bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921304; this version posted January 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

1 # A novel metabarcoding strategy for studying nematode
2 # communities

3

4 Md. Maniruzzaman Sikder[1, 2], Mette Vestergård[1], Rumakanta Sapkota[3], Tina Kyndt[4],
5 Mogens Nicolaisen[1]*

6

7 [1]Department of Agroecology, Aarhus University, Slagelse, Denmark
8 [2]Department of Botany, Jahangirnagar University, Savar, Dhaka, Bangladesh
9 [3]Department of Environmental Science, Aarhus University, Roskilde, Denmark
10 [4]Department of Biotechnology, Ghent University, Ghent, Belgium

11

12
13 *Corresponding author
14 Email: mn@agro.au.dk

15

16 ## Abstract

17 Nematodes are widely abundant soil metazoa and often referred to as indicators of soil health.
18 While recent advances in next-generation sequencing technologies have accelerated
19 research in microbial ecology, the ecology of nematodes remains poorly elucidated, partly due
20 to the lack of reliable and validated sequencing strategies. Objectives of the present study
21 were (i) to compare commonly used primer sets and to identify the most suitable primer set
22 for metabarcoding of nematodes; (ii) to establish and validate a high-throughput sequencing
23 strategy for nematodes using Illumina paired-end sequencing. In this study, we tested four

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921304; this version posted January 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

24 primer sets for amplicon sequencing: JB3/JB5 (mitochondrial, I3-M11 partition);
25 SSU_04F/SSU_22R (18S rRNA, V1-V2 region); Nemf/18Sr2b (18S rRNA, V6-V8 region) from
26 earlier studies; and MMSF/MMSR (18S rRNA, V4-V5 region), a newly developed primer set
27 from this study. In order to test the primer sets, we used 22 samples of individual nematode
28 species, 20 mock communities, 20 soil samples, 20 spiked soil samples (mock communities
29 in soil), and 4 root/rhizosphere soil samples. We successfully amplified the target regions (I3-
30 M11 partition of the COI gene; V1-V2, V4-V8 region of 18S rRNA gene) from these 86 DNA
31 samples with the four different primer combinations and sequenced the amplicons on an
32 Illumina MiSeq sequencing platform. We found that the MMSF/MMSR and Nemf/18Sr2b were
33 efficient in detecting nematode compared to JB and SSU primer sets based on annotation of
34 sequence reads at genus and in some cases at species level. Therefore, these primer sets
35 are suggested for studies of nematode communities in agricultural environments.

36 **Keywords:** nematode diversity, soil, plant, rhizosphere, environmental, NGS, sequencing,
37 primer design.

38 Background

39 Nematodes are highly diverse and abundant metazoans with worldwide distribution [1].
40 Generally, nematologists have relied on classical morphology-based taxonomy along with
41 biochemical or molecular methods for nematode identification [2, 3]. Morphological
42 identification is difficult, requires taxonomic expertise and often becomes challenging when it
43 comes to identifying nematodes at lower taxonomic levels [4]. DNA based identification have
44 eased the task of taxonomic nematode identification in recent years, and most molecular
45 based diagnostic approaches usually target the nuclear ribosomal DNA region. In addition,
46 the mitochondrial cytochrome oxidase I gene (COI gene) has been successfully used for
47 identification of nematodes and for resolving taxonomic relationships among closely related
48 species [5-7]. For certain groups of taxa, the COI gene has been shown to provide greater
49 taxonomic resolution compared to the small subunit (SSU, 18S rRNA) rDNA [8]. The potential

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921304; this version posted January 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

50 of COI gene-based barcoding has been explored for nematode taxa ranging from root-knot
51 nematodes [9], marine nematodes [7], Aphelenchoididae [10] and *Pratylenchus* [11]. Both
52 marker genes, 18S ribosomal DNA and COI, comes with their own limitations and strengths.
53 The reference database for COI sequences is less enriched in comparison to 18S, limiting the
54 implementation of COI barcoding for nematodes. The most inclusive molecular phylogenetic
55 study of nematodes now available comprised 1215 full-fragment sequences of SSU rDNA [12].
56 There as several reports on the use of 18S rRNA based barcodes for successful nematode
57 community analysis, and they resolved several taxonomic issues of identification of several
58 nematodes [13-15]. Consequently, the 18S rRNA gene may remain the most widely used
59 molecular marker for identification of nematodes [16, 17].

60 The field of DNA based identification is transitioning from barcoding individual species to
61 metabarcoding of entire communities. However, the success of metabarcoding approaches
62 largely relies on suitable primers used for amplification of environmental DNA (eDNA).
63 Nematode community studies by earlier workers have relied on nematode extraction [18, 19]
64 to screen out other organisms present in the samples during amplification. This process is
65 time consuming, laborious and may introduce biases [20]. Therefore, in the present study, we
66 compared amplification strategies that avoided such nematode isolation steps. In a previous
67 study, we have already optimized a soil DNA extraction method that we used to evaluate
68 nematode communities from a number of agricultural soils using the Roche 454 platform [21].
69 After alignment of 18S rRNA genes of eukaryotic sequences available in the SILVA database,
70 variable regions V2, V4, and V9 were suggested as the most suitable for biodiversity
71 assessments [22]. The aims of the present study were (i) to compare commonly used primer
72 sets from the literature and a newly designed primer set, and identify the most suitable primer
73 set for metabarcoding of nematodes; (ii) to validate and establish a high-throughput
74 sequencing strategy for nematodes using Illumina paired-end sequencing from individual
75 nematode species as well as bulk DNA from soil. For this, we used single nematodes, mock
76 communities in water and in soil backgrounds, DNA from agricultural fields and from

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921304; this version posted January 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

77 root/rhizosphere samples to validate the primer sets and to test the taxonomic composition of
78 the communities.

79 Materials and Methods

80 Primer sets

81 We selected four primer sets for amplicon sequencing of nematodes (S1 Table). The primer
82 set SSU_04F/SSU_22R (SSU) amplifies the V1-V2 region of the 18S rRNA gene (Fig 1) and
83 was recently used to describe assemblages of free-living soil nematodes using the MiSeq
84 platform [17, 23]. We designed a primer set, MMS (MMSF: 5′-GGTGCCAGCAGCCGCGGTA-
85 3′, MMSR: 5′-CTTTAAGT TTCAGCTTTGC-3′) located in the variable region V4-V5 of the 18S
86 rRNA gene (Fig 1). Furthermore, we included the Nemf/18Sr2b primer set (NEM) covering the
87 V6-V8 regions (Fig 1), which has been used to characterize nematode communities from
88 agricultural soils using the Roche 454 platform [18, 21, 24]. Finally, we tested a mitochondrial
89 primer set JB3/JB5 (JB) targeting the I3-M11 region of the COI gene, which has been used to
90 study nematode communities in agricultural field soils and unmanaged flowerbeds in Japan
91 [8].

92

93 **Fig 1. Location of metabarcoding primers targeting variable regions in 18S rRNA gene**
94 **used in the present study.**

95 Nematode species, mock communities and root/
96 rhizosphere soil samples

97 In order to test the primer sets, we used 22 nematode species obtained from different
98 geographical origins (S2 Table) and artificially assembled 10 mock communities using DNA
99 extracts from these 22 nematodes (named Mock-1, Mock-2 etc.). We combined DNA from the
100 nematode species in different concentrations (S3 Table). We also tested total DNA extracted
101 from soil samples collected from 20 agricultural fields in different parts of Denmark. The field
102 crop history, i.e. the previous crop, stutus, soil type, and pH was recorded (S4 Table). Soil

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921304; this version posted January 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

103 sampling and DNA extraction from the fields were described earlier [24]. Moreover, we spiked
104 20 nematode mock communities in DNA extracted from soil. We pooled mock community DNA
105 with soil DNA in a 1:1 ratio, referred to as soil-mock communities. Furthermore, we included
106 DNA extracted from washed and freeze-dried root knot nematode (*Meloidogyne incognita*)
107 infected tomato (*Solanum lycopersicum* L.) roots, quinoa (*Chenopodium quinoa* Willd.) roots,
108 maize (*Zea mays* L.) roots/rhizosphere soil, and green bean (*Phaseolus vulgaris* L.)
109 roots/rhizosphere soil.

110

111 DNA extraction, PCR and sequencing strategy

112 DNA was extracted from 250 mg of the freeze-dried and ground soil samples using the
113 PowerLyzer soil DNA extraction kit (Qiagen, Germany) according to the manufacturer's
114 instructions, except that samples were homogenized in a Geno/Grinder 2000 at 1500 rpm for
115 3 x 30 seconds. For the root/rhizosphere soil samples, DNA was extracted from 20 mg of
116 ground material with the DNeasy plant mini kit (Qiagen, Germany).

117 To amplify target regions, the first PCR was performed in a reaction mixture of 25 µl consisting
118 of 5 µl of Promega 5X reaction buffer, 1.5 µl of $MgCl_2$ (25 mM), 2 µl dNTPs (2.5 mM), 0.5 µl of
119 each primer (10 µM), 0.125 µl of GoTaq Flexi polymerase (5U, Promega Corporation,
120 Madison, USA) and 2 µl of DNA template (approximate 2 ng/µl). PCR cycles for the JB primer
121 combination were 94ºC for 5 min (94ºC 1 min, 50ºC 30 sec, 72ºC 45 sec) 35 cycles, 72ºC 10
122 min, and 4ºC on hold [25]. Similar PCR cycles were used except that the annealing
123 temperature was 53ºC for MMS and NEM, and 55ºC for the SSU primer set [26]. Each of the
124 primer sets of the first PCR (S1 Table) were tagged with the Illumina adaptor overhang
125 nucleotide sequence, for forward primer 5′-TCGTCGGCAGCGTCAGATGTGTATAAGAGAC
126 AG-3′ and for reverse primer 5′-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG-3′.
127 After this PCR, we pooled and diluted (1:5) the amplicons.

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921304; this version posted January 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

128 A second PCR was performed for dual indexing. The master mix of this PCR was identical to
129 the first PCR except that 2 µl of DNA template and 2 µl of the different combinations of index
130 primers were used. Each index primer consisted of a sequence specific for Illumina
131 sequencing, a unique 8 bp multiplex identifier and the Illumina adapter overhang sequence.
132 The second PCR was performed with the following cycles: 94ºC 5 min, (94ºC 30 sec, 55ºC 30
133 sec, 72ºC 1 min) 13 cycles, 72ºC 10 min, and 4ºC on hold. All amplicons were visualized by
134 gel electrophoresis, pooled (approximately equal amounts), precipitated and the pellet
135 dissolved in DNAase free water. Pooled DNA was run on a gel and amplicons were excised
136 and purified using the QIAquick Gel Extraction kit (QIAGEN, Germany) according the
137 manufacturer's instruction. Finally, the DNA concentrations were measured fluorometrically
138 (Qubit, Thermo Fisher Scientific) and sent for sequencing on an Illumina MiSeq sequencer
139 with PE300 (Eurofins Genomics, Germany).

140

141 Sequence Analysis

142 The paired end reads obtained from the Illumina MiSeq runs were analyzed using VSEARCH
143 version 2.6 [27]. For joining paired-end reads, we used an overlapping minimum read length
144 of 30 base pairs and reads with quality Phred scores <30 were removed. Internal barcodes,
145 forward and reverse primers, and reads less than 200 base pairs were also excluded.
146 Following this, sequences were dereplicated, screened for chimeras and clustered at 99%
147 similarity level using VSEARCH. Taxonomy assignments for the clustered operational
148 taxonomic units (OTUs) were done using the SILVA 132 reference database [28, 29] for
149 eukaryotes in QIIME using assign_taxonomy.py [30]. Moreover, all nematode OTUs were
150 blasted (≥ 98% similarity) against the NCBI GenBank database to reconfirm their taxonomic
151 assignment. Statistics and data visualization were carried out using the statistical package R.

152

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921304; this version posted January 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

153 # Results

154 ## Data characteristics

155 We successfully obtained sequence reads from 22 individual nematodes species, 20 different
156 mock communities with and without a soil background, 20 different soils and 4
157 roots/rhizosphere soil samples using the four primer sets. In total, 18.2 million sequence reads
158 were obtained. After quality control, sequence reads were clustered into 320, 17734, 874 and
159 313 OTUs at 99% similarity for JB, SSU, MMS and NEM primer sets, respectively.

160 ## Sequencing of individual nematode species

161 For the individual nematode species, we could annotate 10 of the 22 samples to species level
162 and nine to genus level with the JB primer set, whereas three species were not amplified with
163 this primer set (Table 1). Using the SSU primer set, only 15 out of the 22 nematode species
164 were amplified. The MMS primer set amplified all nematodes except *Meloidogyne graminicola*
165 (Table 1). This primer set identified *Meloidogyne* at the genus level, and the other nine
166 nematodes were assigned at species level. The NEM primer set successfully amplified all the
167 nematode species used in our study. *Meloidogyne* species were assigned to genus level and
168 cyst nematodes (*Heterodera carotae* and *H. schachtii*) could only be identified at family level
169 (Table 1). The remaining five nematodes were detected at the species level.

170

171 **Table 1. The efficiency of four metabarcoding primer sets in detection of individual**
172 **nematode species.** NCBI Blast tool was used for taxonomic assignments, and top hits with
173 sequence similarities ≥ 99% and coverage 100% were considered for taxonomic assignment
174 at species level.

175

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921304; this version posted January 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Nematode taxa	JB	SSU	MMS	NEM
Meloidogyne incognita	Genus	Genus	Genus	Genus
Meloidogyne arenaria	Genus	Genus	Genus	Genus
Meloidogyne javanica	Genus	Genus	Genus	Genus
Meloidogyne graminicola	Genus	**Not Detected**	**Not Detected**	Genus
Meloidogyne ethiopica	Genus	Genus	Genus	Genus
Meloidogyne inornata	Genus	Genus	Genus	Genus
Meloidogyne ulmi	Genus	Genus	Genus	Genus
Meloidogyne luci	Genus	Genus	Genus	Genus
Root knot nematodes	Genus	Genus	Genus	Genus
Meloidogyne enterolobi	Species	**Not Detected**	Genus	Genus
Meloidogyne chitwoodi	Species	Genus	Genus	Genus
Meloidogyne fallax	Species	Genus	Genus	Genus
Meloidogyne minor	Species	Species	Genus	Genus
Meloidogyne naasi	Species	**Not Detected**	Genus	Species
Pratylenchus penetrans	Species	**Not Detected**	Species	Species
Pratylenchus neglectus	Species	**Not Detected**	Species	Genus
Heterodera carotae	**Not Detected**	**Not Detected**	Species	Family

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921304; this version posted January 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Heterodera schachtii	Species	Species	Species	Family
Belonolaimus longiscaudatus	**Not Detected**	Species	Species	Species
Bursaphelenchus mucronatus	Species	**Not Detected**	Species	Species
Caenorhabditis elegans	Species	Species	Species	Genus
Ditylenchus dipsaci	**Not Detected**	Species	Species	Species

Mock communities and mock communities in soil

In mock communities, JB produced sequences were assigned to genus level within the families Meloidogynidae and Heteroderidae, and to species level within Pratylenchidae and Rhabditidae (Table 2). However, one-third (33%) of the dataset remained unassigned (S1 Fig), and nematodes from Dolichodoridae were not amplified (Table 2). The SSU primer set generated sequences that were assigned to genus level within the Meloidogynidae and Heteroderidae, whereas sequences within Rhabditidae were assigned to species level. The SSU primers failed to amplify nematodes from Pratylenchidae and Dolichodoridae (Table 2). The MMS primer pair generated Meloidogynidae sequences that could be assigned to genus level and for the other three families, Heteroderidae, Dolichodoridae, and Rhabditidae sequences were assigned to species level (Table 2). The NEM primer set was able to amplify and sequence nematodes to the genus level within Meloidogynidae, to species level within Pratylenchidae, Dolichodoridae and Rhabditidae, and to family level within Heteroderidae (Table 2). In the mock communities including diluted DNA of individual nematodes, we observed lower relative abundance of diluted taxa compared to undiluted taxa; however, diluted samples were generally detected in unexpectedly high amounts (S2 and S3 Figs).

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921304; this version posted January 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

194 **Table 2. The efficiency of four metabarcoding primers in detection of nematodes in**
195 **mock communities based on BLAST searches.** Only taxonomic assignments appearing in
196 top hits and with sequence similarities ≥ 99% and coverage 100% were considered.

Mock communities	**JB**	**SSU**	**MMS**	**NEM**
Meloidogynidae	Genus	Genus	Genus	Genus
Heteroderidae	Genus	Species	Species	Family
Pratylenchidae	Species	**Not detected**	**Not detected**	Species
Dolichodoridae	**Not detected**	**Not detected**	Species	Species
Rhabditidae	Species	Species	Species	Species

197

198 In soil-mock combinations, reads from nematodes from the mock communities were generally
199 highly abundant compared to reads from the nematodes derived from the soil background (S5
200 Table). The JB primer pair only detected nematode families from the mock communities and
201 no additional sequences from the soil background were detected. The SSU primer set was
202 able to detect nematodes belonging to three families (Meloidogynidae, Heteroderidae and
203 Rhabditidae) of the mock communities. Both the MMS and NEM primer sets detected
204 nematodes of the families represented in the mock communities and additionally other
205 nematode families from the spiked soil samples.

206 Nematode communities in soil samples

207 For the JB primer set, 4% and 31% of the total number of sequence reads were classified as
208 Nematoda in soil and plant root/rhizosphere soil samples, respectively, while many sequence
209 reads were unassigned (Fig 2). For the SSU primer set, only 1% of the sequence reads were
210 classified as Nematoda, both in soil and plant root/rhizosphere soil samples (Fig 2). This
211 primer set detected a broad spectrum of other eukaryotes such as fungi, plant, Cercozoa and
212 Charophyta. For the newly designed primer pair (MMSF/MMSR), 17% and 34% of total

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921304; this version posted January 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

213 sequence reads belonged to Nematoda (Fig 2), and for the NEM primer set, 74% and 99% of
214 the total sequences belonged to Nematoda in the soil and plant root/rhizosphere soil samples,
215 respectively (Fig 2).

216 We are not presenting any further results for the JB and SSU primer sets due to their poor
217 performance (Fig 2). In the soil samples, NEM and MMS detected a wide range of nematodes
218 from different families (Figs 3 and 4). We recovered 30 nematode families using both the NEM
219 and the MMS primer sets with 6 unique families detected by each primer set (Fig 3). We
220 recorded 14 and 15 different unique genera in soil samples with NEM and MMS primer sets,
221 respectively, and 23 genera were detected by both primer sets (Fig 3). We found 34 and 26
222 unique nematode species with the NEM and MMS primer set, respectively, while 25 species
223 were detected by both primer sets (Fig 3). We observed that the ability of the two primer sets
224 MMS and NEM to detect nematode families were comparable expect for a limited number of
225 nematode families e.g. Rhabditidae, Trichodoridae, Merliniidae, Heteroderidae (Fig 4 and S6
226 Table).

227

228 **Fig 2. Relative distribution of sequence reads in soil and plant root/rhizosphere soil**
229 **samples amplified with primer sets JB, SSU, MMS, and NEM**; percentage in the blue slide
230 indicates the proportion of sequence reads that were assigned to Nematoda.

231

232 **Fig 3. Venn diagrams showing the number of taxa detected in soil samples by the primer**
233 **sets MMS and NEM.** Only taxonomic assignments appearing in top hits of BLAST searches
234 and with sequence similarities ≥ 98% and 100% coverage were considered.

235

236 **Fig 4. Relative distribution of nematode sequence reads in all soil samples amplified**
237 **and sequenced with primer sets MMS and NEM.** Only taxonomic assignments appearing

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921304; this version posted January 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

238 in top hits of BLAST searches and with sequence similarities ≥ 98% and 100% coverage were
239 considered.
240

241 Nematode communities in plant root/rhizosphere soil 242 samples

243 In plant root/rhizosphere soil samples, we recovered 16 families by the NEM primers, followed
244 by 13 families using the MMS primer set (Fig 3). Ten families were detected by both primer
245 sets and we recorded 10 and 9 unique genera with NEM and MMS primer sets, respectively,
246 while 10 genera were detected by both primer sets (Fig 3). We detected 16 unique nematode
247 species using the NEM primer set, and 9 unique species were detected using the MMS
248 primers and 15 species were detected by both primer sets. Both primer sets detected a large
249 variation in nematode presence in the samples, and the two primer pairs showed variation in
250 recovered nematode taxa (Fig 5). The quinoa roots and the root knot nematode infected
251 tomato roots were dominated by Meloidogynidae. Both primer sets detected plant parasitic
252 and free-living nematode taxa in green bean and maize root/rhizospheres soils samples.
253

254 **Fig 5. Relative distribution of sequence reads at family rank in plant root/rhizosphere**
255 **soil samples amplified with primer sets MMS and NEM**. Only taxonomic assignments
256 appearing in top hits of BLAST searches and with sequence similarities ≥ 98% and 100%
257 coverage were considered.

258 Discussion

259 Most protocols for nematode metabarcoding include a nematode extraction step to reduce
260 DNA contamination from other soil-living organisms [17-19, 23]. This extraction step may
261 introduce biases as particular nematode taxa or developmental stages are not necessarily
262 extracted at the same efficiency [20]. Furthermore, extraction steps may not be practical when

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921304; this version posted January 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

263 several groups of organisms such as nematodes, fungi and bacteria are studied in the same
264 samples. To overcome these limitations, we previously developed an amplification strategy
265 for 454 pyrosequencing that selectively amplifies nematode DNA from total soil DNA
266 extractions [21]. In the present study, we have adapted amplification strategies for the Illumina
267 MiSeq platform, and we compared different primer sets for their ability to selectively amplify
268 nematode communities.

269 We observed that the JB primer set only amplified 86% of the individual tested nematode
270 species and did not amplify species that are agronomically important, namely *Heterodera*
271 *carotae*, *Belonolaimus longicaudatus*, and *Ditylenchus dipsaci*. It has been reported that there
272 are not enough reference sequences of the COI target region in the database for effective
273 species identification [16, 17]. As previously suggested by other researchers, the COI gene
274 has high mutation rates. Hence, the primer sequences are poorly conserved throughout the
275 phylum Nematoda [5, 31]. Based on a study of single nematode species, mock communities,
276 and low number of nematode sequence reads in soil and root/rhizosphere soil samples, we
277 found that the JB primer set targeting the I3-M11 partition of the COI gene is not suitable for
278 nematode metabarcoding.

279 In a recent study, the SSU ribosomal DNA marker (SSU_04F/SSU_22R) outperformed the
280 mitochondrial marker (JB3/JB5GED) in terms of nematode species and genus level detection
281 [17, 23]. However, in our study, important nematode species were not amplified and detected
282 by the SSU primer set. Moreover, the amplification strategy using the SSU primer set only
283 resulted in 1% Nematoda reads from soil and plant root/rhizosphere soil DNA samples. Our
284 results corroborates a recent study in which the NF1/18Sr2b primer set provided better
285 taxonomic resolutions compared to the SSU_04F/SSU_22R marker [17]. In other studies, this
286 SSU primer set was found to amplify a large portion of non-nematode reads of environmental
287 marine sediment samples [32-35]. Therefore, this primer set was not considered suitable for
288 nematode diversity studies of environmental samples without an initial nematode extraction
289 step.

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921304; this version posted January 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

290 Results from the analysis of the individual nematodes showed that better taxonomic resolution
291 was achieved with MMS, which targets the V4-V5 region of 18S rRNA gene, compared to JB
292 and SSU primer sets. The efficiency of this primer set was further confirmed using the mock
293 communities as it was able to detect all the nematode taxa in the mock communities, also in
294 a soil background. MMS detected a high diversity of the nematode communities in soil
295 samples, suggesting that this newly designed primer set is well suited for studies of plant
296 parasitic and free-living nematodes. This primer set was also able to detect many nematode
297 families in the plant root/rhizosphere soil samples. Based on these observations, this newly
298 designed MMS primer set is efficient for studies of soil nematode communities, and it clearly
299 outperforms the JB and SSU primer sets.

300 The NEM primer set was previously developed for the 454-sequencing platform using a semi-
301 nested PCR approach. However, in the present study, the second PCR in the nested PCR
302 was omitted, which resulted in a larger PCR product (500bp) including the V6, V7, and V8
303 regions of the 18S ribosomal RNA gene. All individual nematode species in our study were
304 identified using NEM, and all nematode families in the mock communities were detected. In
305 addition, NEM detected a range of diverse nematode taxa in the different soils, reflecting the
306 different crop species that had been grown in the soils, and the different soil parameters. The
307 NEM primer set amplified more nematode taxa in the root/rhizosphere soil samples compared
308 to all the other primer sets tested. This primer set amplified almost 100% nematode DNA in
309 the presence of plant DNA, which indicates that this primer set is highly nematode specific.

310 Although we detected fewer sequence reads using both MMS and NEM primer sets when we
311 used diluted templates in the mock communities, the read counts were not reduced
312 quantitatively. The reason for this is not known.

313 Sequence reads from taxa that belong to the family Rhabditidae were much more prevalent
314 in the MMS than in the NEM-generated data set. This discrepancy is probably due to a three-
315 nucleotide mismatch between 18Sr2b primer of the NEM primer set and the Rhabditidae DNA
316 template. It was reported that the reverse primer sequence (18Sr2b) failed to amplify several

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921304; this version posted January 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

317 Rhabditidae species [36]. In a recent study, a modified version of the primer set Nemf/18Sr2b,
318 named as NemFopt/18Sr2bRopt, was constructed by adding extra nucleotides and by
319 including degenerate bases in both the forward and reverse primer to improve GC content and
320 shift the reverse primer into a more conserved region of Nematoda [18]. The MMS primer set
321 could overcome problems associated with detecting Rhabditidae species. Multiple sequence
322 alignments (S4 Fig), detection of a higher number of taxa belonging to Rhabditidae and the
323 greater relative abundance of Rhabditidae in soil samples in our sequence data confirmed that
324 the MMS primer set efficiently detects Rhabditiae.

325 The MMS primer set did not detect nematode taxa of the families Aporcelaimidae,
326 Diplogastridae, Merliniidae, Neodiplogastridae, Tylenchidae, and Tylenchulidae in
327 root/rhizosphere soil samples, although this group of nematode taxa was detected in our soil
328 samples. This fact could be due to the competition in primer annealing between nematode
329 and plant DNA templates in root/rhizosphere soil samples. The NEM primer set was not able
330 to detect the families Heteroderidae, Dolichodoridae, Telotylenchidae in the root/rhizosphere
331 soil samples. We observed that the NEM primer set could not detect *Heterodera carotae, H.*
332 *schachtii* and *Globodera* spp. neither at genus nor at species level in individual nematode
333 species, mock communities, and mock communities in soil. On the contrary, the MMS primer
334 set was found to be efficient in detecting nematodes belonging to Heteroderidae.

335 We propose to use both primer sets (MMS and NEM) for identification of nematode
336 communities on DNA extracted directly from soil. Together, these two primer sets cover more
337 than 1000 bp of the 18S rRNA gene and they capture a substantial range of the variable
338 regions (V4, V5, V6, V7, and V8) of the 18S rRNA gene in nematodes. Moreover, the
339 assignment of lower Linnaean taxonomies (genus, species) to sequence reads is a very
340 crucial step in the use of DNA markers for biodiversity assessment. We conclude that our two
341 primer sets (MMS and NEM) complement each other in detecting nematode families and can
342 efficiently detect nematodes at the genus level and in some cases at species level.

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921304; this version posted January 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Acknowledgement

We are grateful to Andrea M. Skantar, United State Department of Agriculture; Barbara Geric Stare, Sasa Sirca and Gregor Urek, Department of Plant Protection, Agricultural Institute of Slovenia for contributing DNA samples of nematode species. We would like to thank Susana Santos for data visualization. We would also like to thank Mathilde Schiøtt Dige and Simone Ena Rasmussen for their laboratory technical assistance.

References

1. Lambshead PJD, Boucher G. Marine nematode deep-sea biodiversity - Hyperdiverse or hype? J Biogeogr. 2003;30:475-85. doi: 10.1046/j.1365-2699.2003.00843.x.

2. De Ley P, Tandingan De Ley I, Morris K, Abebe E, Mundo-Ocampo M, Yoder M, et al. An integrated approach to fast and informative morphological vouchering of nematodes for applications in molecular barcoding. Philosophical Transactions of the Royal Society B: Biological Sciences. 2005;360(1462):1945-58. doi: 10.1098/rstb.2005.1726.

3. Agatha S, Strüder-Kypke MC. Phylogeny of the order Choreotrichida (Ciliophora, Spirotricha, Oligotrichea) as inferred from morphology, ultrastructure, ontogenesis, and SSrRNA gene sequences. Eur J Protistol. 2007;43(1):37-63. doi: 10.1016/j.ejop.2006.10.001.

4. Abebe E, Mekete T, Thomas W. A critique of current methods in nematode taxonomy. African Journal of Biotechnology. 2013;10(3):312-23. doi: 10.4314/ajb.v10i3.

5. Blouin MS, Yowell CA, Courtney CH, Dame JB. Substitution bias, rapid saturation, and the use of mtDNA for nematode systematics. Molecular Biology and Evolution. 1998;15(12):1719-27. doi: 10.1093/oxfordjournals.molbev.a025898.

6. Derycke S, Remerie T, Vierstraete A, Backeljau T, Vanfleteren J, Vincx M, et al. Mitochondrial DNA variation and cryptic speciation within the free-living marine nematode Pellioditis marina. Mar Ecol Prog Ser. 2005;300:91-103. doi: 10.3354/meps300091.

7. Derycke S, Vanaverbeke J, Rigaux A, Backeljau T, Moens T. Exploring the use of cytochrome oxidase c subunit 1 (COI) for DNA barcoding of free-living marine nematodes. PLoS One. 2010;5(10):e13716. doi: 10.1371/journal.pone.0013716. PubMed PMID: 21060838.

8. Morise H, Miyazaki E, Yoshimitsu S, Eki T. Profiling nematode communities in unmanaged flowerbed and agricultural field soils in Japan by DNA barcode sequencing. PLoS One. 2012;7(12):e51785. doi: 10.1371/journal.pone.0051785.

9. Janssen T, Karssen G, Verhaeven M, Coyne D, Bert W. Mitochondrial coding genome analysis of tropical root-knot nematodes (Meloidogyne) supports haplotype based diagnostics and reveals evidence of recent reticulate evolution. Sci Rep. 2016;6:22591. doi: 10.1038/srep22591.

10. Sánchez-Monge A, Janssen T, Fang Y, Couvreur M, Karssen G, Bert W. mtCOI successfully diagnoses the four main plant-parasitic Aphelenchoides species (Nematoda: Aphelenchoididae) and supports a multiple origin of plant-parasitism in this paraphyletic genus. European Journal of Plant Pathology. 2017;148(4):853-66. doi: 10.1007/s10658-016-1141-1.

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921304; this version posted January 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

380 11. Janssen T, Karssen G, Orlando V, Subbotin SA, Bert W. Molecular characterization and
381 species delimiting of plant-parasitic nematodes of the genus Pratylenchus from the penetrans group
382 (Nematoda: Pratylenchidae). Mol Phylogen Evol. 2017;117:30-48. doi: 10.1016/j.ympev.2017.07.027.

383 12. Van Megen H, Van Den Elsen S, Holterman M, Karssen G, Mooyman P, Bongers T, et al.
384 A phylogenetic tree of nematodes based on about 1200 full-length small subunit ribosomal DNA
385 sequences. Nematology. 2009;11(6):927-50. doi: 10.1163/156854109X456862.

386 13. Correia MTdS, Rieger TT, Ley PD, Decraemer W, Esteves AM, Genevois VF, et al. Marine
387 nematode taxonomy in the age of DNA: the present and future of molecular tools to assess their
388 biodiversity. Nematology. 2010;12(5):661-72. doi: 10.1163/138855410X500073.

389 14. Tang CQ, Leasi F, Obertegger U, Kieneke A, Barraclough TG, Fontaneto D. The widely
390 used small subunit 18S rDNA molecule greatly underestimates true diversity in biodiversity surveys of
391 the meiofauna. PNAS. 2012;109(40):16208-12. doi: 10.1073/pnas.1209160109.

392 15. Morgan MJ, Bass D, Bik H, Birky CW, Blaxter M, Crisp MD, et al. A critique of Rossberg
393 et al.: noise obscures the genetic signal of meiobiotal ecospecies in ecogenomic datasets. Proceedings
394 of the Royal Society B: Biological Sciences. 2014;281(1783):20133076. doi: 10.1098/rspb.2013.3076.

395 16. Avó AP, Daniell TJ, Neilson R, Oliveira S, Branco J, Adão H. DNA Barcoding and
396 Morphological Identification of Benthic Nematodes Assemblages of Estuarine Intertidal Sediments:
397 Advances in Molecular Tools for Biodiversity Assessment. Frontiers in Marine Science. 2017;4:66. doi:
398 org/10.3389/fmars.2017.00066.

399 17. Ahmed M, Black MA, Prior T, Karssen G, Lawson R, Adams I, et al. Metabarcoding of soil
400 nematodes: the importance of taxonomic coverage and availability of reference sequences in
401 choosing suitable marker(s). Metabarcoding and Metagenomics. 2019;3:77-99. doi:
402 10.3897/mbmg.3.36408.

403 18. Waeyenberge L, de Sutter N, Viaene N, Haegeman A. New insights into nematode DNA-
404 metabarcoding as revealed by the characterization of artificial and spiked nematode communities.
405 Diversity. 2019;11(4):52. doi: 10.3390/d11040052.

406 19. Treonis AM, Unangst SK, Kepler RM, Buyer JS, Cavigelli MA, Mirsky SB, et al.
407 Characterization of soil nematode communities in three cropping systems through morphological and
408 DNA metabarcoding approaches. Sci Rep. 2018;8:2004. doi: 10.1038/s41598-018-20366-5.

409 20. Verschoor BC, de Goede RGM. The nematode extraction efficiency of the Oostenbrink
410 elutriator-cottonwool filter method with special reference to nematode body size and life strategy.
411 Nematology. 2000;2:325-42. doi: Doi 10.1163/156854100509204. PubMed PMID:
412 WOS:000089943000009.

413 21. Sapkota R, Nicolaisen M. High-throughput sequencing of nematode communities from
414 total soil DNA extractions. BMC Ecol. 2015;15(3):1-8. doi: 10.1186/s12898-014-0034-4.

415 22. Hadziavdic K, Lekang K, Lanzen A, Jonassen I, Thompson EM, Troedsson C.
416 Characterization of the 18s rRNA gene for designing universal eukaryote specific primers. PLoS One.
417 2014;9(2):e87624. doi: 10.1371/journal.pone.0087624.

418 23. Macheriotou L, Guilini K, Bezerra TN, Tytgat B, Nguyen DT, Phuong Nguyen TX, et al.
419 Metabarcoding free-living marine nematodes using curated 18S and CO1 reference sequence
420 databases for species-level taxonomic assignments. Ecol Evol. 2019;9(1):1-16. doi:
421 10.1002/ece3.4814.

422 24. Sapkota R, Nicolaisen M. Cropping history shapes fungal, oomycete and nematode
423 communities in arable soils and affects cavity spot in carrot. Agric Ecosyst Environ. 2018;257:120-31.
424 doi: 10.1016/j.agee.2018.01.032. PubMed PMID: WOS:000428005800013.

425 25. Derycke S, Vanaverbeke J, Rigaux A, Backeljau T, Moens T. Exploring the use of
426 cytochrome oxidase c subunit 1 (COI) for DNA barcoding of free-living marine nematodes. PLoS One.
427 2010. doi: 10.1371/journal.pone.0013716. PubMed PMID: 21060838.

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921304; this version posted January 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

428 26. Blaxter ML, De Ley P, Garey JR, Llu LX, Scheldeman P, Vierstraete A, et al. A molecular
429 evolutionary framework for the phylum Nematoda. Nature. 1998. doi: 10.1038/32160.

430 27. Rognes T, Flouri T, Nichols B, Quince C, Mahé F. VSEARCH: a versatile open source tool
431 for metagenomics. PeerJ. 2016;4:e2584. doi: 10.7717/peerj.2584.

432 28. Yilmaz P, Parfrey LW, Yarza P, Gerken J, Pruesse E, Quast C, et al. The SILVA and "all-
433 species Living Tree Project (LTP)" taxonomic frameworks. Nucleic Acids Res. 2014;42(Database):D643–
434 D8. doi: 10.1093/nar/gkt1209.

435 29. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal
436 RNA gene database project: Improved data processing and web-based tools. Nucleic Acids Res.
437 2013;41(Database):D590–D6. doi: 10.1093/nar/gks1219.

438 30. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, et al.
439 QIIME allows analysis of high- throughput community sequencing data Intensity normalization
440 improves color calling in SOLiD sequencing. Nat Methods. 2010;7(5):335–6. doi: 10.1038/nmeth0510-
441 335. PubMed PMID: 20383131.

442 31. Deagle BE, Jarman SN, Coissac E, Pompanon F, Taberlet P. DNA metabarcoding and the
443 cytochrome c oxidase subunit I marker: not a perfect match. Biol Lett. 2014;10(9):20140562. Epub
444 2014/09/12. doi: 10.1098/rsbl.2014.0562. PubMed PMID: 25209199; PubMed Central PMCID:
445 PMCPMC4190964.

446 32. Creer S, Fonseca VG, Porazinska DL, Giblin-Davis RM, Sung W, Power DM, et al.
447 Ultrasequencing of the meiofaunal biosphere: Practice, pitfalls and promises. Mol Ecol.
448 2010;19(Supplementary 1):4–20. doi: 10.1111/j.1365-294X.2009.04473.x.

449 33. Fonseca VG, Carvalho GR, Nichols B, Quince C, Johnson HF, Neill SP, et al. Metagenetic
450 analysis of patterns of distribution and diversity of marine meiobenthic eukaryotes. Global Ecol
451 Biogeogr. 2014;23:1293–302. doi: 10.1111/geb.12223.

452 34. Bik HM, Sung W, De Ley P, Baldwin JG, Sharma J, Rocha-Olivares A, et al. Metagenetic
453 community analysis of microbial eukaryotes illuminates biogeographic patterns in deep-sea and
454 shallow water sediments. Mol Ecol. 2012;21(5):1048-59. doi: 10.1111/j.1365-294X.2011.05297.x.

455 35. Capra E, Giannico R, Montagna M, Turri F, Cremonesi P, Strozzi F, et al. A new primer
456 set for DNA metabarcoding of soil Metazoa. Eur J Soil Biol. 2016;77:53-9. doi:
457 10.1016/j.ejsobi.2016.10.005.

458 36. Darby BJ, Todd TC, Herman MA. High-throughput amplicon sequencing of rRNA genes
459 requires a copy number correction to accurately reflect the effects of management practices on soil
460 nematode community structure. Mol Ecol. 2013;22(21):5456-71. Epub 2013/10/10. doi:
461 10.1111/mec.12480. PubMed PMID: 24103081.

462 37. Sapkota R, Nicolaisen M. High-throughput sequencing of nematode communities from
463 total soil DNA extractions. BMC Ecology. 2015. doi: 10.1186/s12898-014-0034-4.

464 38. Porazinska DL, Giblin-Davis RM, Faller L, Farmerie W, Kanzaki N, Morris K, et al.
465 Evaluating high-throughput sequencing as a method for metagenomic analysis of nematode diversity.
466 Mol Ecol Resour. 2009;9(6):1439-50. doi: 10.1111/j.1755-0998.2009.02611.x.

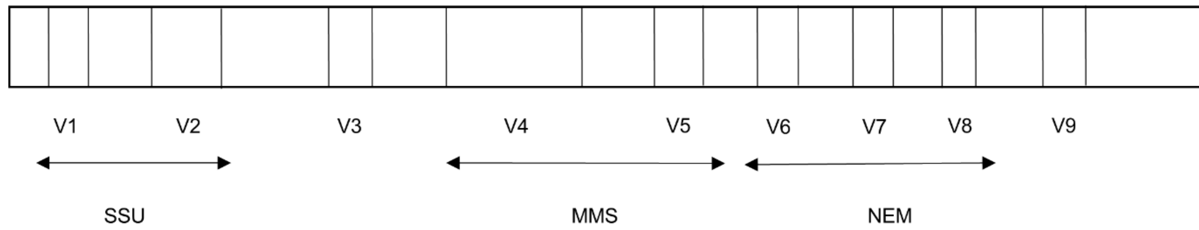
467

468

469

470

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921304; this version posted January 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

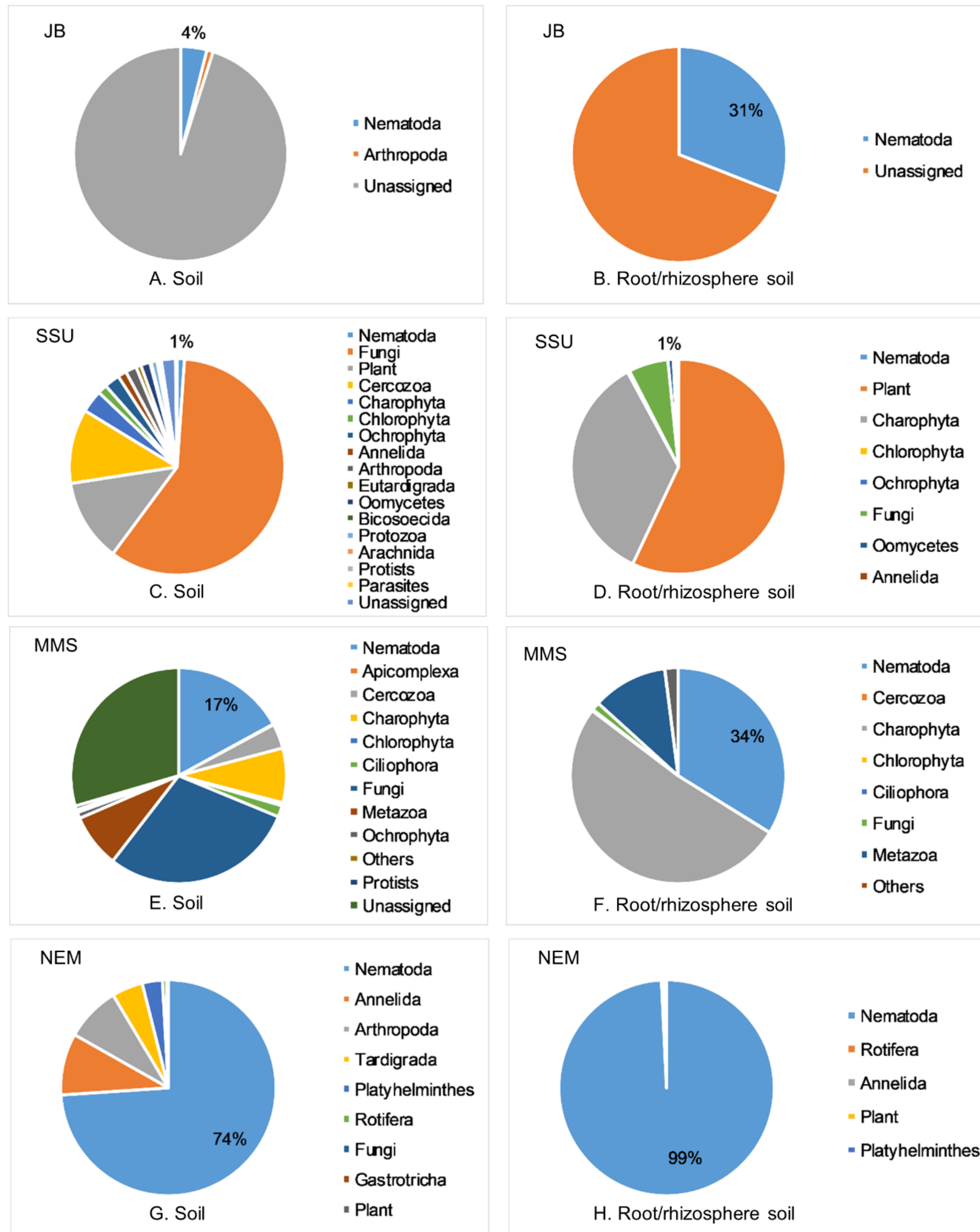


471

472 **Fig 1. Location of metabarcoding primers targeting variable regions in 18S rRNA gene**

473 **used in the present study.**

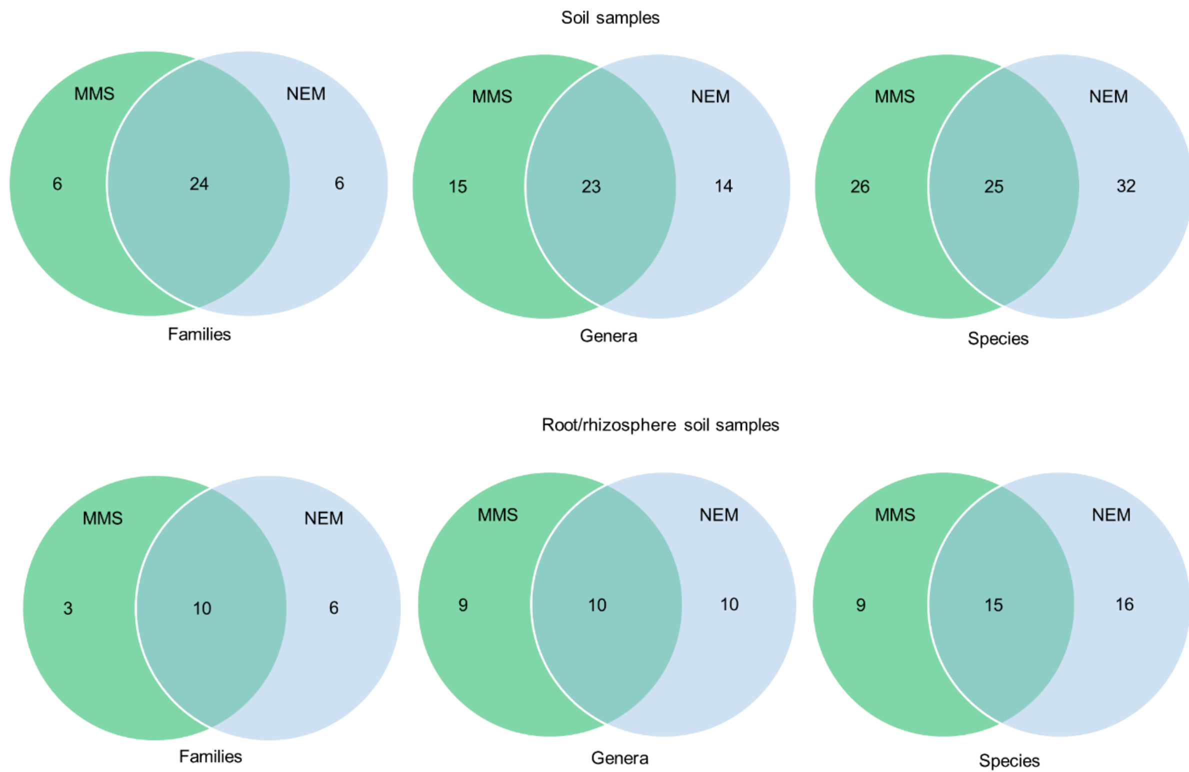
bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921304; this version posted January 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.



474

475 **Fig 2. Relative distribution of sequence reads in soil and plant root/rhizosphere soil**
476 **samples amplified with primer sets JB, SSU, MMS, and NEM**; percentage in the blue slide
477 indicates the proportion of sequence reads that were assigned to Nematoda.

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921304; this version posted January 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

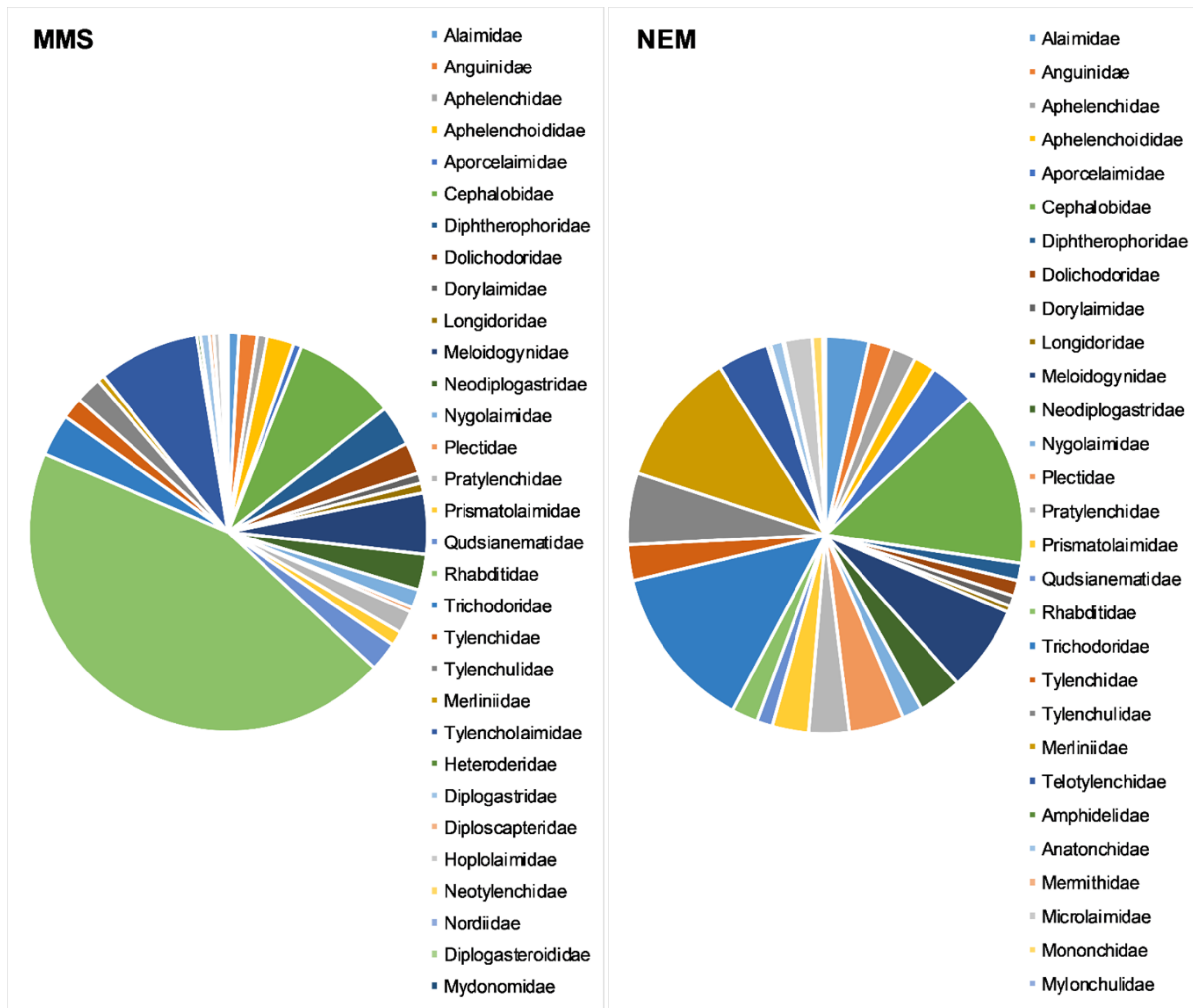


478

479 **Fig 3. Venn diagrams showing the number of taxa detected in soil samples by the primer**
480 **sets MMS and NEM.** Only taxonomic assignments appearing in top hits of BLAST searches
481 and with sequence similarities ≥ 98% and 100% coverage were considered.

482

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921304; this version posted January 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

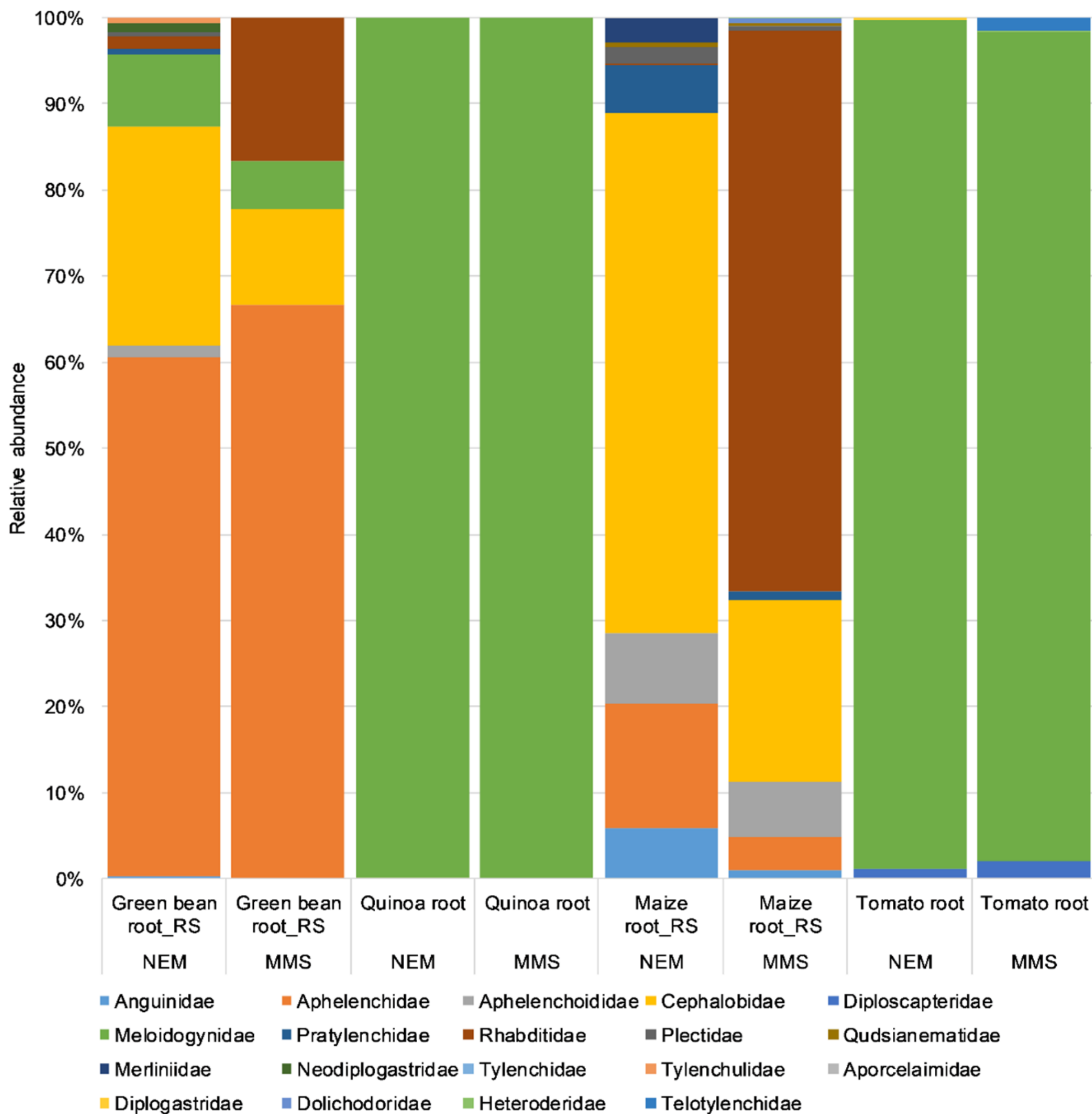


483

484 **Fig 4. Relative distribution of nematode sequence reads in all soil samples amplified**
485 **and sequenced with primer sets MMS and NEM.** Only taxonomic assignments appearing
486 in top hits of BLAST searches and with sequence similarities ≥ 98% and 100% coverage were
487 considered.

488

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921304; this version posted January 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.



489

490 **Fig 5. Relative distribution of sequence reads at family rank in plant root/rhizosphere**
491 **soil samples amplified with primer sets MMS and NEM.** Only taxonomic assignments
492 appearing in top hits of BLAST searches and with sequence similarities ≥ 98% and 100%
493 coverage were considered.

494

495

496

497

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921304; this version posted January 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

498 # Supporting information

499

500 **S1 Table. Metabarcoding primer sets used in the present study**

Primer code	Nucleotide sequence (5′-3′)	Amplicon length (bp)	Reference
Nemf	GGGGAAGTATGGTTGCAAA	500	[37]
18Sr2b	TACAAAGGGCAGGGACGTAAT		[38]
MMSF	GGTGCCAGCAGCCGCGGTA	550	This study
MMSR	CTTTAAGTTTCAGCTTTGC		This study
SSU_04F	GCTTGTCTCAAAGATTAAGCC	360	[26]
SSU_22R	GCCTGCTGCCTTCCTTGGA		
JB3	TTTTTTGGGCATCCTGAGGTTTAT	400	[25]
JB5	AGCACCTAAACTTAAAACATAATGAAAATG		

501
502
503
504
505
506
507
508
509
510
511
512
513

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921304; this version posted January 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

514 **S2 Table. List of individual nematode species, DNA extraction method and order of the**
515 **species used in the study**

	Nematode taxa	DNA extraction method	Order
1.	*Meloidogyne incognita*	Qiasafe	Tylenchida
2.	*Meloidogyne arenaria*	Qiasafe	Tylenchida
3.	*Meloidogyne nassi*	DNeasy	Tylenchida
4.	*Meloidogyne minor*	Qiasafe	Tylenchida
5.	*Meloidogyne javanica*	Worm lysis buffer	Tylenchida
6.	*Meloidogyne graminicola*	Worm lysis buffer	Tylenchida
7.	*Meloidogyne fallax*	Qiasafe	Tylenchida
8.	*Meloidogyne chitwoodi*	DNeasy	Tylenchida
9.	*Meloidogyne ulmi*	DNeasy	Tylenchida
10.	Mixed *Meloidogyne* spp.	Worm lysis buffer	Tylenchida
11.	*Meloidogyne enterolobi*	Worm lysis buffer	Tylenchida
12.	*Meloidogyne inornata*	Worm lysis buffer	Tylenchida
13.	*Meloidogyne ethiopica*	Worm lysis buffer	Tylenchida
14.	*Meloidogyne luci*	Worm lysis buffer	Tylenchida
15.	*Belonolaimus longicaudatus*	Qiasafe	Tylenchida
16.	*Pratylenchus penetrans*	DNeasy	Tylenchida
17.	*Pratylenchus neglectus*	Worm lysis buffer	Tylenchida
18.	*Heterodera schachtii*	Qiasafe	Tylenchida
19.	*Heterodera carotae*	Qiasafe	Tylenchida
20.	*Caenorhabditis elegans*	DNeasy	Rhabditida
21.	*Ditylenchus dipsaci*	Worm lysis buffer	Tylenchida
22.	*Bursaphelenchus mucronatus*	Worm lysis buffer	Aphelenchida

516
517
518
519
520

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921304; this version posted January 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

521 **S3 Table. Composition of mock communities used in the study**

	Nematode taxa	Nematode taxa	Nematode taxa	Nematode taxa	Nematode taxa	Nematode taxa
Mock1	*Meloidogyne hapla*	*Globodera* spp.	*Belonolaimus longicaudatus*	*Pratylenchus penetrans*	*Heterodera carotae*	*Caenorhabditis elegans*
Mock2	*Meloidogyne incognita*	*Globodera* spp.	*B. longicaudatus*	*P. penetrans*	*Heterodera schachtii*	*C. elegans*
Mock3	*Meloidogyne arenaria*	*Globodera* spp.	*B. longicaudatus*	*P. penetrans*	*H. carotae*	*C. elegans*
Mock4	*Meloidogyne minor*	*Globodera* spp	*B. longicaudatus*	*P. penetrans*	*H.schachtii*	*C. elegans*
Mock5	*Meloidogyne fallax*	*Globodera* spp	*B. longicaudatus*	*P. penetrans*	*H. carotae*	*C. elegans*
Mock6	*Meloidogyne chitwoodi*	*Globodera* spp	*B. longicaudatus*	*P. penetrans*	*H.schachtii*	*C. elegans*
Mock7	*Meloidogyne ulmi+ Meloidogyne nassi*	*Globodera spp*	*B. longicaudatus*	*P. penetrans*	*H. carotae*	*C. elegans*
Mock8	*Meloidogyne inornata* +Mixed *Meloidogyne* spp.	*Globodera* spp	*B. longicaudatus*	*P. penetrans*	*H.schachtii*	*C. elegans*
Mock9	*Meloidogyne ethiopica* +Mixed *Meloidogyne* spp.	*Globodera spp*	*B. longicaudatus*	*P. penetrans*	*H. carotae*	*C. elegans*
Mock10	*Meloidogyne luci+ Meloidogyne inornata*	*Globodera* spp	*B. longicaudatus*	*P. penetrans*	*H.schachtii*	*C. elegans*
Mock11	*Meloidogyne hapla*	*Globodera* spp	*B. longicaudatus*	*P. penetrans*	*H. carotae*	*C. elegans*
Mock12	*Meloidogyne incognita*	*Globodera* spp	*B. longicaudatus*	*P. penetrans*	*H.schachtii*	*C. elegans*
Mock13	*Meloidogyne arenaria*	*Globodera* spp	*B. longicaudatus*	*P. penetrans*	*H. carotae*	*C. elegans*
Mock14	*Meloidogyne minor*	*Globodera* spp	*B. longicaudatus*	*P. penetrans*	*H.schachtii*	*C. elegans*
Mock15	*Meloidogyne fallax*	*Globodera* spp	*B. longicaudatus*	*P. penetrans*	*H. carotae*	*C. elegans*
Mock16	*Meloidogyne chitwoodi*	*Globodera* spp	*B. longicaudatus*	*P. penetrans*	*H. schachtii*	*C. elegans*
Mock17	*Mixed Meloidogyne spp. + Meloidogyne nassi+ Meloidogyne inornata*	*Globodera* spp	*B. longicaudatus*	*P. penetrans*	*H. carotae*	*C. elegans*
Mock18	*Meloidogyne inornata*	*Globodera* spp	*B. longicaudatus*	*P. penetrans*	*H. schachtii*	*C. elegans*
Mock19	*Meloidogyne ethiopica*	*Globodera spp*	*B. longicaudatus*	*P. penetrans*	*H. carotae*	*C. elegans*
Mock20	*Meloidogyne luci*	*Globodera* spp	*B. longicaudatus*	*P. penetrans*	*H. schachtii*	*C. elegans*

522 Blue font indicates 1:10 dilution of template DNA

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921304; this version posted January 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

523 **S4 Table. Cropping history and soil properties of twenty different soils used in the study**

Soil ID	Previous crops	Year of sampling	Status	Soil type	pH
Soil-1	Carrot	2012	NA	NA	NA
Soil-2	Strawberry	2012	NA	NA	NA
Soil-3	Spinach	2014	NA	NA	NA
Soil-4	Maize	2018	Conventional	*clayey sand*	6.0
Soil-5	Beans	2018	Conventional	*clayey sand*	5.9
Soil-6	Triticale	2013	Conventional	*heavy clay*	5.8
Soil-7	Corn	2014	Conventional	*heavy clay*	6.3
Soil-8	Iceberg lettuce	2018	Conventional	*clayey sand*	6.3
Soil-9	Rye	2013	Conventional	*heavy clay*	5.9
Soil-10	Rye	2014	Organic	*clay*	6.0
Soil-11	Barley	2013	Organic	*clayey sand*	5.2
Soil-12	Barley	2014	Conventional	*heavy clay*	6.5
Soil-13	Wheat	2013	Conventional	*heavy clay*	6.5
Soil-14	Wheat	2014	Conventional	*coarse sand*	6.1
Soil-15	Potato	2013	Conventional	*heavy clay*	5.4
Soil-16	Potato	2014	Conventional	*heavy clay*	5.5
Soil-17	Clover	2013	Organic	*heavy clay*	5.7
Soil-18	Clover	2014	Organic	*coarse sand*	6.1
Soil-19	Oat	2013	Organic	*heavy clay*	5.9
Soil-20	Oat	2014	Organic	*heavy clay*	6.0

524 NA indicates not analysed

525

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921304; this version posted January 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

526 **S5 Table. The efficiency of four metabarcoding primers in the detection of nematodes**
527 **from different families in twenty different soil mock communities.**

Families	JB	SSU	MMS	NEM
Alaimidae	ND	D	D	D
Amphidelidae	ND	ND	ND	D
Anatonchidae	ND	ND	ND	D
Anguinidae	ND	ND	D	D
Aphelenchidae	ND	ND	D	D
Aphelenchoididae	ND	ND	D	D
Aporcelaimidae	ND	ND	D	D
Axonolaimidae	ND	D	ND	ND
Bastianiidae	ND	D	ND	ND
Cephalobidae	ND	D	D	D
Diplopeltidae	ND	D	ND	ND
Diphterophoridae	ND	D	D	D
Diplogasteridae	ND	ND	ND	D
***Dolichodoridae**	ND	ND	D	D
Dorylaimidae	ND	ND	D	D
***Heteroderidae**	D	D	D	D
Longidoridae	ND	ND	ND	D
***Meloidogynidae**	D	D	D	D
Merliniidae	ND	ND	ND	D
Mermithidae	ND	ND	ND	D
Microlaimidae	ND	ND	ND	D

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921304; this version posted January 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Monhysteridae	ND	D	ND	ND
Mononchidae	ND	ND	ND	D
Mydonomidae	ND	ND	D	ND
Mylonchulidae	ND	ND	ND	D
Neodiplogastridae	ND	ND	D	D
Nygolaimidae	ND	ND	D	D
Plectidae	ND	D	ND	D
***Pratylenchidae**	D	ND	D	D
Prismatolaimidae	ND	D	D	D
Qudsianematidae	ND	ND	D	D
***Rhabditidae**	D	D	D	D
Telotylenchidae	ND	ND	D	D
Trichodoridae	ND	D	ND	D
Tylenchidae	ND	ND	D	D
Tylenchulidae	ND	ND	D	D
Unassigned	Above 70%	-	Negligible	Negligible

Here, D denotes amplified and detected, ND denotes not detected, * denotes the families were abundant in soil-mock samples.

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921304; this version posted January 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

542 **S6 Table. Efficiency of four metabarcoding primers in detection of nematodes from**
543 **twenty different soils at lower taxonomic rank than family level.**

Taxa	JB	SSU	MMS	NEM
Achromadoridae	ND	D	ND	ND
***Achromadora* sp.**		G		
Achromadora ruricola		S		
Alaimidae	ND	D	D	D
***Alaimus* sp.**		G	G	G
Alaimus arcuatus				S
Alaimus parvus				S
Anatonchidae	ND	ND	ND	D
Anatonchus tridentatus				S
Anguinidae	ND	ND	D	D
***Ditylenchus* sp.**			G	G
Ditylenchus dipsaci				S
Ditylenchus destructor				S
Aphelenchidae	ND	ND	D	D
Aphelenchus avenae			S	S
Aphelenchoididae	ND	ND	D	D
***Aphelenchoides* sp.**			G	
Aphelenchoides parietinus			S	S
Aphelenchoides bicaudatus			S	S
Seinura demani				S
Aporcelaimidae	ND	ND	D	D

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921304; this version posted January 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

***Aporcelaimellus* sp.**			G	G
Bastianiidae	ND	D	ND	ND
Bastiania gracilis		S		
Cephalobidae	ND	D	D	D
Acrobeles complexus		S	S	S
Acrobeles ctenocephalus				S
Acrobeles ciliatus		S		S
***Acrobeloides* sp.**		G	G	G
Acrobeloides thornei		S		
Acrobeloides apiculatus		S		
Acrobeloides varius				S
Chiloplacus propinquus		S	S	
Eucephalobus sp.				G
Eucephalobus striatus		S	S	S
Eucephalobus oxyuroides		S	S	
Heterocephalobus elongatus			S	
***Pseudacrobeles* sp.**				G
Diplopeltidae	ND	D	ND	ND
Cylindrolaimus communis		S		
Diphterophoridae	ND	D	D	D
***Diphtherophora* sp.**		G	G	G
Diphterophora communis		S	S	
Tylolaimophorus typicus		S		

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921304; this version posted January 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Diplogasteridae	ND	D	ND	ND
Butlerius butleri		S		
Diplogasteroididae	ND	ND	D	ND
Diplogasteroides sp.			G	
Diploscapteridae	ND	ND	D	ND
Diploscapter sp.			G	
Dolichodoridae	ND	ND	D	D
Merlinius sp				G
Merlinius nanus			S	
Dorylaimidae	ND	ND	D	D
Mesodorylaimus bastiani			S	
Prodorylaimus sp.				G
Thonus circulifer			S	S
Heteroderidae	D	ND	D	ND
Globodera pallida	S		S	
Globodera rostochiensis	S			
Hoplolaimidae	ND	ND	D	ND
Helicotylenchus minzi			S	
Longidoridae	ND	ND	D	D
Longidorus sp.			G	
Longidorus attenuatus				S
Meloidogynidae	D	D	D	D
***Meloidogyne* sp.**	G	G	G	G

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921304; this version posted January 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Meloidogyne hapla			S	S
Merliniidae	ND	ND	*D	ND
Microlaimidae	ND	ND	ND	D
Prodesmodora circulata				S
Mylonchulidae	ND	ND	ND	D
Mylonchulus hawaiiensis				S
Mononchidae	ND	ND	ND	D
Clarkus papillatus				S
Monhysteridae	ND	D	ND	ND
***Eumonhystera* sp.**		G		
Eumonhystera vulgaris		S		
Eumonhystera hungarica		S		
***Geomonhystera* sp.**		G		
Mydonomidae	ND	ND	D	ND
Dorylaimoides micoletzkyi			S	
Neodiplogastridae	ND	ND	D	D
Mononchoides americanus			S	
***Pristionchus* sp.**			G	G
Pristionchus lheritieri				S
Neotylenchidae	ND	ND	D	ND
***Rubzovinema* sp.**			G	
Nordiidae	ND	ND	D	ND
Pungentus sp.			G	

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921304; this version posted January 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Nygolaimidae	ND	ND	D	D
Nygolaimus brachyuris			S	S
Plectidae	ND	D	ND	D
***Plectus* sp.**				G
Plectus minimus		S		
Plectus aquatilis		S		
Anaplectus porosus		S		S
Tylocephalus auriculatus		S		
Wilsonema otophorum		S		
Pratylenchidae	D	ND	D	D
Pratylenchus thornei			S	S
Pratylenchus crenatus			S	S
Pratylenchus penetrans			S	S
Pratylenchus neglectus			S	S
Pratylenchus fallax	S			
Prismatolaimidae	ND	D	D	D
***Prismatolaimus* sp.**				G
Prismatolaimus dolichurus		S	S	
Prismatolaimus intermedius		S	S	
Qudsianematidae	ND	ND	D	D
Microdorylaimus miser				S
Ecumenicus monohystera			S	S
Rhabditidae	D	ND	D	D

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921304; this version posted January 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

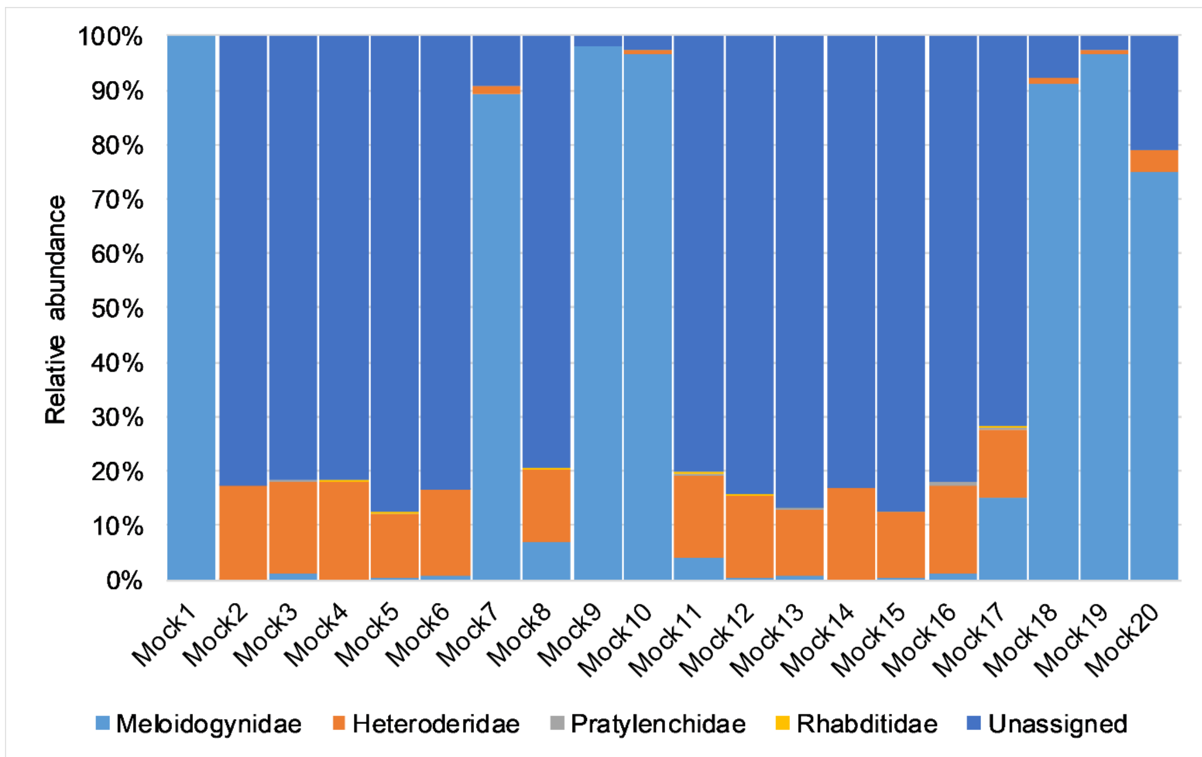
***Rhabditis* sp.**			G	G
***Pelodera* sp.**			G	
Pelodera teres				S
***Pellioditis* sp.**				G
Pellioditis marina	S			
***Mesorhabditis* sp.**			G	
Mesorhabditis belari			S	
Caenorhabditis elegans	S			
Steinernematidae	ND	D	ND	ND
Steinernema affine		S		
Telotylenchidae	ND	D	D	D
Tylenchorhynchus maximus			S	S
Tylenchorhynchus dubius			S	S
Tylenchorhynchus teeni		S		
Trichodoridae	ND	D	D	D
Trichodorus primitivus		S	S	S
Trichodorus viruliferus		S		
Paratrichodorus pachydermus			S	S
Paratrichodorus allius		S		
Tylenchidae	ND	ND	D	D
***Filenchus* sp.**				G
Filenchus aquilonius				S
Coslenchus turkeyensis			S	

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921304; this version posted January 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Basiria **sp.**				G
Basiria duplexa			S	
Tylenchulidae	ND	ND	D	D
Paratylenchus sp.				G
Paratylenchus conicephalus				S
Paratylenchus similis				S
Paratylenchus nanus			S	
Paratylenchus projectus			S	
Tylencholaimidae	ND	ND	D	ND
Tylencholaimus sp.			G	

544 Here, G: genus level detection, S: species level detection, D: detected at lower taxonomic
545 level, *D: detected at family level only, ND: not amplified and detected at any taxonomic level.
546 Classification was based on Nemaplex database (http://nemaplex.ucdavis.edu/ Accessed on
547 11/11/2019); NCBI Blast tool was used for taxonomic assignments, and top hits with sequence
548 similarities ≥ 99% and coverage 100% were considered for taxonomic assignment at species
549 level.
550
551

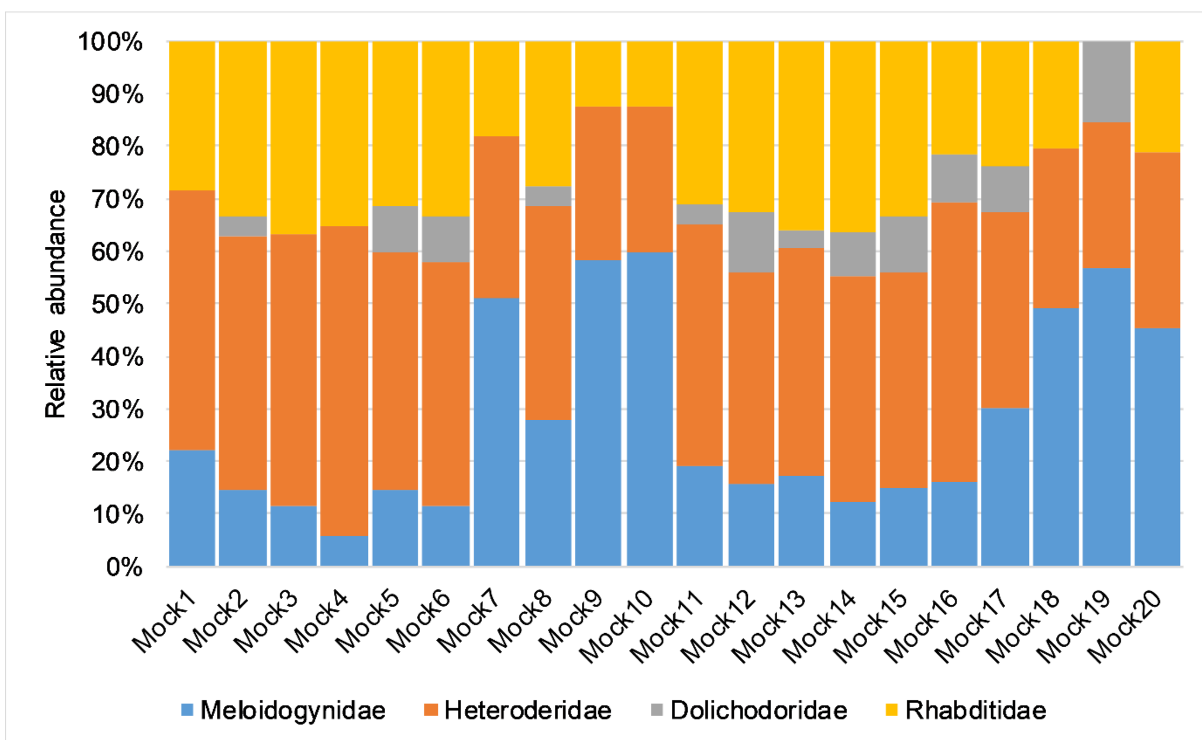
bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921304; this version posted January 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.



552

553 **S1 Fig. Relative abundance of sequence reads at family rank in mock samples amplified**
554 **and sequenced using JB primer set.** Only taxonomic assignments appearing in top hits and
555 with sequence similarities ≥ 99% and coverage 100% were considered.

556

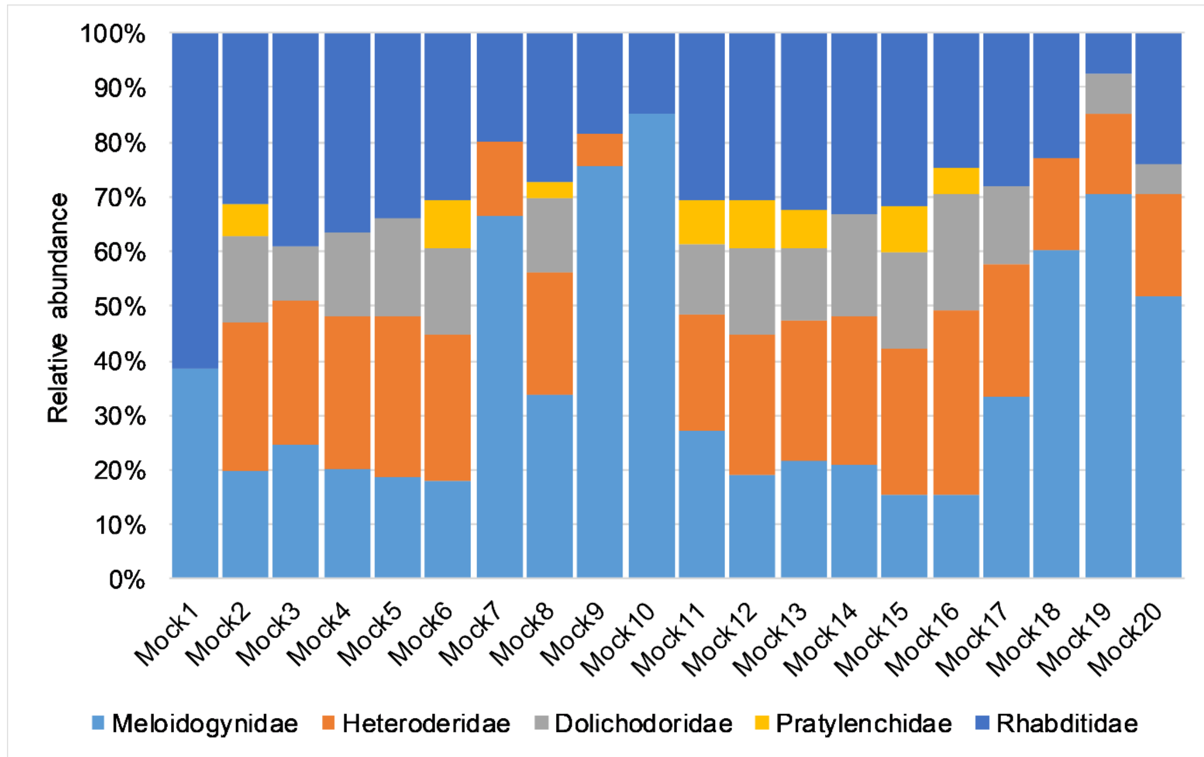


557

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921304; this version posted January 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

558 **S2 Fig. Relative abundance of sequence reads at family rank in mock samples amplified**
559 **and sequenced using MMS primer set.** Only taxonomic assignments appearing in top hits
560 and with sequence similarities ≥ 99% and coverage 100% were considered.

561



562

563 **S3 Fig. Relative abundance of sequence reads at family rank in mock samples amplified**
564 **and sequenced using NEM primer set.** Only taxonomic assignments appearing in top hits
565 and with sequence similarities ≥ 99% and coverage 100% were considered.

566

567

568

569

570

571

572

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921304; this version posted January 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

573 **S4 Fig. Multiple sequence alignment of MMS and NEM primer sets and representative**
574 **taxa of Rhabditidae**

575

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.27.921304; this version posted January 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

576

577

578