# Using high-throughput sequence data to predict trophic-group composition and community indices of nematodes

## Introduction

The role of nematodes as indicators of soil quality is amongst the most researched and reviewed sub-disciplines of nematology (Bongers, 1990; Bongers and Bongers, 1998; Bongers and Ferris, 1999; Ferris and Bongers, 2006; Okada and Harada, 2007; Porazinska et al., 1999, 1998). Nematode community composition and structure provide an excellent tool for determining whether the soil has been disturbed or not (Bongers, 1990). This typically entails identifying the different taxa and recording their abundance in soil samples. This information is then used to generate community indices such as the maturity index (MI), plant parasitic index (PPI) (Bongers, 1990), enrichment index (EI), basal index (BI), structure index (SI) and channel index (CI) (Ferris et al., 2001) of a sample, which if determined correctly allows inferences about the condition of the soil the sample was taken from. The taxonomic aspect of this process has traditionally been carried out using morphological characteristics of nematodes. Except for channel index, which only requires trophic group classifications, all the above-mentioned indices require identifying the individual nematodes to the genus level or at least to the family level (Bongers, 1990; Ferris et al., 2001). Although identifying most nematode taxa to the family level may not present significant difficulty to a non-expert, the sheer number of specimens in each extract and the number of extracts requiring examination may preclude carrying out such study beyond local scales.

The usual workaround to this constraint caused by nematode densities and sample size is to identify only a given number of randomly selected specimens, usually at least 100 (Yeates and Bongers, 1999), from each sample while placing a certain limit to the number of samples taken for each study (Neher and Campbell, 1996). Although this approach of identifying a stipulated number of nematodes also serves as a means of normalising the data, it is a step that should better be done at the statistical analysis stage rather than during the identification because leaving the investigator to decide what specimens to identify can introduce potential bias, which can then lead to false representation of the community. Especially in samples that are taxonomically rich, the sampling depth of 100 may not capture the diversity sufficiently. The outcome of this is a study deficient both in terms of sampling depth for taxa and scale. This can also result in the exclusion of some rare but important taxa in the analysis especially in agricultural fields, where samples can be dominated by one or two plant parasitic taxa, thus resulting in a very low representation of free-living taxa needed for the computation of indices such as MI, EI, BI, SI and CI. By reducing the resolution of the classification to trophic groupings, it is also possible to facilitate the identification process without picking only a fixed number of individuals per sample. Despite some favourable arguments put forward to support this approach as providing ecologically and functionally relevant information (Ritz and Trudgill, 1999), adopting it completely precludes the use of most of the indices mentioned earlier since they all require at least family level identification.

The use of metabarcoding has the potential to allow for identifying unlimited number of individuals from several samples all in just a single sequencing run. Moreover, because of the high sensitivity of some next generation sequencing (NGS) technology, even those taxa that are represented in low numbers can be detected. Identification to the genus level is routinely obtainable with this technology, assuming all the necessary bioinformatics tools and expertise are available (Porazinska et al., 2009). This method can, therefore, not only increase the speed with which hundreds of environmental samples can be analysed but also reduce the requirement for expert knowledge on how to identify nematode genera or species while still ensuring highly accurate identification.

Despite the promise that metabarcoding holds with regards to nematode community analysis, there is still one important aspect of it that makes its full implementation thus far highly challenging. All the aforementioned indices used as indicators of soil quality require abundance information for their computation. To fully substitute the traditional method, a new tool needs to detect the taxa present in a sample, and also correctly predict the abundances of these taxa. Successful prediction of relative abundance of taxa in mixed samples of any group of organisms studied by metabarcoding is missing (Amend et al., 2010; Edgar, 2017; Porazinska et al., 2009). Due to this, some have suggested that all PCR-dependent assessment of biodiversity should be based solely on presence/absence, and not abundance (Elbrecht and Leese, 2015).

Criticism of the ability of metabarcoding approaches to quantify taxa within mock or real communities have largely been based on specific taxa, species or genera, and how their read frequencies deviate from their relative abundance (Porazinska et al., 2009; Yu et al., 2012). Although all the indices used to assess nematode communities depend on abundance data (Bongers, 1990; Bongers et al., 1995; Ferris et al., 2001; Yeates, 1994), the required abundance information does not necessarily have to be of the individual taxa. In the case of MI for instance, it is the frequencies of the colonizer-persister groups that are needed for its computation. Similarly, the EI, BI, SI and CI each utilizes abundance information at the functional guild level. In simple terms, a functional guild is a group of taxa with the same feeding habit and life history characteristics (Neher et al., 2004). In view of this, there is a possibility that read numbers obtained from metabarcoding would correlate better at these functional and trophic group levels with their respective actual abundances from morphological identification compared to those obtained for the genera or species. Additionally, transformation of the read counts into binary (presence/absence) records have been recommended and successfully used as a means circumventing the lack of correlation between read frequencies and relative abundance (Ji et al., 2013; Yu et al., 2012).

This study is based on the experiment on how tillage and traffic treatments influence the structure of nematode communities in the soil. In the current chapter, the objective was to examine to what extent indices obtained using relative read frequencies of taxa reflect those obtained for the same samples using standard methods of analysis which involved morphology-based identification. Additionally, the impact of different tillage and traffic regimes on nematode community are analysed through metabarcoding. Finally, metabarcoding data obtained from samples taken from the same experiment a year later were used to assess changes in the nematode community during that period.

## Materials and Methods

### Sample preparation

Soil samples were collected in February and March 2016 from the controlled traffic farming experiment at Harper Adams University (52°46.7899’N, 002°25.5236’W). The experiment was set up in four randomised complete blocks of nine treatments. Each treatment is a combination of one of three tillage and one of three traffic systems. The tillage treatments were zero, shallow or deep tillage and the traffic treatments were controlled, random and low ground pressure traffic. Details of the treatment types have been described in Chapter 6. Samples were taken at three different depths (0-5 cm, 5-15 cm and 15-30 cm) for each of the 36 plots, making a total of 108 samples. Nematodes were extracted from sub-samples of 200g and the extracts concentrated to 10 ml as described in chapter 6.

To determine if there were changes in the nematode community a year later following the sampling that was carried out from February and March 2016, new samples were taken in February 2017 from the top 0-5 cm from the same plots as in 2016. Samples in 2017 were analysed using only metabarcoding. Sampling was limited to a single depth because of time constraint. The 0-5 cm depth was chosen, since prior morphology-based analysis (Chapter 6) recorded the highest diversity within this profile.

### DNA extraction

For the DNA extraction 1 ml subsample was taken from each of the 10 ml sample extracts and stored at -20°C. DNA extractions were performed on 100 µl aliquots using the Qiagen DNeasy Blood and Tissue Kit. To each sample, 800 µl of Qiagen ATL buffer and 100 µl proteinase K were added before incubation overnight at 56ºC and shaking at 109 rpm. The genomic DNA extraction followed the manufacturers guide. To maintain the same ratio of proteinase K and other reagents as described in the manufacturer’s protocol, five times the recommended volumes AL and ethanol were added.

### PCR and Library Preparation

Polymerase chain reaction was performed on each sample using the primer pairs NF1 5’- TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGGTGGTGCATGGCCGTTCTTAGTT-3’) and 18Sr2b (5’-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGTA CAAAGGGCAGGGACGTAAT-3’) (Porazinska et al., 2009). The underlined sections of the primer sequences here represent the Illumina adapter sequence. Each reaction tube contained a total of 25 µl reaction mix consisting of 1X Phusion HF buffer (New England Biolabs, Ipswich, MA, USA), 0.2 mM dNTPs (New England Biolabs), 0.5 µM each of adapter-ligated forward and reverse primers, 1U of Phusion DNA polymerase (New England Biolabs) and 5 µl of template DNA. The PCR programme was set at 98ºC for 2 min, 30 cycles of 98ºC for 20 secs, 66ºC for 45 secs, 72ºC for 45 secs before a final extension step at 72ºC for 5 mins.

Amplicons were purified using Ampure XP Beads (Bechman Coulter, Inc. USA) and quantified using a Thermo Scientific™ Fluoroskan Ascent™ Microplate Fluorometer (Thermo Fisher Scientific, Wilmington, DE, USA). This was then followed by the index PCR step where unique dual indices and the sequencing adapters were attached to each amplicon using Nextera XT index primers (Illumina inc. San Diego, CA, USA) for amplification (Illumina’s 16S Metagenomic Sequencing Library Preparation protocol). The reaction conditions for this were: 98ºC for 3 min, 8 cycles of 98ºC for 30 secs, 55ºC for 30 secs, 72ºC for 30 secs and a final extension step at 72ºC for 5 mins. The products were purified again, quantified using the Microplate Fluorometer and pooled according to their molarity. Length of the amplicons in the pooled sample was verified on the Agilent 2200 TapeStation (Agilent Technologies, Santa Clara, CA, USA). The pool was quantified and diluted to 4 nM concentration. Using the Illumina protocol, the pool was denatured by mixing with 0.2N NaOH. 10% denatured PhiX control library was added to the denatured pool to introduce diversity. The mixture was incubated for 2 min at 96ºC and immediately put on ice, before being loaded on a MiSeq machine (Illumina) for sequencing. The sample was sequenced at Fera (York, UK) in a paired-end approach using a V3 run kit and 2 x 300 cycles.

### Bioinformatic Analysis

The sequence reads were demultiplexed by the MiSeq Reporter software on the MiSeq sequencer into the individual samples based on the paired nextera indexes used- applying default settings. All subsequent analyses were carried out using the USEARCH pipeline (Edgar, 2010). Initial attempts at merging the paired reads using the command *fastq\_mergepairs* in USEARCH resulted in a significant amount of data failing to merge. The forward reads alone were used for subsequent analysis due to superior quality similar to the approach adopted by Unterseher et al. (2016). For each of the samples, reads were first trimmed to 200 bp from the 3’ end, while at the same time removing primer sequences with the command *fastx\_truncate*. The truncated reads were then used to pass the command, *fastq\_filter* which filtered out reads with expected errors of more than 1 (Edgar and Flyvbjerg, 2015) and at the same time converted the files from fastq to fasta formats. The filtered reads of all samples were then combined into one file. The command, *fastx\_uniques* was used to reduce the sequences to unique individual sequences. Clustering was done on the unique sequences at 97% sequence similarity using *cluster\_otus*, which also functions as chimera removal step. Otus were assigned taxonomy based on the *utax* algorithm using PR2 v4.72 (Guillou et al., 2012) as reference database. The command, *usearch\_global* was used for making an otu table.

The otu table was converted into a format similar to the species table obtained from the morphological identification with the taxa names arranged in columns and the samples names in rows. Firstly, all nematodes-only otus with their records in each sample were selected and copied onto a separate worksheet in Microsoft Excel. For otus whose assignments did not include genus names, blast search (Zhang et al., 2000) was performed against the NCBI reference database on 5th February 2018 to confirm their identities at the genus level or species if possible. After this, reads of otus with the same species names were summed to produce a single total for each unique genus/species.

The steps involved in obtaining the species table for the standard morphological approach are detailed in Chapter 6. In short, from 1 ml subsample of each sample extract, all individuals were identified to the genus level, although for some individuals, only family-level identification was possible. The number of each taxon (genus or family) in each sample was recorded in a species table similar to one described for the metabarcoding approach.

### Statistical analysis

For calculating disturbance indices, both the metabarcoding species table and the morphology species table were analysed in three different formats. First analysis involved computing all indices and trophic classifications using the unrarefied data (raw data). The second analysis involved the computation of these indices and trophic classifications on rarefied data from both the metabarcoding and morphological identification. The standard and metabarcoding datasets were rarefied to 50 individuals and 100 reads per sample, respectively. The third analysis used presence-absence data. The three different analyses were performed to determine if the format of the data had any impact on to the comparison. In each of these analysis classes, difference in means of metabarcoding- and morphology-based indices were tested using the Welch’s t-test (Welch, 1951) in RStudio development environment for R (RStudio Team, 2015).

All computations of nematode indices, trophic group classifications and food web diagnostics were performed using NINJA (Sieriebriennikov et al., 2014), an online tool for nematode faunal analyses. The indices obtained for all samples were added to the species table together with some measured environmental variables. Using NINJA, c-p triangle (De Goede et al., 1993a), which graphically presents nematode community structure based on coloniser-persister (c-p) classifications of the community were also obtained.

The effects of treatments and depth of sampling on nematode community compositions were demonstrated using non-metric multidimensional scaling (NMDS) ordination of Jaccard dissimilarity metrics. Most of the analysis involving NMDS were performed following the workflow described in (Ji et al., 2013). For this procedure alone, only presence-absence data from the metabarcoding was compared with the unrarefied data from the standard morphological data to conform with the procedure described by Ji et al. (2013). An NMDS plot comparing the presence-absence metabarcoding data with the rarefied standard data instead of the unrarefied one did not change the plot. The decision to use the unrarefied standard data was, therefore, to avoid loss of data that may result from the rarefaction. The function *vegdist* of the vegan package (Oksanen et al., 2015) was used to calculate distances, and *metaMDS* for performing NMDS. For the metabarcoding data alone, the parameter ‘binary’ in *vegdist* was set to true because the data was in a binary (presence/absence) format. Correlations between the metabarcoding and morphological data were performed using *mantel* and *protest* correlation tests in *vegan*.

Effects of treatments on community compositions were analysed using *mvabund* (Wang et al., 2012) package. Differences in the effects of treatment levels were calculated using *summary.manyglm*, and *p.adjust* function based on the ‘fdr’ method (Benjamini and Yekutieli, 2001) was used to correct for multiple testing. The ‘fdr’ method controls the expected proportion of false discoveries amongst the rejected hypotheses.

### Comparisons between 2016 and 2017 samples

Sequence files for the 0-5 cm depth of the samples from 2016 were combined with the sample data collected in 2017 (which consisted only of 0-5 cm depth samples) using USEARCH. Analyses inside USEARCH were carried following the same steps, using the same procedure as described in the section, 7.2.4. Species tables for the two sampling years were separated before performing statistical analysis. A possible outcome when data are separated in this manner is the occurrence of species with zero records for all samples in one part of the data because they only occurred in the other. Therefore, an additional code was added to remove species with zero records in all samples for each of the separate years. Distances were calculated using *vegdist* for this comparison. Treatment effects on community compositions were also calculated using *summary.manyglm* function with *p.adjust* to correct for multiple testing as was done in the previous section.

## Results

### Recovered taxa

Standard morphological analysis recovered 83 unique taxa identified to the genus level, plus a few taxa that were identified only to the family level for individuals whose genus identification could not be made. The metabarcoding data set yielded 194 otus at 97% sequence similarity that were assigned to 55 unique species belonging to 41 different genera. Although the 97% similarity cut-off produced otus in excess of the number of species recovered by taxonomy assignment, using lower cut-offs failed to recover some of the taxa. And since taxonomic identities to at least family level was needed for nematode community indices and trophic classifications, taxonomic diversity instead of otus richness was used.

### Trophic Groupings

Apart from the relative dominance of herbivorous nematodes in the metabarcoding data set, changes in the relative abundance of the different trophic groups followed similar patterns in both standard and metabarcoding data sets (Figure 7.1). Both data sets also revealed the numerical dominance of the bacterivorous and herbivorous nematodes under all treatments. Both showed that zero tillage favours herbivores most while deep tillage tends to be favourable for bacterivores. Across the traffic treatments, both data sets show minimal to non-existent difference in the distribution of the trophic types. Although predacious and omnivorous nematodes had low representations in the two data sets, this appeared to be more extreme in the metabarcoding data where they were rarely recovered from the samples.

### Nematode Community indices

Two terminologies used in this section and the subsequent ones need clarification. The term **data format** is used to refer to unrarefied, rarefied or presence-absence data while **data type** refers to the standard morphological or metabarcoding data. For both the standard and metabarcoding data types, indices were calculated from three different data formats: the unrarefied data, data rarefied to 100 and 50, respectively and binary data. Attention was focussed only on MI and MI-derived indices alone, namely, MI, MI2-5 PPI and Sigma MI. Structure, channel and basal indices between the two data types were too different, hence their exclusion. Of the three data formats, the binary data revealed the most deviations between indices from the standard and metabarcoding data. By using rarefied or unrarefied data, most of the indices were not different between the two data types (Table 7.1). Plant-parasitic index was the only index that showed no significant divergence between the two data types across all data formats. Maturity index 2-5 varied significantly between the two data types for all treatments in at least one data format. There was generally more evidence to support the fact that indices computed from the three data types did not differ than there was to support otherwise.

Figure 7.1. Nematode trophic groups’ distribution across traffic and tillage treatments from standard and metabarcoding data sets. STD represents standard data set and MTB represents metabarcoding data sets.

Table 7.1. Comparison between MI-family indices calculated from standard and metabarcoding data types from raw, rarefied and the binary data formats under different traffic and tillage treatments. Numbers are means calculated for each treatment group. Insignificant P values are highlighted in grey.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Index | Unrarefied | | Rarefied | | Pres-Abs | | Welch’s t-test *P* value | | |
| Traffic |  |  | |  | |  | |  | | |
| CTF |  | MTB | MOR | MTB | MOR | MTB | MOR | UNR | RAR | PrAb |
|  | MI | 1.75 | 1.99 | 1.79 | 2.03 | 1.77 | 2.43 | 0.100 | 0.163 | <0.001 |
|  | MI2-5 | 2.05 | 2.62 | 2.00 | 2.62 | 2.06 | 2.73 | 0.003 | <0.001 | <0.001 |
|  | PPI | 2.82 | 2.63 | 2.75 | 2.65 | 2.81 | 2.81 | 0.132 | 0.472 | 1.000 |
|  | SigMI | 2.41 | 2.22 | 2.49 | 2.24 | 2.31 | 2.57 | 0.255 | 0.297 | 0.001 |
| LGP |  |  |  |  |  |  |  |  |  |  |
|  | MI | 1.86 | 1.86 | 1.76 | 1.86 | 1.89 | 2.22 | 0.978 | 0.735 | 0.060 |
|  | MI2-5 | 2.12 | 2.36 | 2.13 | 2.63 | 2.12 | 2.56 | 0.018 | 0.222 | 0.018 |
|  | PPI | 2.68 | 2.66 | 2.66 | 2.66 | 2.71 | 2.75 | 0.898 | 0.984 | 0.610 |
|  | SigMI | 2.32 | 2.19 | 2.33 | 2.19 | 2.36 | 2.46 | 0.172 | 0.159 | 0.153 |
| RTF |  |  |  |  |  |  |  |  |  |  |
|  | MI | 1.71 | 1.76 | 1.74 | 1.74 | 1.70 | 2.31 | 0.399 | 1.000 | 0.001 |
|  | MI2-5 | 2.00 | 2.49 | 2.00 | 2.46 | 2.00 | 2.66 | 0.005 | 0.002 | 0.005 |
|  | PPI | 2.75 | 2.68 | 2.86 | 2.66 | 2.72 | 2.68 | 0.424 | 0.222 | 0.269 |
|  | SigMI | 2.39 | 2.11 | 2.43 | 2.10 | 2.31 | 2.45 | 0.201 | 0.154 | 0.045 |
|  |  |  |  |  |  |  |  |  |  |  |
| Tillage |  |  |  |  |  |  |  |  |  |  |
| Deep |  |  |  |  |  |  |  |  |  |  |
|  | MI | 1.83 | 1.92 | 1.90 | 1.95 | 1.80 | 2.32 | 0.617 | 0.827 | 0.003 |
|  | MI2-5 | 2.07 | 2.51 | 2.13 | 2.50 | 2.04 | 2.65 | 0.008 | 0.093 | 0.001 |
|  | PPI | 2.73 | 2.66 | 2.72 | 2.65 | 2.72 | 2.76 | 0.700 | 0.701 | 0.631 |
|  | SigMI | 2.42 | 2.14 | 2.50 | 2.15 | 2.32 | 2.48 | 0.184 | 0.159 | 0.049 |
| Shallow |  |  |  |  |  |  |  |  |  |  |
|  | MI | 1.77 | 1.89 | 1.80 | 1.89 | 1.80 | 2.37 | 0.148 | 0.303 | 0.004 |
|  | MI2-5 | 2.07 | 2.48 | 2.00 | 2.48 | 2.08 | 2.66 | 0.021 | 0.022 | 0.002 |
|  | PPI | 2.78 | 2.62 | 2.82 | 2.62 | 2.72 | 2.73 | 0.133 | 0.317 | 0.825 |
|  | SigMI | 2.31 | 2.12 | 2.32 | 2.12 | 2.34 | 2.51 | 0.107 | 0.242 | 0.048 |
| Zero |  |  |  |  |  |  |  |  |  |  |
|  | MI | 1.72 | 1.80 | 1.59 | 1.79 | 1.75 | 2.28 | 0.253 | 0.268 | 0.038 |
|  | MI2-5 | 2.04 | 2.47 | 2.00 | 2.46 | 2.06 | 2.64 | 0.038 | 0.041 | 0.021 |
|  | PPI | 2.74 | 2.70 | 2.73 | 2.71 | 2.80 | 2.74 | 0.578 | 0.841 | 0.480 |
|  | SigMI | 2.39 | 2.25 | 2.43 | 2.27 | 2.32 | 2.48 | 0.290 | 0.230 | 0.039 |

Hypothesis testing to determine the impact of tillage and traffic treatments on the indices using Kruskal-Wallis test revealed some similarities between the two data types. Traffic was the only treatment that had significant effect and it affected only the MI25. This effect was presented as significant in the standard data set (*p* = 0.015) and marginally significant in the metabarcoding data (*p* = 0.083). The other indices showed no response to any of the treatments.

### Depiction of Disturbance conditions

Using the c-p triangle to visualize a snapshot of level of stress in the soil, the two data types gave different depictions of the nematode communities (Figure 1). While both data types mostly indicated communities with low stability, this condition was highly exaggerated in the metabarcoding data where only very few communities were situated away from the zero-stability border. This was even worse in the rarefied data which not only showed no stability but also depicted most of the communities as being highly stressed. The use of the binary data for the c-p triangle gave the best representation of the community through metabarcoding if compared with the rarefied or unrarefied standard data but not with the binary-formatted standard data.

### Beta Diversity

Differences between nematode communities as a result of tillage and traffic treatment as well depths of sampling across field compaction gradients were depicted similarly by the NMDS ordinations in both the standard and metabarcoding data types. There was no clear separation of the treatment levels in the NMDS ordinations for both data types (Figure 7.3 A and B). The ordinations involving the different depths, on the other hand, reveal clear separations again in both the standard and metabarcoding data types (Figure 7.4 A and B). The Procrustes tests of correlation between ordinations of the standard and metabarcoding data sets was highly significant (*r* = 0.42, *P* = 0.001). Between their dissimilarity matrices, the Mantel test also showed a highly significant correlation (*r* = 0.18, *P* = 0.001).

Tests of significance of the treatments using *summary*.*manyglm* test showed that in both data types, treatments had no effect on the nematode community. But as was evident from the NMDS ordinations, difference in the community structure across the sampling depths were highly significant in both the standard data and metabarcoding data.

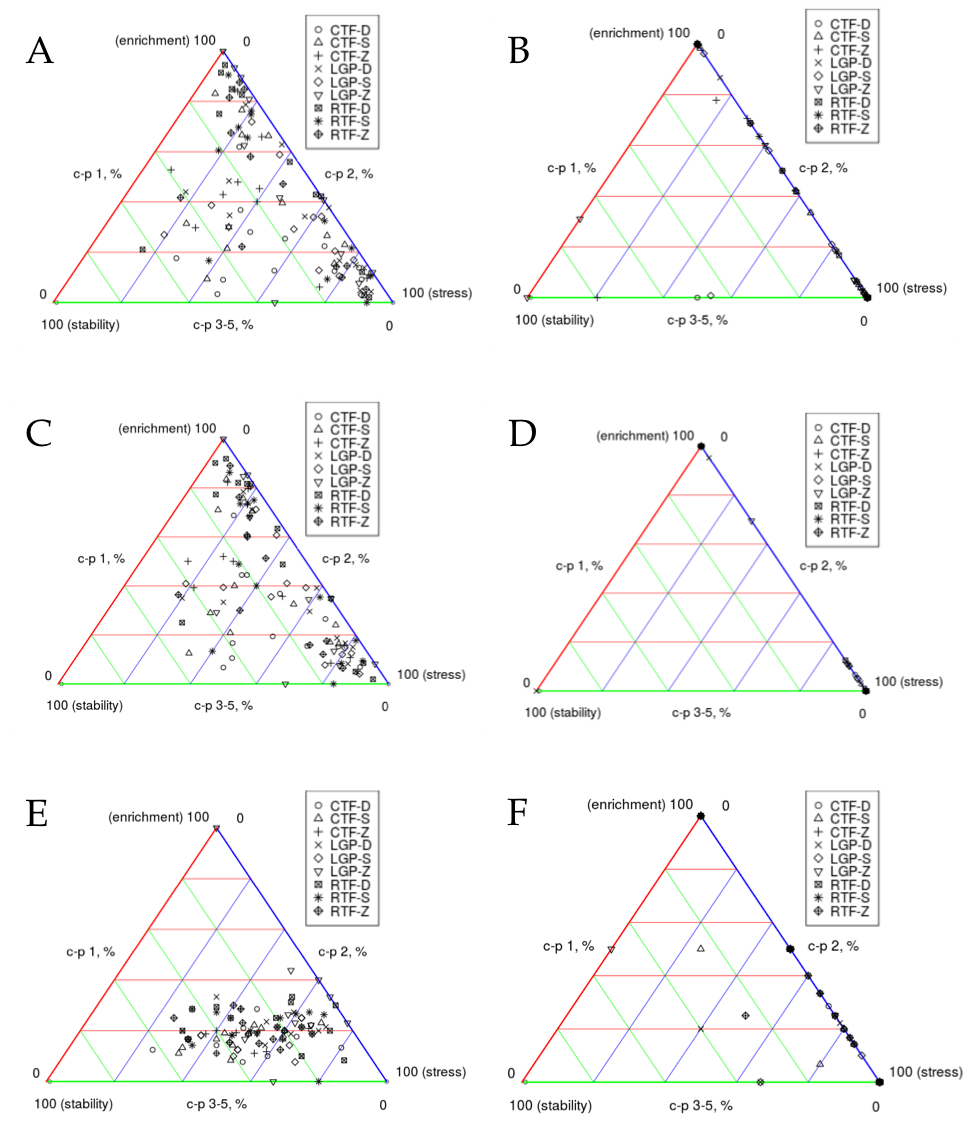


Figure 7.2. Graphical representation of nematode communities using c-p triangles. Each marker represents one of the 108 samples. Samples within the same treatment categories have identical marker shapes. A and B depict the status of the communities based on unrarefied data of the standard and metabarcoding approaches, respectively. C and D depict the status of the communities based on rarefied data of the standard and metabarcoding approaches, respectively. E and F depict the status of the communities based on presence-absence data of the standard and metabarcoding approaches, respectively

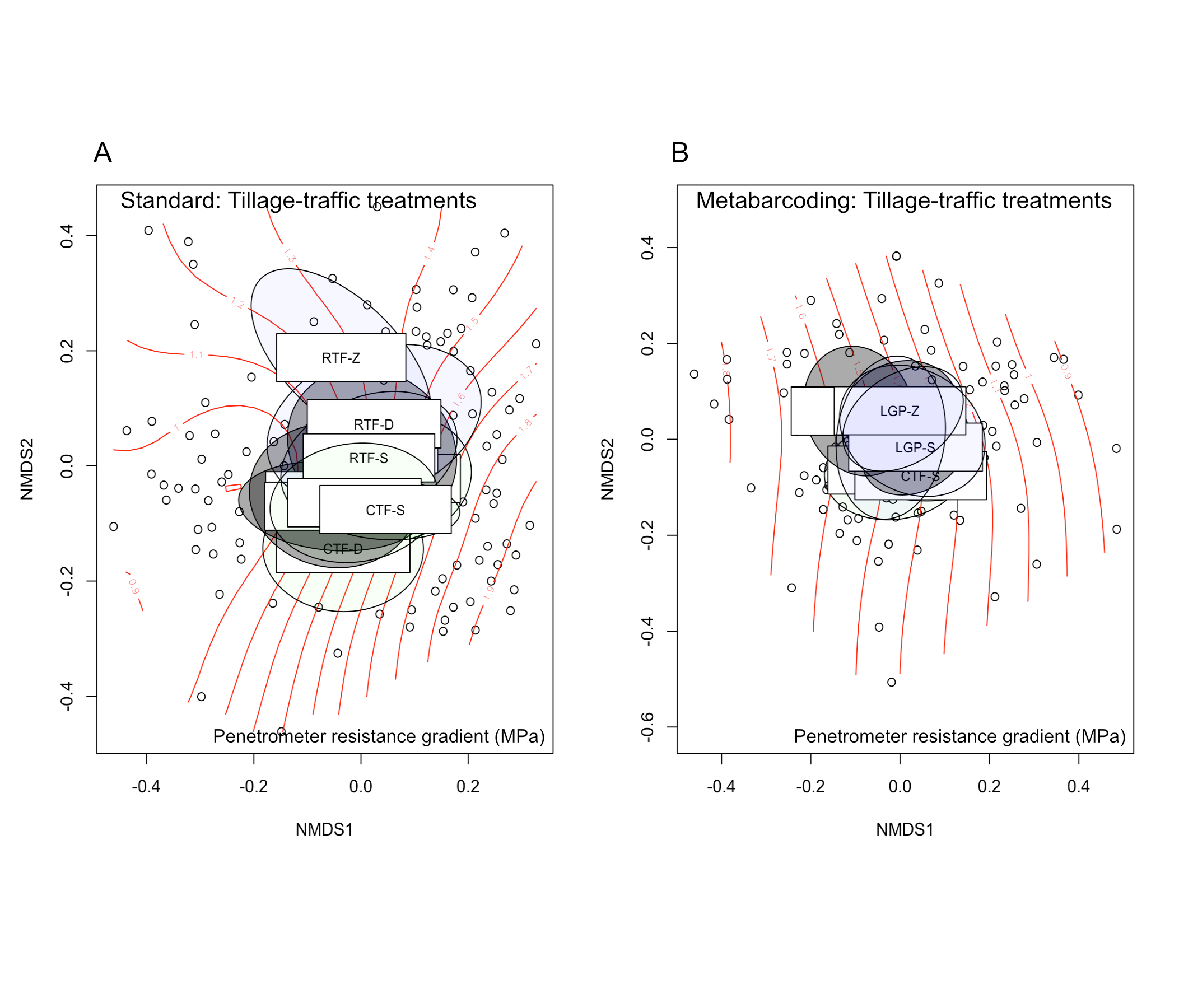
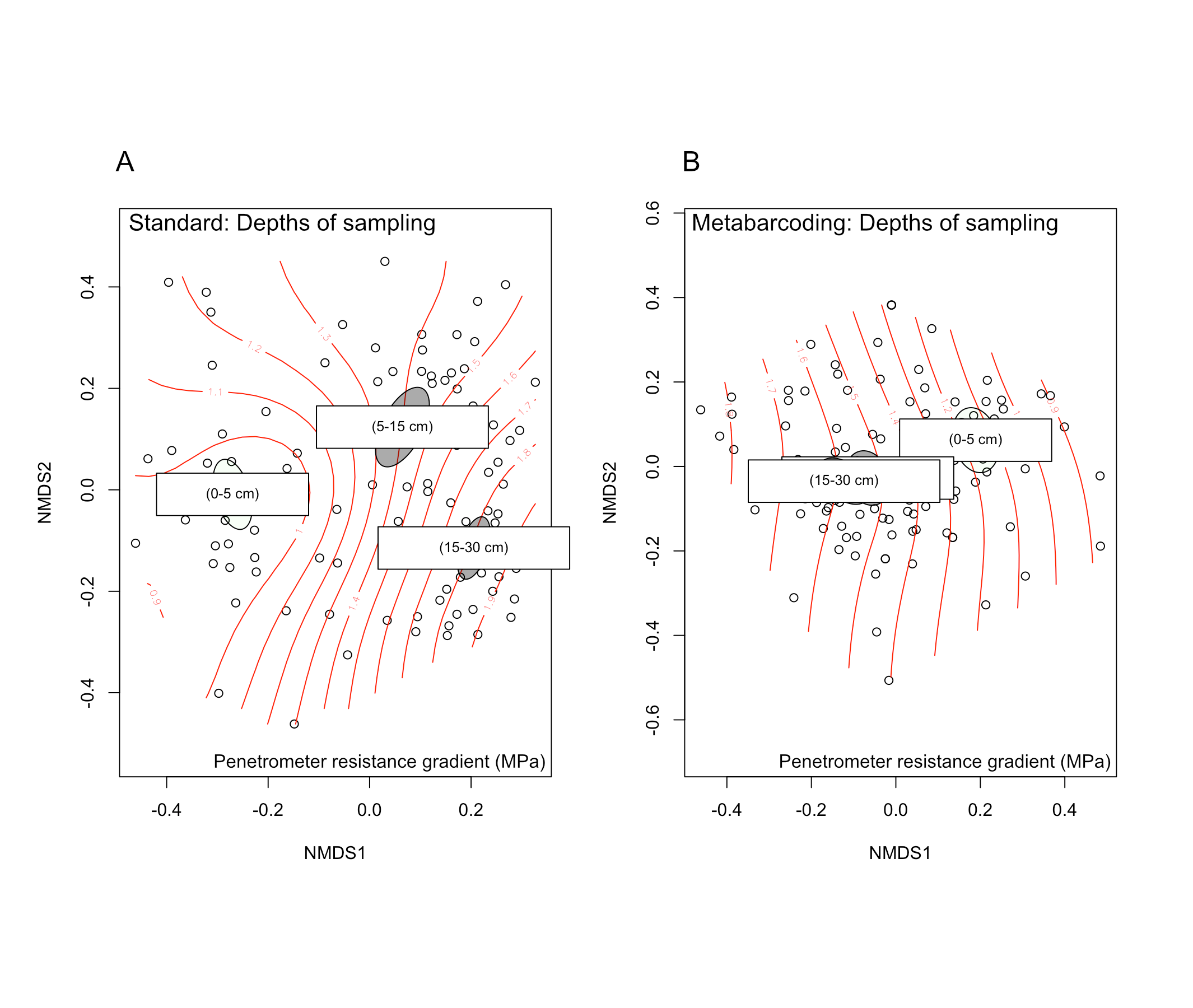


Figure 7.3. Non-metric multidimensional scaling (NMDS) ordinations based on morphology (A) and metabarcoding (B) based data for samples collected in 2016. Points are experimental plots, and coloured ellipses are 95% confidence intervals of species centroids for each treatment (ellipses generated using the function, ‘ordiellipses’ inside vegan). Plots along the same line segments have the same level of compaction. Colour assignments to ellipses in the tillage-traffic treatments ordination is based on the different traffic systems.



*Figure 7.4. Non-metric multidimensional scaling (NMDS) ordinations based on morphology (A) and metabarcoding (B) based data for samples collected in 2016. Points are experimental plots, and coloured ellipses are 95% confidence intervals of species centroids for each depth (ellipses generated using the function, ‘ordiellipses’ inside vegan). Plots along the same line segments have the same level of compaction. Colour assignments to ellipses in the sampling depths ordination is based on the different depths of sampling (A and B)*

### Seasonal Difference

Between the sampling times, most indices did not change significantly under the different tillage and traffic treatment conditions (Table 7.2). Plant parasitic index (PPI) changed significantly and it did so under most treatments. In all cases where PPI changed, it increased from years 2016 to 2017. Apart from the PPI, MI was the only other index that changed significantly from 2016 to 2017 and it did so under deep tillage alone.

Hypothesis testing of treatment effects on the indices using Kruskal-Wallis test showed that neither tillage nor traffic had any effect on the indices in the 2016 samples. However, there was a significant difference in the median MI of the tillage treatments in the 2017 (*P* = 0.023). Dunn (1964) test for multiple comparison showed that the difference in median was between deep and shallow tillage alone, with the latter having the higher value. Sigma MI and PPI were only marginally affected in 2017 by traffic (*P* = 0.073) and tillage (P = 0.077), respectively.

Table 7.2. Comparison between MI-family indices of the two metabarcoding data obtained from sampling years, 2016 and 2017 under different traffic and tillage treatments. The data for the two sampling years were both rarefied to 100 reads per sample. Numbers are means calculated for each treatment group.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Index** | **Sampling Year** | | **Welch’s t-test *P* value** |
|  |  | **2016** | **2017** |  |
| **Traffic** |  |  |  |  |
| CTF |  |  |  |  |
|  | MI | 1.94 | 1.67 | 0.150 |
|  | MI2-5 | 2.00 | 2.10 | 0.339 |
|  | PPI\* | 2.62 | 4.25 | <0.001 |
|  | SigMI | 2.44 | 1.80 | 0.055 |
| LGP |  |  |  |  |
|  | MI | 1.69 | 1.77 | 0.630 |
|  | MI2-5 | 2.00 | 2.14 | 0.168 |
|  | PPI | 2.85 | 3.55 | 0.074 |
|  | SigMI | 2.32 | 2.24 | 0.769 |
| RTF |  |  |  |  |
|  | MI | 1.92 | 1.72 | 0.442 |
|  | MI2-5 | 2.00 | 2.22 | 0.217 |
|  | PPI\* | 2.25 | 3.38 | 0.012 |
|  | SigMI | 2.12 | 1.96 | 0.613 |
|  |  |  |  |  |
| **Tillage** |  |  |  |  |
| Deep |  |  |  |  |
|  | MI\* | 1.83 | 1.43 | 0.013 |
|  | MI2-5 | 2.00 | 2.12 | 0.151 |
|  | PPI\* | 2.33 | 4.06 | <0.001 |
|  | SigMI | 2.17 | 1.92 | 0.449 |
| Shallow |  |  |  |  |
|  | MI | 1.80 | 1.98 | 0.451 |
|  | MI2-5 | 2.00 | 2.17 | 0.336 |
|  | PPI\* | 2.67 | 3.89 | 0.007 |
|  | SigMI | 2.30 | 2.14 | 0.676 |
| Zero |  |  |  |  |
|  | MI | 1.92 | 1.76 | 0.421 |
|  | MI2-5 | 2.00 | 2.17 | 0.166 |
|  | PPI | 2.73 | 3.23 | 0.138 |
|  | SigMI | 2.42 | 1.95 | 0.025 |

### Beta Diversity

NMDS ordinations on Jaccard binary distance matrices for the two sampling periods, although similar, showed no difference in community composition across the different treatment levels (Figure 3). Hypothesis testing using the *summary.manyglm* in the r package *mvabund* showed that treatments had no significant effect on the community composition of the top 0-5 cm profile in both 2016 and 2017. The c-p triangle showed a slight increase in stability of some communities in the period between the two years. Surprisingly, some of these plots were ones that actually received disturbance treatments such as traffic and tillage.

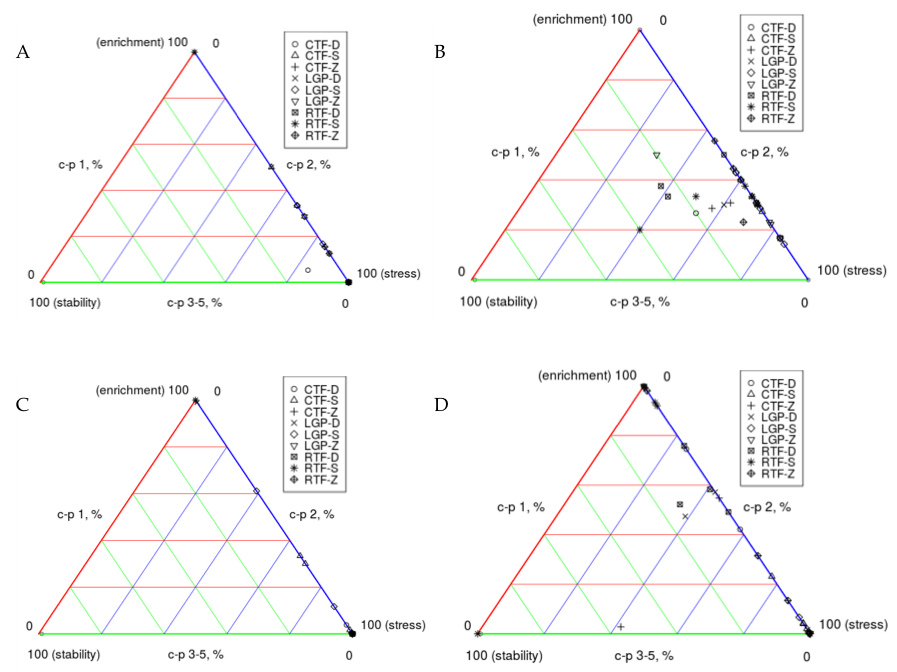


Figure 7.5. Graphical representation of the change in nematode communities in an interval of one year using c-p triangles. Each marker represents one of 36 samples collected within the top 0-5 cm depth in the years, 2016 and 2017. Samples within the same treatment categories have identical marker shapes. A and B depict the status of the communities based on presence-absence data for the 2016 and 2017 samples, respectively. C and D depict the statuses of the communities based on unrarefied data for the 2016 and 2017 approaches, respectively.

# Discussion

The results of this study demonstrate the possibility of determining the condition of a soil using metabarcoding of nematode communities in the soil. This was true for rarefied, unrarefied and to some extent the binary community data. It also shows that metabarcoding data can be expected to provide a reliable representation of the difference in community composition between treatments, times or possibly localities. Correlation between the morphological data and the binary-transformed metabarcoding data was highly significant with the correlations coefficient *r*, based on *Mantel* and *Procrustes* tests showing low and medium effect sizes, respectively.

The standard morphological data set produced a much higher diversity in terms of the number of unique taxa recovered than the metabarcoding data set did. The opposite would have been the case had diversity for the latter been based on the number of unique otus as is often done with metabarcoding data (Ji et al., 2013). It was necessary to utilize the taxonomic identities of the otus in this study in order to correctly assess how the two data sets describe disturbance conditions in the nematode communities. The relatively low taxonomic coverage of the metabarcoding data set concurred with some previous studies that showed that metabarcoding is not always able to detect all resident taxa in samples (Cowart et al., 2015; Zimmermann et al., 2015). In Zimmermann et al. (2015), it was only at the genus level where fewer taxa were recovered by the metabarcoding method (28 versus 30). At the species level, over twice the number recovered by the standard approach were recovered by the metabarcoding method (265 versus 102). In Cowart et al. (2015), the metabarcoding approach based on both the COI and 18 rDNA could only recover 36% of the morphologically-determined diversity. This may be for a number of reasons ranging from primer mismatch to dominant taxa leading to the masking of rare taxa from the data.

The importance of nematode trophic groups to soil processes such as decomposition and primary production pathways has been well documented (Beare et al., 1992; Wardle et al., 1995). This study showed that metabarcoding data sets could depict how anthropogenic disturbances such as tillage affects their relative abundance in the soil based on the relative abundance of the various trophic groups. The similarity was strong, particularly for the dominant groups, herbivores and bacterivores. Lejzerowicz et al. (2015), in their comparison of morphological and metabarcoding-based assessment of metazoans in benthic environments, also observed such high congruence in the abundance of the dominant taxa.

Sufficient molecular data exists that supports the phylogenetic origins of most of the current nematode classifications (Blaxter et al., 1998; Holterman et al., 2006; van Megen et al., 2009). The same studies have demonstrated that features such as herbivory may have arisen multiple times independently in the course of evolution. Similar evidence exists for at least fungivory and predation. One of the things these independent evolutions of feeding behaviour may point to is the discrepancies between trophic classification and taxonomic classification, at least based on the 18S rDNA region. This may therefore explain why metabarcoding data based on the same region was able to correctly quantify trophic groups even though most have shown the contrary for actual taxa in communities (Amend et al., 2010; Edgar, 2017; Porazinska et al., 2009).

Much like the trophic classification, the colonizer-persister groupings and functional guilds used to determine nematode indices do not depend on the phylogenetic positions of their constituent taxa, but rather the life history characteristics. And the same as it was for the trophic compositions, almost all indices obtained using metabarcoding data showed no deviation from those obtained from the standard data set. The implications of this are that most information needed to understand the soil environment can reliably and fairly accurately be obtained from metabarcoding even if the same data misrepresent taxonomic compositions of individual taxa.

The c-p triangle presents graphically three aspects of the condition of the soil, how enriched, stressed or stable it is based on the proportion of the community constituted by c-p1, c-p 2 or c-p 3-5 nematodes, respectively (De Goede et al., 1993a). Although the main goal of c-p triangles is to demonstrate patterns in community change over time, they can also give a snapshot of the state of soil at any point in time. The two data sets in this study provided different depiction of the nematode communities. The format of the data used [rarefied, unrarefied or presence-absence] did not matter much. The metabarcoding data showed communities highly dominated by c-p 1 and c-p 2 nematodes, thus either highly enriched or stressed. Overrepresentation of a few dominant taxa in sequence reads can occur in metabarcoding data (Sogin et al., 2006); and this may explain the exacerbated stress condition in the metabarcoding data.

The only significant factor according to the NMDS ordination of dissimilarity matrices was sampling depth and both the standard and metabarcoding data sets demonstrated this. In a previous study that looked at the difference between nematode communities within different depths, Sánchez-Moreno et al. (2006) also found that some taxa were significantly higher in some parts of the profile than they were in others. They observed a significantly higher number of bacterivores and fungivores, particularly those belonging to the c-p 1 and c-p 2 groups in the top soil layers (0-15 cm) of the profile. The lower depths had significantly higher number of herbivores such as *Pratylenchus* and Tylenchidae. Similar observations were made regarding the vertical distribution of herbivores versus bacterivores in both standard and metabarcoding data sets (Table 7.3 and Table 7.4). Bacterivores were most abundant in 0-15 cm in both data sets regardless of whether the data were rarefied or unrarefied. Herbivore percentage of total nematodes, on the other hand, was highest in the lower 15-30 cm for both data sets. In the metabarcoding data set, *Rhabditis* was lowest in the top 0-5 cm depth, which contradicts what was expected as this genus also belongs to the bacterivore trophic group. It also contradicts the result from the standard data set. *Rhabditis* is amongst the bacterivorous taxa known to form dauer, a non-feeding dormant stage. And this unexpected outcome with the metabarcoding data may have had to do with the difference in how the two approaches, standard and metabarcoding, detect juveniles in the dauer stage. Because it is nearly impossible to identify the genus or even the family of a dauer through standard morphological identification, they are usually classified as a separate group. This is however different with the metabarcoding approach which can easily identify the genus or even species that a dauer belongs. Some of the *Rhabditis* in the metabarcoding data may have been among what were classified as dauer in the standard data. And since a significant number of dauer juveniles were detected in the lower depths in the standard identification method, they most likely added to the number of *Rhabditis* in the metabarcoding data within this depth.

Community succession is a very important aspect of environmental monitoring (De Goede et al., 1993b; Ettema and Bongers, 1993; Ferris and Matute, 2003; Háněl, 2010; Wardle et al., 1995). Even though this study compared only two separate time points, analysed using only metabarcoding, there were some significant changes in some of the community indices. Increased perturbation in the form of deep tillage caused a significant decrease in maturity index in the second sampling year. Plant parasitic index is the MI equivalent designated exclusively for the plant-parasitic nematodes (Bongers, 1990). Previous studies have shown that its relationship with MI can be either direct (Neher and Campbell, 1994) or inverse (Bongers et al., 1997). This therefore explains why the PPI in this study increases between two sampling dates while MI decreases. Another reason could be the availability of a suitable host in the interval between the two sampling since this could have resulted in the multiplication of herbivores such as *Meloidogyne* or *Pratylenchus*. If that was the case, it would certainly have increased the PPI since both of these taxa have plant parasitic indices of three.

The c-p triangle’s representation of pattern of change within the nematode communities showed some mild increases in both stability and enrichment for both rarefied and binary data formats. It has to be said that, based on the earlier comparison between c-p triangles from the standard and metabarcoding data sets, metabarcoding based c-p triangles may be presenting completely exaggerated conditions of stress. More testing may be required to increase its reliability.

In summary, this study demonstrates that metabarcoding data is capable of depicting significant aspects of a community’s condition and do match in many respects the conditions presented by data obtained from the standard morphology-based approach. However, there are a few acknowledgeable limitations to this study that if addressed in future studies, may improve the utility of metabarcoding for nematode community analysis.

Firstly, the DNA extraction methodology could be improved. DNA extractions were carried out using 100 µl of suspensions making it a tenth of what was used for morphological analysis and a hundredth of the whole extracts from 200 g. This could have resulted in loss of taxa that were already rare in the soil sample. Moreover, the volume of suspension containing nematodes also contained other eukaryotes which were also amplified by the universal primers used as was reported in other studies (Sapkota and Nicolaisen, 2015). In fact, a significant percentage of the otus generated for the 2016 samples were from these non-target taxa (82.43%). To maximize the percentage of nematode reads it may be important to perform DNA extraction on either the whole suspension of nematode extracts or at least concentrate the suspension better through centrifugation. This may not necessarily exclude other taxa but would certainly increase the total nematode recovery. To exclude some of the non-target taxa, the nested PCR approach reported by (Sapkota and Nicolaisen, 2015) could be used to first exclude non-nematode targets before the actual amplification.

The second limitation is to do with the extent to which nematode extracellular DNA could have made it into nematode extracts and eventually, into the recovered reads. Extracellular nematode DNA has been shown to occur in bulk nematode extracts (Peham et al., 2017), and the implications of this could be the detection of taxa that are no longer an active part of the community. Theoretically, this can be resolved by using RNA instead of DNA for detection, an area that may require some testing and developing in the future. For now, extracting nematodes from the soil first before DNA extraction can help reduce the amount of extracellular DNA. Finally, as has been shown in a number of tests conducted in this study, both rarefied and unrarefied metabarcoding data gave similar outcomes in terms of nematode MI-based indices and trophic composition. The exception may be for agricultural soils where some taxa, particularly the sensitive ones, are rarely recovered in the samples. In such cases, one might opt for unrarefied data to avoid losing these taxa for computing indices and trophic composition. While this study points to the right direction in terms of the potential of using metabarcoding for monitoring soil health through nematode community indices, further studies may be needed to that focus on soils from natural undisturbed environments in order to better establish the reliability of such approach.

Table 7.3. Abundance of herbivorous and bacterivorous nematodes recovered from the three depths according to the unrarefied data. Figures within the rows labelled herbivores and bacterivores are mean and standard errors of percentages of the total reads associated with herbivores and bacterivores, respectively. Figures within rows of taxa are actual read numbers associated with each of those taxa. Test of statistical significance was performed using robust ANOVA as implemented in the R function, t1waybt within the WRS2 package (Wilcox, 2012).

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Metabarcoding** | | |  | **Standard** | | |  |
|  | 0-5 cm | 5-15 cm | 15-30 cm | *P* values | 0-5 cm | 5-15 cm | 15-30 cm | *P* values |
| Herbivores (%) | 36.69 ± 7.11 | 62.98 ± 6.40 | 79.69 ± 4.77 | <0.01 | 35.19 + 3.10 | 35.37 ± 3.15 | 61.39 ± 2.91 | < 0.001 |
| *Pratylenchus* | 0 ± 0 | 1261.50 ± 1258.30 | 429.39 ± 403. 25 | NA | 2.17 ± 0.30 | 3.14 ± 0.48 | 2.58 ± 0.36 | 0.447 |
| *Tylenchus* | 0 ± 0 | 453.31 ± 452.76 | 0.50 ± 0.33 | NA | 1.22 ± 0.24 | 3.72 ± 0.55 | 8.17 ± 1.24 | < 0.001 |
| *Paratylenchus* | 36.19 ± 35.88 | 48.638 ± 24.578 | 2568.94 ± 1207.82 | NA | 7.47 ± 1.52 | 5.77 ± 0.99 | 13.47 ± 1.46 | < 0.001 |
| Bacterivores (%) | 52.67 ± 7.15 | 27.28 ± 5.83 | 13.42 ± 4.12 | <0.01 | 50.62 ± 2.64 | 54.47 ± 3.21 | 29.86 ± 3.09 | < 0.001 |
| *Rhabditis* | 5.94 ± 3.22 | 725.06 ± 389.89 | 1017.08 ± 696.66 | 0.427 | 5.64 ± 0.86 | 1.14 ± 0.64 | 0.03 ± 0.03 | NA |
| *Eucephalobus* | 3658.81 ± 1815.22 | 1151.56 ± 1016.83 | 1.05 ± 0.40 | 0.643 | 11.06 ± 1.14 | 0.53 ± 0.13 | 3.36 ± 0.48 | < 0.001 |

Table 7.4. Abundance of herbivorous and bacterivorous nematodes recovered from the three depths according to the rarefied data. Figures within the rows labelled herbivores and bacterivores are means and standard errors of percentages of the total reads associated with herbivores and bacterivores, respectively. Figures within rows of taxa are actual read numbers associated with each of those taxa. Test of statistical significance was performed using robust ANOVA as implemented in the R function, t1waybt within the WRS2 package (Wilcox, 2012).

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Metabarcoding** | | |  | **Standard** | | |  |
|  | 0-5 cm | 5-15 cm | 15-30 cm | *P* values | 0-5 cm | 5-15 cm | 15-30 cm | *P* values |
| Herbivores (%) | 36.91 ± 9.38 | 60.50 ± 9.58 | 81.04 ± 6.37 | 0.018 | 35.27 + 3.49 | 35.27 ± 3.20 | 60.19 ± 3.06 | <0.001 |
| *Pratylenchus* | 0 ± 0 | 5.00 ± 5.00 | 9.04 ± 5.86 | NA | 1.00 ± 0.14 | 1.11 ± 0.23 | 1.74 ± 0.34 | 0.109 |
| *Tylenchus* | 0 ± 0 | 0.17 ± 0.17 | 2.40 ± 2.40 | NA | 0.70 ± 0.24 | 1.17 ± 0.26 | 6.13 ± 0.74 | <0.01 |
| *Paratylenchus* | 0.35 ± 0.31 | 9.70 ± 6.57 | 23.39 ± 7.91 | NA | 2.78 ± 0.55 | 2.11 ± 0.43 | 8.87 ± 0.97 | <0.001 |
| Bacterivores (%) | 31.65 ± 8.17 | 23.14 ± 8.06 | 37.17 ± 8.61 | 0.315 | 50.61 ± 3.06 | 55.22 ± 3.37 | 29.61 ± 3.39 | <0.001 |
| *Rhabditis* | 0.30 ± 0.23 | 8.60 ± 5.17 | 7.30 ± 4.61 | NA | 3.00 ± 0.62 | 0.38 ± 0.16 | 0 ± 0 | NA |
| *Eucephalobus* | 15.91 ± 6.10 | 7.45 ± 4.90 | 0.04 ± 0.04 | NA | 5.42 ± 0.64 | 1.27 ± 0.41 | 0.26 ± 0.10 | NA |

# References

Amend, A.S., Seifert, K.A., Bruns, T.D., 2010. Quantifying microbial communities with 454 pyrosequencing: does read abundance count? Mol. Ecol. 19, 5555–5565.

Beare, M.H., Parmelee, R.W., Hendrix, P.F., Cheng, W., Coleman, D.C., Crossley, D.A., 1992. Microbial and faunal interactions and effects on litter nitrogen and decomposition in agroecosystems. Ecol. Monogr. 62, 569–591.

Benjamini, Y., Yekutieli, D., 2001. The control of the false discovery rate in multiple testing under dependency. Ann. Stat. 1165–1188.

Blaxter, M.L., De Ley, P., Garey, J.R., Liu, L.X., 1998. A molecular evolutionary framework for the phylum Nematoda. Nature 392, 71.

Bongers, T., 1990. The maturity index: an ecological measure of environmental disturbance based on nematode species composition. Oecologia 83, 14–19.

Bongers, T., Bongers, M., 1998. Functional diversity of nematodes. Appl. Soil Ecol. 10, 239–251. doi:10.1016/S0929-1393(98)00123-1

Bongers, T., De Goede, R.G.N., Korthals, G.W., Yeates, G.W., 1995. Proposed changes of cp classification for nematodes. Russ. J. Nematol.

Bongers, T., Ferris, H., 1999. Nematode community structure as a bioindicator in environmental monitoring. Trends Ecol. Evol. 14, 224–228.

Bongers, T., van der Meulen, H., Korthals, G., 1997. Inverse relationship between the nematode maturity index and plant parasite index under enriched nutrient conditions. Appl. Soil Ecol. 6, 195–199.

Cowart, D.A., Pinheiro, M., Mouchel, O., Maguer, M., Grall, J., Miné, J., Arnaud-Haond, S., 2015. Metabarcoding is powerful yet still blind: a comparative analysis of morphological and molecular surveys of seagrass communities. PLoS One 10, e0117562.

De Goede, R.G.M., Bongers, T., Ettema, C., 1993a. Graphical presentation and interpretation of nematode community structure: cp triangles. Meded. van Fac. Landbouwwet. Rijksuniv. Gent 58, 743–750.

De Goede, R.G.M., Georgieva, S.S., Verschoor, B.C., Kamerman, J.W., 1993b. Changes in nematode community structure in a primary succession of blown-out areas in a drift sand landscape. Fundam. Appl. Nematol. 16, 501–513.

Edgar, R.C., 2017. UNBIAS: An attempt to correct abundance bias in 16S sequencing, with limited success. bioRxiv 124149.

Edgar, R.C., 2010. Search and clustering orders of magnitude faster than BLAST. Bioinformatics 26, 2460–2461. doi:10.1093/bioinformatics/btq461

Edgar, R.C., Flyvbjerg, H., 2015. Error filtering, pair assembly and error correction for next-generation sequencing reads. Bioinformatics 31, 3476–3482.

Elbrecht, V., Leese, F., 2015. Can DNA-Based Ecosystem Assessments Quantify Species Abundance? Testing Primer Bias and Biomass—Sequence Relationships with an Innovative Metabarcoding Protocol. PLoS One 10, 1–16.

Ettema, C.H., Bongers, T., 1993. Characterization of nematode colonization and succession in disturbed soil using the Maturity Index. Biol. Fertil. Soils 16, 79–85.

Ferris, H., Bongers, T., 2006. Nematode indicators of organic enrichment. J. Nematol. 38, 3–12.

Ferris, H., Bongers, T., De Goede, R.G.M., 2001. A framework for soil food web diagnostics: extension of the nematode faunal analysis concept. Appl. soil Ecol. 18, 13–29.

Ferris, H., Matute, M.M., 2003. Structural and functional succession in the nematode fauna of a soil food web. Appl. Soil Ecol. 23, 93–110.

Guillou, L., Bachar, D., Audic, S.S., Bass, D., Berney, C.C., Bittner, L., Boutte, C., Burgaud, G.G., De Vargas, C., Decelle, J., Del Campo, J., Dolan, J.R., Dunthorn, M., Edvardsen, B., Holzmann, M., Kooistra, W.H.C.F., Lara, E., Le Bescot, N., Logares, R., Mah??, F., Massana, R., Montresor, M., Morard, R., Not, F., Pawlowski, J., Probert, I., Sauvadet, A.L., Siano, R., Stoeck, T., Vaulot, D., Zimmermann, P., Christen, R., 2012. The Protist Ribosomal Reference database (PR2): a catalog of unicellular eukaryote small sub-unit rRNA sequences with curated taxonomy. Nucleic Acids Res. 41, D597–D604. doi:10.1093/nar/gks1160

Háněl, L., 2010. An outline of soil nematode succession on abandoned fields in South Bohemia. Appl. soil Ecol. 46, 355–371.

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. Mol. Biol. Evol. 23, 1792–1800.

Ji, Y., Ashton, L., Pedley, S.M., Edwards, D.P., Tang, Y., Nakamura, A., Kitching, R., Dolman, P.M., Woodcock, P., Edwards, F.A., 2013. Reliable, verifiable and efficient monitoring of biodiversity via metabarcoding. Ecol. Lett. 16, 1245–1257.

Lejzerowicz, F., Esling, P., Pillet, L., Wilding, T.A., Black, K.D., Pawlowski, J., 2015. High-throughput sequencing and morphology perform equally well for benthic monitoring of marine ecosystems. Sci. Rep. 5, 13932.

Neher, D., Bongers, T., Ferris, H., 2004. Computation of nematode community indices, in: Society of Nematologists Workshop. p. 33.

Neher, D.A., Campbell, C.L., 1996. Sampling for regional monitoring of nematode communities in agricultural soils. J. Nematol. 28, 196.

Neher, D.A., Campbell, C.L., 1994. Nematode communities and microbial biomass in soils with annual and perennial crops. Appl. Soil Ecol. 1, 17–28.

Okada, H., Harada, H., 2007. Effects of tillage and fertilizer on nematode communities in a Japanese soybean field. Appl. Soil Ecol. 35, 582–598. doi:10.1016/j.apsoil.2006.09.008

Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O’hara, R.B., Simpson, G.L., Solymos, P., Stevens, M.H.H., Wagner, H., 2015. vegan: community ecology package. R package version 2.2–1. 2015. There is no Corresp. Rec. this Ref.

Peham, T., Steiner, F.M., Schlick‐Steiner, B.C., Arthofer, W., 2017. Are we ready to detect nematode diversity by next generation sequencing? Ecol. Evol.

Porazinska, D.., Duncan, L.., McSorley, R., Graham, J.., 1999. Nematode communities as indicators of status and processes of a soil ecosystem influenced by agricultural management practices. Appl. Soil Ecol. 13, 69–86. doi:10.1016/S0929-1393(99)00018-9

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. Mol. Ecol. Resour. 9, 1439–1450. doi:10.1111/j.1755-0998.2009.02611.x

Porazinska, D.L., McSorley, R., Duncan, L.W., Graham, J.H., Wheaton, T.A., Parsons, L.R., 1998. Nematode Community Composition under Various Irrigation Schemes in a Citrus Soil Ecosystem. J. Nematol. 30, 170–178.

Sánchez-Moreno, S., Minoshima, H., Ferris, H., Jackson, L.E., Sanchez-Moreno, S., Minoshima, H., Ferris, H., Jackson, L.E., 2006. Linking soil properties and nematode community composition: effects of soil management on soil food webs. Nematology 8, 703–715. doi:Doi 10.1163/156854106778877857

Sapkota, R., Nicolaisen, M., 2015. High-throughput sequencing of nematode communities from total soil DNA extractions. BMC Ecol. 15, 1–8. doi:10.1186/s12898-014-0034-4

Sieriebriennikov, B., Ferris, H., de Goede, R.G.M., 2014. NINJA: an automated calculation system for nematode-based biological monitoring. Eur. J. Soil Biol. 61, 90–93.

Sogin, M.L., Morrison, H.G., Huber, J.A., Welch, D.M., Huse, S.M., Neal, P.R., Arrieta, J.M., Herndl, G.J., 2006. Microbial diversity in the deep sea and the underexplored “rare biosphere”. Proc. Natl. Acad. Sci. 103, 12115–12120.

Team, Rs., 2015. RStudio: integrated development environment for R. Boston (MA): RStudio, Inc.

Unterseher, M., Siddique, A.B., Brachmann, A., Peršoh, D., 2016. Diversity and composition of the leaf mycobiome of beech (Fagus sylvatica) are affected by local habitat conditions and leaf biochemistry. PLoS One 11, e0152878.

van Megen, H., van den Elsen, S., Holterman, M., Karssen, G., Mooyman, P., Bongers, T., Holovachov, O., Bakker, J., Helder, J., 2009. A phylogenetic tree of nematodes based on about 1200 full-length small subunit ribosomal DNA sequences. Nematology 11, 927–950.

Wang, Y.I., Naumann, U., Wright, S.T., Warton, D.I., 2012. mvabund–an R package for model‐based analysis of multivariate abundance data. Methods Ecol. Evol. 3, 471–474.

Wardle, D.A., Yeates, G.W., Watson, R.N., Nicholson, K.S., 1995. Development of the decomposer food-web, trophic relationships, and ecosystem properties during a three-year primary succession in sawdust. Oikos 155–166.

Welch, B.L., 1951. On the comparison of several mean values: an alternative approach. Biometrika 38, 330–336.

Wilcox, R., 2012. Chapter 7 - One-Way and Higher Designs for Independent Groups BT - Introduction to Robust Estimation and Hypothesis Testing (Third Edition), in: Statistical Modeling and Decision Science. Academic Press, Boston, pp. 291–377. doi:https://doi.org/10.1016/B978-0-12-386983-8.00007-X

Yeates, G.W., 1994. Modification and qualification of the nematode maturity index. Pedobiologia (Germany).

Yeates, G.W., Bongers, T., 1999. Nematode diversity in agroecosystems. Agric. Ecosyst. Environ. 74, 113–135.

Yu, D.W., Ji, Y., Emerson, B.C., Wang, X., Ye, C., Yang, C., Ding, Z., 2012. Biodiversity soup: metabarcoding of arthropods for rapid biodiversity assessment and biomonitoring. Methods Ecol. Evol. 3, 613–623.

Zhang, Z., Schwartz, S., Wagner, L., Miller, W., 2000. A greedy algorithm for aligning DNA sequences. J. Comput. Biol. 7, 203–214.

Zimmermann, J., Glöckner, G., Jahn, R., Enke, N., Gemeinholzer, B., 2015. Metabarcoding vs. morphological identification to assess diatom diversity in environmental studies. Mol. Ecol. Resour. 15, 526–542.