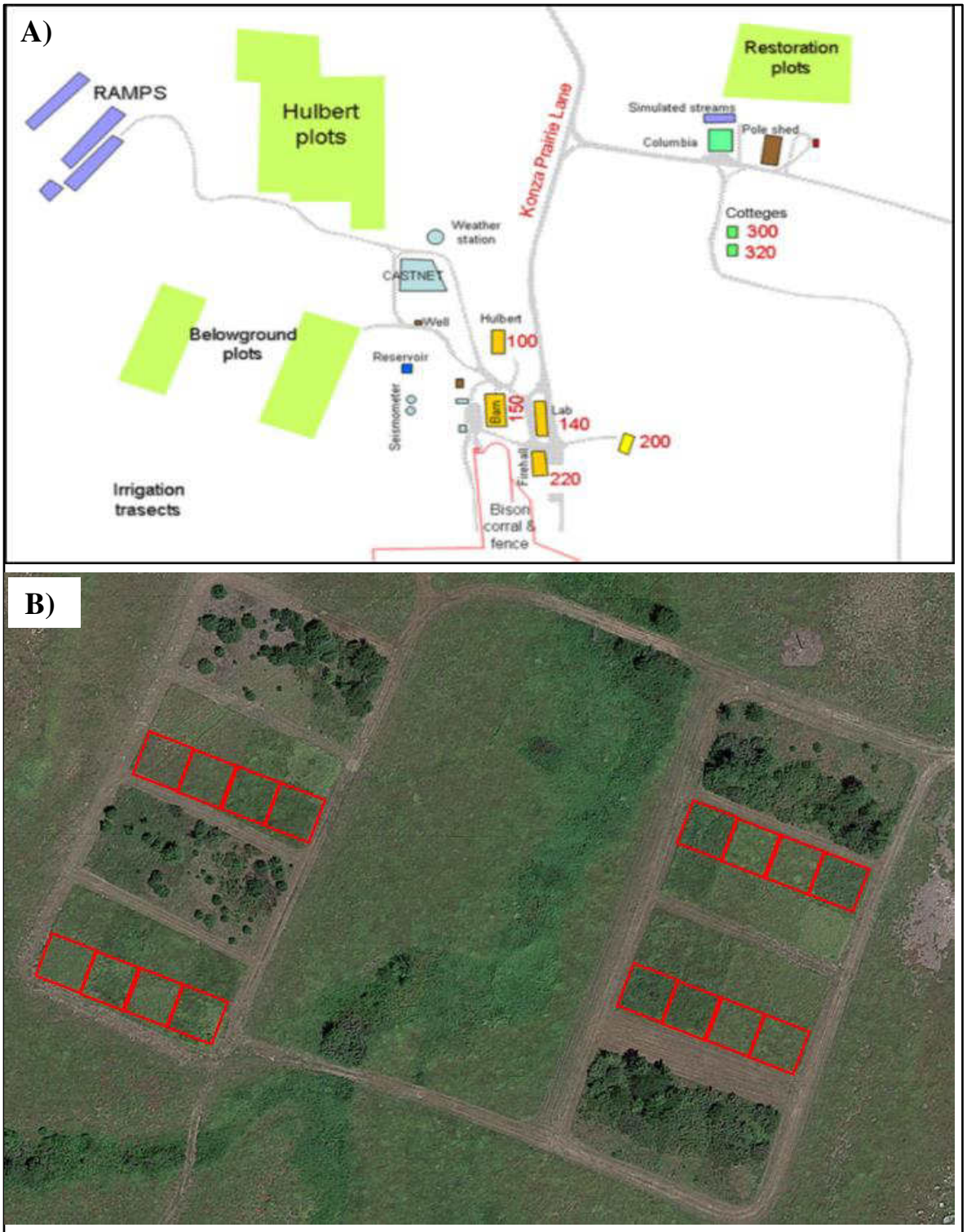

Figure 4.1: Konza Prairie. A) Overall map of Konza Prairie Biological Station, B) Satellite image of Belowground Plots, with plots sampled from in red.

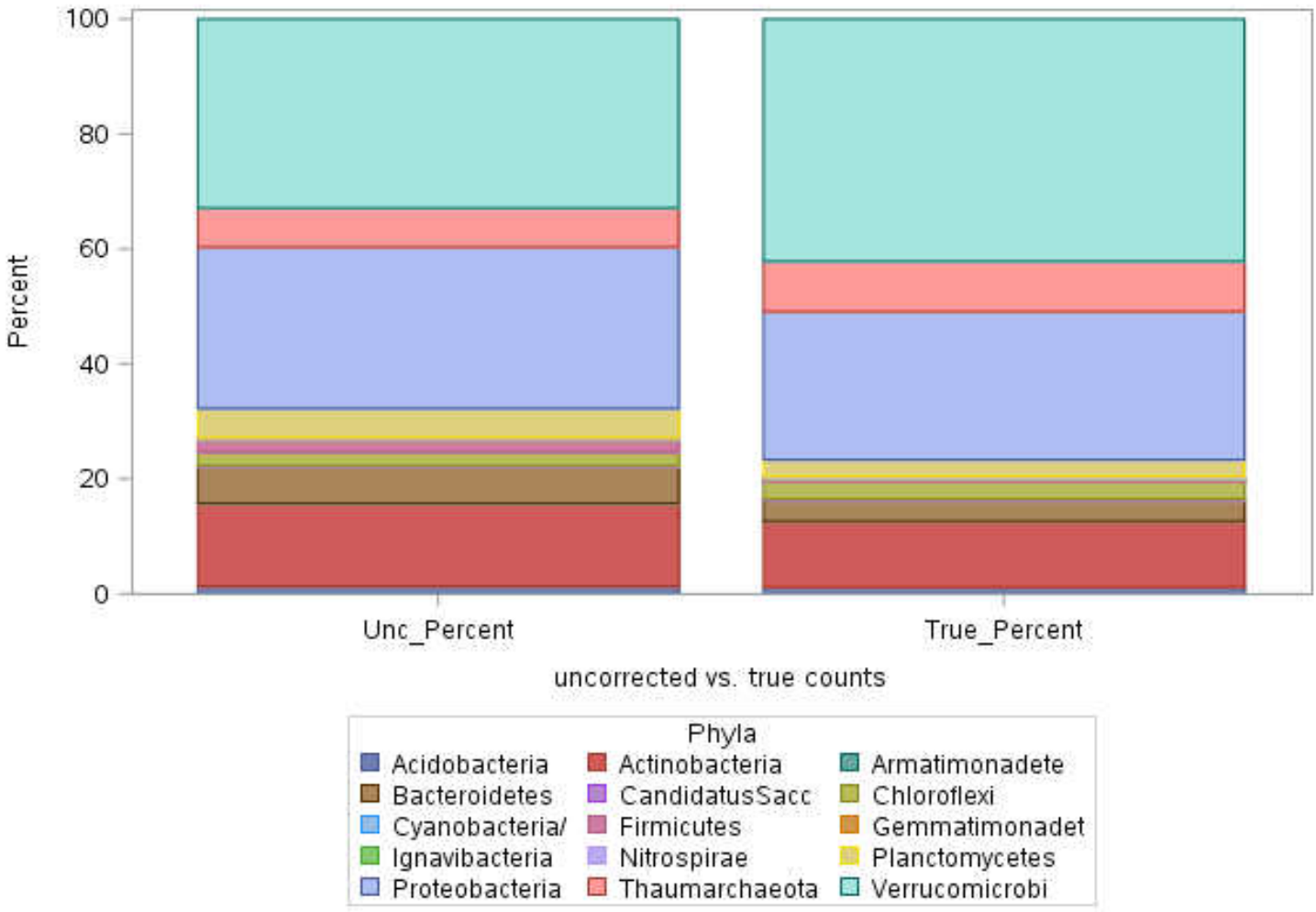

Figure 4.2: Bacterial Phyla Copy Number Corrections. Proportion bar plots of bacterial phyla across prairies. The left bar plot shows the uncorrected proportions, while the right plot shows the copy number corrected

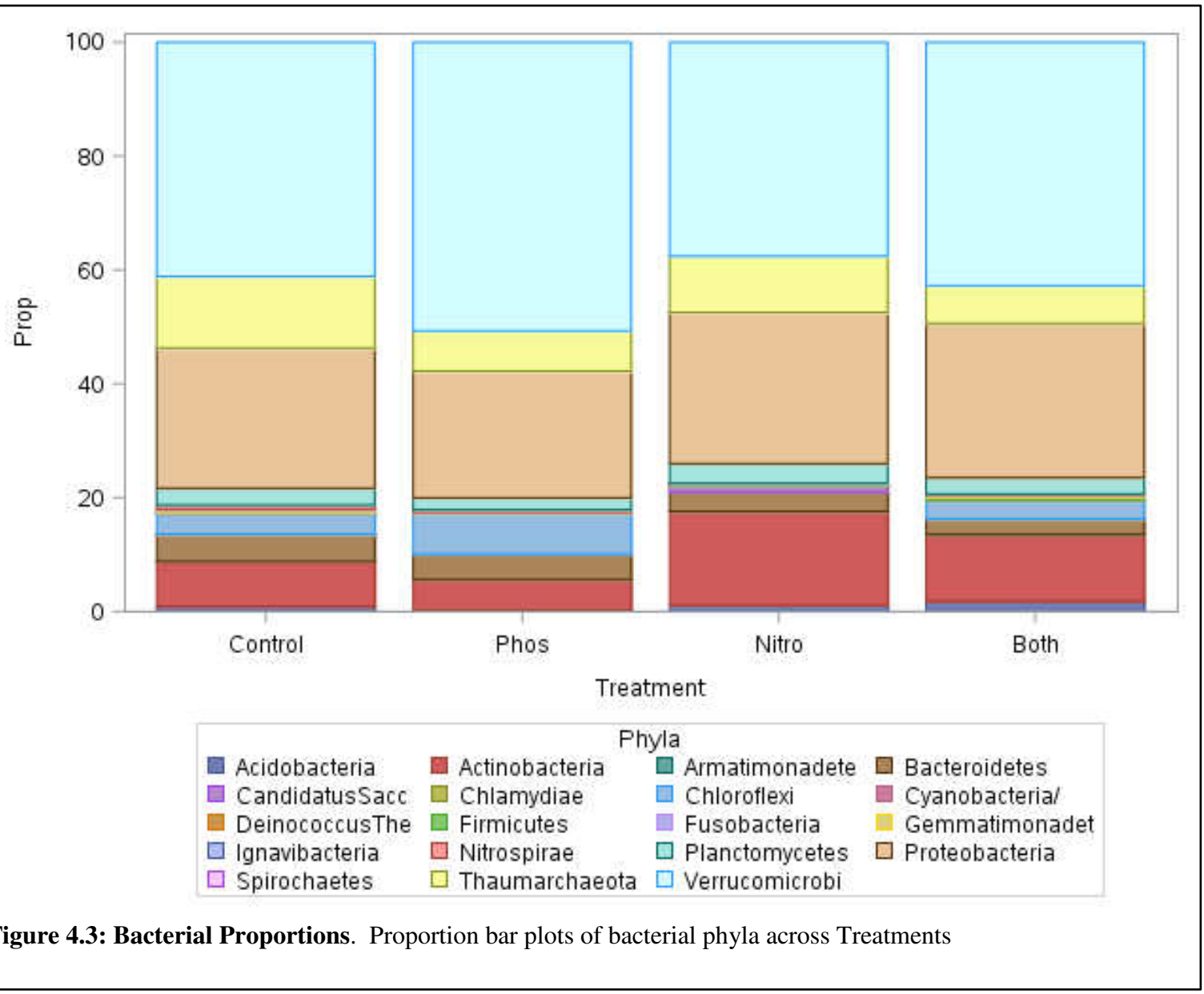

Figure 4.3: Bacterial Proportions. Proportion bar plots of bacterial phyla across Treatments

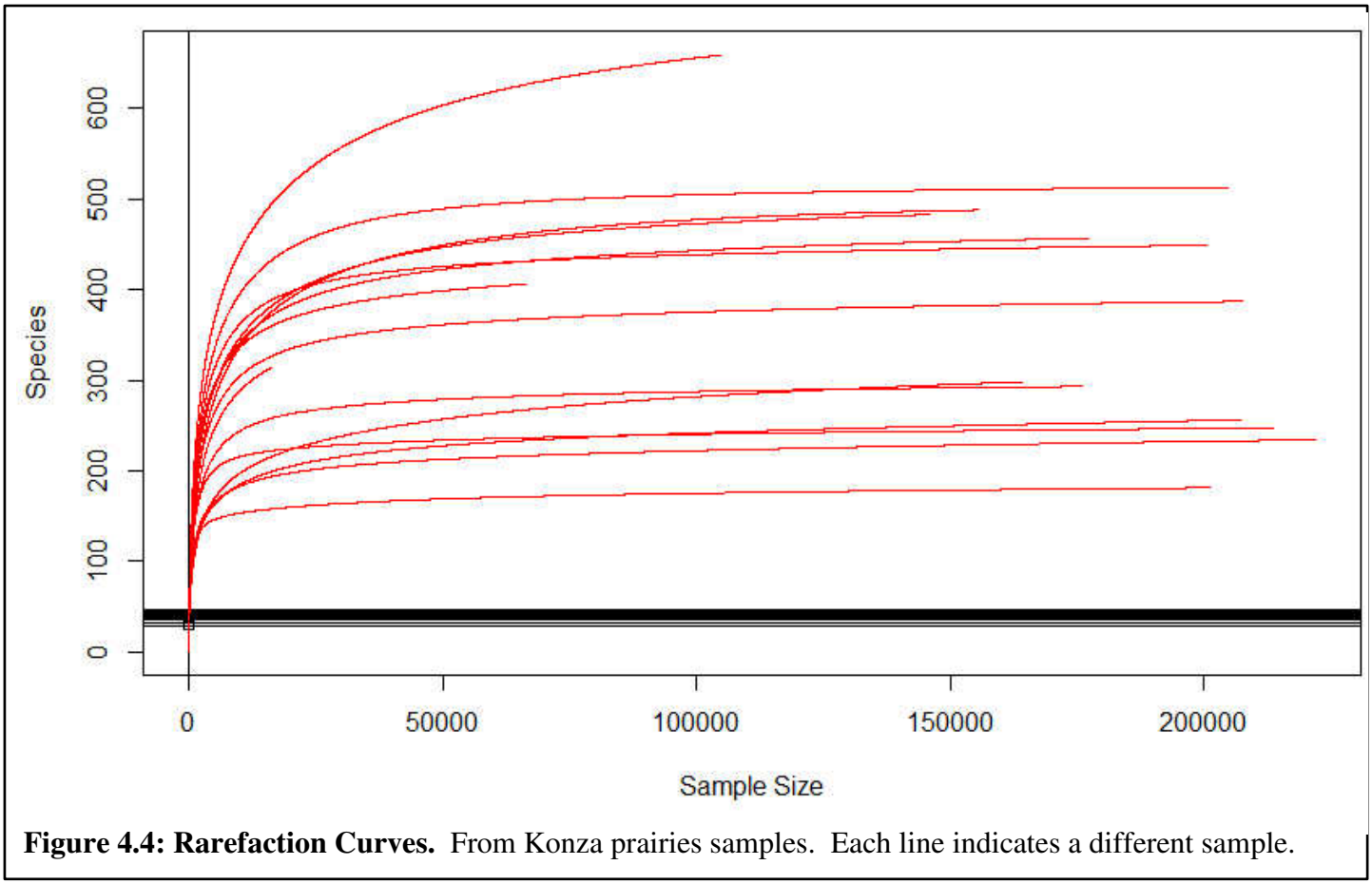

Figure 4.4: Rarefaction Curves. From Konza prairies samples. Each line indicates a different sample.

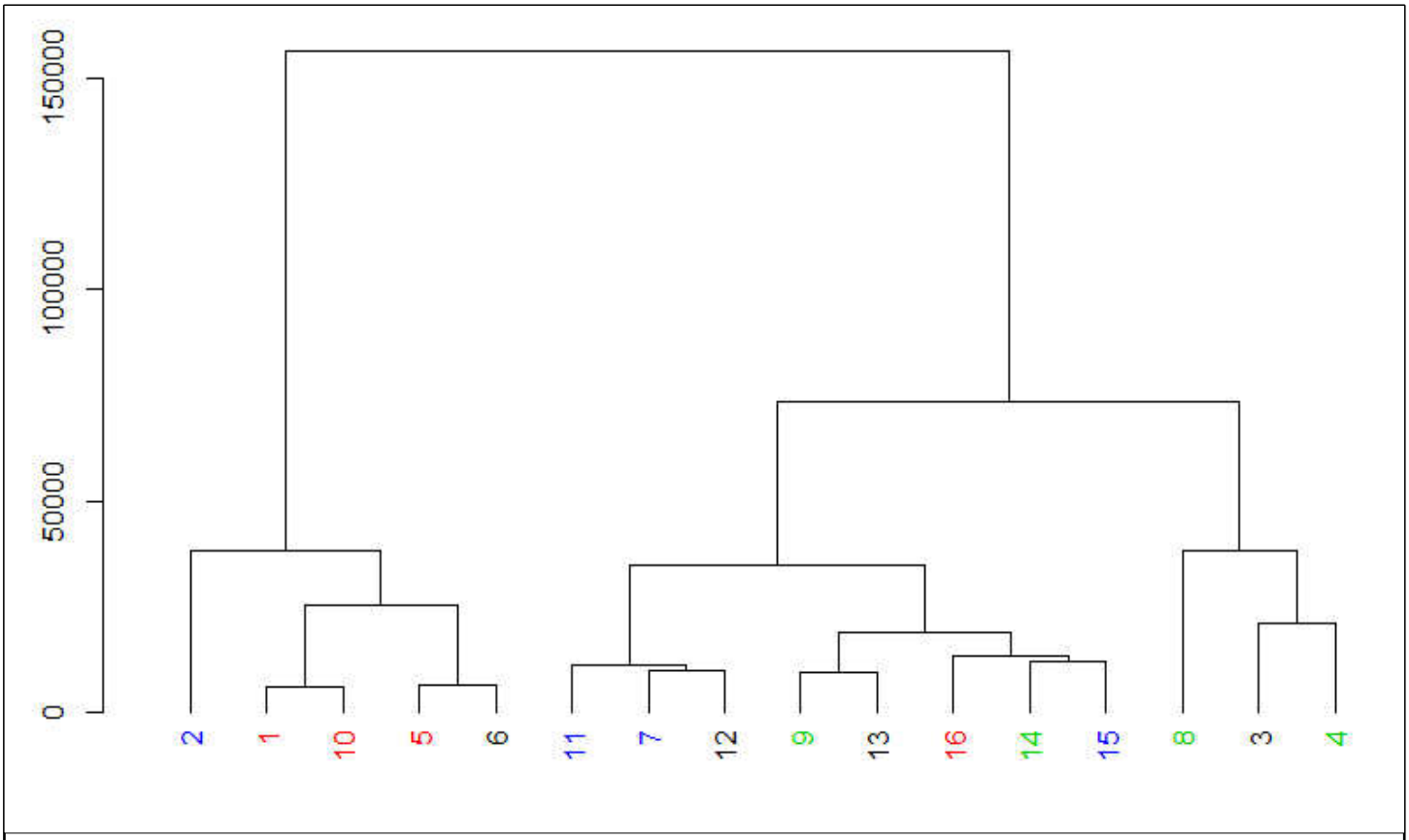

Figure 4.5: Cluster Analysis by Treatment. Konza Prairie soil samples. Control=RED, Phos=BLUE, Both=GREEN, and Nitro=BLACK

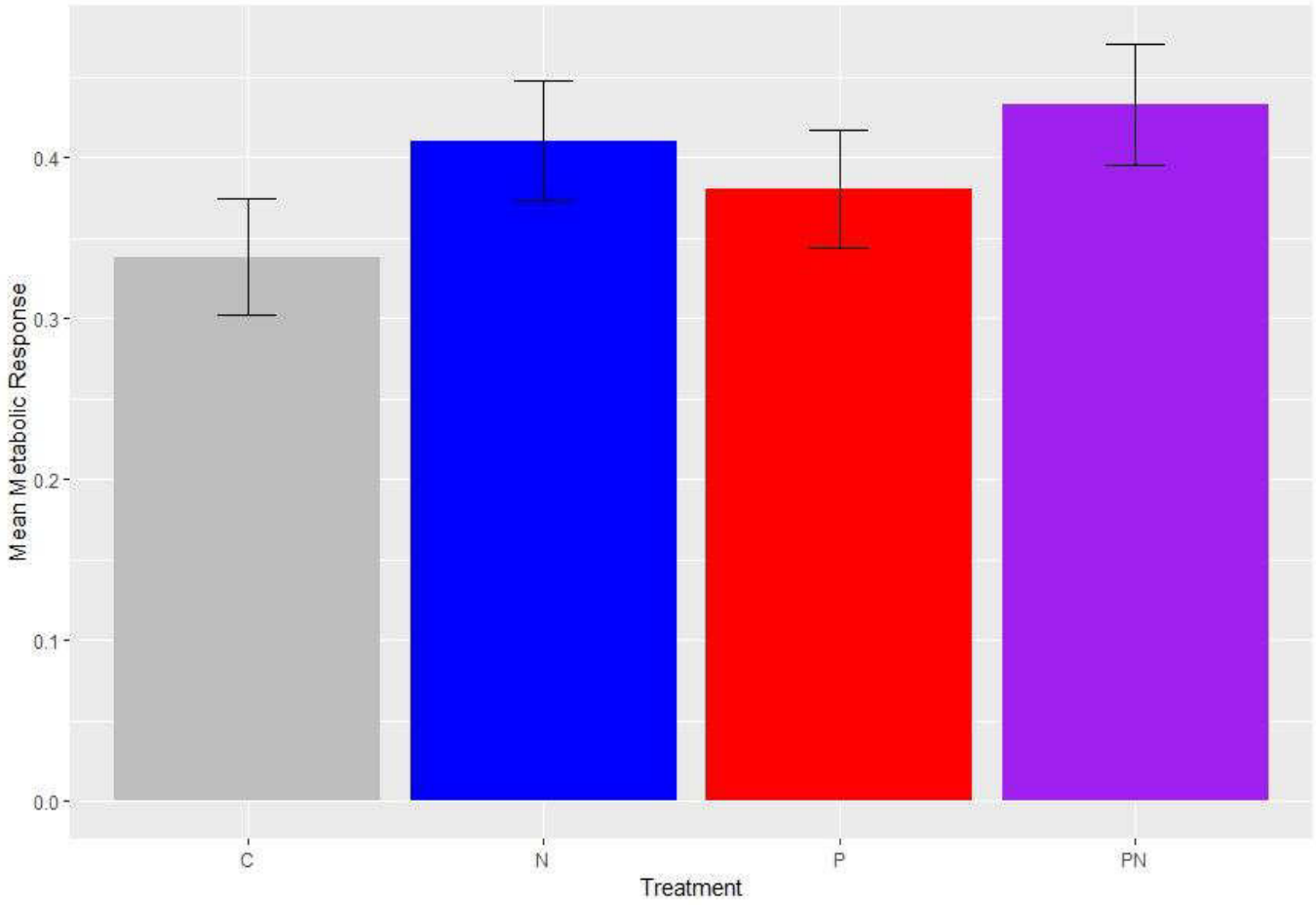

Figure 4.6: Konza Metabolic Response. Average metabolic response for Konza prairie across treatment.

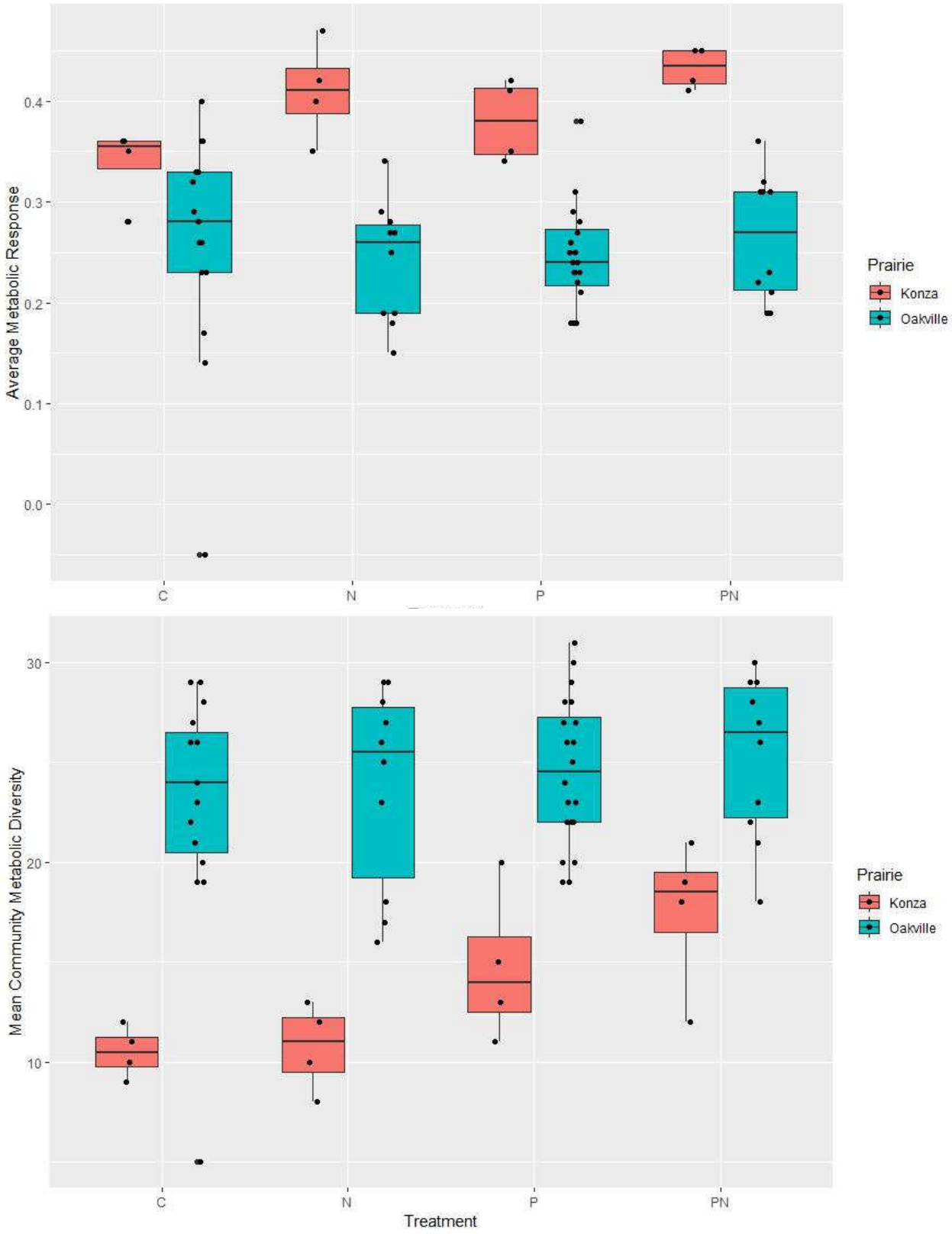

Figure 4.7: CMD and AMR Plots. Boxplots of Community Metabolic Diversity and Average Metabolic Response across treatment and prairie (Oakville and Konza).

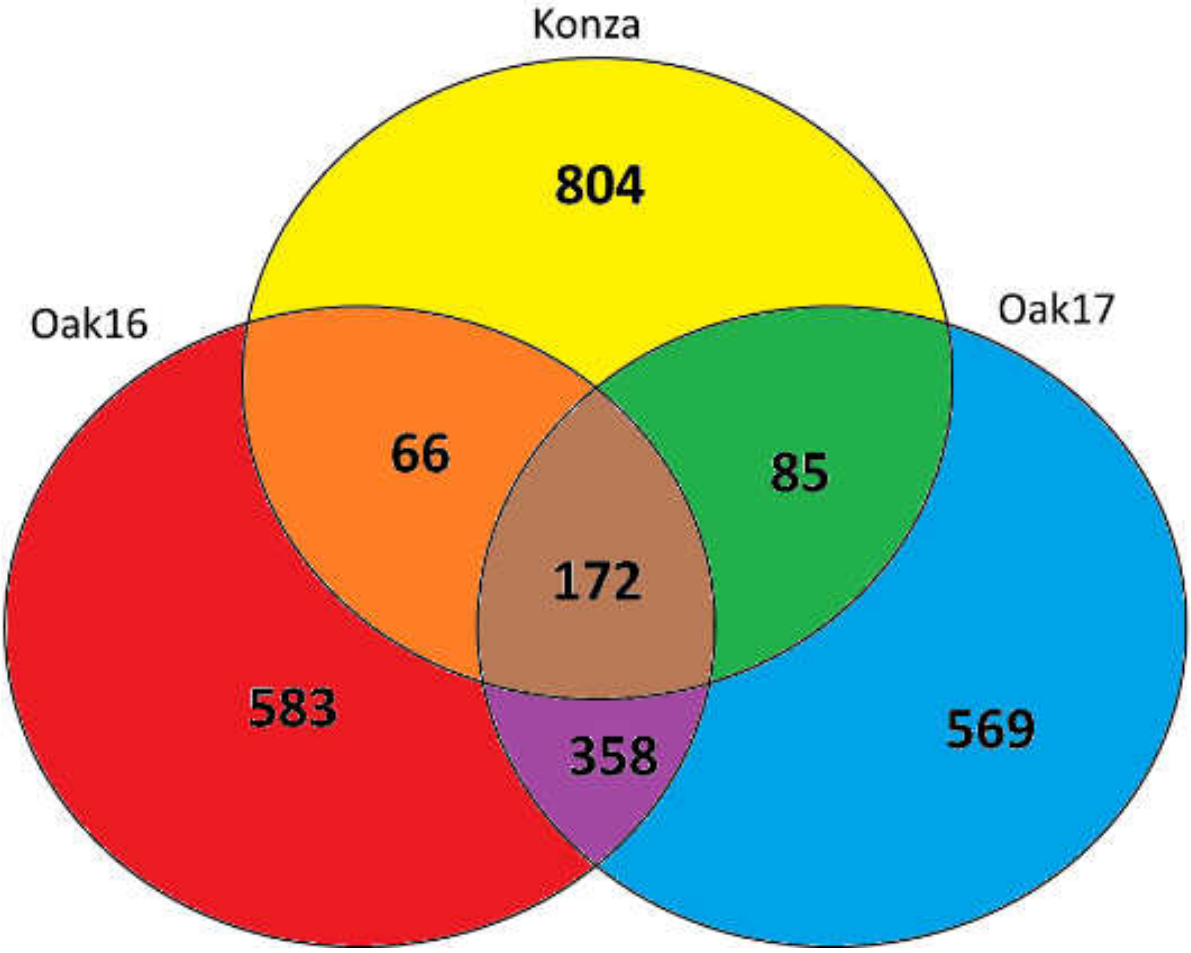

Figure 4.8: Venn Diagram across Prairies. Numbers indicate bacterial species.

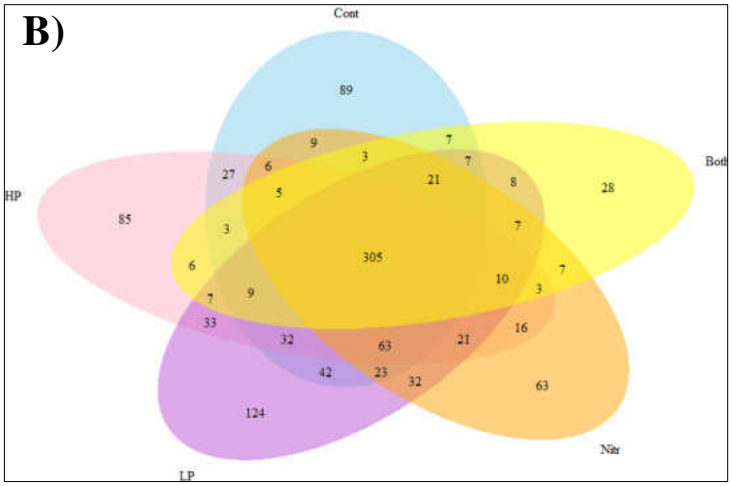

Figure 4.9: Venn Diagrams across Treatments. A) Konza Bacteria, B) Oakville 2016 Bacteria, C) Oakville 2017 Bacteria

Table 4.1: Treatment for Konza Prairie Belowground Experimental Soil Plots

Block 2		Block 4		Block 6		Block 7	
Plot #	Treatment	Plot #	Treatment	Plot #	Treatment	Plot #	Treatment
13	B, UM, C	25	B, UM, C	41	B, UM, N	53	B, UM, N+P
14	B, UM, P	26	B, UM, N+P	42	B, UM, C	54	B, UM, N
15	B, UM, N+P	27	B, UM, P	43	B, UM, P	55	B, UM, P
16	B, UM, N	28	B, UM, N	44	B, UM, N+P	56	B, UM, C

- B= burned, U=non-mowed, C=control (not treatment), P=phosphorus, N=nitrogen

Table 4.2: Most abundant prokaryotic species in Konza

Rank	rrnDB species ID	Phylum	Proportion
1	Candidatus *Xiphinematobacter* sp. Idaho Grape	Verrucomicrobia	0.41998
2	*Nitrososphaera viennensis* EN76	Thaumarchaeota	0.08726
3	*Rhodoplanes* sp. Z2-YC6860	Proteobacteria	0.06357
4	*Conexibacter woesei* DSM 14684	Actinobacteria	0.03442
5	*Brevefilum fermentans*	Chloroflexi	0.02914
6	*Pseudorhodoplanes sinuspersici*	Proteobacteria	0.02749
7	*Steroidobacter denitrificans*	Proteobacteria	0.02573
8	*Bradyrhizobium diazoefficiens* USDA 110	Proteobacteria	0.01691
9	Actinobacteria bacterium IMCC26256	Actinobacteria	0.01646
10	*Microlunatus phosphovorus* NM-1	Actinobacteria	0.01551

Chapter 5: Nematode Life History Traits and 18S rRNA Copy Number

Abstract

This chapter tests if growth rate and rRNA copy number (CN) are positively correlated in soil bacterivore nematodes. This would connect copy number to a phosphorus-mediated growth response in soil cores to determine if the Growth Rate Hypothesis was valid in soil nematodes and rRNA copy number is the genetic mechanism. Nematodes were obtained from local soil sources and the Caenorhabditis Genetics Center (CGC) and grown on NGM plates. Life-history traits were performed to obtain growth metrics. DNA was isolated from species for identification and copy number estimation. Relative copy number was estimated by sequencing equal numbers of all or partial species in a single Illumina sequence run. Full species assemblages found no correlation between copy number and growth rate. For partial assemblages, rank copy number was positively correlated with growth rate. This chapter's results combined with the nematode results of Chapter 4 connects the link between phosphorus enrichment, growth rate, and copy number. Therefore, there is some evidence of copy number as the genetic mechanism for the Growth Rate Hypothesis in bacterivore nematodes. The limitations of copy number estimation via bulk Illumina analysis include lack of equal amplification and inability to resolve some closely related species.

Introduction

Background:

The experiments from Chapter 4 tested the correlation between copy number and growth rate across nutrient-amended soil cores. Fast-growing nematodes were found to be more abundant in plots enriched with phosphorus. But unlike prokaryotes, nematodes do not have extensive copy number estimates. So, growth rate and copy number need to be examined to compliment the link between growth rate and phosphorus enrichment. It is, therefore, appropriate to turn to controlled-laboratory experiments to estimate growth rate and copy number.

The standard method for measuring growth rate characteristics is a life history table (Begon, Harper, & Townsend, 1990). A group of individuals of the same age--a cohort--is monitored across their lifespan. A table of their life history information can be used to calculate variables such as lifespan, generation time, and fecundity. Basic reproductive rate is a standard life history variable to use as a measure of growth rate.

Bacterivore nematodes are an ideal group for this study. They can be easily coaxed from the soil, kept alive, and grown on nematode growth medium (NGM) plates with OP-50 *E. coli* as a food source. They are expected to be the most affected by nutrient enrichment because they feed directly on bacteria. It is expected that fast-growing, high-fecundity nematodes will have a similarly high copy number.

Growth Rate Studies:

The nematode *Caenorhabditis elegans* was developed by Sydney Brenner as a model organism for genetics studies (Corsi, Wightman, & Chalfie, 2015). It has a number of useful traits: small size, fast life history, transparency, and small cell count. It was the first multicellular organism to its genome sequenced (C. elegans Sequencing Consortium, 1998) and the only one to have its entire cell lineage mapped (Sulston & Horvitz, 1977). It is also used as an animal model for microevolutionary studies (Delattre & Félix, 2001). A further selection of discoveries can be found in Table 2 of Corsi, Wightman, &Chalfie (2015).

Despite classification as a soil bacterivore, more recent data show that *C. elegans* is a colonizer of nutrient-rich environments, such a decaying plants (Félix & Braendle, 2010). In the laboratory, *C. elegans* is typically grown on NGM plates and fed a bacterial lawn of *E.coli* (Brenner, 1974). However, other methods, such as semi-fluid culture medium (Muschiol et al., 2009) have been used. *C. elegans* preferentially feeds on bacteria, resulting in a growth pattern that reflects the best prey of the mix (Darby and Herman, 2014). Yu et al (2015) found a trade-off between offspring number and lifespan between active and non-active bacterial prey.

While the majority of cultures studies have focused on *C. elgans*, a number of other species have been used, such as *Pelodera teres, Diplogaster nudicapitatus, Rhabditis terricola* and *Mesodiplogaster biform* for development (Sohlenius,1969), *C. briggsae* and *Plectus palustris* for food dependence and energetics (Schiemer, 1983), *Acrobeloides* for life cycle and reproductive output (Sohlenius, 1973), and *Pelodera* for energy budgeting (Marchant & Nicholas, 2012). Nematode assemblages are also used. Respiration and metabolic activity for

eight bacterivore nematodes revealed different peaks at different temperatures (Ferris, Lau & Venette, 1995). Rhabditidae has faster development than Cephalobidae (Ferris et al, 1996). Examining 6 species, Venette and Ferris (1997) found that growth rate was affected by both bacterial type and density. Bacterial density also affected interactions between different species of nematodes (Dos Santos et al., 2009). For *Mesorhabditus* and *Acrobeloides*, both had a specific bacterial prey preference, based on chemical properties of the prey and a similar offspring/lifespan tradeoff (Liu et al., 2017). This preference, furthermore, was not a symptom of familiarity (Liu et al, 2018).

Copy Number Studies:

The review of repeated genes, copy number, began with Long and Dawid (1980). In their seminal work, they examined the phenomenon of repeated genes across eukaryotes, including rRNA genes. Of the wide assortment of species compiled with rRNA copy number estimates, there were three nematodes, the most important being *C. elegans* (Sulston & Brenner, 1974; Ellis, Sulston, & Coulson, 1986). Much later, Bik et al (2013) estimated copy number for six nematodes from whole-genome shotgun data. Darby et al (2013) used a genetic algorithm to estimate relative copy number for nematodes from Konza Prairie. Pereira and Baldwin (2016) used cloning from vectors to get copy number estimates for a variety of *Cephalenchus* species. As far as known, there are no other copy number estimates for nematodes. Table 5.1 gives all known estimates of non-relative copy number.

Rhabditida is a diverse order that includes species that inhabit every environment, though rarely marine. They also parasitize plants and animals (Figure 5.1). Soil genera used in this study are phylogenetically distinct from obligate parasitic groups that include notable genera

such as *Ascaris* and *Brugia* (Figure 5.2). A more in-depth phylogeny shows the wide coverage of *Rhabidida* used in this study (Figure 5.3).

Objectives:

The objective of this chapter is to test the hypothesis that copy number is positively correlated to growth rate in bacterivore nematodes. To accomplish this, we obtained nematodes from soil samples and the CGC, determined life-history traits from cultured isolates, and then estimated relative copy number to test the correlation between copy number and growth rate metrics.

Methods

Culturing Work:

Soil Extraction and Identification:

Nematode growth medium (NGM) plates were seeded with a lawn of 100 µl broth solution of OP-50 *E. coli* and allowed to develop. Soil from Oakville Prairie, Mekanock Field Station, and the University of North Dakota was added in a ring and moistened with water (Figure 5.4). Plates were checked each day for nematodes and individual nematodes were picked off and added to a clean NGM plate. Each isolate was monitored saved if viable eggs were found. When a saved isolate reached a successful colony, 15-20 nematodes were plated into the cap of a PCR tube with 20 µL of worm lysis (alternate solution modified by Darby). The tubes were spun down and subjected to 3 cycles of freezing and thawing (from -80 °C) to facility better breakdown. GoTaq clear buffer and Protease K was added to the resulting lysis. The PCR tubes were then incubated in a thermocycler on a 60-95-4 (60 minutes at 60°C, 15 minutes at 95°, and holding at 4°) setting.

Upon successful lysis, PCR amplification was run using DreamTaq polymerase and the NR1/NR2 primers. PCR product was cleaned up and run through the Big dye Terminator cycle sequencing kit. The resulting product was cleaned again and the prepared isolate DNA was submitted for Sanger sequencing. Resulting sequences were blasted to NCBI. Only isolates that were either different species or the same species from a different location were retained.

Lab Isolates:

Twenty-four strains of nematodes were ordered from the Caenorhabditis Genetics Center (CGC). Upon arrival, adults from the small axenic NGM plates were replated to larger NGM plates with a bacterial lawn and allowed to grow up.

General Culturing:

Nematodes were grown in monoxenic cultures (single species of nematodes with *E. coli* as a food source) on NGM plates. The plates had 10 ml of agar and 100 µl of bacteria lawn. In general, 2 plates with 2 or more generations of viable eggs were kept for each isolate.

Life History Traits:

First, synchronized eggs were obtained from either bleach-synchronized plates or from clean adults that had laid eggs within the same day. Once juveniles hatched, they were added to new plates, for life-history trait (LHT) collection. The number of plates ranged from one to two and individuals ranged from five to twenty depending on the growth rate and reproductive strategy of the species. Slower growing and non-hermaphroditic species had more initial individuals plated. After the LHT plates were started, they were placed in a 25° incubator and only removed to count and replate.

Each day, the plates were checked for number of nematodes and the presence of eggs. If there were eggs present, the adults were picked off on a new plate and the eggs were counted

upon hatching. If eggs did not hatch, the plates were saved until at least ten days after being laid before being labeled as a failure. If a species failed to have viable eggs, or the eggs count was very low, the LHT round was considered a failure. That species was rerun until it had a successful LHT round. For some species which had a chronic issue of digging into the auger, thin NGM plates (only 4 ml of agar) were used. These plates were wrapped in Parafilm to slow desiccation.

Life History Tables:

When all the nematodes from an LHT round had died and all the egg plates were counted, a variety of life-history traits were calculated. The day (x), number of adults (nx), number of eggs (mx) and the number of eggs per adult (mx/nx) were computed. Next, the survivorship (lx) was calculated by dividing the number of adults still living by the starting number. Then the contribution to net reproductive rate during interval (lxmx) followed by the average number of offspring per individual during interval (xlxmx). From there, the lxmx and the xlxms were summed. The Ro (summed lxmx), was the net reproductive rate per capita. The generation time was the Ro divided by the summed xlxms. Finally, the R (ln Ro /Generation time) was the net reproductive rate per capita per generation.

Molecular Work:

Relative Copy Number Estimation:

Because of the difficulties with qPCR, absolute copy number estimation was switched to relative copy number estimation by Illumina sequencing of pooled samples. Each nematode isolate was first bleach synchronized. Then for each isolate, the earliest juveniles were isolated, visibly inspected by microscopy, and added to a 2 ml microfuge tube. This was continued until

the tube contained 10, day-one juveniles from all species from both local soil and the CGC.

This procedure was repeated for two more age classes. The second was adults that had laid eggs recently. The third was elder nematodes that had stopped laying eggs. Care was taken to avoid any individuals who exhibited endotokia.

Once the three tubes (juvenile, adult, and elder) were complete, they were centrifuged at high speed and aspirated. The tubes were run through DNA extraction (Zymo Microprep), and 18S DNA was amplified with PCR for 20 cycles using NF1/NR2 miseq primers and DreamTaq. The resulting 18S amplicons were cleaned and concentrated (DNA Clean and Concentrator-5, Zymo), then ran through indexing PCR. Each tube was given a unique forward and reverse primer and run for 20 cycles with DreamTaq.

Mixed Assemblages:

To assess how effective pooled sequencing was at recovering true individual count, 9 tubes were created of a mixed number of adult nematodes. The row categories measured the effect of evenness while the column categories measured the effect of relatedness. Each of the 9 tubes was run through DNA extraction, PCR, cleanup, and indexing PCR the same as the full tubes, with the exception that first-round PCR used 30 cycles.

Illumina Sequencing:

All twelve tubes from both the full and mixed assemblage experiments were pooled using 5 µL of each sample. The concentration of the mixed sample was determined using a Qubit assay (broad range, double-stranded DNA). The tube was added to other amplicon experiments and submitted to the Genomics Core at the University of North Dakota School of Medicine and Health Sciences for Illumina Sequencing using a MiSeq Reagent Kit v2 (600 cycles).

Bioinformatics:

Sequence Database construction:

Based on preliminary table construction and sequence analysis, a database of nematode sequences was created with twenty-one sequences (see Appendix's for additional information) *A. thornei* was added as an Acrobeloides_complex and an environmental *Mesorhabiditis* sample was added as Mesorhabodius_enviormental_complex. *O. tipulae* was treated as a complex of 3 species, *Ceph*. Sp ILVO-C as a complex of 4 species, *O. myriophilus* as a complex of 2 species, *Pangroliamus* sp. PS1169 as a complex of 2 species, and *P. teres* as a complex of 2 species. This accounted for 33 of the 34 species. The local *Rhabditophanes sp*. had no matches above 75% identity, so was dropped from the study. The 21 sequences were made into a database and ran with USEARCH's –otutab function at IDs of 1.00, 0.99, 0.98, and 0.97, with the 0.98 as the final ID selected.

Table Creation:

Using the USEARCH software (version 11, Edgar 2010) and custom Python scripts, the Illumina sequence files were unzipped, renamed, and merged. Then they were demultiplexed using the NF1 primer to separate nematodes from bacterial samples. Next, they were uniquified and pooled. They were then blasted to the custom database above using blastn command line using an id of 0.98.

Basic Relative Copy Number Estimation:

For the full assemblages (10 specimens of each culture from juveniles, adults, and elders), the number of sequences per individual was calculated by dividing the sequence abundances by the number of individuals. For most sequences, the number was 10, but the species complexes had 20-40. From there, each sequence number per individual was divided by the lowest sequence number per individual (*O. tipulae* for juveniles, *Meso-environmental* for adults, and *P. americanus* for elders) to get relative copy number. This gave the lowest copy number as one. Partial assemblages were done in the same way.

Genetic Algorithm Copy Number Estimation:

A genetic algorithm (Darby et al., 2013) was run in MatLab (The MathWorks Inc) was used for both partial and full assemblages as an alternative method of estimating relative copy number. The algorithm took in sequence count and specimen counts to perform a 'guess-and-test' over thousands of generations.

Life History and Copy Number Correlation:

Life history metrics (Ro, R, G, and eggs/adult) were taken from the life history experiments and natural log-transformed to approximate normality. Pearson's correlation tests were performed between the life history metrics and log(relative_CN) in R (R Core Team) for juveniles, adults, and elders. For the partial assemblages, relative copy number was averaged across treatments before log transformation and correlation. Correlation tests were also run using copy number estimates from the genetic algorithm for both partial and full assemblages.

For both full and partial assemblages, rank copy number was also calculated and used in correlation. First, copy number was ranked for each species within an assemblage. The highest copy number got a rank of 1, while the lowest got a rank of n, where n was the total number of species present in that assemblage. Second, rank was averaged across multiple samples; juveniles, adults, and elders for the full assemblages, and assemblage 1-9 for the partial assemblages. Then, the inverse average rank was calculated by the equation (n/n)*(t-n), where n was the rank and t was the total number. This made higher ranking species have larger numbers. Finally, the correlation between inverse average rank and the natural log of Ro in was calculated in R using a non-parametric Spearman Rank correlation test. Significant results were plotted.

Other Methods:

Survivorship curves were calculated for each isolate by plotting the percent survivorship each day for the full length of the experiment. The highest day was the day the last species had the last individual die. Local and CGC species were done separately.

A phylogenetic tree was constructed for the sequences used in the custom database. Sequences were aligned with ClustalW in Mega7 (Kumar, Stecher, and Tamura 2015) and a Maximum Likelihood tree was made using default settings. A subtree using only sequences from the partial assemblages was constructed as well. Phylogenetic signal of copy number was tested using the R package phytools (Revell 2012) on each nexus tree using the lambda method. for full and partial assemblages.

Results

Cultured Nematode Isolates:

For the soil isolates, there were a total of 17 isolates representing 10 species from local soil sources. The majority came from Oakville Prairie, but a few came from other locations (Mekanock Field Station and University of North Dakota). Isolates were duplicated if they came from different soil sources. Based on previous samplings from Oakville Prairie and Konza Biological Station (a tallgrass prairie in the Flint Hills region of Kansas), multiples species from the genera *Oscheius*, *Rhabditis*, *Mesorhabditis*, *Protorhabditis*, *Panagrolaimus*, *Pristionchus*, and *Acrobeloides* were expected to be found. Of those, only *Protorhabditis* isolates were not found.

Twenty-four species were ordered from the CGC, chosen for their feasibility of growing on plates and gathering life-history traits. Species that were found locally were not ordered, save for possible overlaps when specific species names were not known.

Of the 35 species, 15 genera were represented. The families of *Cephalobidae, Rhabditidae, Panagroliamidae, Neodiplogasteridae,* and *Alloinematidae* were all represented by at least one species. All families resided in order *Rhabditida*, class *Chromadorea*. The NCBI taxonomic browser and the CGC directed the genus *Dolichorhabditis* to *Oscheius*, though WORMbase (citation) listed it as its own genus. *Operculrhabditis* was directed to *Mesorhabditis*, but LabBio (citation) listed it as a sub-genus under Rhabditis.

Relative Copy Number:

The full and partial assemblages had copy number estimates for 21 species/species complexes (Tables 5.2 and 5.3). There was no correlation between copy number and net reproductive rate per capita per generation (R), intrinsic rate of natural growth (Ro), generation time (G), or eggs-per-adult in juveniles, adults, or elders. Similarly, there was no correlation between rank copy number and Ro. For nematodes in the partial assemblages, there was no correlation between copy average copy number across treatments and R, Ro, G, and eggs-per-adult. The genetic algorithm estimated copy number for both full and partial assemblages (Figures 5.5 and 5.6). Using the copy number estimates from the genetic algorithm, there was no correlation for both partial and full assemblages.

For nematodes in the partial assemblages, there was a significant, positive correlation between inverse rank copy number and Ro (rho = 0.594, p-value = 0.04168). Log Ro increases as a function of rank copy number (Figure 5.7).

Survivorship Curves:

The average survival time across all species was 21.4 days. There was a high degree of variation, with the shortest-lived species lasting 9 days and the longest-lived species lasting 42 days. In earlier trials (not shown), there were individuals in some species that lived almost three months. Overall, most species tended towards Type III survivorship (Figure 5.8), characterized by low survivorship in young and high survivorship in old individuals. However, some species approached Type II, steady survivorship. Given the artificial environment, it is expected that Type III is the true mode for all species.

Phylogeny:

The final sequence database resolved into a tree (Figure 5.9) that accurately captured most expected phylogenetic relationships. *Rhabiditis* was non-monophyletic, interspersed with *Rhabditella, Oscheius*, and *Pellioditis. Acrobeloides* and *Cephalobus* were difficult to distinguish, which accounts for the difficulty of resolving their sequences in the OTU tables. Copy number had no phylogenetic signal across the sequence tree for both full and partial assemblages.

Partial Assemblages:

Abundances were recovered for most samples across the partial assemblages (Table 5.3). Assemblages with high evenness had enough sequence capture for all species except *P. rigidus*. Moderate evenness had similar success, except in the phylogenetically diverse assemblage, where only *O. akosreti* and *P. pacificus* were obtained. Low evenness had difficulty in all three phylogenetic categories. *P. rigidus* had low capture overall and no sequences directly matched to it. There were 24 sequences that matched to *Panagrolaimus_sp_PS1159* in A9, so were attributed to *P. rigidus*.

Behavior:

Nematodes were observed during the LHT trials and general behavior was noted (supplemental Table S5.1). All species were able to grow and have successful egg development, though some were notably harder than others and required more attempts.

Discussion

The results of the full assemblages were negative. There was no correlation between copy number and any of the growth rate metrics. Because of the difficulty of getting sequence matches to the known species that went in, the quality of the resulting OTU tables is questionable. Attempting copy number estimation on many closely related species at once with Illumina sequencing was not accurate. This was likely due to the number of species used at once, ambiguities in species identification (some species from the CGC only had genus designation or had a species name used nowhere else in the literature), or species not being able to be resolved differently by the portion of the 18S used.

Three different life stages were used for copy number estimation. Juveniles were attempted because they should be close to the same size and not have developed gametes. However, they were very small and hard to obtain. Elders, chosen at a stage that should no longer have gametes, were hard to obtain. Many would die or exhibit endotokia before they could be picked. Furthermore, there was no guarantee that they did not still have gametes. Finally, some species had long-lived elders that got very large, while others did not. They would further complicate copy number estimation, as it was copies per individual. Large elder species would have significantly higher copy number just by size alone.

The correlation curve for juveniles and elders was flat. However, for adults, the curve was positive if not significant. This follows with the trend with the absolute copy number and rank copy number in the partial assemblages. Adults, despite producing eggs, are the best choice for copy number estimation trials.

Because of this, the selection of adults for the partial assemblages worked. While copy number estimation was not initially the main objective of these partial assemblages, the failure of good copy number estimation from the full assemblages warranted an attempt. For each assemblage (1-9), species output was linked to the species hit as best as possible. All species had a good match, except *A. thornei* and *P. rigidus*.

Highly conserved species were accurately pulled out from the partial assemblages. However, this was due to selecting an appropriate genus *Ochesius*, which can readily be distinguished with the 18S. If *Acrobeloides* had been selected, it would have failed.

Similarly, the general trouble in gaining sequences in the phylogenetically diverse assemblages appears to be due to species choice. *A. thornei* and *P. rigidus* were found in only one assemblage each, at the lowest abundance. *M. lonspiculosa* was found in two, both at low abundance.

High evenness modeled the full assemblage for a small sample number. Medium evenness modeled expected environmental samples. Low evenness modeled rare species. Phylogenetically conserved species were all closely related, within the same genus. Phylogenetically intermediate species had a mix and dispersed had as highly disparate samples as the study could get.

The moderate intermediate assemblage, which was meant to mimic natural abundances, was successful outside the phylogenetically diverse assemblages. Therefore, this method should retrieve species accurately form environmental assemblages. However, the low evenness samples, which also may be found in some environments had difficulty recovering many of the

rare species. Therefore, while Illumina 18S amplification will recover a good deal of species diversity, it is not guaranteed to obtain all rare species.

The genetic algorithm gave different copy number estimates than the basic calculations. Those estimations were not correlated with growth rate either. The effectiveness of the genetic algorithm is not known because copy number estimates were not known beforehand.

Behavior:

The selected nematodes were a diverse group (Table 5.4) that ranged across 17 general and 5 families. Eggs per adult ranged from 0.64 to 508.80. Some species exhibited endokia or digging. Others were particularly active, even at low nutrient levels. Because the species were phylogenetically and behaviorally diverse, it is not surprising that there was a high variation in copy number and growth rate. Despite the variation, growth rate still predicted copy number to some degree. The results are more robust given the species diversity.

Conclusion

Copy number was positively correlated with growth rate in bacterivore nematodes in a limited capacity. Relative copy number estimation worked best with adult nematodes in small assemblages. The genetic algorithm was expected to give better copy number estimation than raw counts alone but requires accurate data to begin with. This study completes the link between copy number, growth rate, and phosphorus enrichment. Fast-growing nematodes were more abundant in phosphorus-enriched soil cores from Oakville Prairie. Fast-growing nematodes had higher copy numbers in the culture studies. This provides evidence for both the Growth Rate

Hypothesis in soil nematode communities and rRNA copy number variation as the genetic mechanism.

Copy number estimation for nematodes and other metazoans remains an unsolved technical problem. In the future, copy number should be estimated by using more precise qPCR methods if the issues with amplification efficiency can be mediated. Furthermore, as only nematodes at with Colonizer-Persister values of one to two were examined, obtaining growth rate and copy number estimates of the full range of soil nematodes is important to access their correlation as a general aspect of nematode genetics and ecology. This too, is a technical problem, as the nematodes that can easily be cultured are bacterivores with low CP-values.

Figures and Tables

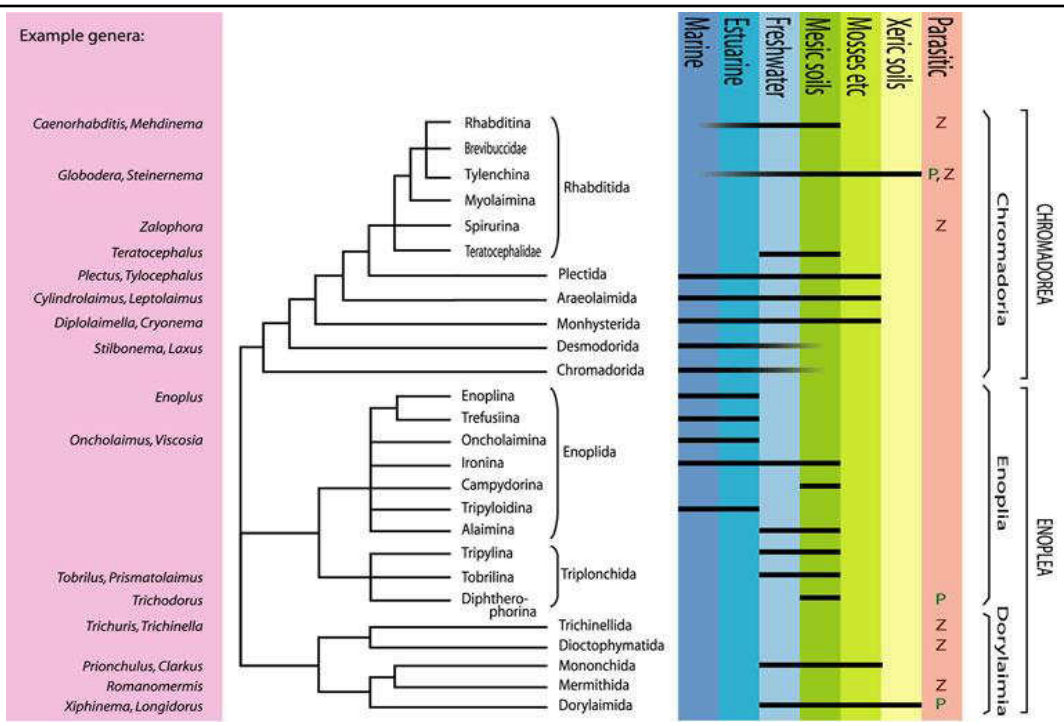

Figure 5.1: **Nematode Phylogeny**. Summarized SSU phylogeny of Nematoda with example taxa, ecological range and higher classification (from De Lay, 2006; adapted from De Ley & Blaxter, 2002). P = plant-parasitic, Z = zoo-parasitic.

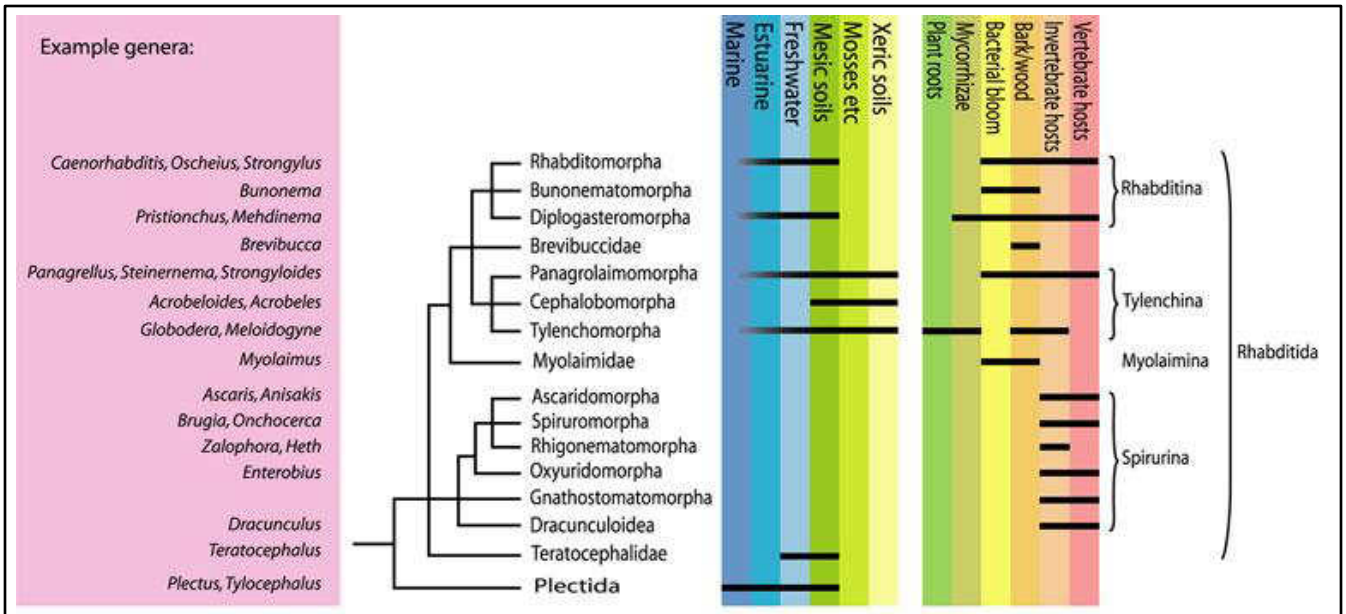

Figure 5.2: Summary Rhabditid Phylogeny. Summarized SSU phylogeny of Rhabditida with example taxa, ecological range and higher classification (from De Lay, 2006; adapted from De Ley & Blaxter,2002). Note the use of infraorder names (ending in -omorpha).

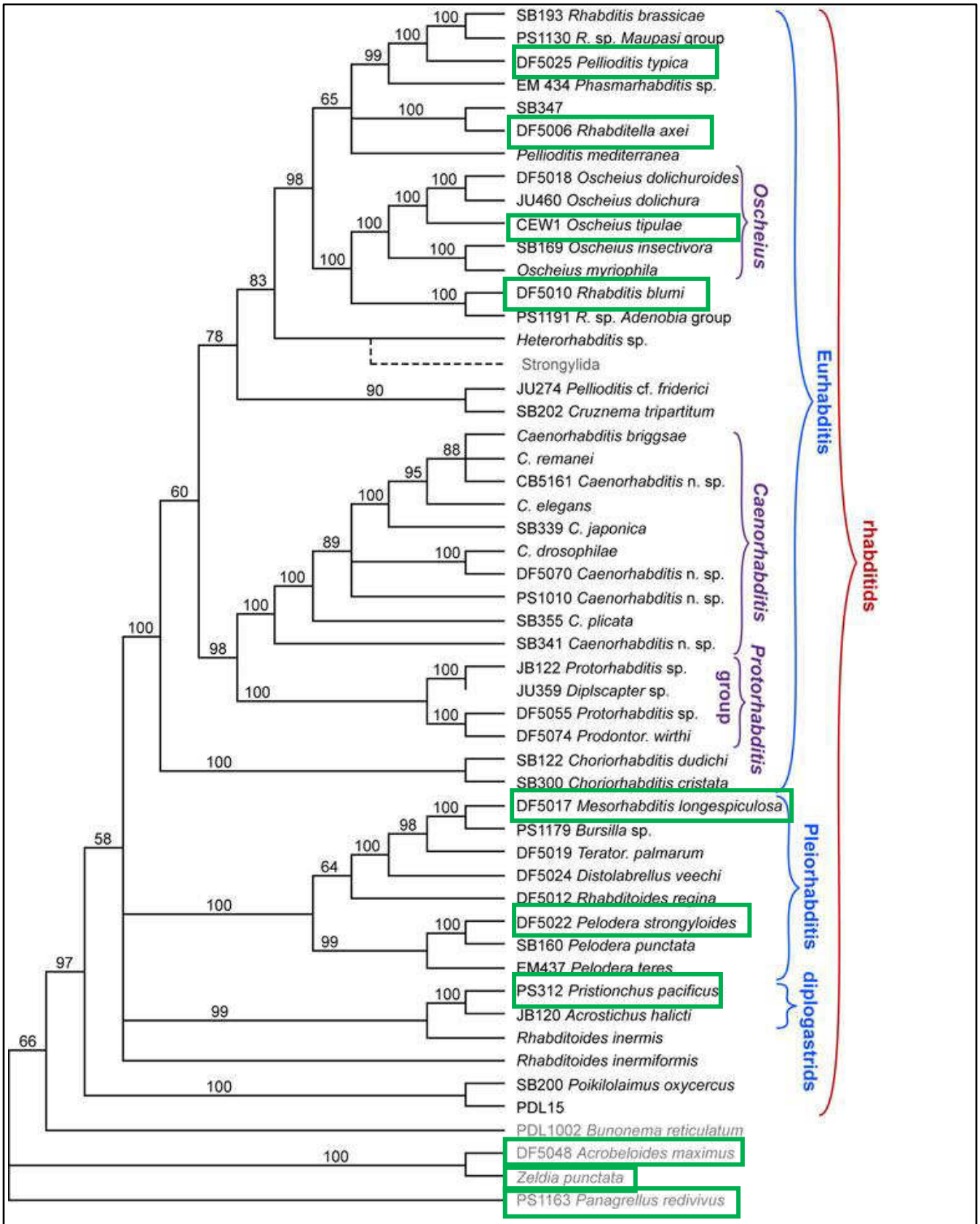

Figure 5.3 Detailed Rhabditid Phylogeny. Rhabditid phylogeny as inferred by weighted parsimony jackknife analysis using DNA sequences from three genes (SSU and LSU rRNA). From Kiontke & Fitch, 2005. Green outlines show genera used in current study (10/15).

Figure 5.4: Nematode Growth Medium Plate with Soil.

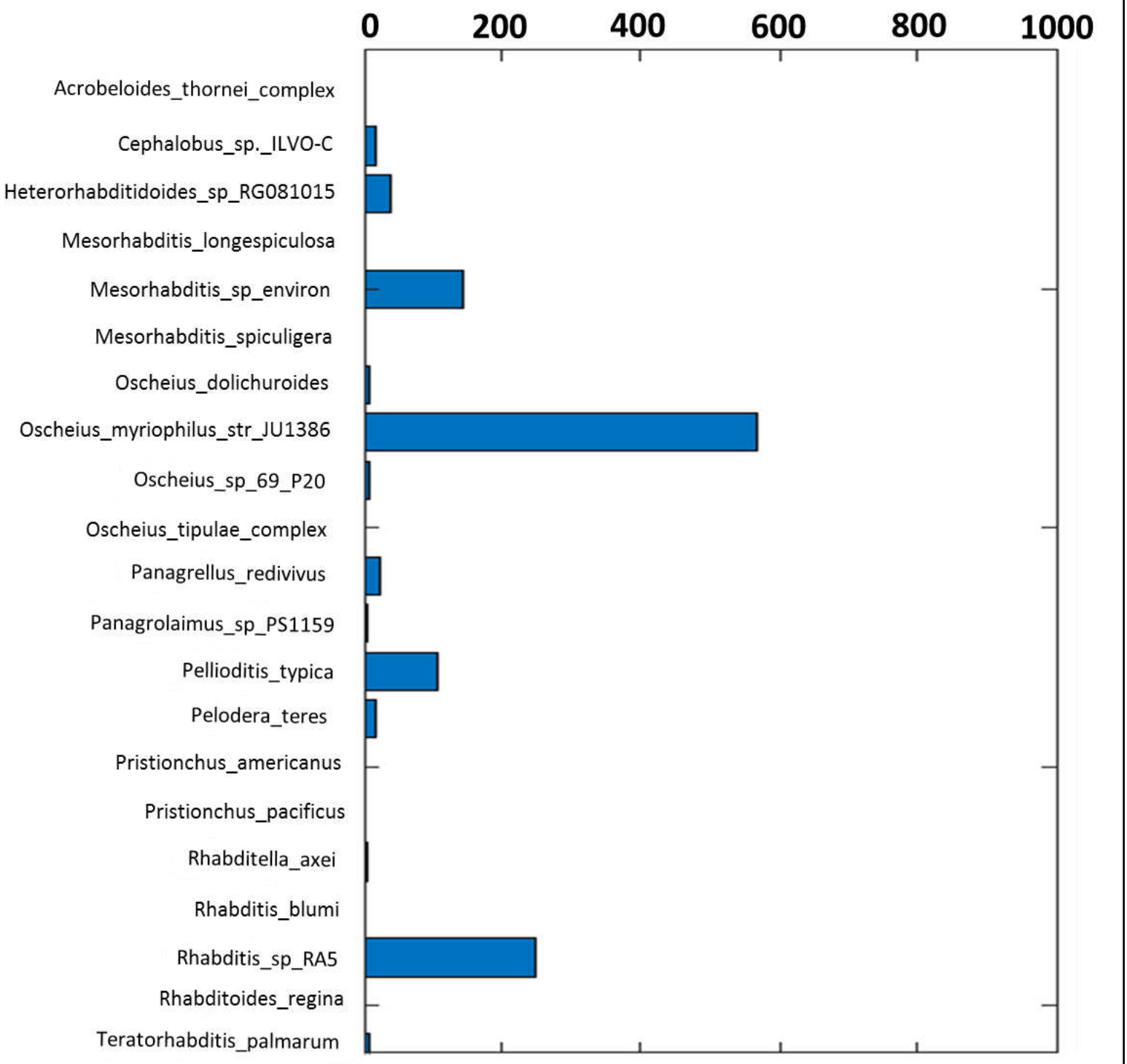

Figure 5.5: Full Assemblage Genetic Algorithm Copy Number Estimates. Relative copy number estimates using the genetic algorithm for full assemblages. Entries without names were calculated to have relative copy number of 1.

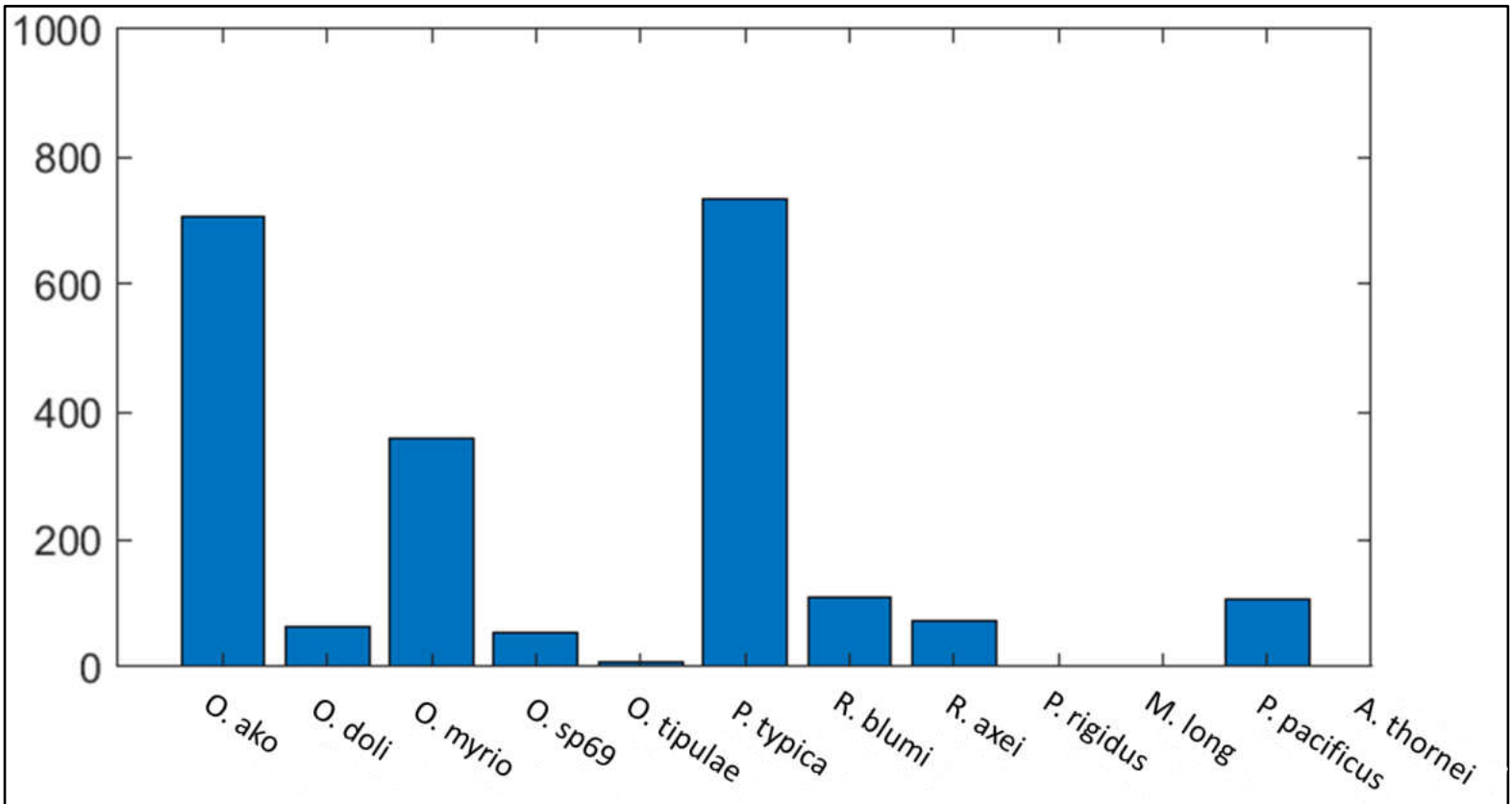

Figure 5.6: Partial Assemblage Genetic Algorithm Copy Number Estimates. Relative copy number estimates using the genetic algorithm for partial assemblages. Entries with a bar height that was no visible had relative copy

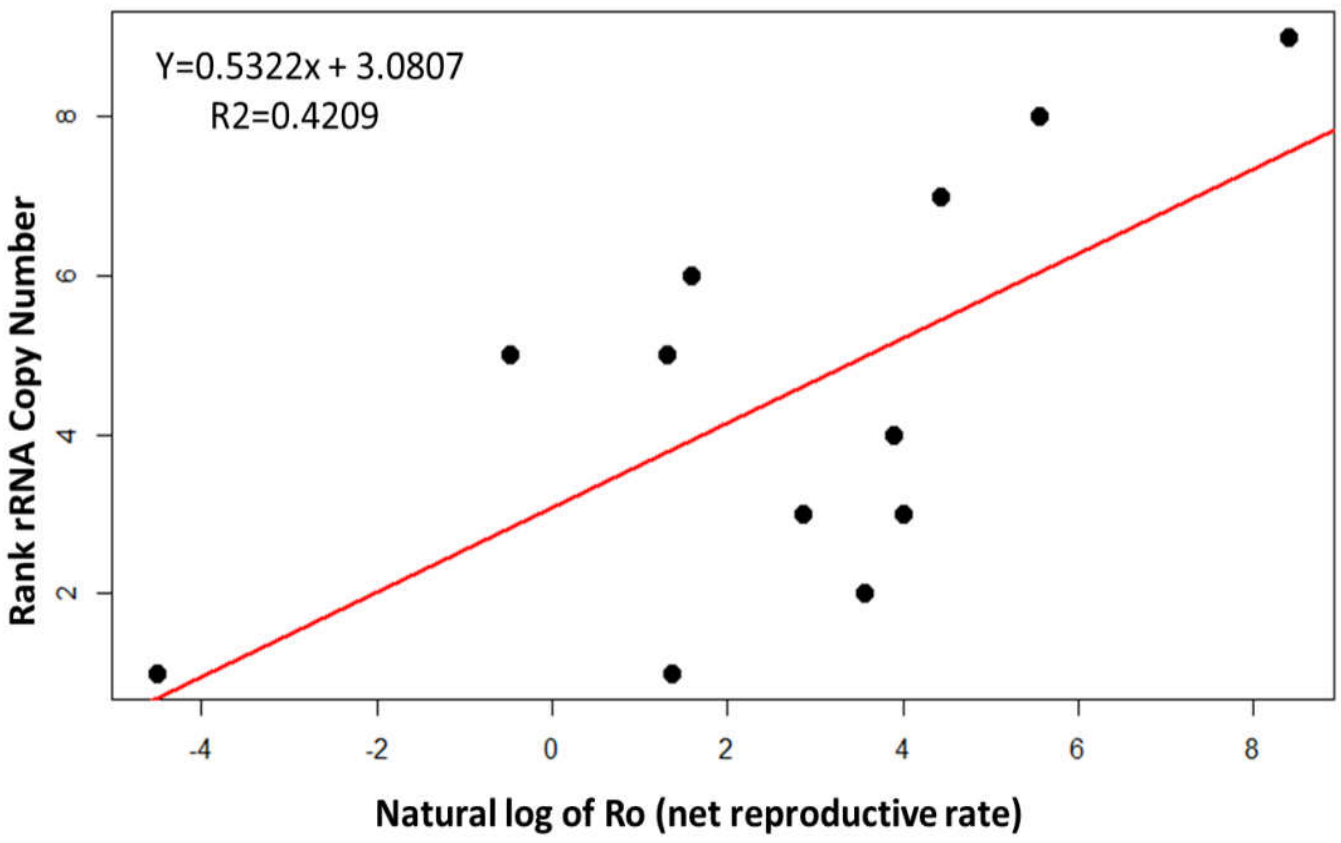

Figure 5.7: Rank Copy Number Correlation. Graph of copy number to growth rate correlation in partial assemblages.

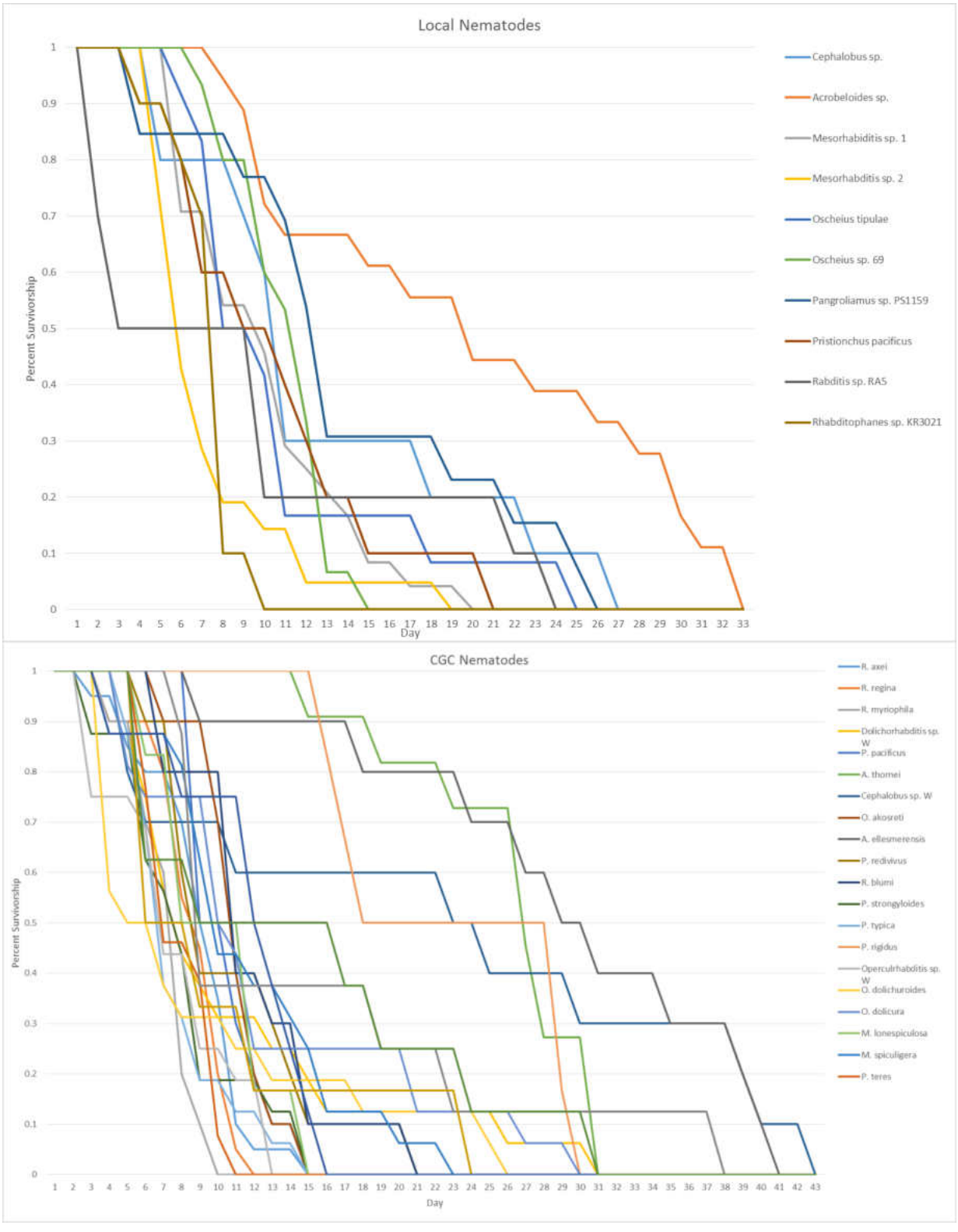

Figure 5.8: Survivorship Curves for Nematode Taxa.

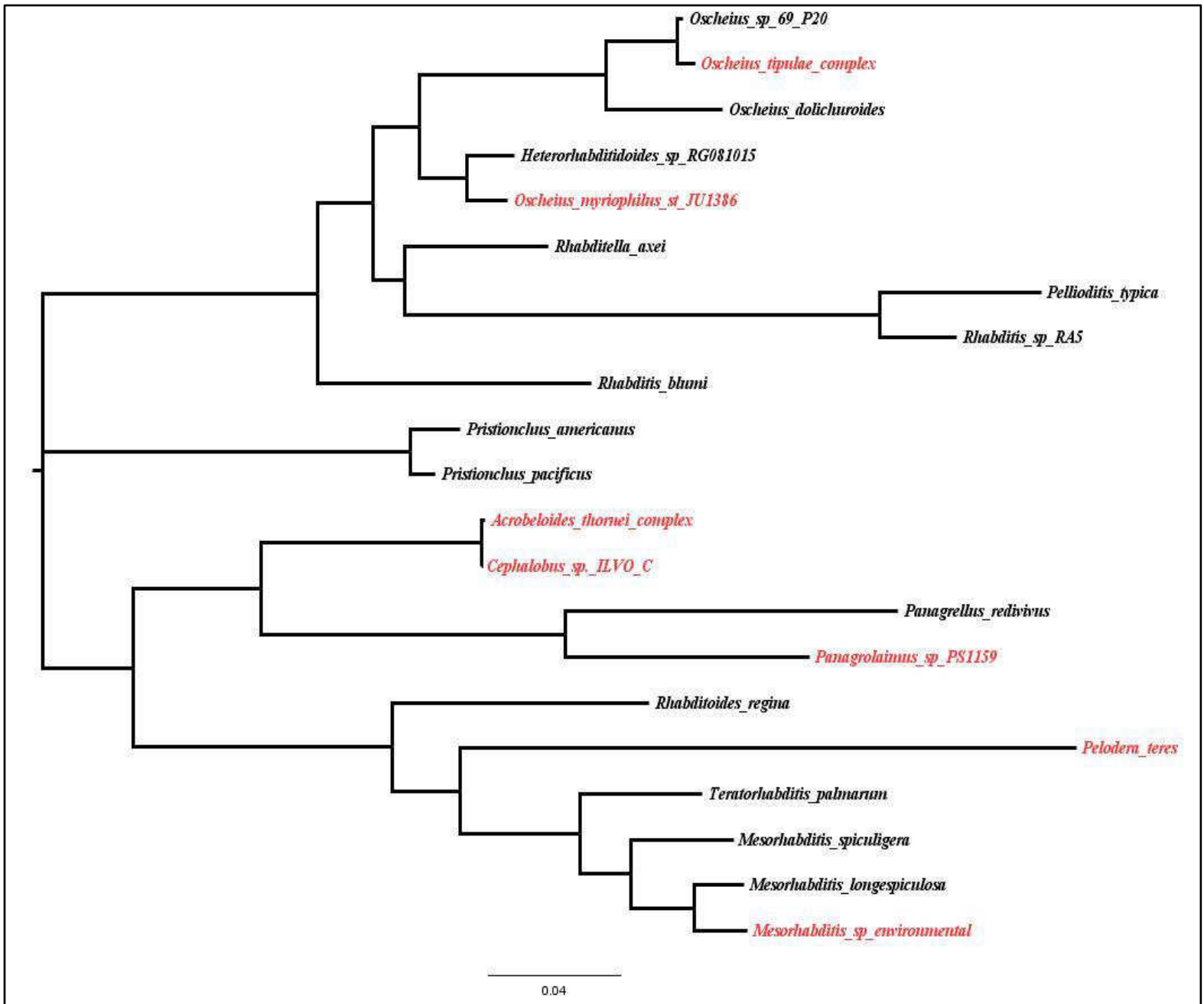

Figure 5.9: Nematode Study Phylogeny. Phylogeny of nematode species used in final database. Tree created in Mega7 using Maximum Likelihood method based on the Tamura-Nei model. Red branches indicate complexes where two or more species had to be pooled in analysis.

Table 5.1: Nematode Species with rRNA Copy Number Estimates

Species	Copy Number Estimate	Source
Ascaris lumbricoides	300	Long and Dawid (1980)
Caenorhabditis elegans	55	Long and Dawid (1980)
Panagrellus silusiae	280	Long and Dawid (1980)
Caenorhabditis brenneri	323	Bik et al (2013)
Caenorhabditis remanei	183	Bik et al (2013)
Caenorhabditis briggsae	56	Bik et al (2013)
Caenorhabditis japonica	115	Bik et al (2013)
Brugia malayi	150	Bik et al (2013)
Pristionchus pacificus	205	Bik et al (2013)
Cephalenchus sp1	58	Pereira and Baldwin (2016)
Cephalenchus sp.	10	Pereira and Baldwin (2016)
Cephalenchus cylindricus	19	Pereira and Baldwin (2016)
Cephalenchus cephalodiscus	34	Pereira and Baldwin (2016)
Cephalenchus daisuce	30.5 *	Pereira and Baldwin (2016)
Cephalenchus sp2	29	Pereira and Baldwin (2016)
Cephalenchus nemoralis	12	Pereira and Baldwin (2016)
Cephalenchus hexalineatus	3	Pereira and Baldwin (2016)

* indicated averaged value from 4 isolates

Table 5.2: Copy Number Estimates for Full Assemblages.

#OTU ID	Juvenile CN	Adult CN	Elder CN
Acrobeloides_thornei_complex	9.1	11.4	N/A
Cephalobus_sp._ILVO-C	161.1	184.8	15.1
Heterorhabditidoides_sp_RG081015	20.6	791.1	17.2
Mesorhabditis_longespiculosa	N/A	3.6	N/A
Mesorhabditis_sp_environ	N/A	1.0	N/A
Mesorhabditis_spiculigera	9.4	4.3	5.0
Oscheius_dolichuroides	10.9	118.7	24.5
Oscheius_myriophilus_str_JU1386	5694.4	9008.4	2544.1
Oscheius_sp_69_P20	N/A	29.1	37.1
Oscheius_tipulae_complex	1.0	6.1	0.7
Panagrellus_redivivus	N/A	423.6	N/A
Panagrolaimus_sp_PS1159	75.4	N/A	N/A
Pellioditis_typica	N/A	18.9	2.4
Pelodera_teres	10.8	318.1	1.5
Pristionchus_americanus	7.7	16.1	1.0
Pristionchus_pacificus	N/A	N/A	1.6
Rhabditella_axei	69.4	N/A	N/A
Rhabditis_blumi	19.1	9.2	1.2
Rhabditis_sp_RA5	98.8	N/A	1460.5
Rhabditoides_regina	14.4	N/A	5.1
Teratorhabditis_palmarum	4.3	138.2	1.4

Table 5.3: Partial Assemblages Results

	A1 (High Evenness, Phylogen. Conserved)			A2 (Mod. Evenness, Phylogen. Conserved)			A3 (Low Evenness, Phylogen. Conserved)		
Species	**Num**	**Seqs**	**CN**	**Num**	**Seqs**	**CN**	**Num**	**Seqs**	**CN**
O. akosreti	10	188829	235.15	10	174651	2910.85	10	190930	6.71
O. dolichuroides	10	31803	39.61	5	3704	123.47	1	0	0.00
O. myriphila	10	6629	8.26	2	30783	2565.25	1	5373	1.89
O. sp	10	803	1.00	2	12	1.00	1	2845	1.00
O. tipulae	10	3728	4.64	1	161	26.83	1	0	0.00
	A4 (High Evenness, Phylogen. Intermediate)			**A5 (Mod. Evenness, Phylogen. Intermediate)**			**A6 (Low Evenness, Phylogen. Intermediate)**		
Species	**Num**	**Seqs**	**CN**	**Num**	**Seqs**	**CN**	**Num**	**Seqs**	**CN**
O. akosreti	10	56684	11.30	10	189938	412.9087	10	241901	456.42
O. dolichuroides	10	6582	1.31	5	8835	38.41304	1	388	7.32
P. typica	10	77906	15.53	2	1058	11.5	1	0	0.00
R. blumi	10	5017	1.00	2	11839	128.6848	1	213	4.02
R. axei	10	7054	1.41	1	46	1	1	53	1.00
	A7 (High Evenness, Phylogen. Diverse)			**A8 (Mod. Evenness, Phylogen. Diverse)**			**A9 (Low Evenness, Phylogen. Diverse)**		
Species	**Num**	**Seqs**	**CN**	**Num**	**Seqs**	**CN**	**Num**	**Seqs**	**CN**
O. akosreti	10	98750	790.00	10	219388	38.76	10	181093	754.55
P. rigidus	10	0	0.00	5	0	0.00	1	24	1.00
M. longspiculosa	10	125	1.00	2	0	0.00	1	36	1.50
P. pacificus	10	19778	158.22	2	1132	1.00	1	2439	101.63
A. thornei	10	234	1.87	1	0	0.00	1	0	0.00

Table 5.4: Taxonomy and Characteristics of Nematodes Used in Study

Identification			Life History				Behavior			
Species	**Family**	**Origin**	**Ro**	**G**	**R**	**eggs/adult**	**digging**	**endotokia**	**activity**	**reproduction**
Acrobeloides sp.	Cephalobidae	Local	14.12	17.32	0.15	38.46	no	no	low	hermaphrodite
Cephalobus sp.	Cephalobidae	Local	32.20	38.43	0.09	25.20	no	no	low	hermaphrodite
Mesorhabditis sp. 1	Rhabditidae	Local	1.17	6.67	0.02	6.75	no	no	average	male/female
Mesorhabditis sp. 2	Rhabditidae	Local	1.32	5.34	0.05	3.43	no	no	average	male/female
Oscheius sp. (69)	Rhabditidae	Local	0.01	2.23	-2.02	0.64	no	no	average	hermaphrodite
Oscheius tipulae	Rhabditidae	Local	23.58	8.01	0.39	26.08	no	no	average	hermaphrodite
Pangroliamus sp. (PS1159)	Panagrolaimidae	Local	7.50	6.85	0.29	16.46	no	no	high	hermaphrodite
Pristionchus sp. (americanus)	Neodiplogasteridae	Local	308.69	2.78	2.06	98.90	low	no	high	hermaphrodite
Rhabditiophane sp. (KR3021)	Alloionematidae	Local	40.80	2.87	1.29	40.80	no	no	average	unknown
Rhabditis sp. (RA5)	Rhabditidae	Local	31.30	5.25	0.66	31.30	no	no	average	hermaphrodite
Acrobeloides butschlii	Cephalobidae	CGC	0.86	7.03	-0.02	1.88	no	no	low	hermaphrodite
Acrobeloides ellesmerensis	Cephalobidae	CGC	3.77	25.71	0.05	14.10	low	no	low	hermaphrodite
Acrobeloides thornei	Cephalobidae	CGC	140.96	11.54	0.43	142.18	low	no	low	hermaphrodite
Cephalobus sp. Wild isolate	Cephalobidae	CGC	1.77	26.38	0.02	6.50	low	no	low	hermaphrodite
Dolichorhabditis s. Wild isolate	Rhabditidae	CGC	0.10	5.86	-0.38	1.25	no	no	high	male/female
Mesorhabditis longespiculosa	Rhabditidae	CGC	49.32	6.18	0.63	30.00	no	no	low	male/female
Mesorhabditis spiculigera	Rhabditidae	CGC	85.24	12.09	0.37	46.75	no	no	high	male/female
Operculrhabditis sp. Wild isolate	Rhabditidae	CGC	8.30	6.43	0.33	10.31	no	no	high	unknown
Oscheius akosreti	Rhabditidae	CGC	4468.90	3.53	2.38	508.80	no	no	high	male/female
Oscheius dolichuroides	Rhabditidae	CGC	4.88	8.23	0.19	4.94	no	yes	high	male/female
Oscheius myriophila	Rhabditidae	CGC	165.99	5.11	1.00	51.00	no	yes	high	hermaphrodite
Oscheius dolicura	Rhabditidae	CGC	82.18	7.51	0.59	50.75	no	yes	high	male/female
Panagrellus redivivus	Panagrolaimidae	CGC	18.28	5.10	0.57	19.40	no	yes	high	male/female
Panagrolaimus rigidus	Panagrolaimidae	CGC	0.39	15.76	-0.06	1.33	high	no	high	male/female
Pellioditis typica	Rhabditidae	CGC	0.62	5.80	-0.08	3.63	no	no	average	male/female

Pelodera strongyloides	Rhabditidae	CGC	13.25	4.58	0.56	22.63	no	no	low	male/female
Pelodera teres	Rhabditidae	CGC	0.06	4.52	-0.61	1.31	no	no	average	male/female
Pristionchus pacificus	Neodiplogasteridae	CGC	257.03	2.32	2.39	77.70	high	no	high	hermaphrodite
Rhabditella axei	Rhabditidae	CGC	54.65	3.07	1.30	53.40	no	no	average	male/female
Rhabditis blumi	Rhabditidae	CGC	3.72	6.31	0.21	10.10	no	no	average	male/female
Rhabditis myriophila	Rhabditidae	CGC	1.18	2.46	0.07	5.20	no	no	high	hermaphrodite
Rhabditoides regina	Rhabditidae	CGC	8.66	2.57	0.84	24.00	no	yes	high	male/female
Teratorhabditis palmarum	Rhabditidae	CGC	73.42	7.94	0.54	19.83	high	no	high	male/female
Zeldia sp. Wild isolate	Cephalobidae	CGC	22.03	21.09	0.15	6.25	high	no	low	hermaphrodite

Epilogue

Summary

This dissertation work spanned three major disciplines of biology: ecology, evolution, and genetics. In an attempt to ground the genetic basis of the Growth Rate Hypothesis, I explored the phenomenon of copy number variation of the rRNA operon. For copy number variation to be a viable genetic mechanism, it needs to evolve in a stepwise, heritable manner. This would show that it is under selection and that a species' copy number comes from its lineage's genetic heritage of environmental adaptation, not random fluctuation. Furthermore, since rRNA is phosphorus-limited, fast-growing species with high copy number should respond to phosphorus enrichment more than other species. So, in nutrient-enriched soil, there should be a copy number specific shift in species composition. Finally, copy number should be correlated with growth rate under ideal culturing conditions. Enrichment-type bacterivore nematode growth rate with unlimited food should be correlated with their copy number.

Phylogenetic Autocorrelation Study

I found that copy number is phylogenetically autocorrelated in bacteria, which proves both that it is a viable genetic mechanism and that it can be estimated by phylogeny. However, autocorrelation falls off at high sequence dissimilarity, which is a problem because many prokaryote species with no copy number estimates have few close relatives with copy number estimates. But, by examining copy number estimation accuracy by phylum, some phyla have much better estimation success than others. Overall, copy number estimation can be better achieved by focusing on a series of best practices, rather than ignoring or broadly construing copy number.

Nutrient Enrichment Study

High copy number prokaryotes increased in nutrient-enriched soil cores in the fall but not the spring. Enrichment-type nematodes increased in phosphorus treatments in the spring and nitrogen treatments in the fall. This showed that high copy number species responded to the removal of phosphorus limitation. But it also revealed more complicated interactions. It is possible that nutrient limitations in grassland soils differ between spring and fall. Also, high copy number prokaryotes growth rate in nutrient-enriched soils may be masked by predation by enrichment-type bacterivores.

Life History Study

In culturing studies, I found that fast-growing species of nematodes did have high copy number in a limited case. Relative copy number was not accurate using pooled Illumina sequencing, so better copy number estimation is needed for the future. Getting accurate copy number estimates for metazoans remains an unsolved technical problem.

Synthesis

The Growth Rate Hypothesis (GRH) was upheld in bacterivore nematodes and soil prokaryotes. Fast-growing species were more abundant in phosphorus-rich soil environments under certain conditions. Fast-growing species also tended to have higher rRNA copy number. GRH. Copy number was under strong positive selection and phylogenetically autocorrelated in bacteria but not in bacterivore nematodes. Nematodes had poor copy number estimates and smaller sample sizes. As opposed to most studies that did not show phosphorus affecting

nematode community abundances, we found fast-growing nematodes to be overabundant in phosphorus-rich soil cores.

Implications

This research developed a number of tools (rRNA copy numbers, analysis pipeline) that are needed in order to make high-throughput amplicon sequencing of soil faunal communities accurate (Kembel et al. 2012, Darby et al. 2013), such as for regional and national surveys or monitoring programs. It also developed useful culturing and life-history trait techniques (thin nematode plates) for nematodes that are less amiable to traditional culturing.

Future Research

While prokaryotic copy number estimates continue to grow, work still needs to be done to increase the coverage, especially for taxa with few culturable isolates. Nematodes also need more copy number estimates, and more complete sequences of single-copy genes in order to do so. Methods of accurate copy number correction still need to move forward to transition the qualitative assessment that most metagenomics tools possess to truly quantitative. Furthermore, the problem of order and taxonomic cutoffs should be explored in datasets outside the rrnDB. Finally, a large scale analysis of copy number distribution and the prevalence of constitutive expression in rRNA operons would grow the understanding of how copy number shapes evolution. A meta-study of other nutrient-enrichment experiments or environments with naturally differing phosphorus (such as the NEON sites) would solidify this one way or another. More work on how phosphorus is actually taken up by bacteria would help solve some issues as well.

We have over two dozen nematodes in active cultures, so a variety of interesting questions could still be asked. We could put differing copy number species together on a single plate and vary the amount of food and see how they interact and grow. We could look at *Rhabditiophane*s and confirm if they do indeed grow best at low temperatures. More broadly, we could test the effect of temperature on growth rate and life history. Further work could also be done to get cell count so that the copy number estimates per individual can be made into copy number estimates per cell.

Appendix A: Supplemental Methods

Chapter 3 Supplemental Methods

Merging Diagnostics

After merging, I pulled out merge percentage and average length for the five levels of max diffs (10, 20, 30, 40, 50). Other than a few samples, the merge rates hit an acceptable percentage at 30/40 max difference (Table S3.1). The average length was fairly consistent across max difference levels. The bacterial only samples were all close to the expected amplicon size. The rest of the samples were systematically higher, but it was not as clear due to the mix of nematode and bacteria (Table S3.2).

Amplicon length information

For nematodes, the size of the 18S should be around 358bp (ungapped) and a total size with primers around 403bp. The bacterial 16S should be around 255bp ungapped, with a total size with primers around 294bp.

Primer hit rate

When demultiplexing into nematodes and bacteria (using the initial primers as to separate), I was able to calculate the percentage of missed sequences (not labeled as bacteria or nematodes). The percentage was consistent within samples; meaning the max difference number did not bias the percentage. The maximum missed percentages were around 2% (Table S3.3). There were a total of 40951 missed sequences. The number of unique sequences as 8957, with around 22% of all missed sequences being unique.

Stripping primers, trimming, quality filtering, uniquifying

The bacteria had 8 left and 8 right bases stripped while nematodes had 24 left and 20 right bases stripped. For trimming, I ran eestats and found that 250 was a good length to trim down to for both bacteria and nematodes. The best max diff score for this was 30 for nematodes, representing a peak. The best for bacteria was 10, falling off from there. For quality filtering, I used the default max ee (1.0). The max difference with the best pass rate was 10 diffs for both. (Table S3.4). Both were uniquified using a minimum unique size of 2.

Zotus and zotu tables

Zotus were created as well as zotu tables. The best percentage mapped was 50 diffs for nematodes and 30 diffs for bacteria (Table S3.5). Looking at the number of species, total abundances, and the number of singlets (zotus found in only one of the samples), we see an increase as max diffs increases in a plateauing fashion. (Figure S3.1).

Chapter 5 Supplemental Methods

Sequence Identification

Several species from the CGC had no known sequences, either because the isolate was only identified by the CGC to genus or the species name used was designated nowhere else. Of the 25 CGC nematode cultures and the 10 local nematode cultures (34 total), 20 had full sequences between the NF1 and NR2 primer regions. An additional 10 had at least a forward or reverse Sanger sequence.

So first, we tried to get a matching sequence for species that had only a genus. These genera were Oscheius, Operculrhabidits (heterotype of Mesorhabdiditis), Acrobelodies, and Zeldia. Databases for each of the genera were created by taking the 18S sequence of a representative species of the genus from NCBI and blasting that sequence to its genus and taking all (or the top 250 in the case of Acrobeloides) and downloading it as a fasta file. Then the using the local blastn, the unique file was blasted to the respective genus databases (making the genus fastas a database in blastn first).

Zeldia:

This approach failed for Zeldia, with zero matches at ids of 1.00, 0.99, 0.95, 0.70, and 0.50. I only got matches at 0.02 percent id, so there was some big issue. So, I went to Silva and made a database using all genus samples from there (n=4). While there was no hit on 1.00, I got a hit on 0.99 to Zeldia sp. JH-2004.

Acrobeloides:

Using the NCBI database and matching otus from the unique file garnered three matches (presumably at 1.00), and none of them were my species (*A. nanus, A. apiculatus*, and *A. bordenheineri*). So, I made a smaller Silver database (*A. buetschlii*, and *A. thornei* only, as *A. ellesmerensis,* only had 23S in either NCBI or Silva). I also looked at the sequences in Geneious and aligned them. There was a single nucleotide difference (R in one) between the two and the databasing cannot distinguish the two (would find some of both and it would label reads from assemblages that could have only contained *A. thornei* as *A. buetschlii*).

To see if the other two possible Acrobeloides species, one from local and *A. ellesmerensis*, could be parsed out, I examined them. I looked at the partial sanger sequence for the local Acrobeloides. Blasting to NCB matched them to environmental and Cephalobus for *A. nanus, thornei* and *buetschlii.* I got a good *A. nanus* sequence and aligned in Geneious. It had one nucleotide difference from the other two, but no match for *A. nanus* when added and run in the database. For *A. ellesmerensis*, I looked at the closest matches to the 28S by blasting and found *A. bordenheimeri* to be close. Since *A. bordenheimeri* had also been a hit in the earlier NCBI database, I checked alignment and found it to be different from the other Silva Acrobeloides by several bases. But, adding it to the database and matching was not successful.

From there, I created a database from all unique Acrobeloides species in Silva with a size >1000 (n=11) and ran the otutab with Usearch. At identity 1.00, *A. apiculatus* and *A. buetchlii* came out. At 0.99, both came out along with A. sp. PS1146. At 0.98, *A. apiculatus* and *A. buetchlii* again came out. *A. buetchlii* was then found in the partial assemblages (1-9), *despite A. thornei* being the only Acrobeloides added to the partial assemblages.

This work showed that Acrobeloides could not be distinguished down to species.

Oscheius and Mesorhabdidis:

Based on the failure of Acrobeloides, did not try to pursue this method with the other two genera.

Sanger Sequence Matching:

There were 10 species which had no full sequence matched from NCBI or Silva but had at least a forward or reverse sequence from sanger sequencing. I went through all of them (and some others) by using a method to blasts the sanger sequence to the best Unique sequence

First, the Unique sequence file was made into a database for the blastn suite. Then the individual sanger sequences were blasted to the Unique DB. The top hit was extracted and the full sequence for that hit was found in NCBI. Then, the NCBI sequence was tested backwards by blasting to the Unique Database (from 0.97 to 1.00 identity). This resulted in obtaining a 100 sequence for Cephalobus sp. and Pristionchus sp. The two local Mesorhabidits species both went to an environmental sample. The local Acrobeloides species and *Cephalobus sp. wild isolate* blasted to the same sequence as the *Cephalobus sp. Dolichorhabidis sp. wild isolate* blasted to *O. tipulae strain CEW1* and *O. myriophyla* blasted to *R. myriophila*.

Finally, I blasted the uniques to the full Nematode database again, at different levels of identity (1.00, 0.99, 0.98, 0.97) and found 16 of the species would show up with an ID of 0.98. Dropping the id lower did not include any further species.

Taxonomic Information:

The following sections will outline general information on the various genera, as well as specifics for species within the general, as available. Table S5.1 summarizes colonizer-persister (CP) values and feeding type. Unless otherwise indicated, all CP values come from Bongers (1990), while feeding types come from Yeates et al. (1993).

Acrobeloides (Cobb, 1924): This genus is small with a spool-shaped oesophagus corpus (Bongers, 1994). It is typically considered to be a bacterivore (Yeates et al., 1993) but at least some species may graze fungal hyphae (Ruess and Dighton, 1996). They have a CP value of 2 but are sensitive to heavy metal toxicity at acute levels over and above similar CP-valued Aphelenchus (Li et al. 2005). They are found in rivers, lakes, and other semi-aquatic habitats, as well as caves (Eyualem-Abebe et al, 2006) and sand. In the soil, they are mostly found around roots or decaying plant matter (Goodey, 1963) and males are rare (hermaphrodites). On agar plates, they have been observed (Dean, from Croll, 1970) to aggregate together, either coiling around or parallel.

Acrobeloides butchlii, Acrobeloides ellesmerensis, and *Acrobeloides thornei* were all obtained from the CGC. *A. butchlii*'s original sample came from a peach orchard at Lodi, CA from B. Jaffee and identified by E.M. Noffsinger. *A. ellesmerensis* was also isolated from CA and sent by E.M. Noffsinger, as was *A. thornei* (does not match original description of *A. thornei* well, based on observation by Paul De Ley). *Acrobeloides* sp., isolated from ND, was identified by sequencing and had a low match to *A. nanus* and various other species, including *A. butchlii* and *A. thornei*.

Cephalobus (Bastian, 1865): Genus of typical soil nematode with a wide distribution. Three lips and blunt tail (Bongers, 1994). They are found in bromeliad water, caves, rivers, semi-aquatic and non-specified aquatic habitats (Eyualem-Abebe et al, 2006). Likely saprophagous/microbivorous and found around plant roots (Goodey, 1963; Rashid, Geraert, & Sharma, 1984). They are described as sluggish (Eyualem-Abebe et al, 2006). *C. litoralis* known to survive desiccation well, reproduce parthogenically and lay eggs around 3-4 days, and have endotokia, which is when juveniles hatch inside a dead female (Saeed et al. 1988). *C. brevicauda* showed bacteria choice preference via chemotaxis and is hermaphrodite (Salinas et al., 2007).

Cephalobus sp. was obtained from the CGC; the isolate was isolated at Fort Collins, CO and identified by Lynn Carta. Another *Cephalobus* sp. was isolated from ND and identified by sequencing. The top sequence match was *Cephalobus* sp. ILVO-C.

Dolichorhabditis (Andrássy, 1983): Genus is considered a synonym of *Oscheius* (Ident of Free-living Nematodes). Hermaphroditic species (males rare) found in the Dolicura group, and sp. CEW1 one of the most common free-living (Felix, Delattre, and Dichtel 2000), but is likely just *O. tipulae*. They have different patterns of vulva development than *C. elegans* (Dichtel et al, 2000).

Dolichorhabditis sp. was obtained from the CGC, which was isolated from Brazil.

Mesorhabditis (Osche, 1952): Genus of small nematodes with well-developed lips (Eyualem-Abebe et al, 2006). Reproduces through pseudogamy and males are rare (Nigon & Félix, 2017). They occur where there are abundant bacteria, as they need to ingest live bacteria (Goodey, (1963) and can establish colonies on a wide range of bacteria (Wood, 1973). Found

eurytopically in a study of streams and rivers, irrelevant of nutrient enrichment (Niemann et al, 1996).

Mesorhabditis longespiculosa and *Mesorhabditis spiculigera* were obtained from the CGC. *M. longespiculosa* was isolated by W. Sudhaus from tunnels of beetle larvae in Kenya and is a male/female strain. *M. spiculigera* was isolated in PA. Its original description was *M. mioki*, but was corrected by K. Kiontke. *Mesorhabditis* sp1. and sp2. isolated from ND and identified by sequencing. Sp1. was matched to several species (sp. 339Meso2; sp. MR1; sp. WB-2009; JH-2004), the first several likely incorrect (Darby, personal communication). Sp2. was may not be different form Sp1.

Oscheius (Andrássy, 1976): This genus contains the most abundant nematodes found in non-desert soils and is hermaphrodite (self-fertile hermaphrodites and facultative males). This genus is closely related to *Caenorhabditis elegans* and the species *Oscheius tipulae* exhibits similar behavior in culture, albeit being smaller and slower. It has also received recent model attention (Félix, 2006). Previous research in the Darby lab involved the genome sequencing and epigenetic examination of *O. tipulae* (McGlynn, 2018).

Oscheius akosreti, dolichuroides, dolichura, and myriophila were all obtained from the CGC. *O akosreti* was isolated from WI. There are no hermaphrodites and dauers are SDS resistant. *O. dolichuroides* was isolated near a tree in Kenya and is a male/female strain. *O. dolichura* was isolated from the Galapagos archipelago and is a male/female strain. O. myriophila was isolated from a compost heap in NY and is a hermaphroditic strain and previously called *Rhabditis sp*. *Oscheius* sp2. and *O. tipulae* were isolated from ND and

identified with sequencing. *Oscheius* sp2. had several hits (sp. 69 P20; sp. 67 P20; sp. FDL-2014; sp. FVV-2) while *O. tipulae* matched to strain CEW1.

Operculrhabditis (Khera, 1969): The only thing known from the literature about this genera is related to the genus *Butlerius* but is distinguished by a flap-like, cuticular operculum covering the vulva (Khera, 1969).

Operculrhabditis sp. was obtained from the CGC, identified by E. M. Noffsinger, and is a wild isolate originally from the UCR collection.

Panagrellus (Thorne, 1938): This genus is found in fermentative environments (acid produced), such a cider and beer spill, from which the common name of the representative *Pangagrellus redivivus*, the sour paste eelworm, originates. Their probable natural environment is tree exudate (Goodey, 1963), but have been found in polluted freshwater. This makes it useful for aquatic toxicity bioassays (Sherry et al., 1997). It is used in locomotion experiments and highly active (Croll, 1971), as well as commercially for fish food (Santiago et al., 2004). Recently, it had its genome sequenced (Srinivasan, 2013).

Pangagrellus redivivus was obtained from the CGC and listed as non-hermaphroditic (sexes separate) and slow-growing.

Pangroliamus (Fuchs, 1930): Genera of long nematodes with a body that narrows after the vulva (Bongers, 1994). They are hardy, very active, partially freshwater, and are capable of anhydrobiosis. They are opportunistic and have been found in polar regions, bromeliad water, non-specific aquatic environments, and caves (Eyualem-Abebe et al, 2006). In the soil, they are bacterial feeders or saprophages and found in decaying plant matter, associated often with bark beetles (Goodey, 1963). They were found eurytopically in a study of streams and rivers,

irrelevant of nutrient enrichment (Niemann et al, 1996). They have been studied in regard to sex attraction (Greet, 1964), reproductive development (as the genus contains gonochoristic, hermaphroditic and parthenogenetic reproductive modes), and cold tolerance (Lewis et al., 2009).

Panagroliamus rigidus was obtained from the CGC and were males and females. *Pangroliamus* sp. was isolated from ND and identified by sequencing. It matched *Panagroliamus* sp. PS1159, sp. JB115, and sp. JB051.

<u>*Pellioditis*</u> (Dougherty, 1953): This genus is very close to *Rhabditis* (Goodey, 1963). They are bacterial feeders, both in the soil and in estuarine environments (Moens and Vincx, 1997; Yeates et al., 1993) and the genus includes a complex of marine species. They've been used in absorption rate studies (Moens et al., 1996; Moens & Vincx, 2000) and showed that efficiencies of absorption depended on food density.

Pellioditis typica was obtained from the CGC and was isolated by W. Sudhaus from feces in Kenya. It is a male/female strain.

<u>*Pelodera*</u> (Schneider, 1866): A genus of large, stout, and moderately active nematodes with the same biology as *Rhabditis* (Goodey, 1963). In a study of streams and rivers, they were most abundant in organically enriched ones (Niemann et al, 1996). More broadly, species of *Pelodera* are often found in organically enriched environments (Eyualem-Abebe et al, 2006). Dauers of *P. coarctata* are phoretic on dung beetle (Croll, 1971). *P. strongyloides* larvae respond to electric currents and will migrate to anodes (Whittaker, 1969). Medically, they can cause dermatosis, though rarely (Saari & Nikander, 2006).

Pelodera teres and *Pelodera strongyloides* were obtained from the CGC. *P. teres* is a male/female strain and previously called DF5016. *P. strongyloides* is also a male/female strain and originally came from Sudhaus.

<u>*Pristionchus*</u> (Kries, 1932): Pristionchus is a genus of large-bodied nematodes that are self-fertilizing hermaphrodites with frequent males and includes the model organism *Pristionchus pacificus*. *P. pacificus* is used as a satellite model organism, in comparison to C. elegans. Its growth rate on agar is similar, but has a different molting pattern (Sommer, 2006) and feeds on bacteria, fungi, and can also be predatory (Fürst von Lieven & Sudhaus, 2000). *Pristionchus* species, in general, have been found to associate with beetles (Sommer, 2006).

Pristionchus pacificus was obtained from the CGC and is a hermaphroditic strain. *Pristionchus* sp. was isolated from ND and identified by sequencing. It was matched at lower levels to *P. americanus, P.* sp. 3 EJR-2013, *P. pseudaerivorus, P. maupasi*, and *P. aerivorus*.

<u>*Rhabditella*</u> (Cobb, 1929): They are a genus of nematodes closely related to *Rhabditis* and morphologically similar to *Cephaloboides*, but with slimmer and longer tails (University of Nebraska Lincoln Nematology). They are present where bacteria are plentiful and require living bacteria for breeding. They have been found in sediment and waters (Eyualem-Abebe et al, 2006), as well as feces (el-Azay, el-Gawady, & Nada, 1968).

Rhabditella axei was obtained from the CGC and is a male/female strain. It was isolated by W. Sudhaus from a compost heap in France and easily grown in culture.

<u>*Rhabditis*</u> (Dujardin, 1845): Is the typical soil bacterivore genus and contains the *Rhabditis terricola*, one of the most common soil nematodes (Bongers, 1994). It also used to contain *C. elegans*, before the subgenus *Caenorhabditis* was made and then subsequently raised to proper

genus (Univesity of Nebraska Lincoln Nematology). They are present where bacteria are abundant (Goodey, 1963), and colonies have been found utilizing many different bacterial species (Wood, 1973) and are considered generalists. They have been found in a variety of freshwater habitats, including springs and estuaries (Eyualem-Abebe et al, 2006). They exhibit swarming behavior (McBride & Hollis, 1966), especially to the presence of light (Staniland, 1957), and responds to mechanical stimuli (Croll and Smith, 1970) and exhibit endotokia.

Rhabditis blumi and *Rhabditis myriophila* were obtained from the CGC. *R. blumi* was isolated by J.P. Blum in a dung heap in Spain and will form dauer larvae on overcrowded and starved plates. *R. myriophila* was isolated by G. O. Poiner. *Rhabditis* sp. was isolated from ND and identified with sequencing. It matched to sp. RA5, sp. DF5059, and *terricola* JH-2004.

<u>*Rhabditoides*</u> (Goody, 1929): This is another genus closely related to *Rhabditis*. In a study of streams and rivers, they were limited to organically enriched ones (Niemann et al, 1996), and has been associated with beetle larvae (Shulte & Poiner, 1991). Not much else is known about the genera besides its similarity to *Rhabditis*.

Rhabditoides regina was obtained from the CGC and is a male/female strain. It was isolated by M. Velasquez from a beetle larva from Guatemala.

<u>*Rhabditiophanes*</u> (Fuchs, 1930): Genus of poorly-known nematodes. Males are very rare, seem to be saprophages and found in abundant bacteria (Goodey, 1963). They are a possible free-living model of *Strongyloides stercoralis*, a parasitic nematode (Cadavid, 2017). They also may exhibit optimal growth at low temperatures (Darby, personal communication).

Rhabditophanes sp. was isolated from ND and identified with sequencing as *Rhabditophanes* sp. KR3021.

Teratorhabditis (Osche, 1952): This genus is hermaphroditic and otherwise little known, though presumably follow similar biology to *Rhabidtis*. Species are associated with palm weevil (Gerber & Giblin-Davis, 1990) and have been found in sewage and manure (Tahseen et al., 2007).

Teratorhabditis palmarum was obtained from the CGC and is a male/female strain. It was isolated by R. Giblin-Davis from the palm weevil in Florida

Zeldia (Thorne, 1937): This is a bacterial feeding genus with very rare males that do not have a bursa (Allen & Noffsinger, 1972). They are found in soil and decaying plants, where they are presumably saprophagous/microbivorous (Goodey, 1963).

Zeldia sp. was obtained from the CGC, which was isolated from CO and identified by Lynn Carta.

Appendix B: Supplemental Figures and Tables

Chapter 3 Supplemental Figures and Tables

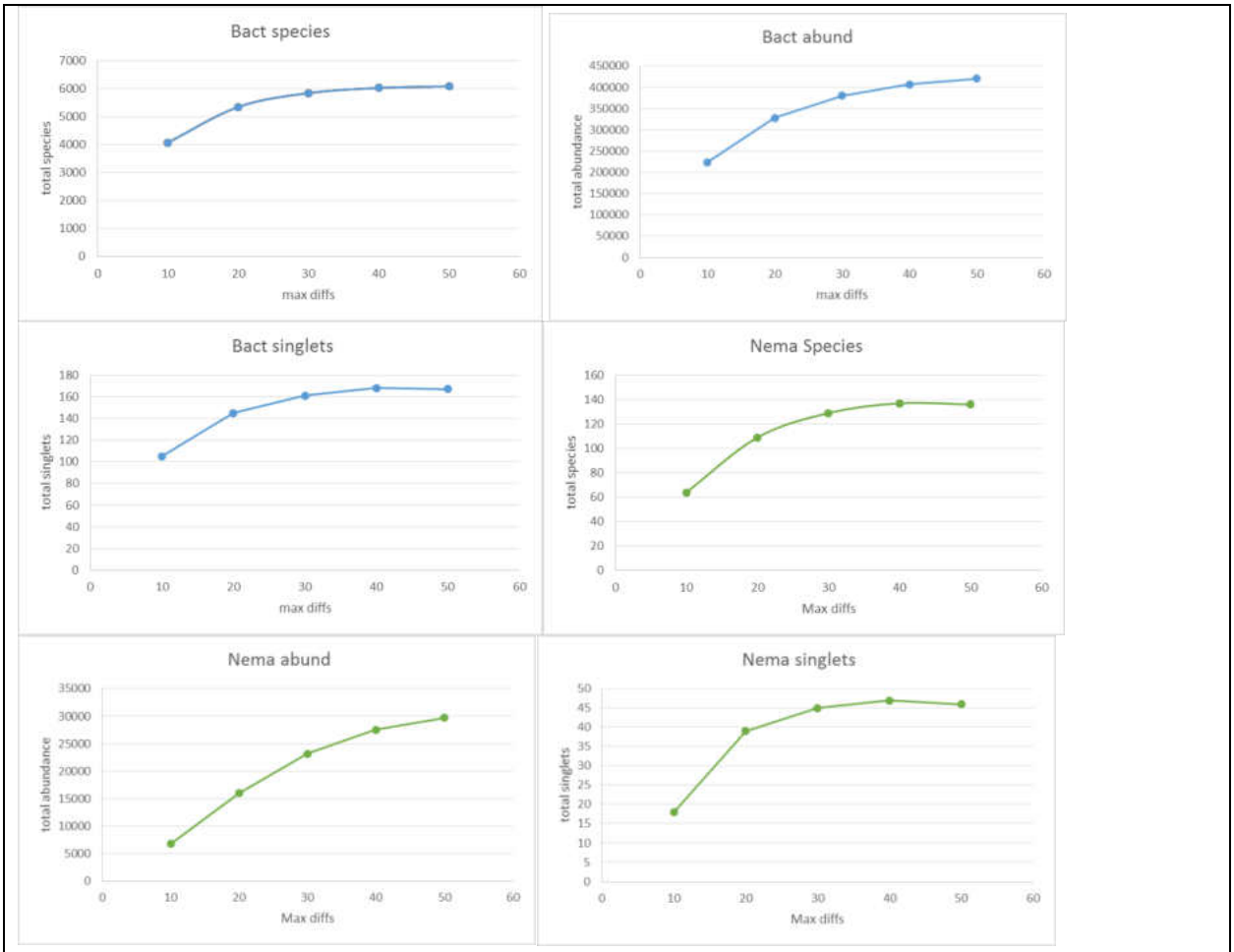

Figure S3.1: Accumulation curves. For species, abundance, and singlets for nematodes and bacteria across different max differences.

Table S3.1: Merging Percent Success Across Max Diff Levels.

Sample	Type	md_10	md_20	md_30	md_40	md_50
25	N + B	58.51%	78.32%	87.45%	92.20%	94.61%
26	N + B	58.99%	76.21%	84.27%	88.60%	90.94%
27	N + B	28.77%	55.54%	71.93%	81.61%	87.13%
28	N + B	57.59%	78.10%	87.71%	92.73%	95.35%
29	N + B	5.39%	21.84%	42.00%	59.60%	72.74%
30	N + B	56.07%	75.89%	85.17%	90.17%	92.73%
avg	N/A	0.455968	0.6619	0.770601	0.830519	0.865355

Table S3.2: Average Length of Amplicons after Merging

Sample	Type	md_10	md_20	md_30	md_40	md_50
26	N + B	268.06	272.96	274.65	275.45	275.87
27	N + B	299.96	302.78	303.97	304.49	304.87
28	N + B	293.46	293.71	293.77	293.82	293.84
29	N + B	304.04	306.92	306.91	306.14	305.58
30	N + B	292.16	292.36	292.44	292.48	292.52

Table S.3.3: Statistics of Missed Sequences after Demultiplexing.

Sample	nema	bact	both	all	all_hand	leftover	leftover%
S25__10diffs	389	92954	0	94137	93343	794	0.843452
S25__20diffs	525	124444	0	126006	124969	1037	0.822977
S25__30diffs	595	138947	0	140703	139542	1161	0.825142
S25__40diffs	641	146439	0	148337	147080	1257	0.847395
S25__50diffs	669	150247	0	152222	150916	1306	0.857957
S26__10diffs	13644	85020	0	100004	98664	1340	1.339946
S26__20diffs	14290	113306	0	129208	127596	1612	1.247601
S26__30diffs	14581	126507	0	142874	141088	1786	1.250052
S26__40diffs	14719	133591	0	150207	148310	1897	1.262924
S26__50diffs	14810	137406	0	154185	152216	1969	1.277037
S27__10diffs	5305	42732	0	48677	48037	640	1.314789
S27__20diffs	12268	80417	0	93991	92685	1306	1.389495
S27__30diffs	17029	102931	0	121717	119960	1757	1.443512
S27__40diffs	19877	116171	0	138104	136048	2056	1.488733
S27__50diffs	21697	123485	1	147442	145183	2259	1.532128
S28__10diffs	1557	94328	0	96855	95885	970	1.001497
S28__20diffs	2405	127630	0	131352	130035	1317	1.002649
S28__30diffs	2764	143214	0	147520	145978	1542	1.045282
S28__40diffs	2994	151269	0	155952	154263	1689	1.083026
S28__50diffs	3100	155477	0	160360	158577	1783	1.111873
S29__10diffs	874	4657	0	5647	5531	116	2.054188
S29__20diffs	3847	18568	1	22880	22416	464	2.027972
S29__30diffs	7248	35934	2	44005	43184	821	1.865697
S29__40diffs	9693	51611	3	62448	61307	1141	1.82712
S29__50diffs	11306	63470	4	76212	74780	1432	1.878969
S30__10diffs	849	105666	0	107540	106515	1025	0.953134
S30__20diffs	1309	142846	0	145552	144155	1397	0.959794
S30__30diffs	1510	160256	0	163345	161766	1579	0.966666
S30__40diffs	1622	169612	0	172932	171234	1698	0.981889
S30__50diffs	1703	174365	0	177846	176068	1778	0.999741

Table S3.4: Trimming Length Max Diff Scores

	10 diffs	20 diffs	30 diffs	40 diffs	50 diffs
Nematodes	99.8	97.4	89.4	80.3	73.7
Bacteria	99.6	96.8	92.8	89.3	86.6

Table S3.5. Zotu Success by Max Diff Rate.

File	Ratio Mapped to OTUs	Percentage Mapped
10 diffs nema	6744 / 21930	30.8%
20 diffs nema	15966 / 33832	47.2%
30 diffs nema	23159 / 42906	54.0%
40 diffs nema	27509 / 48720	56.5%
50 diffs nema	29718 / 52457	56.7%
10 diffs bact	224028 / 425200	52.7%
20 diffs bact	327856 / 607012	54.0%
30 diffs bact	380197 / 707566	53.7%
40 diffs bact	406825 / 768452	52.9%
50 diffs bact	420757 / 804201	52.3%

Chapter 5 Supplemental Figures and Tables

Table S5.1: Metrics of nematode genera used in this study.

Genus	**CP Value**	**Feeding Type**
Acrobeloides	2	Bacterial feeding
Cephalobus	2	Bacterial feeding
Dolichorhabditis	1	Bacterial feeding
Mesorhabditis	1	Bacterial feeding
Oscheius	1	Bacterial feeding *
Operculrhabditis	1	Bacterial feeding, animal predation *
Panagrellus	1	Bacterial feeding
Panagrolaimus	1	Bacterial feeding
Pellioditis	1	Bacterial feeding
Pelodera	1	Bacterial feeding
Pristionchus	1	Bacterial feeding, animal predation (ingestion)
Rhabditella	1	Bacterial feeding *
Rhabditoides	1	Bacterial feeding
Rhabditiophanes	1	Bacterial feeding
Rhabditis	1	Bacterial feeding
Teratorhabditis	1	Bacterial feeding
Zeldia	2	Bacterial feeding

*-presumed, given closest relative feeding group and literature

Table S5.2: Nematode Behavioral Observations	
Species	**Behavior Notes**
Local Specimens	
Acrobeloides sp.	Laid eggs daily and had a lot of replating; often went missing; lived shorter at 25 degrees; average life=44.5
Cephalobus sp.	Room temp plate had a very long life of last adult, which laid eggs for a long time; eggs spread out across long stable middle period; lived shorter at 25 degrees, but only due to longevity of last adult on room temp; average life=56.5
Mesorhabditis sp. 1	Either replated with no eggs hatching or just remained for a long time without eggs; average life=21
Mesorhabditis sp. 2	Dropped off very quickly in mortality; eggs but none hatched; average life=19.5
Rhabditiophane sp.	Failed at 25 degrees; average life=8
Rhabditis sp.	Large bodied; lots of eggs at the beginning, then dropped to zero, average life=20.5
Oscheius sp.	Had eggs but did always hatch (replated many times without any offspring); eggs at beginning, then dropped to zero; dug in a little at 25 degrees; lived shorter at 25 degrees; average life=17
Oscheius tipulae	Sporadic eggs; lived shorter in 25 degrees average life=17
Pangroliamus sp.	Large and long; often went missing; sporadic eggs despite regular replating; average life=29.5
Pristionchus sp.	Large bodied with lots of eggs right away; room temp plates had a very long life of last adult; some dug in; lived shorter at 25 degrees, but only due to longevity of last few adults on room temp; average life=38

CGC Specimens	
Acrobeloides butchlii	Small, thick adults that were long-lived. Not as transparent as other Acrobeloides species.
Acrobeloides ellesmerensis	Small and slow-growing. They had a longer life span and produced few but steady streams of eggs. Not as thick as and more transparent than *A. buchlii*. Occasionally dug in.
Acrobeloides thornei	Small and slow-growing. Had a longer life span and produced steady stream of eggs. More active and dug in more than *A. ellesmerensis*
Cephalobus sp.	Small and slow-growing. Have a long life span with few eggs, but steady. Similar to Acrobeloides species, but slightly larger and more fecund. Sometimes dug in.
Dolichorhabditis sp.	Eggs grew very fast. Adults mobile and female with vulva sometimes extended.
Mesorhabditis longespiculosa	Fragile, slow-moving juveniles. Moderate egg number, short life. Male spiracles attached to female vulva sideways.
Mesorhabditis spiculigera	Longer bodies have dark intestines/reproductive tract. Much faster growing and fecund than *M. longespiculosa*. But not as active as *P. rigidus*.
Operculrhabditis sp.	Liked the edges of the agar plate. Few eggs early on and a single adult had a long life.
Oscheius akosreti	Very fast-growing and fecund. Had to restart with fewer individuals. Produced most eggs per plate yet observed and died quickly.
Oscheius dolichulroides	Long, active, and wide-ranging. Females with extended vulva. Would decompose fast after death and some endotokia observed.
Oscheius myriophila	Large, long, active, and roamed around a lot, disappearing and reappearing. Thick body with darker insides. Elders could die fast, sometimes in parallel, and exhibited endotokia. Large, old females could persist for a few days. Juveniles could hatch a grown very fast, to the same size as the adults that gave rise to them in just a day or two.
Oscheius sp.	Decent egg output (short timeframe) and longevity. Longer body and active motion. Enotokia observed.
Panagrellus redivivus	Long, vigorous, and wide-ranging. Observed grouping together occasionally. Juveniles were usually seen rather than eggs. Endotokia very common in older females. Would often coil when dying, but laid straight when dead.

Panagrolaimus rigidus	Long and usually gave birth to juveniles. Active and very prone to dig into agar.
Pellioditis typica	A small number of eggs and short-lived.
Pelodera strongyloides	Thicker and slower with a blunt tail. Dark insides very visible and extended vulva. Eggs may come out maladapted; small and dead juveniles from them. Liked to curl up but went straight when dead and could decompose fast. Some adults died and decomposed while full of eggs.
Pelodera teres	Medium size and moderate activity. Seen moving backward often. Died and broke apart fast. Few eggs and short life.
Pristionchus pacificus	Big and fast-growing with final eggs not hatching. Short life span, though elders would often get short and stout and lived longer in observation than on LHT. Would dig in a lot.
Rhabditella axei	Thin and could have a blunter tail; many eggs right away before dying within a short time
Rhabditis blumi	Moderate size, with offspring that could hatch in a day. A similar number of eggs and longevity to *R. myriophila*.
Rhabditis myriophila	Adults grew fast. Much fewer eggs than similar species, but same longevity.
Rhabditoides regina	Large adults, some having blunt tails (sex-specific, males probably). Adults liked to range outside the bacteria lawn. Several times, found smaller adults wrapped around larger at the middle; presumably mating. Many eggs right away before busting quickly. Endotokia occurred in old females.
Teratorhabditis palmarum	Thicker bodied and very prone to digging into the agar. Observed adult tails fused end to end, presumably mating. Moderate offspring. Often very active on low nutrient plates and would thrash heads about.
Zeldia sp.	Long, slow-growing, and prone to digging into agar. Few offspring.

References

"Research reported in this publication was supported by the National Institute of General Medical Sciences of the National Institutes of Health under Award Number U54GM128729 and Award number 2P20GM104360-06A1."

Aiyar, S. E., Gaal, T., & Gourse, R. L. (2002). rRNA Promoter Activity in the Fast-Growing Bacterium Vibrio natriegens rRNA Promoter Activity in the Fast-Growing Bacterium Vibrio natriegens. *Journal of Bacteriology, 184*(5), 1349–1358. http://doi.org/10.1128/JB.184.5.1349

Allen, M. W. & Noffsinger, E. M. (1972). A Review of the Genus Zeldia Thome, 1937 (Nematoda: Cephalobidae) with Descriptions of Seven New Species. *Proc. Of the Helmin. Society of Wash., 30*(2), 206-223.

Anderson, T., Elser, J.J., & Hessen, D.O. (2004). Stoichiometry and population dynamics. *Ecology Letters, 7*: 884-900

Angly, F. E., Dennis, P. G., Skarshewski, A., Vanwonterghem, I., Hugenholtz, P., & Tyson, G. W. (2014). CopyRighter: A rapid tool for improving the accuracy of microbial community profiles through lineage-specific gene copy number correction. *Microbiome*, *2*(1), 1–13. http://doi.org/10.1186/2049-2618-2-11

Arharya, K., Kyle, M., & Elser, J.J. (2004). Biological stoichiometry of Daphnia growth: An ecophysiological test of the growth rate hypothesis. *Limno. Oceanogr., 49*(3): 656-665

Azzalini, A. & Genz, A. (2016). The R package 'mnormt': The multivariate normal and 't' distributions (version 1.5-5). URL http://azzalini.stat.unipd.it/SW/Pkg-mnormt

Bardgett, R. D. (2005). The biology of soil: A community and ecosystem approach. Oxford: Oxford university press.

Barns, S. M., Fundyga, R. E., Jeffries, M. W., & Pace, N. R. (1994). Remarkable archaeal diversity detected in a Yellowstone National Park hot spring environment. *Proceedings of the National Academy of Sciences of the United States of America, 91*(5), 1609–1613. http://doi.org/10.1073/pnas.91.5.1609

Beardsley, C., Pernthaler, J., Wosniok, W., & Amann, R. (2003). Are readily culturable bacteria in coastal north sea waters suppressed by selective grazing mortality? *Applied and Environmental Microbiology, 69*(5), 2624–2630. 166

Begon, M., Harper, J. L., & Townsend, C. R. (1990). Ecology: Individuals, Populations, and Communities. 2nd edition. Cambridge, Massachusetts: Blackwell Scientific Publications, Inc.

Bell, K. L., Burgess, K. S., Botsch, J. C., Dobbs, E. K., Read, T. D., & Brosi, B. J. (2018). Quantitative and qualitative assessment of pollen DNA metabarcoding using constructed species mixtures. *Molecular Ecology*, (July). http://doi.org/10.1111/mec.14840

Bentjen, S. A., Fredrickson, J. K., Van Voris, P., & Li, S. W. (1989). Intact Soil-Core Microcosms for Evaluating the Fate and Ecological Impact of the Release of Genetically Engineered Microorganisms. *Applied and Environmental Microbiology*, 55(1), 198-202

Bik, H. M., Fournier, D., Sung, W., Bergeron, R. D., & Thomas, W. K. (2013). Intra-Genomic Variation in the Ribosomal Repeats of Nematodes. *PLoS ONE*, *8*(10), 1–8. http://doi.org/10.1371/journal.pone.0078230

Bintrim, S. B., Donohue, T. J., Handelsman, J., Roberts, G. P., & Goodman, R. M. (1997). Molecular phylogeny of Archaea from soil. Proceedings of the National Academy of Sciences of the United States of America, 94(1), 277–282. http://doi.org/10.1073/pnas.94.1.277

Bongers, T. (1990). The Maturity Index: An Ecological Measure of Environmental Disturbance Based on Nematode Species Composition. *Oecologia, 83*(1), 14–19.

Bongers, T. (1994). The Nematodes of the Netherlands. 2nd edition. Stichting Uitgeverij Koninklijke Nederlandse Natuurhistorische Vereniging.

Bongers, T., & Bongers, M. (1998). Functional Diversity of Nematodes. *Applied Soil Ecology*, *10*, 239–251.

Boscaro, V., Felletti, M., Vannini, C., Ackerman, M. S., Chain, P. S. G., Malfatti, S., Vergez, L. M., Shin, M., Doak, T. G., Lynch, M, & Petroni, G. (2013). Polynucleobacter necessarius, a model for genome reduction in both free-living and symbiotic bacteria. *Proceedings of the National Academy of Sciences of the United States of America, 110*(46), 18590–5. http://doi.org/10.1073/pnas.1316687110

Bowen, J. L., Ward, B. B., Morrison, H. G., Hobbie, J. E., Valiela, I., Deegan, L. A, & Sogin, M. L. (2011). Microbial community composition in sediments resists perturbation by nutrient enrichment. *The ISME Journal*, *5*(9), 1540–8. http://doi.org/10.1038/ismej.2011.22

Bowman, J. S., & Ducklow, H. W. (2015). Microbial communities can be described by metabolic structure: A general framework and application to a seasonally variable, depth-stratified microbial community from the coastal West Antarctic Peninsula. *PLoS ONE, 10*(8), 1–18. http://doi.org/10.1371/journal.pone.0135868

Brenner, S. (1974). The Genetics of Caenorhabditis Elegans. *Genetics*, *77*, 71–94.

Brock, E. H., Ketterings, Q. M., & Kleinman, P. J. A. (2007). Phosphorus leaching through intact soil cores as influenced by type and duration of manure application. *Nutrient Cycling in Agroecosystems, 77*(3), 269–281. http://doi.org/10.1007/s10705-006-9065-3 167

Broszat, M., Nacke, H., Blasi, R., Siebe, C., Huebner, J., Daniel, R., & Grohmanna, E. (2014). Wastewater irrigation increases the abundance of potentially harmful Gammaproteobacteria in soils in Mezquital Valley, Mexico. *Applied and Environmental Microbiology, 80*(17), 5282–5291. http://doi.org/10.1128/AEM.01295-14

Buckley, D. H., Graber, J. R., & Schmidt, T. M. (1998). Phylogenetic analysis of nonthermophilic members of the kingdom Crenarchaeota and their diversity and abundance in soils. *Applied and Environmental Microbiology, 64*(11), 4333–4339.

C. elegans Sequencing Consortium (1998). Genome Sequence of the Nematode C. elegans: A Platform for Investigating Biology. *Science, 282*(5396), 2012–2018.

Cadavid, V., & Stoltzfus, J. (2017). Transgenesis in Rhabditophanes sp. KR3021: a model for human parasitic nematodes. Society for Develop. Biology Mid-Atlantic Regional Meeting, https://pdfs.semanticscholar.org/baa7/dffb1272708ef0a0d02ab68f36c2c1f40abc.pdf

Capra, E., Giannico, R., Montagna, M., Turri, F., Cremonesi, P., Strozzi, F., Leone, P., Gandini, G., & Pizzi, F. (2016). A new primer set for DNA metabarcoding of soil Metazoa. *European Journal of Soil Biology*, *77*, 53–59. http://doi.org/10.1016/j.ejsobi.2016.10.005

Carter, M.R. & Gregorich, E.G. (2007) Soil Sampling and Methods of Analysis. 2nd Edition, CRC Press, Boca Raton.

Ceulemans, T., Stevens, C. J., Duchateau, L., Jacquemyn, H., Gowing, D. J. G., Merckx, R., Wallace, H., Rooijen, N. V., Goethem, T., Bobbink, R., Dorland, E., Gaudnik, C., Alard, D., Corcket, E., Muller, S., Dise, N. B., Dupre, C., Diekmann, M, & Honnay, O. (2014). Soil phosphorus constrains biodiversity across European grasslands. *Global Change Biology*, *20*(12), 3814–3822. https://doi.org/10.1111/gcb.12650

Cilia, V., Lafay, B., & Christen, R. (1996). Sequence heterogeneities among 16S ribosomal RNA sequences, and their effect on phylogenetic analyses at the species level. *Molecular Biology and Evolution, 13*(3), 451–461. https://doi.org/10.1093/oxfordjournals.molbev.a025606

Chen, H. (2018). VennDiagram: Generate High-Resolution Venn and Euler Plots. R package version 1.6.20. https://CRAN.R-project.org/package=VennDiagram

Chun, J., Huq, A, & Colwell, R. R. (1999). Analysis of 16S-23S rRNA intergenic spacer regions of Vibrio cholerae and Vibrio mimicus. *Applied and Environmental Microbiology*, *65*(5), 2202–2208.

Cock, P. A., Antao, T., Chang, J. T., Chapman, B. A., Cox, C. J., Dalke, A., Friedberg, I., Hamelryck, T., Kauff, F., Wilczynski, B. & de Hoon, M. J. L. (2009). Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics, 25*, 1422-1423

Coolon, J. D., Jones, K. L., Todd, T. C., Blair, J. M., & Herman, M. A. (2013). Long-Term Nitrogen Amendment Alters the Diversity and Assemblage of Soil Bacterial Communities in Tallgrass Prairie. *PLoS ONE, 8*(6). http://doi.org/10.1371/journal.pone.0067884

Cooper, G.M. (2000). The Cell: A Molecular Approach. 2nd edition. Sunderland (MA): Sinauer Associates. RNA Processing and Turnover. Available from: https://www.ncbi.nlm.nih.gov/books/NBK9864/

Condon, C., Philips, J., Fu, Z. Y., Squires, C., & Squires, C. L. (1992). Comparison of the expression of the seven ribosomal RNA operons in Escherichia coli. *The EMBO Journal, 11*(11), 4175–4185.

Condon, C., Liveris, D., Squires, C., Schwartz, I., & Squires, C. L. (1995). rRNA operon multiplicity in Escherichia coli and the physiological implications of rrn inactivation. *Journal of Bacteriology, 177*(14), 4152–4156. http://doi.org/10.1128/jb.177.14.4152-4156.1995

Corsi, A.K., Wightman, B., & Chalfie, M. A. (2015) Transparent window into biology: A primer on Caenorhabditis elegans. *Wormbook.*

Cox, R. A. (2004). Quantitative relationships for specific growth rates and macromolecular compositions of Mycobacterium tuberculosis, Streptomyces coelicolor A3(2) and Escherichia coli B/r: an integrative theoretical approach. *Microbiology, 150*(5), 1413– 1426. http://doi.org/10.1099/mic.0.26560-0 168

Croll, N. A., & Smith, J. M. (1970). The sensitivity and responses of Rhabditis sp. to peripheral mechanical stimulation. *Proc. Helminth. Soc. Wash., 37*, 1-5.

Croll, N. A. (1971). The Behaviour of Nematodes: Their Activity, Senses, and Responses. New York: St. Martin's..

Csuros, M. & Miklos, I. (2009). Streamlining and large ancestral genomes in archaea inferred with a phylogenetic birth-and-death model. *Molecular Biology and Evolution*, *26*(9), 2087–2095. http://doi.org/10.1093/molbev/msp123

Cuypers, T. D., & Hogeweg, P. (2012). Virtual genomes in flux: An interplay of neutrality and adaptability explains genome expansion and streamlining. *Genome Biology and Evolution*, *4*(3), 212–229. http://doi.org/10.1093/gbe/evr141

Darby, B. J., Todd, T. C., & Herman, M. A. (2013). High-throughput amplicon sequencing of rRNA genes requires a copy number correction to accurately reflect the effects of management practices on soil nematode community structure. *Molecular Ecology*, *22*(21), 5456–5471. http://doi.org/10.1111/mec.12480

Darby, B. J., & Herman, M. A. (2014). Effect of prey richness on a consumer's intrinsic growth rate. *Oecologia, 175*(1), 243–250. https://doi.org/10.1007/s00442-014-2883-0

Delagado-Baquerizo, M., Oliverio, A. M., Brewer, T. E., Benavent-González, A., Eldridge, D. J., Bardgett, R. D., Maestre, F. T., Singh, B. J, & Fierer, N. (2018) A global atlas of the dominant bacteria found in the soil. *Science, 359*: 320-325

Delattre, M., & Félix, M. A. (2001). Microevolutionary studies in nematodes: A beginning. *BioEssays*, *23*(9), 807–819. https://doi.org/10.1002/bies.1116

DeLong E. F. (2009). The microbial ocean from genomes to biomes. *Nature, 459*(14). doi:10.1038/nature08059

Dichtel, M., Louvet-Vallée, S., Viney, M. E., Félix, M., & Sternberg, P. W. (2000). Control of Vulval Cell Division Number in the Nematode *Oscheius/Dolichorhabditis* sp. CEW1. *Genetics 157*, 183-197.

Dodd, R. J., McDowell, R. W., & Condron, L. M. (2014). Manipulation of fertiliser regimes in phosphorus enriched soils can reduce phosphorus loss to leachate through an increase in pasture and microbial biomass production. *Agriculture, Ecosystems and Environment, 185*, 65–76. http://doi.org/10.1016/j.agee.2013.12.018

Dos Santos, G. A. P., Derycke, S., Genevois, V. G. F., Coelho, L. C. B. B., Correia, M. T. S., & Moens, T. (2009). Interactions among bacterial-feeding nematode species at different levels of food availability. *Marine Biology, 156*(4), 629–640. https://doi.org/10.1007/s00227-008-1114-9

Edgar, R. C. (2010) Search and clustering orders of magnitude faster than BLAST, *Bioinformatics 26*(19), 2460-2461.

Edgar, R. C. (2017). UNBIAS: An attempt to correct abundance bias in 16S sequencing, with limited success. *BioRxiv*, 124149. http://doi.org/10.1101/124149 169

Eilers, H., Pernthaler, J., & Amann, R. (2000). Succession of pelagic marine bacteria during enrichment: A close look at cultivation-induced shifts. *Applied and Environmental Microbiology, 66*(11), 4634–4640. http://doi.org/10.1128/AEM.66.11.4634-4640.2000

Ekschmitt, K., & Korthals, G. (2009). Molecular Markers, Indicator Taxa, and Community Indices: the Issue of Bioindication Accuracy. Nematodes as Environmental Indicators. Cambridge: CAB International.

el-Azay, O. M., el-Gawady, H.M., & Nada, M. S. (1968). The occurrence of Rhabditis (Rhabditella) axei in the faeces of a chicken in Egypt. *Journal of Helminthology, 62*(3), 219-220.

Elser, J. J., Dobberfuhl, D. R., MacKay. N. A., & Schampel, J. H. (1996). Organism Size, Life History, and N:P Stoichiometry. *BioScience, 46*(9): 674-684

Elser, J. J., Fagan, W. F., Denno, R. F., Dobberfuhl, D. R., Folarin, A., Huberty, A., Interlandi, S., Kilham, S. S., McCauley, E., Schulz, K. L., Siemann, E. H., & Sterner, R. W. (2000a). Nutritional constraints in terrestrial and freshwater food webs. *Nature, 408*(30): 578-580

Elser, J. J., O'Brien, W. J., Dobberfuhl, D. R., & Dowling, T. E. (2000b). The evolution of ecosystem processes: growth rate and elemental stoichiometry of a key herbivore in temperate and arctic habitats. *J. Evol. Biol., 13*: 845-853

Elser, J. J., Sterner, R. W., Gorokhova, E., Fagan, W. F., Markow, T. A., Cotner, J. B., Harrison, J. F., Hobbie, S. E., Odell, G. M., & Weider, L. J. (2000c). Biological stoichiometry from genes to ecosystems. *Ecology Letters, 3*: 540-550

Elser, J. J., Arharya, K., Kyle, M., Cotner, J., Makino, W., Markow, T, Hobbie, S., Fagan, W., Schade, J., Hood, J., & Sterner, R. W. (2003). Growth rate-stoichiometry couplings in diversie biota. *Ecology Letters, 6*: 936-943

Erkosar, B., Yashiro, E., Zajitschek, F., Friberg, U., Maklakov, A. A., van der Meer, J. R., & Kawecki, T. J. (2018). Host diet mediates a negative relationship between abundance and diversity of *Drosophila* gut microbiota. *Ecology and Evolution*, (July), 1–12. http://doi.org/10.1002/ece3.4444

Espejo, R. T., & Plaza, N. (2018). Multiple Ribosomal RNA operons in bacteria; Their concerted evolution and potential consequences on the rate of evolution of their 16S rRNA. *Frontiers in Microbiology, 9*(JUN), 1–6. http://doi.org/10.3389/fmicb.2018.01232

Eyualem-Abebe, Andrassy, I. & Traunspurger, W. (2006). Freshwater Nematodes: Ecology and Taxonomy. CAB International.

Félix M. A., Delattre, M., & Dichtel, M. L. (2000). Comparative developmental studies using Oscheius/Dolichorhabditis sp. CEW1 (Rhabditidae). *Nematology, 2*:89–98.

Félix, M.A. (2006) Oscheius tipulae, *WormBook*.

Félix, M. A., & Braendle, C. (2010). The natural history of Caenorhabditis elegans. *Current Biology, 20*(22), R965–R969. https://doi.org/10.1016/j.cub.2010.09.050 170

Ferris, H., Lau, S., & Venette, R. (1995). Population energetics of bacterial-feeding nematodes: Respiration and metabolic rates based on CO 2 production. *Soil Biology and Biochemistry, 27*(3), 319–330. https://doi.org/10.1016/0038-0717(94)00186-5

Ferris, H., Eyre, M., Venette, R. C., & Lau, S. S. (1996). Population energetics of bacterial-feeding nematodes: Stage-specific development and fecundity rates. *Soil Biology and Biochemistry, 28*(3), 271–280. https://doi.org/10.1016/0038-0717(95)00127-1

Ferris, H., Bongers, T., & De Goede, R. G. M. (2001). A framework for soil food web diagnostics: Extension of the nematode faunal analysis concept. *Applied Soil Ecology, 18*(1), 13–29. http://doi.org/10.1016/S0929-1393(01)00152-4

Ferris, H., & Bongers, T. (2006). Nematode indicators of organic enrichment. *Journal of Nematology*, *38*(1), 3–12.

Ferris, H., & Matute, M. M. (2003). Structural and functional succession in the nematode fauna of a soil food web. *Applied Soil Ecology*, *23*(2), 93–110. http://doi.org/10.1016/S0929-1393(03)00044-1

Ferris, H. & Bongers, T. (2009). Indices Developed Specifically for Analysis of Nematode Assemblages. Nematodes as Environmental Indicators. Cambridge: CAB International.

Fierer, N., & Jackson, R. B. (2006). The diversity and biogeography of soil bacterial communities. *Proceedings of the National Academy of Sciences of the United States of America, 103*(3), 626–631. http://doi.org/10.1073/pnas.0507535103

Fierer, N., Bradford, M. A., & Jackson, R. B. (2007). Toward an ecological classification of soil bacteria. *Ecology, 88*(6), 1354–1364. http://doi.org/10.1890/05-1839

Fierer, N., Lauber, C. L., Ramirez, K. S., Zaneveld, J., Bradford, M. A., & Knight, R. (2012). Comparative metagenomic, phylogenetic and physiological analyses of soil microbial communities across nitrogen gradients. *The ISME Journal*, *6*(5), 1007–17. http://doi.org/10.1038/ismej.2011.159

Floyd, R., Abebe, E., Papert, A., & Blaxter, M. (2002). Molecular barcodes for soil nematode identification. *Molecular Ecology, 11*(4), 839–850. http://doi.org/10.1046/j.1365-294X.2002.01485.x

Franzo, A., & Del Negro, P. (2019). Functional diversity of free-living nematodes in river lagoons: can biological traits analysis (BTA) integrate traditional taxonomic-based approaches as a monitoring tool? *Marine Environmental Research*, *145*(March), 164–176. http://doi.org/10.1016/j.marenvres.2019.02.015

Fuhrman, J. A., McCallum, K., & Davis, A. A. (1993). Phylogenetic diversity of subsurface marine microbial communities from the Atlantic and Pacific Oceans. *Applied and Environmental Microbiology, 59*(5), 1294–1302.

Fürst von Lieven, A., & Sudhaus, W. (2000). Comparative and functional morphology of the buccal cavitiy of Diplogastrina (Nematoda) and a first outline of the phylogeny of this taxon. *J. Zoolog. Syst. Evol. Res. 38*, 37–63. 171

Gagliardi, J. V., Angle, J. S., Germida, J. J., Wyndham, R. C., Chanway, C. P., Watson, R. J., Greer, C., W., McIntyre, T., Yu, H. H., Levin, M. A., Russek-Cohen, E., Rosolen, S., Nairn, J., Seib, A., Martin-Heller, T., & Wisse, G. (2001). Intact soil-core microcosms compared with multi-site field releases for pre-release testing of microbes in diverse soils and climates. *Canadian Journal of Microbiology, 47*(3), 237–252. http://doi.org/10.1139/cjm-47-3-237

Galili, T. (2015). dendextend: an R package for visualizing, adjusting, and comparing trees of hierarchical clustering. *Bioinformatics*. DOI: 10.1093/bioinformatics/btv428\

Gao, Z. M., Wang, Y., Tian, R. M., Wong, Y. H., Batang, Z. B., Al-Suwailem, A. M., Bajic, V. B., & Qian, P. Y. (2014). Symbiotic Adaptation Drives Genome Streamlining of the Cyanobacterial Sponge Symbiont "Candidatus Synechococcus spongiarum". *MBio, 5*(2), 1–11. http://doi.org/10.1128/mBio.00079-14

Gerber, K., & Giblin-Davis, R. M. (1990). Teratorhabditis palmarum n. sp. (Nemata: Rhabditidae): An Associate of Rhynchophorus palmarum and R. cruentatus. *Journal of nematology, 22*(3), 337–347.

Ghani, A., Cox, N., Sarathchandra, S., Yeates, G., & Burch, G. (2002). Effect of nitrogen and phosphate fertilizers on microbial and nematode diversity in pasture soils. *Soil Biology and Biochemistry*, *33*(7–8), 953–964. http://doi.org/10.1016/s0038-0717(00)00245-5

Giovannoni, S.J., & Rappe, M.S. (2003). The uncultured microbial majority. *Annual Review of Microbiology, 57*: 369-394

Giovannoni, S. J., Cameron Thrash, J., & Temperton, B. (2014). Implications of streamlining theory for microbial ecology. *The ISME Journal*, *8*(8), 1–13. http://doi.org/10.1038/ismej.2014.60

Goodey, T. (1963). Soil and Freshwater Nematodes. CAB International.

Gorokhova, E., Dowling, T. E., Weider, L. J., Crease, T. J., & Elser, J. J. (2002). Functional and ecological significance of rDNA intergenic spacer variation in a clonal organism under divergent selection for production rate. *Proceedings. Biological Sciences / The Royal Society*, *269*(1507), 2373–2379. http://doi.org/10.1098/rspb.2002.2145

Graham, T. A., Golsteyn-Thomas, E. J., Thomas, J. E., & Gannon, V. P. J. (2009). Inter- and Intraspecies Comparison of the 16S-23S rRNA Operon Intergenic Spacer Regions of Six Listeria spp. *International Journal of Systematic Bacteriology, 47*(3), 863–869. http://doi.org/10.1099/00207713-47-3-863

Greet, D. N. (1964). Observation on sexual attraction and copulation in the nematode, *Panagroliamus rigidus*, Schneider. *Nature,204*, 96.

Griffiths, B. S., de Groot, G. A., Laros, I., Stone, D., & Geisen, S. (2018). The need for standardization: Exemplified by a description of the diversity, community structure and ecological indices of soil nematodes. *Ecological Indicators*, *87*(December 2017), 43–46. http://doi.org/10.1016/j.ecolind.2017.12.002

Grote, J., Thrash, J. C., & Huggett, M. J. (2012). Streamlining and Core Genome Conservation among Highly Divergent Members of the SAR11 Clade. *mBio 3*(5), 1–13. http://doi.org/10.1128/mBio.00252-12.Editor 172

Gürtler, V., & Stanisich, V. A. (1996). New approaches to typing and identification of bacteria using the 16S-23S rDNA spacer region. *Microbiology 142*:3–16

Güsewell, S. (2004). N: P ratios in terrestrial plants: Variation and functional significance. *New Phytol 164*:243–266.

Gyorfy, Z., Draskovits, G., Vernyik, V., Blattner, F. F., Gaal, T., & Posfai, G. (2015). Engineered ribosomal RNA operon copy-number variants of E. coli reveal the evolutionary trade-offs shaping rRNA operon number. *Nucleic Acids Research*, *43*(3), 1783–1794. http://doi.org/10.1093/nar/gkv040

Hayat, R., Ali, S., Amara, U., Khalid, R., & Ahmed, I. (2010). Soil beneficial bacteria and their role in plant growth promotion: A review. *Annals of Microbiology*, *60*(4), 579–598. http://doi.org/10.1007/s13213-010-0117-1

Hessen, D. O., Ventura, M., & Elser, J. J. (2008). Do phosphorus requirements for RNA limit genome size in crustacean zooplankton? *Genome 51*(9): 685-691.

Hessen, D. O., Jeyasingh, P. D., Neiman, M., & Weidre, L. J. (2010a). Genome streamlining and the elemental costs of growth. *Trends in Ecology and Evolution, 25*(2): 75-80

Hessen, D. O., Jeyasingh, P. D., Neiman, M., & Weidre, L. J. (2010b). Genome streamlining in prokaryotes versus eukaryotes. *Trends in Ecology and Evolution, 25*(6) 320-321

Holterman, M., Van Der Wurff, A., Van Den Elsen, S., Van Megen, H., Bongers, T., Holovachov, O., Bakker, J., & Helder, J. (2006). Phylum-wide analysis of SSU rDNA reveals deep phylogenetic relationships among nematodes and accelerated evolution toward crown clades. *Molecular Biology and Evolution*, *23*(9), 1792–1800. http://doi.org/10.1093/molbev/msl044

Jakobsen, I., Gazey, C., & Abbott, L. K. (2001). Phosphate transport by communities of arbuscular mycorrhizal fungi in intact soil cores. *New Phytologist, 149*(1), 95–103. http://doi.org/10.1046/j.1469-8137.2001.00006.x

Janezic, S., Indra, A., Rattei, T., Weinmaier, T., & Rupnik, M. (2014). Recombination drives evolution of the Clostridium difficile 16S-23S rRNA intergenic spacer region. *PLoS ONE*, *9*(9). http://doi.org/10.1371/journal.pone.0106545

Jenkins, W. R. (1964). A rapid centrifugal-flotation technique for separating nematodes from soil. *Plant Disease Reporter 48*:692. 173

Jones, K. L., Todd, T. C., Wall-Beam, J. L., Coolon, J. D., Blair, J. M., & Herman, M. A. (2006). Molecular approach for assessing responses of microbial-feeding nematodes to burning and chronic nitrogen enrichment in a native grassland. *Molecular Ecology*, *15*(9), 2601–2609. http://doi.org/10.1111/j.1365-294X.2006.02971.x

Kay, A. D., Ashton, I. W., Gorohkova, E., Kerkhoff, A. J., Liess, A., & Litchman, E. (2005). Toward a stoichiometric framework for evolutionary biology. *Oiko, 109*: 6-17

Kearns, P. J., Angell, J. H., Howard, E. M., Deegan, L. A., Stanely, R. H. R., & Bowen, J. L. (2016). Nutrient enrichment induces dormancy and decreases diversity of active bacteria in salt marsh sediments. *Nature communications 7*: 12881-12881.

Kembel, S. W., Wu, M., Eisen, J. A., & Green, J. L. (2012). Incorporating 16S Gene Copy Number Information Improves Estimates of Microbial Diversity and Abundance. *PLoS Computational Biology, 8*(10), 16–18. http://doi.org/10.1371/journal.pcbi.1002743

Klappendbach, J. A., Dunbar, J. M., & Schmidt, T. M. (2000). rRNA operon copy number reflects ecological strategies of bacteria. *Applied and Environmental Microbiology*, 1328-1333

Kumar, S., Stecher, G., & Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for Bigger Datasets. *Molecular Biology and Evolution, 33*(7):1870-4 180

Kuo S. (1996) Phosphorus. *Methods of soil Analysis. Part 3*. Chemical Methods-SSSA Book Series no. 5.

Kurm, V., Van Der Putten, W. H., De Boer, W., Naus-Wiezer, S., & Gera Hol, W. H. (2017). Low abundant soil bacteria can be metabolically versatile and fast growing. *Ecology, 98*(2), 555–564. http://doi.org/10.1002/ecy.1670

Kwan, J. C., Donia, M. S., Han, A. W., Hirose, E., Haygood, M. G., & Schmidt, E. W. (2012). Genome streamlining and chemical defense in a coral reef symbiosis. *Proceedings of the National Academy of Sciences of the United States of America, 109*(50), 20655–20660. http://doi.org/10.1073/pnas.1213820109

Larsson, J., Nylander, J. A., & Bergman, B. (2011). Genome fluctuations in cyanobacteria reflect evolutionary, developmental and adaptive traits. *BMC Evolutionary Biology, 11*(1), 187. http://doi.org/10.1186/1471-2148-11-187

Lan, R., Pérez Luz, S., Reeves, P. R., & Rodríguez-Valera, F. (1998). Variation of the ribosomal operon 16S-23S gene spacer region in representatives of Salmonella enterica subspecies. *Journal of Bacteriology*, *180*(8), 2144–2151.

Lane, D. J., Pace, B., Olsen, G. J., Stahlt, D. A., Sogin, M. L., & Pace, N. R. (1985). Rapid determination of 16S ribosomal RNA sequences for phylogenetic analyses (reverse transcriptase/dideoxynudeotide). *Evolution*, *82*(October), 6955–6959.

Langille, M. G. I., Zaneveld, J., Caporaso, J. G., McDonald, D., Knights, D., Reyes, J. A., Clemente, J. C., Burkpile, D. E., Vega-Thurber, R. L., Knight, R., Beiko, R. J. & Huttenhower, C. (2013). Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nature Biotechnology*, *31*(9), 814–821. http://doi.org/10.1038/nbt.2676 174

Lauber, C. L., Strickland, M. S., Bradford, M. A., & Fierer, N. (2008). The influence of soil properties on the structure of bacterial and fungal communities across land-use types. *Soil Biology and Biochemistry*, *40*(9), 2407–2415. http://doi.org/10.1016/j.soilbio.2008.05.021

Lee, C., Lee, S., & Shin, S. G. (2008). Real-time PCR determination of rRNA gene copy number: absolute and relative quantification assays with *Escherichia coli*. *Appl Microbiol Biotechnol, 78* 371-376

Leprince, A., van Passel, M. W. J., & dos Santos, V. A. P. M. (2012). Streamlining genomes: Toward the generation of simplified and stabilized microbial systems. *Current Opinion in Biotechnology*, *23*(5), 651–658. http://doi.org/10.1016/j.copbio.2012.05.001

Lewis, S. C., Dyal, L. A., Hilburn, C. F., Weitz, S., Liau, W., LaMunyon, C. W., & Denver, D. R. (2009). Molecular evolution in Panagrolaimus nematodes: origins of parthenogenesis, hermaphroditism and the Antarctic species *P. davidi*. *BMC Evolutionary Biology, 9*(15), 1-13

Li, F., Neher, D., Darby, B., & Weicht, T. (2005). Observed Differences in Life History Characteristics of Nematodes Aphelenchus and Acrobeloides upon Exposure to Copper and Benzo(a)pyrene. *Ecotoxicology* (London, England). 14. 419-29. 10.1007/s10646-004-1347-4.

Li, Y., Cao, Z., Hu, C., Li, J., & Yang, H. (2014). Response of nematodes to agricultural input levels in various reclaimed and unreclaimed habitats. *European Journal of Soil Biology*, *60*, 120–129. http://doi.org/10.1016/j.ejsobi.2013.12.001

Liang, L. Z., Zhao, X. Q., Yi, X. Y., Chen, Z. C., Dong, X. Y., Chen, R. F., & Shen, R. F. (2013). Excessive application of nitrogen and phosphorus fertilizers induces soil acidification and phosphorus enrichment during vegetable production in Yangtze River Delta, China. *Soil Use and Management, 29*(2), 161–168. http://doi.org/10.1111/sum.12035

Liu, T., Yu, L., Xu, J., Yan, X., Li, H., Whalen, J. K., & Hu, F. (2017). Bacterial traits and quality contribute to the diet choice and survival of bacterial-feeding nematodes. *Soil Biology and Biochemistry, 115*, 467–474. https://doi.org/10.1016/j.soilbio.2017.09.014

Liu, T., Yu, L., Li, M., Wu, J., Li, H., Whalen, J. K., & Hu, F. (2018). Food familiarity does not change nematode feeding behavior. *Soil Biology and Biochemistry, 125*(July), 136–143. https://doi.org/10.1016/j.soilbio.2018.07.011

Lofgren, L. A., Uehling, J. K, Branco, S., Bruns, T. D., Martin, F., & Kennedy, P. G. (2019). Genome-based estimates of fungal rDNA copy number variation across phylogenetic scales and ecological lifestyles. *Molecular Ecology 28*(4): 721-730.

Long, E.O. & Dawid, I.B. (1980). Repeated genes in eukaryotes. *Ann. Rev. Biochem., 49*, 727±764.

Longa, C. M. O., & Pertot, I. (2009). An intact soil-core microcosm method to evaluate the survival and vertical dispersal of Trichoderma atroviride SC1. *Letters in Applied Microbiology, 49*(5), 609–614. http://doi.org/10.1111/j.1472-765X.2009.02715.x

Louca, S. & Doebeli, M. (2017). Efficient comparative phylogenetics on large trees. *Bioinformatics* htpp://doi.org/10.1093/bioinformatics/btx70

Louca, S., Doebeli, M., & Parfrey, L. W. (2018). Correcting for 16S rRNA gene copy numbers in microbiome surveys remains an unsolved problem. *Microbiome, 6*(1), 1–12. http://doi.org/10.1186/s40168-018-0420-9

MacArthur, R. A., & Wilson, E. O. (1967) The Theory of Island Biogeography. Princeton University Press.

Main, T. M., Dobberfuhl, D. R., & Elser, J. J. (1997). N:P stoichiometry and ontogeny of crustacean zooplankton: A test of the growth rate hypothesis. *Limnol. Oceanogr., 42*(6): 1474-1478

Maniloff, J. (1996). The minimal cell genome: "on being the right size". *Proceedings of the National Academy of Sciences of the United States of America, 93*(19), 10004–6. http://doi.org/10.1073/pnas.93.19.10004

Mantel, N. (1967). The detection of disease clustering and a generalized regression approach. *Cancer Research. 27* (2): 209–220

Marchant, R., & Nicholas, W. L. (2012). An Energy Budget for the Free-Living Nematode Pelodera (Rhabditidae). *Oecoligia; 16*(3): 237-252

Marschner, P., Kandeler, E., & Marschner, B. (2003). Structure and function of the soil microbial community in a long-term fertilizer experiment. *Soil Biology and Biochemistry*, *35*(3), 453–461. http://doi.org/10.1016/S0038-0717(02)00297-3

Martnez-Cano, D. J., Reyes-Prieto, M., Martnez-Romero, E., Partida-Martnez, L. P., Latorre, A., Moya, A., & Delaye, L. (2015). Evolution of small prokaryotic genomes. *Frontiers in Microbiology*, *6*(JAN), 1–23. http://doi.org/10.3389/fmicb.2014.00742

Matyášek, R., Renny-Byfield, S., Fulneček, J., Macas, J., Grandbastien, M.-A., Nichols, R., Leitch, A., & Kovařík, A. (2012). Next generation sequencing analysis reveals a relationship between rDNA unit diversity and locus number in Nicotiana diploids. *BMC Genomics*. http://doi.org/10.1186/1471-2164-13-722

McBride, J. M., & Hollis, J. P. (1966). The phenomenon of swarming in nematodes. *Nature., 211*, 545-546.

Mira, A., Ochman, H., & Moran, N. A. (2001). Deletional bias and the evolution of bacterial genomes. *Trends in Genetics*, *17*(10), 589–596. http://doi.org/10.1016/S0168-9525(01)02447-7

Moens, T. & Vincx, M. (1997). Observations on the feeding ecology of estuarine nematodes. *Journal of the Marine Biology Assoc. U.K. 77*:211-227.

Moens, T., Vierstraete, T. M. & Vincx, M. (1996). Life Strategies in Two Bacterivorous Marine Nematodes: Preliminary Results. *Marine Ecology, 17*: 509-518.

Moens, T. & Vincx, M. (2000). Temperature, salinity and food thresholds in two brackish-water bacterivorous nematode species: assessing niches from food absorption and respiration experiments. *Journal of Experimental Marine Biology and Ecology, 243*(1), 137-154.

Morris, R. M., Rappé, M. S., Connon, S. A., Vergin, K. L., Siebold, W. A., Carlson, C. A., & Giovannoni, S. J. (2002). SAR11 clade dominates ocean surface bacterioplankton communities. *Nature, 420*(6917), 806–810. http://doi.org/10.1038/nature01240

Münkemüller, T., Lavergne, S., Bzeznik, B., Dray, S., Jombart, T., Schiffers, K., & Thuiller, W. (2012). How to measure and test phylogenetic signal. *Methods in Ecology and Evolution 3*(4): 743-756.

Munson, M. A., Nedwell, D. B., & Embley, T. M. (1997). Phylogenetic diversity of Archaea in sediment samples from a coastal salt marsh. *Applied and Environmental Microbiology, 63*(12), 4729–4733.

Muschiol, D., Schroeder, F., & Traunspurger, W. (2009). Life cycle and population growth rate of Caenorhabditis elegans studied by a new method. *BMC Ecology, 9*, 1–13. https://doi.org/10.1186/1472-6785-9-14 176

MyGlynn, R. D. (2018). Genomic Regulation of Abiotic Stress Response in the Soil Nematode Oscheius tipulae. Dissertation, University of North Dakota, Grand Forks, ND.

Nagy, P. (2009). Case Studies Using Nematode Assemblage Analysis in Terrestrial Habitats. Nematodes as Environmental Indicators. Cambridge: CAB International.

Neher, D. A. (1999). Nematode communities in organically and conventionally managed agricultural soils. *Journal of Nematology, 31*(2), 142–154.

Neher, D. A. & Darby, B. J. (2009) General Community Indices that can be used for Analysis of Nematode Assemblages. Nematodes as Environmental Indicators. Cambridge: CAB International.

Neilson, R., Donn, S., Griffiths, B., Daniell, T., Rybarczyk-Mydlowska, K., Elsen, S. J. J., Mooijman, P., & Johannes, H. (2009). Molecular Tools for Analyzing Nematode Assemblages. Nematodes as Environmental Indicators. Cambridge: CAB International.

Nelson, K. E., Cheng, G., He, X., Lux, R., Yooseph, S., Esquenazi, E., Hunter, R. C., Cheng, G., Nelson, K. E., & Shi, W. (2014). Cultivation of a human-associated TM7 phylotype reveals a reduced genome and epibiotic parasitic lifestyle. *Proceedings of the National Academy of Sciences, 112*(1), 244–249. http://doi.org/10.1073/pnas.1419038112

Nemergut, D. R., Knelman, J. E., Ferrenberg, S., Bilinski, T., Melbourne, B., Jiang, L, Violle, C., Darcy, J. L., Prest, T., Schmidt, S. K., & Townsend, A. R.. (2016). Decreases in average bacterial community rRNA operon copy number during succession. *ISME Journal 10*(5): 1147-1156.

Niederdorfer, R., Besemer, K., Battin, T. J. & Peter, H. (2017). Ecological strategies and metabolic trade-offs of complex environmental biofilms. *npj Biofilms and Microbiomes 3*(1).

Niemann, R., Arens, M., Koczwara, K., & Sturhan, D. (1996). Studies on the potential of nematodes for evaluating the quality of flowing waters. *Mitteilungen aus der Biologischen Bundesansalt für Land – und Forstwirschaft Berlin-Dahlem* 317, 195-208.

Nigon, V. M. & Félix, M. A. (2017). History of research on C. elegans and other free-living nematodes as model organisms. *WormBook*.

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P. R., O'Hara, R. B., Simpson, G. L., Solymos, P., Stevens, M. H. H., Szoecs, E., & Wagner, H. (2018). vegan: Community Ecology Package. R package version 2.5-2. https://CRAN.R-project.org/package=vegan

Olatunji, O. A., Gong, S., Tariq, A., Pan, K., Sun, X., Chen, W., Zhang, L., Dakhil, M. A., Huang, D, & Tan, X. (2019). The effect of phosphorus addition, soil moisture, and plant type on soil nematode abundance and community composition. *Journal of Soils and Sediments*, *19*(3), 1139–1150. http://doi.org/10.1007/s11368-018-2146-5

Osorio, C. R., Collins, M. D., & Toranzo, A. E. (2005). Variation in 16S-23S rRNA Intergenic Spacer Regions in *Photobacterium damselae*: a Mosaic-Like Structure. *Applied and Environmental Microbiology, 71*(2), 636–645. http://doi.org/10.1128/AEM.71.2.636

Ortiz-Álvarez, R., Fierer, N., de los Rios, A., Casamayor, E., & Barberan, A. (2018). Consistent changes in the taxonomic structure and functional attributes of bacterial communities during primary succession. *ISME Journal 12*(7): 1658-1667.

Paradis E., Claude J. & Strimmer K. (2004). APE: analyses of phylogenetics and evolution in R language. *Bioinformatics 20*: 289-290.

Tarroso, P., Velo-Anton, G., & Carvalho, S. (2015). phylin: Spatial Interpolation of Genetic Data. R package version 1.1.1. https://CRAN.R-project.org/package=phylin 177

Peham, T., Steiner, F. M., Schlick-Steiner, B. C., & Arthofer, W. (2017). Are we ready to detect nematode diversity by next generation sequencing? *Ecology and Evolution*, (March), 4147–4151. http://doi.org/10.1002/ece3.2998

Pereira, T. J., & Baldwin, J. G. (2016). Contrasting evolutionary patterns of 28S and ITS rRNA genes reveal high intragenomic variation in Cephalenchus (Nematoda): Implications for species delimitation. *Molecular Phylogenetics and Evolution, 98*, 244–260. https://doi.org/10.1016/j.ympev.2016.02.016

Perisin, M., Vetter, M., Gilbert, J. A., & Bergelson, J. (2016). 16Stimator: statistical estimation of ribosomal gene copy numbers from draft genome assemblies. *The ISME Journal*, *10*(4), 1020–1024. http://doi.org/10.1038/ismej.2015.161

Porazinska, D. L., Giblin-Davis, R. M., Faller, L., Farmerie, W., Kanzaki, N., Morris, K., Powers, T. O., Tucker, A. E., Sung, W., & Thomas, W. K. (2009). Evaluating high-throughput sequencing as a method for metagenomic analysis of nematode diversity. *Molecular Ecology Resources, 9*(6), 1439–1450. http://doi.org/10.1111/j.1755-0998.2009.02611.x

Porazinska, D. L., Giblin-Davis, R. M., Powers, T. O., & Thomas, W. K. (2012). Nematode Spatial and Ecological Patterns from Tropical and Temperate Rainforests. *PLoS ONE, 7*(9), 1–11. http://doi.org/10.1371/journal.pone.0044641

Preston, C. M., Wu, K. Y., Molinski, T. F., & DeLong, E. F. (1996). A psychrophilic crenarchaeon inhabits a marine sponge: Cenarchaeum symbiosum gen. nov., sp. nov. *Proceedings of the National Academy of Sciences of the United States of America, 93*(13), 6241–6246. http://doi.org/10.1073/pnas.93.13.6241

Prokopowich, C. D., Gregory, T. R., & Crease, T.J. (2002). The correlation between rDNA copy number and genome size in eukaryotes. *Genome; 45*(1): 48-50

Providenti, M. A., Mautner, S. I., Chaudhry, O., Bombardier, M., Scroggins, R., Gregorich, E., & Smith, M. L. (2004). Determining the environmental fate of a filamentous fungus, Trichoderma reesei, in laboratory-contained intact soil-core microcosms using competitive PCR and viability plating. *Canadian Journal of Microbiology, 50*(8), 623–31. http://doi.org/10.1139/w04-053

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., & Glöckner, F. O. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucl. Acids Res. 41* (D1): D590-D596.

Qudsia, T., Razia, S., & Rahmat, K. (2007). Species of Teratorhabditis (Osche, 1952) Dougherty, 1953 (Nematoda:Rhabditida) from sewage and manure samples in India including a discussion on relationships within the genus. *Hydrobiologia. 583*. 127-140.

R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rashid, F., Geraert, E., & Sharma, R. D. (1984). Morphology, Taxonomy, and Morphometry of Some Cephalobidae (Nematoda: Rhabditida) from Brazil, with Descriptions of Two New Genera and Four New Species. *Nematologica, 30*, 251-299. 178

Rastogi, R., Wu, M., Dasgupta, I., & Fox, G. E. (2009). Visualization of ribosomal RNA operon copy number distribution. *BMC Microbiology, 9*, 208. http://doi.org/10.1186/1471-2180-9-208

Redfield, A. (1934). On the proportions of organic derivatives in sea water and their relation to the composition of plankton. In Daniel, R.J. (ed James Johnstone Memorial Volume). University Press of Liverpool, 177–192.

Revell, L. J. (2012). phytools: An R package for phylogenetic comparative biology (and other things). *Methods Ecol. Evol.* 3217-223. doi:10.1111/j.2041-210X.2011.00169.x

Reznick, D., Bryant, M. J., & Bashey, F. (2002). r- and K-Selection Revisited: The Role of Population Regulation in Life-History Evolution. *Ecology. 83*. 1509-1520. 10.2307/3071970.

Roller, B. R. K., Stoddard, S. F., & Schmidt, T. M. (2016). Exploiting rRNA operon copy number to investigate bacterial reproductive strategies. *Nature Microbiology, 1*(September), 1–7. http://doi.org/10.1038/nmicrobiol.2016.160

Ruess, L. & Dighton, J. (1996). Cultural studies on soil nematodes and their fungal hosts. *Nematological 42*, 330-346.

Saari, S. A., & Nikander, S. E. (2006). Pelodera (syn. Rhabditis) strongyloides as a cause of dermatitis—a report of 11 dogs from Finland. *Acta veterinaria Scandinavica, 48*(1), 18. Doi:10.1186/1751-0147-48-18

Sabath, N., Ferrada, E., Barve, A., & Wagner, A. (2013). Growth temperature and genome size in bacteria are negatively correlated, suggesting genomic streamlining during thermal adaptation. *Genome Biology and Evolution*, *5*(5), 966–977. http://doi.org/10.1093/gbe/evt050

Saeed, I. A. M., MacGuidwin, A. E., & Rouse, D. I. (1998). Effect of initial nematode population density on the interaction of Pratylenchus penetrans and Verticillium dahliae on 'Russet Burbank' potato. *Journal of Nematology 30*, 100–7.

Salinas, K. A., Edenborn, S.L., Sexstone, A. J., & Kotcon, J.B. (2007). Bacterial preferences of the bacterivorous soil nematode Cephalobus brevicauda (Cephalobidae): Effect of bacterial type and size. *Pedobiologia 51*: 55–64. 10.1016/j.pedobi.2006.12.003

Samad, M. S., Johns, C., Richards, K. G., Lanigan, G. J., de Klein, C. A. M., Clough, T. J., & Morales, S. E. (2017). Response to nitrogen addition reveals metabolic and ecological strategies of soil bacteria. *Molecular Ecology, 26*(20), 5500–5514. http://doi.org/10.1111/mec.14275

Santiago, C. B., Ricci, M. & Reyes-Lampa, A. (2004). Effect of nematode Panagrellus redivivus density on growth, survival, feed consumption and carcass composition of bighead carp Aristichthys nobilis (Richardson) larvae. *Journal of Applied Ichthyology, 20*: 22-27. Doi:10.1111/j.1439-0426.2004.00483.x

Sapkota, R., & Nicolaisen, M. (2015). High-throughput sequencing of nematode communities from total soil DNA extractions. *BMC Ecology*, *15*(1), 3. http://doi.org/10.1186/s12898-014-0034-4

Schiemer, F. (1983). Comparative Aspects of Food Dependence and Energetics of Freeliving Nematodes. *Oikos, 41*(1), 32–42.

Schulte, F & Poinar, George. (1991). Description of Rhabditis (Rhabditoides) regina n. sp. (Nematoda : Rhabditidae) from the body cavity of beetle larvae in Guatemala.

Sechi, V., De Goede, R. G. M., Rutgers, M., Brussaard, L., & Mulder, C. (2018). Functional diversity in nematode communities across terrestrial ecosystems. *Basic and Applied Ecology*, *30*, 76–86. http://doi.org/10.1016/j.baae.2018.05.004 179

Seesao, Y., Gay, M., Merlin, S., Viscogliosi, E., Aliouat-Denis, C. M., & Audebert, C. (2016). A review of methods for nematode identification. *Journal of Microbiological Methods*, *138*, 37–49. http://doi.org/10.1016/j.mimet.2016.05.030

Shah, P., Ding, Y., Niemczyk, M., Kudla, G., & Plotkin, J. B. (2013). Rate-limiting steps in yeast protein translation. *Cell, 153*(7), 1589–601. http://doi.org/10.1016/j.cell.2013.05.049

Shaw, E. A., Boot, C. M., Moore, J. C., Wall, D. H., & Baron, J. S. (2019). Long-term nitrogen addition shifts the soil nematode community to bacterivore-dominated and reduces its ecological maturity in a subalpine forest. *Soil Biology and Biochemistry*, *130*(November 2018), 177–184. http://doi.org/10.1016/j.soilbio.2018.12.007

Sherry, J., Scott, B. & Dutka, B. (1997). Use of various acute, sublethal and early life-stage tests to evaluate the toxicity of refinery effluents. *Environmental Toxicology and Chemistry, 16*: 2249-2257. Doi:10.1002/etc.5620161108

Sohlenius, B. (1969). Monoxenic Cultivation of Some Rhabditid Nematodes. *Oikos, 20*(2), 287–293.

Sohlenius, B. (1973). Growth and Reproduction of a Nematode Acrobeloides sp . Cultivated on Agar. *Oikos, 24*(1), 64–72.

Sommer, R.J. (2006). Pristionchus pacificus. *WormBook*.

Srinivasan, J., Dillman, A. R., Macchietto, M. G., Heikkinen, L., Lakso, M., Fracchia, K. M., Antoshechkin, I., Mortazavia, A., Wong, G., & Sternberg, P. W. (2013). The draft genome and transcriptome of Panagrellus redivivus are shaped by the harsh demands of a free-living lifestyle. *Genetics, 193*(4), 1279–1295. Doi:10.1534/genetics.112.148809

Srivastava, A. K. & D. Schlessinger (1990). Mechanism and regulation of bacterial ribosomal RNA processing. *Annu Rev Microbiol 44*: 105-129.

Staniland, L. N. (1957). The swarming of Rhabditid eelworms in mushroom houses. *Pl. Path., 6*, 61-62.

Stevenson, B. S., & Schmidt, T. M. (2004). Life History Implications of rRNA Gene Copy Number in Escherichia coli. *Society*, *70*(11), 6670–6677. http://doi.org/10.1128/AEM.70.11.6670

Stoddard S.F, Smith B.J., Hein R., Roller B.R.K. & Schmidt T.M. (2015). rrnDB: improved tools for interpreting rRNA gene abundance in bacteria and archaea and a new foundation for future development. *Nucleic Acids Research 2014*; doi: 10.1093/nar/gku1201

Strehl, B., Holtzendorff, J., Partensky, F., & Hess, W. R. (1999). A small and compact genome in the marine cyanobacterium Prochlorococcus marinus CCMP 1375: Lack of an intron in the gene for tRNA(Leu)(UAA) and a single copy of the rRNA operon. *FEMS Microbiology Letters*, *181*(2), 261–266. http://doi.org/10.1016/S0378-1097(99)00548-0

Sulston, J. E., & Brenner, S. (1974). The DNA of Caenorhabditis elegans. *Genetics, 77*(1), 95–104.

Sulston, J. E. & Horvitz, H. R. (1977). Post-embryonic cell lineages of the nematode, Caenorhabditis elegans. *Dev. Biol. 56*: 110-156.

Swan, B. K., Tupper, B., Sczyrba, A., Lauro, F. M., Martinez-Garcia, M., González, J. M., Luo, H., Wright J. J., Landry, Z. C., Hanson, N. W., Thompson, B. P., Poulton, N. J., Schwientek, P., Acinas, S. G., Giovannoni, S. J., Moran, M. A., Hallam, S. J., Cavicchioli, R., Woyke, T. & Stepanauskas, R. (2013). Prevalent genome streamlining and latitudinal divergence of planktonic bacteria in the surface ocean. *Proceedings of the National Academy of Sciences of the United States of America, 110*(28), 11463–8. http://doi.org/10.1073/pnas.1304246110

Tripp, H. J., Bench, S. R., Turk, K. A, Foster, R. A, Desany, B. A, Niazi, F., Affourtit, J. P., & Zehr, J. P. (2010). Metabolic streamlining in an open-ocean nitrogen-fixing cyanobacterium. *Nature, 464*(7285), 90–94. http://doi.org/10.1038/nature08786

Valdivia-Anistro, J. A., Eguiarte-Fruns, L. E., Delgado-Sapién, G., Márquez-Zacarías, P., Gasca-Pineda, J., Learned, J., Elser, J. J., Olmedo-Alvarez, G., & Souza, V. (2016). Variability of rRNA operon copy number and growth rate dynamics of bacillus isolated from an extremely oligotrophic aquatic ecosystem. *Frontiers in Microbiology, 6*(JAN), 1–15. http://doi.org/10.3389/fmicb.2015.01486

Větrovský, T., & Baldrian, P. (2013). The Variability of the 16S rRNA Gene in Bacterial Genomes and Its Consequences for Bacterial Community Analyses. *PLoS ONE, 8*(2), 1– 10. http://doi.org/10.1371/journal.pone.0057923

Vieira-Silva, S., Touchon, M., & Rocha, E. P. C. (2010). No evidence for elemental-based streamlining of prokaryotic genomes. *Trends in Ecology and Evolution, 25*(6): 319-320

Vonk, J. A., Breure, A. M., & Mulder, C. (2013). Environmentally-driven dissimilarity of trait-based indices of nematodes under different agricultural management and soil types. *Agriculture, Ecosystems and Environment*, *179*, 133–138. http://doi.org/10.1016/j.agee.2013.08.007 181

Vrede, T. (1998). Elemental composition (C:N:P) and growth rates of bacteria and Rhodomonas grazed by Daphnia. *Journal of Plankton Research*, *20*(3), 455–470. http://doi.org/10.1093/plankt/20.3.455

Vrede, T., Anderson, T., & Hessen, D. O. (1999). Phosphorus distribution in three crustacean zooplankton species. *Limnol. Oceanogr., 44*(1): 225-229

Wang, J., Pan, F. Soininen, J., Heino, J., & Shen, J. (2016). Nutrient enrichment modifies temperature-biodiversity relationships in large-scale field experiments. *Nature communications 7*(1): 13960.

Watts, T., Woods, H. A., Hargand, S., Elser, J. J., & Markow, T. A. (2006). Biological stoichiometry of growth in Drosophila melanogaster. *Journal of Insect Physiology, 52*(2), 187–193. http://doi.org/10.1016/j.jinsphys.2005.09.012

Weber, K. & Legge, R. (2010). Community-Level Physiological Profiling. *Methods Mol Biol; 599*: 263-81. doi:10.1007/978-1-60761-439-5_16.

Whittaker, F. W. (1969). Galvanotaxix of *Pelodera strongyloides* (Nematoda, Rhabiditidae). *Proceedings of the Helminthological Society of Washington 36*, 40.

Wilson, K. H., Blitchington, R. B., & Greene, R. C. (1990). Amplification of bacterial 16S ribosomal DNA with polymerase chain reaction. *Journal of Clinical Microbiology*, *28*(9), 1942–1946.

Wilson, M. J., & Kakouli-Duarte, T. (2009). Nematodes as Environmental Indicators. Cambridge: CAB International.

Woese, C. R., & Fox, G. E. (1977). Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proceedings of the National Academy of Sciences of the United States of America*, *74*(11), 5088–5090. http://doi.org/10.1073/pnas.74.11.5088

Woese, C. R. (1987). Bacterial Evolution Background. *Microbiology Reviews, 51*(2), 221–271.

Woese, C. R., Kandler, O., & Wheelis, M. L. (1990) Towards a natural system of organisms: Proposal for the domains Archaea, Bacteria, and Eucarya. *Proceedings of the National Academy of Sciences of the United States of America*, *87*, 4576-4579.

Wolf, Y. I., & Koonin, E. V. (2013). Genome reduction as the dominant mode of evolution. *BioEssays*, *35*(9), 829–837. http://doi.org/10.1002/bies.201300037

Wood, F.H. 1973. Nematode feeding relationships: Feeding relationships of soil-dwelling nematodes. *Soil Riol. Biochem. 5*: 593-601.

Yang, B., Zhang, T., Huang, J., Bhusal, D. R., & Pang, X. (2019). Response of soil nematode community to phosphorous amendment in a subalpine spruce plantation. *CLEAN - Soil, Air, Water*, 1800202. http://doi.org/10.1002/clen.201800202

Yao, Z., Wolf, B., Chen, W., Butterbach-Bahl, K., Brüggemann, N., Wiesmeier, M., Dannenmann, M., Blank, B., & Zheng, X. (2010). Spatial variability of N2O, CH4 and CO2 fluxes within the Xilin River catchment of Inner Mongolia, China: A soil core study. *Plant and Soil, 331*(1), 341–359. http://doi.org/10.1007/s11104-009-0257-x 182

Yeates, G. W., Bongers, T., De Goede, R. G., Freckman, D. W., & Georgieva, S. S. (1993). Feeding habits in soil nematode families and genera-an outline for soil ecologists. *Journal of Nematology, 25*(3), 315–331.

Yokokawa, T., Nagata, T., Cottrell, M. T., & Kirchman, D. L. (2004). Growth rate of the major phylogenetic bacterial groups in the Delaware estuary. *Limnology and Oceanography, 49*(5), 1620–1629. http://doi.org/10.4319/lo.2004.49.5.1620

Yu, L., Yan, X., Ye, C., Zhao, H., Chen, X., Hu, F., & Li, H. (2015). Bacterial respiration and growth rates affect the feeding preferences, brood size and lifespan of Caenorhabditis elegans. *PloS ONE, 10*(7), 1–13.

Zhang, T., Shao, M. F., & Fang H. H. P. (2009). A qRT-PCR-based method for the measurement of rrn operon copy number. *Letters in Applied Microbiology,* doi:10.1111/j.1472-765X.2009.02613.x