

¹ **The effect of *Ascophyllum nodosum* extracts on tomato
2 and pepper plant productivity and their associated
3 fungal and bacterial communities.**

⁴ **Sébastien Renaut^{1,2}, Jacynthe Masse^{1,2}, Mengxuan Kong^{1,2}, Mohamed Hijri^{1,2}** ¹Département
⁵ de Sciences Biologiques, Institut de Recherche en Biologie Végétale, Université de Montréal, 4101 Sherbrooke Est,
⁶ Montreal, H1X 2B2, Quebec, Canada. ²Quebec Centre for Biodiversity Science, Montreal, Quebec, Canada

⁷ Abstract to be written here

⁸ *Keywords:* Stella Marris, 16S, ITS, microbial diversity, Illumina MiSeq

9 INTRODUCTION

10 Liquid extracts of marine macroalga are used as biostimulants in agriculture. These extracts con-
11 tain phytohormones that can influence physiological processes even at very low concentrations
12 Craigie (2011). Stella Maris® is derived from fresh *Ascophyllum nodosum* algae harvested from the
13 nutrient-laden waters of the North Atlantic off the Eastern Coast of Canada.

14

15 The aim of this project was to develop a better understanding of the effects of *A. nodosum* extracts
16 on plant growth. We tested the effect of these extract on two commonly used plants (Tomato -
17 *Solanum lycopersicum* and Pepper - *Capsicum annuum*) using different measures of productivity. In
18 addition, we tested how the bacterial and fungal communities responded to the addition of *A.
19 nodosum* extracts.

20

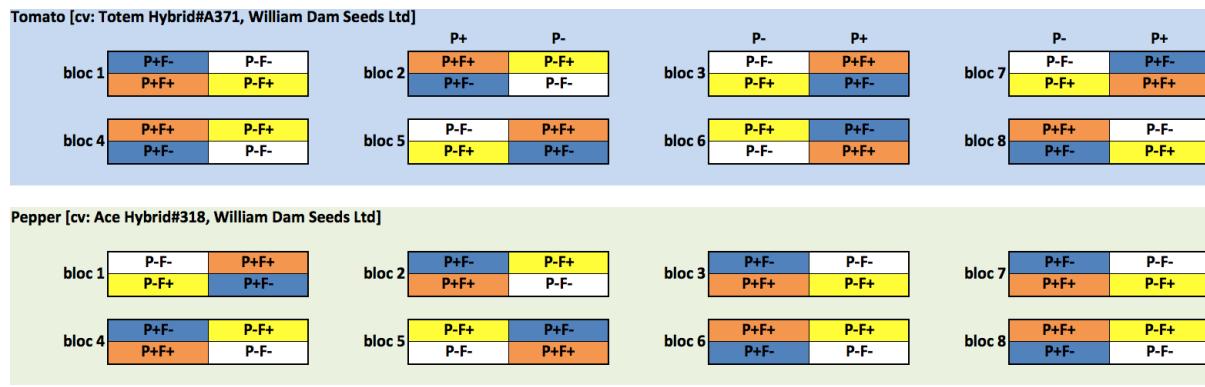
21 MATERIAL AND METHOD

22 Study design

23 Two greenhouse experiment were set up in large trays (60x30x18 cm) in November (tomato [cv:
24 Totem Hybrid#A371, William Dam Seeds Ltd]) and December (Pepper [cv: Ace Hybrid#318,
25 William Dam Seeds Ltd]) 2015. Soil was collected from an agricultural field under organic regime
26 at the IRDA research station in St-Bruno (Qc, Canada) on October 7th 2015 (loamy sand soil, 15 cm
27 top layer collected). Soil characteristics (pH, conductivity, nutrients, see Table 1) were measured
28 by AgriDirect (Longueuil, Qc, Canada).

29

30 For each species tested (Tomato - *Solanum lycopersicum*, Pepper - *Capsicum annuum*), a randomized
31 split block design (Figure 1) was used with four trays set up per block (eight blocks). Half of
32 the trays were fertilized (fertilization treatment), as described below. Half of the trays were also
33 planted with four replicate plants each, while the other trays were left bare. This allowed a direct
34 comparison of the fungal and bacteria soil communities with respect to the fertilization and plant-
35 ing treatment.



36

37 **Figure 1: experimental design**

38

39 Half of the tomato plants were fertilized using multipurpose organic fertilizer (pure hen manure, 18 g per tray repeated every 4 weeks, 5-3-2) from Acti-sol (Notre-Dame-du-Bon-Conseil, 40 Qc, Canada) in addition to Stella Maris® (3.5 ml per 1L, each tray received 250 ml, repeated every 41 2 weeks) for the duration of the experiment. The other half were unfertilized. Stella Maris® is a 42 registered trademark from Acadian Seaplants Ltd. (Darmouth, NS, Canada). It is primarily com- 43 posed of *Ascophyllum nodosum* seaweed and is advertised as a natural activator of the crops' own 44 growth and defense mechanisms to improve root growth and resist temperature, drought, and 45 salinity stress in order to maximize yield and crop qualities (Acadian Seaplants Ltd. 2018). Half 46 of the pepper plants were treated using solely Stella Maris (3.5 ml per 1L, each tray received 250 47 ml, repeated every 2 weeks) for the duration of the experiment. The other half were untreated. 48

49 Thrips were managed with *Neoseiulus cucumeris* (syn. *Amblyseius cucumeris*) (100 bags), Fungus 50 gnat and thrips were also controlled using predatory mite *Gaeolaelaps gillespiei* (1L). Plants were 51 treated once a week with Oidium Milstop to control the fungus.

52

53

54 *Plant productivity*

55 At the end of the experiment, plant productivity was assessed by measuring four different traits 56 (fruit number, average fruit weight, shoots fresh weight, roots fresh weight) on three plants chosen 57 randomly per tray (for each treatment [fertilization/control], species [tomato/pepper] and block 58 [eight blocks]) for a total of 96 samples. In addition, both shoots and roots were dried in a 70 59 degrees drying oven, and dry weights were measured after 48 hours. Together, these traits are

60 expected to represent well the plant overall productivity.

61

62

63 *Sample preparation, DNA extraction and High throughput sequencing*

64 We sampled both the microbial and fungal communities from soil and root samples. Soil DNA
65 was extracted using XXX DNA isolation kit with YYY g of soil. Roots were first washed with
66 sterile water and DNA was extracted using XXX DNA isolation kit with YYY g of root samples.
67 Amplicon sequencing targeting 16S rRNA gene (bacteria) and ITS (fungi) was performed on both
68 root and soil samples.

69

70 In order to target fungi, we used fungal primers ITS3_KYO2 (5'-ACACTGACGA CATGGTTCT
71 ACAGATGAAGAAC GYAGYRAA-3') and ITS4_KYO3 (5'-TACGGT AGCAGAGACTT GGTCTCTBTTV
72 CCKCTTCACTCG-3') to produce a final amplicon size of ~430bp. This primer pair should target
73 the Internal transcribed spacer and inhibit the amplification of plant sequences and enable the
74 selective amplification of fungal communities from soil, mycorrhizal and other environmental
75 samples (Toju *et al.* 2012).

76

77 Bacterial primers 341F (5'-CCTACGGG NGGCWGCAG-3') and 805R (5'-GACTACC AGGGTATC
78 TAATC-3') producing a final amplicon size of ~464b and targeting specifically the bacterial V3-V4
79 region of the 16S ribosomal gene were chosen given that they has been used extensively in high-
80 throughput sequencing studies in a range of environments Toju et al. (2012). This primer pair was
81 shown to be the least biased among 512 primer pairs evaluated in silico for bacterial amplification
82 Klindworth *et al.* (2013).

83

84 DNA samples were then barcoded, pooled and sequenced (2X300bp, paired-end) using an Illu-
85 mina MiSeq (San Diego, CA, USA) sequencer at the Genome Quebec Innovation Centre (Montreal,
86 Canada). Sequences were demultiplexed by the sequencing facility (Genome Quebec Innovation
87 Centre) and further processed as described below.

88

89

90 *Bioinformatics*

91 All bioinformatics, statistical, and graphical analyses further described were performed in R 3.4.4
92 (R Core Team 2018) and detailed scripts are available here (https://github.com/seb951/Acadian_SeaPlants).
93

94

95 We used the R package dada2 Callahan et al. (2016) to infer *Amplicon Sequence Variants* (ASVs).
96 Dada2 offers accurate sample inference from amplicon data with single-nucleotide resolution in
97 an open source (R) environments. Unlike the Operational Taxonomic Unit (OTU) approach (e.g.
98 Schloss et al. (2009), Caporaso et al. (2010)), ASV are not treated as cluster of sequences defined
99 with an *ad hoc* sequence similarity threshold. Instead, after sequences are quality trimmed and
100 error-corrected, dada2 reveals the unique members of the sequenced community, thus allowing
101 sequences and abundance counts to be compared among studies Callahan et al. (2016).

102

103 First, sequences were trimmed following strict quality thresholds (removing primers and low
104 quality nucleotides, see parameter details in the accompanying R scripts). Following this, we
105 applied the error model algorithm of dada2 which incorporates quality information after filtering,
106 unlike other OTU based methods. Then dereplication, sample inference, merging of paired end
107 reads and removal of chimera reads were performed in order to obtain a sequence (ASVs) table
108 of abundance per sample. Taxonomy was also assigned using the Ribosomal Database Project
109 (RDP) Naive Bayesian Classifier algorithm from Wang et al. (2007). Depending on support (min-
110 imum bootstrap support of 80), we assigned taxonomy from Kingdom to species. We used the
111 silva database formatted for dada2 to infer bacterial taxa Callahan (2018). We used the Com-
112 munity (2018) fasta release (including singletons) to infer fungal taxa after formatting it to the
113 dada2 format using a custom R script. The pipeline was run on a multithreaded (48 CPUs) com-
114 puter infrastructure provided by Westgrid (<https://www.westgrid.ca/support/systems/cedar>)
115 and Compute Canada (www.computecanada.ca). Note that the pipeline was run separately for
116 fungal-root, fungal-soil, bacteria-soil and bacteria-root samples given the markedly different type
117 of amplicons, taxa and error models of each dataset.

118

119 *Statistical analyses - plant productivity*

120 We tested for the effect of species (tomato vs pepper), fertilization and their interaction on six plant
121 productivity measures (fruit number, average fruit weight, shoots fresh weight, roots fresh weight,
122 shoots dry weight, roots dry weight). We used linear mixed effect models (LMM) in the R package
123 `nlme` Pinheiro et al. (2017), which are more appropriate than an Analysis of Variance (ANOVA)
124 given the current block design (blocks and replicates nested within a block were treated as random
125 variables). All six plant productivity measures were either square root or log transformed in or-
126 der to help satisfy the assumption of normality of the residuals in the LMM statistical framework.
127 For the variables *fruit number* and *average fruit weight*, we also used a permutation-based 2-way
128 ANOVA (Anderson & Legendre (1999)) given that the residuals of the LMM were not normally
129 distributed, and results were similar.

130

131

132 *Statistical analyses - microbial and fungal diversity*

133 We analysed separately fungal-root, fungal-soil, bacterial-root and bacterial-soil ASV diversity.
134 For each of these four datasets, we removed samples that showed poor sequencing output and
135 contained few ASVs. In order to do this, we summed the abundance of all ASVs for each sam-
136 ple ($\sum_{i=1}^n ASV$) and eliminated samples that had fewer than the mean sum ($\overline{\sum_{i=1}^n ASV}$) - 4σ (four
137 standard deviations). In addition, we removed ASVs from our dataset that were present in fewer
138 than 5% of the samples (less than ten individuals in the soil samples, and less than five in the root
139 samples). This was done to remove very rare ASVs which were unique to a block or replicate, but
140 not found in the majority of a treatment.

141

142 We then conducted community-based analyses looking at the effect of the fertilization treatment
143 on the abundance ASV taxa in the tomato and pepper experiments. To reduce the complexity of
144 the datasets, relative abundance of all taxa were calculated per family using the R package `dplyr`
145 Wickham et al. (2015). Barplots were drawn using `ggplot2` Wickham (2016) to vizualize com-
146 munities. ASV (*a*)-diversity was calculated for each sample using the inverse Simpson diversity
147 index in `vegan` Oksanen et al. (2013). The effect of fertilization treatment, species (and planting
148 for soil communities) were assessed using a linear mixed-effect (LMM) model in the R package
149 `nlme` Pinheiro et al. (2017), given the unbalanced, replicated block design. Alpha diversity was

150 log transformed in order to help satisfy the assumption of normality of the residuals of the LMM
151 statistical framework.

152

153 Using the community matrix data of ASVs abundance, we performed PERmutational Multivariate
154 ANalysis Of VAriance tests (PERMANOVA; Anderson (2001)) to identify relationships between
155 the communities according to the experimental design. ASV abundance data was Hellinger-
156 transformed and significance was assessed using 10,000 permutations in vegan Oksanen et al.
157 (2013). Blocks and replicates nested within blocks were factored as strata (blocks) in the model.

158

159 We also performed canonical correspondence analyses (CCAs) using Hellinger-transformed ASV
160 abundance data in vegan Oksanen et al. (2013) to visually assess (*species* scaling based on ASV ma-
161 trix) the grouping of samples, ASVs and their association with productivity variables. Data were
162 analysed separately for fungal-root, fungal-soil, bacterial-root and bacterial-soil, but also accord-
163 ing to species (tomato/pepper), given that analyses of *a* diversity showed that tomato and pepper
164 were markedly different. This gave a total of eight CCAs. Data were constrained based on four
165 of the productivity measures (fruit number, average fruits weight, shoots fresh weight, roots fresh
166 weight). We excluded the shoot & root dry weights as constraints to simplify the model and given
167 that they were highly correlated with the fresh weighth already included as constraints ($r^2=0.98$
168 and 0.76 for shoot dry/fresh weights and root dry/fresh weights, respectively).

169

170 Finally, we attempted to identify candidate ASVs positively associated with productivity. As such,
171 we identified the ten ASVs most positively associated with the measures of fruit number, shoots
172 fresh weight and roots fresh weight from each canonical correspondence analysis for a total of 40
173 fungal and 40 bacterial candidates ASVs. We aligned candidate sequences from these candidates
174 ASVs using the Bioconductor R package decipher Wright (2016) and build pairwise distances ma-
175 trices using a JC69 substitution models of DNA sequence evolution (equal base frequencies, Jukes
176 & Cantor (1969)) in phangorn Schliep (2010). Phylogenetic trees for bacteria and fungi were plotted
177 using ape Paradis, Claude & Strimmer (2004). This permitted to identify if similar candidate ASVs
178 were found under different experimental conditions (soil/root, pepper/tomato), thus reinforcing
179 their role in productivity increase, and decreasing the chance that these are false positive.

¹⁸¹ **RESULTS**

¹⁸² *Soil characteristics* In Table 1, we present the soil characteristics.

¹⁸³

Table 1: Table 1. Soil characteristics of the soil used for the experiment

Soil Characteristics	Average value
pH	6.01
Conductivity (mmhos/cm)	0.68
Nitrate (ppm N)	62.40
Ammonium (ppm)	0.09
Phosphorus (ppm)	0.41
Potassium (ppm)	29.30
Calcium (ppm)	64.40
Magnesium (ppm)	13.80
Chloride (ppm)	28.50
Sulfate (ppm)	19.30
Sodium (ppm)	17.80
Zinc (ppm)	0.12
Manganese (ppm)	0.06
Cooper (ppm)	0.81
Iron (ppm)	0.90
Aluminium (ppm)	1.66

¹⁸⁴

¹⁸⁵ *productivity*

¹⁸⁶ We tested the effect of the fertilization treatment on six measures of overall plant growth and pro-
¹⁸⁷ ductivity (fruit number, average fruit weight, shoots fresh weight, shoots dry weight, roots fresh
¹⁸⁸ weight, roots dry weight) for both tomato and peppers. Visually, both above ground and below

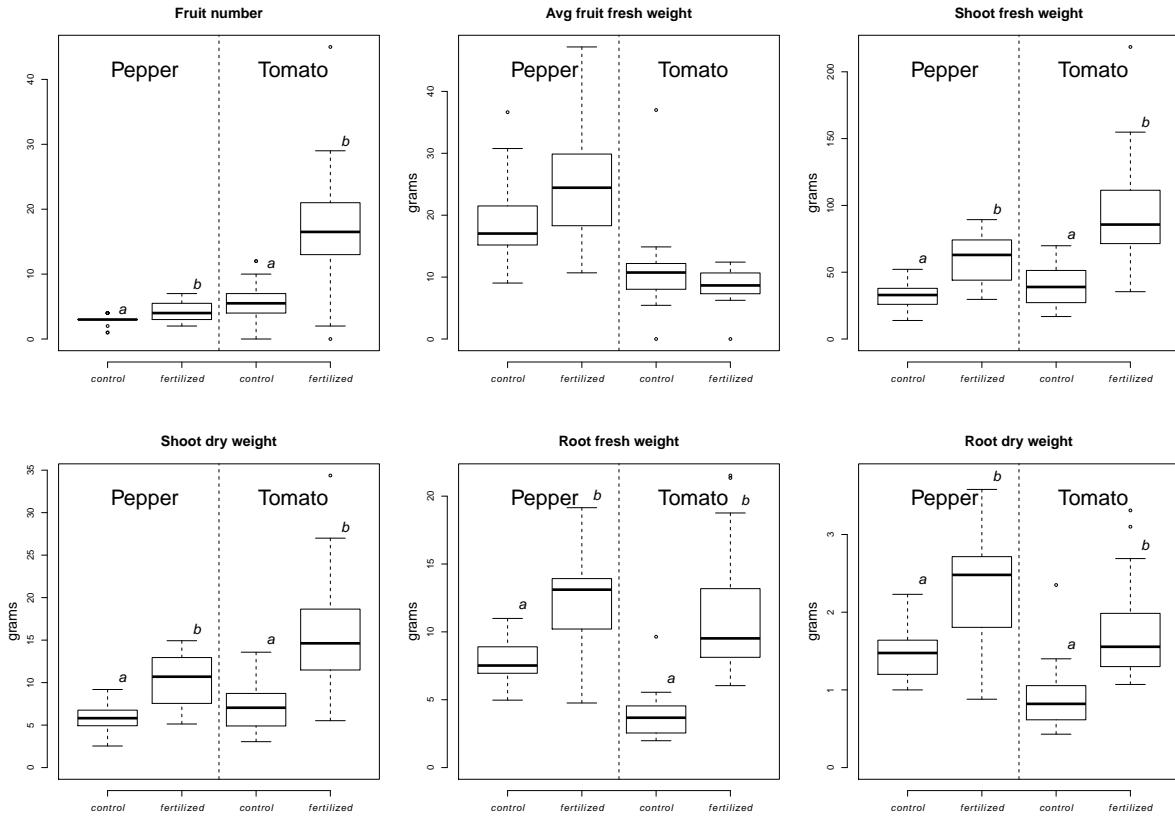
189 ground plant structure grew larger in fertilized plants, in addition to producing more fruits (see
190 Figure 2 for some examples of the striking difference between fertilized and unfertilized plants).



191
192 **Figure 2: Plant productivity.** Photos were taken at the end of the experimental treatment. In
193 each photo, fertilized plants are on the left. **A:** pepper plants, **B:** pepper roots, **C:** pepper fruits
194 and **D:** tomato fruits.

195
196 Statistically, all six productivity measures significantly differed according to species, and five of
197 those were significantly different according to the fertilization treatment. The only exception was
198 the average fruit weight which did not differ between fertilized and control plants (LMM, $F_{(1,69)}$
199 = 1.27, p -value=0.26). However the model did reveal a significant interaction between treatment
200 and plant ($F_{(1,69)} = 9.6$, p -value=0.0028). In fact, when testing only the pepper plants, the effect of
201 fertilization on average fruit weight was significantly higher in the fertilized pepper plants ($F_{(1,23)}$
202 = 10.84, p -value=0.0032).

203



204

205 **Figure 3: measures of plant productivity.**

206

207

208 *Sequencing*

209 A total of 2.7 million paired-end raw reads were obtained for all samples combined (976,000 for
210 fungi-soil, 920,000 for fungi-root, 309,000 for bacteria-soil and 535,000 for bacteria-root, Table 2).

211 Note that sequencing samples were analysed separately for fungal-soil, fungal-root, bacteria-soil
212 and bacteria-root conditions. On average, 46,965 paired-end reads were obtained per sample. Af-
213 ter quality filters were applied, including removing chimeras, and paired-end reads were merged,
214 an average of 18,435 sequences remained. While 192 soil samples for fungi and bacteria, and 92
215 root samples for fungi and bacteria were sequenced, three fungi-soil samples, 13 fungi-root sam-
216 ples and two bacteria-root samples were removed because they had too few reads based on our
217 strict quality thresholds.

218

219 The dada2 pipeline inferred, on average, 112 Amplicon Sequence Variants per sample (average of
220 163 fungal-soil ASV, 49 fungal-root ASVs, 112 bacterial-soil ASVs and 122 bacterial-root ASVs).
221 Many of those were unique to one of a few samples (total number of 6,178 fungal-soil, 930 fungal-
222 root, 10,120 bacterial-soil and 3,143 bacterial-roots ASVs). After quality filtering ASVs that were
223 found in fewer than 10% of the samples, we retained 418, 169, 206 and 250 ASVs and which
224 comprised 91%, 88%, 50% and 89% of all reads in the fungal-soil, fungal-root, bacterial-soil and
225 bacterial-root samples, respectively.

226

227

Table 2: Sequencing and ASV summary

	fungi-soil	fungi-root	bacteria-soil	bacteria-root
Nb_seq_sum	976,000	309,000	920,000	535,000
Nb_seq_mean	51,381	32,208	47,907	56,365
Nb_seq_mean_filtered	38,045	14,635	38,287	46,081
Nb_seq_mean_filt_merged	32,014	13,335	13,780	41,058
Nb_seq_mean_filt_merg_non_chimeras	24,737	8,505	12,049	28,451
Nb_samples	192	96	192	96
Nb_samples_trimmed	189	83	192	94
ASV_sum	6,178	930	10,120	3,143
ASV_sum_trimmed	418	169	206	250
ASV_persample	163	49	112	122

228

229

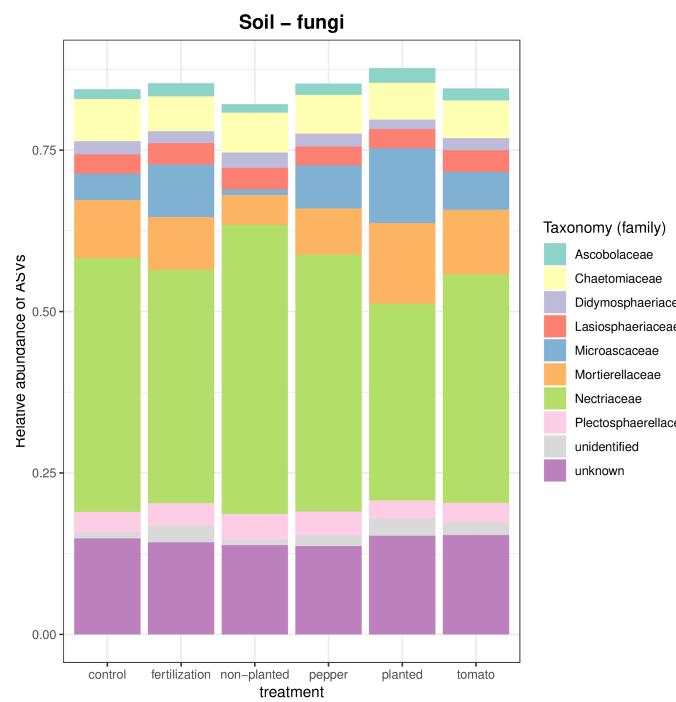
230 *Root, soil, microbial and bacterial diversity*

231 We then analysed the whole community structure and report the relative abundance of taxa (fam-
232 ily) for the fungal-soil, fungal-root, bacteria-soil and bacteria-root conditions (Figure 4). Fungal
233 communities were dominated by Nectriaceae, both in the root and soil samples. Bacterial root

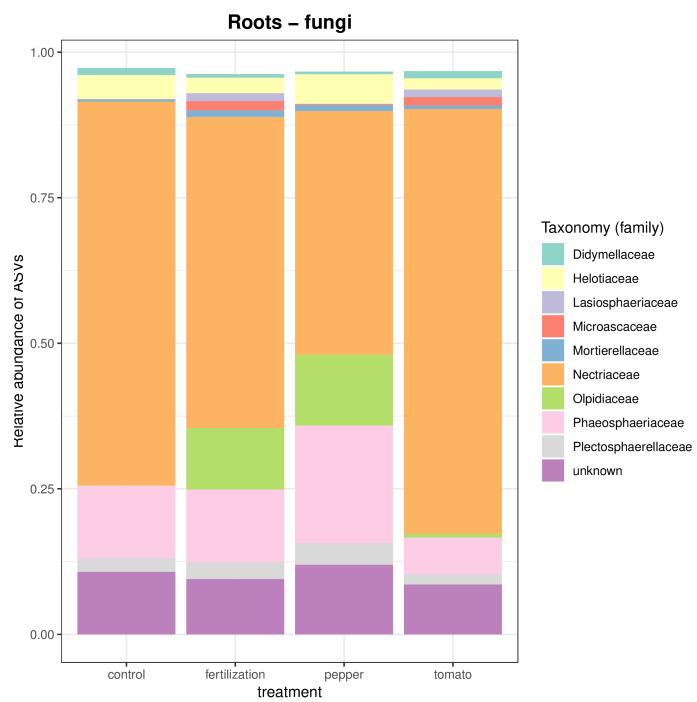
²³⁴ communities were largely dominated by the Cyanobacteria phylum (identified as *chloroplast* ac-
²³⁵ cording to the Ribosomal Database Project Naive Bayesian Classifier and the silva database). In
²³⁶ fact, these ASVs are likely sequenced chloroplasts from the plants themselves, despite the fact that
²³⁷ the primer pair used should have primarily targeted the bacterial V3-V4 region of the 16S riboso-
²³⁸ mal gene. The bacterial family Bacilaceae dominated to a lesser extent the soil communities.

²³⁹

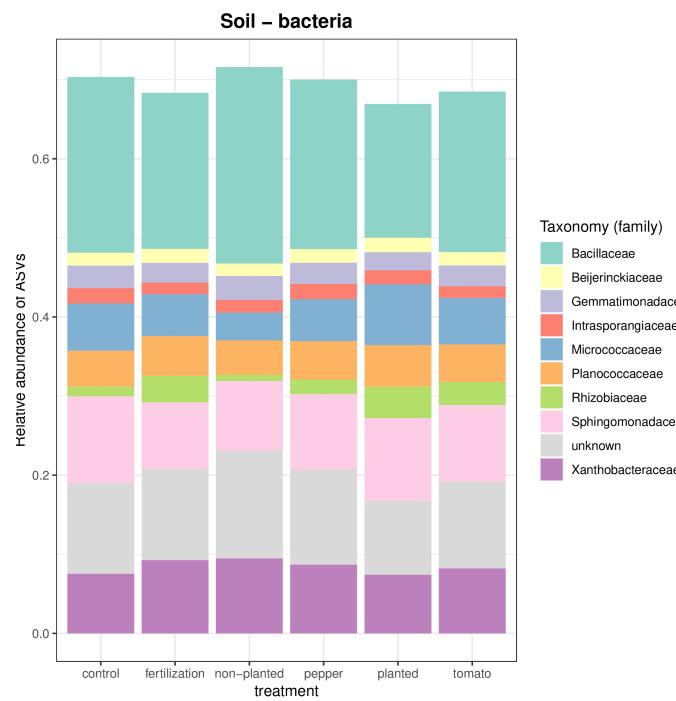
A



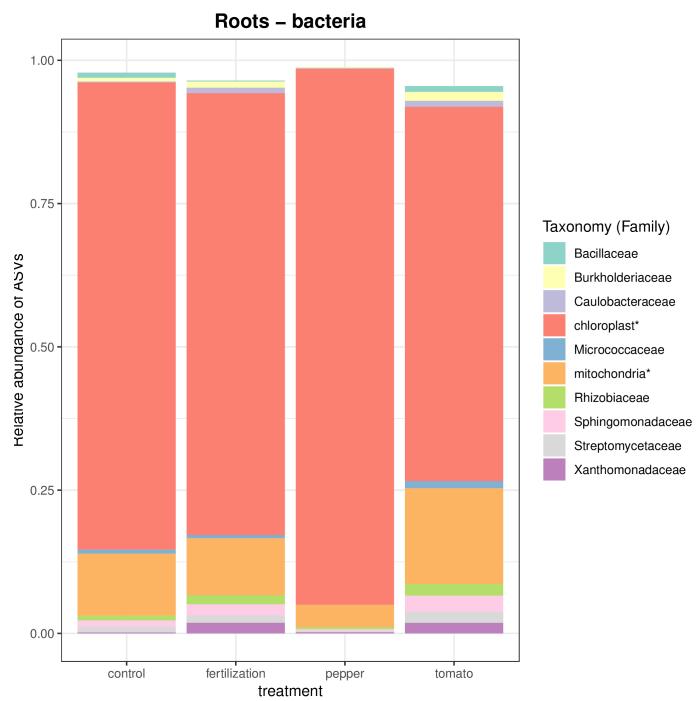
B



C



D



240

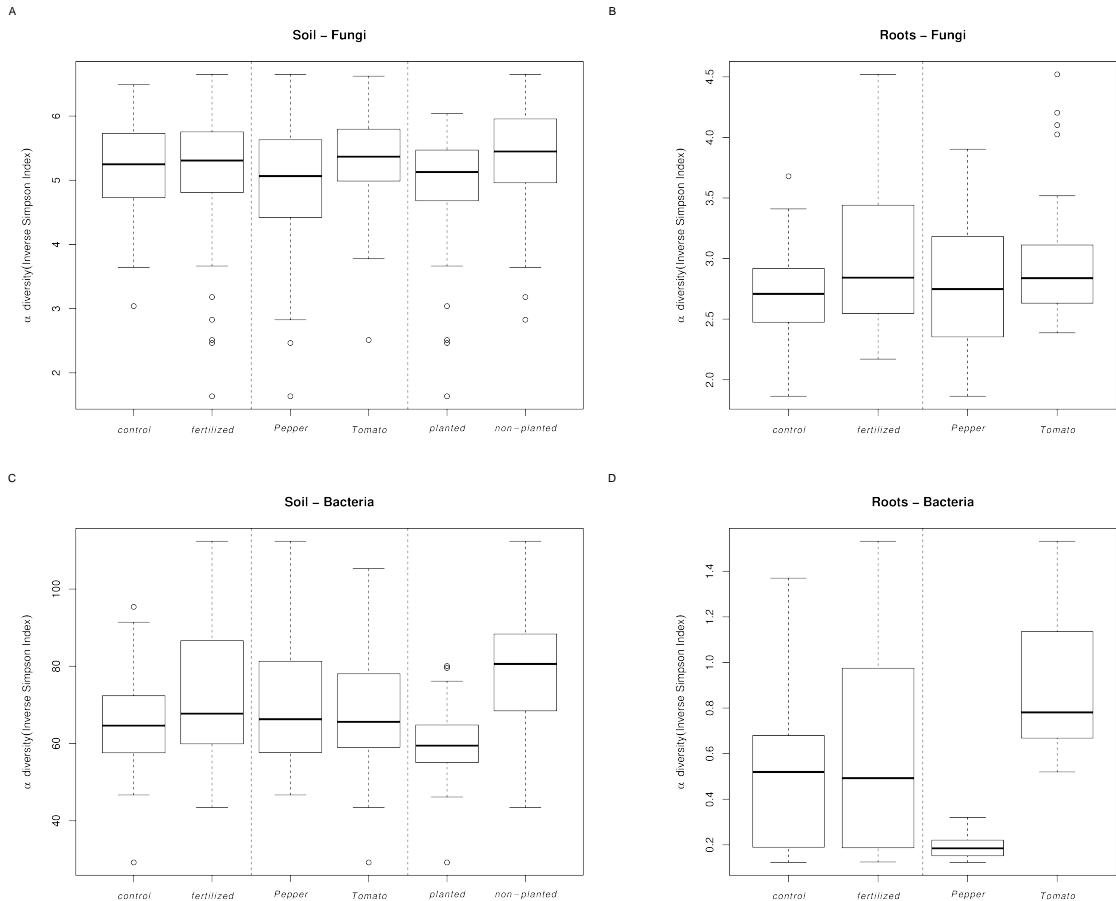
241 **Figure 4: Barplots fo the relative abundance of ASVs for fungal-root, fungal-soil, bacteria-soil**242 **and bacteria-root**

243

244

245 *Local (a-diversity)*

246 The diversity of each site (*a*-diversity) was calculated separately for each sample and under each
247 experimental conditions (fungi-soil, fungi-root, bacteria-soil and bacteria-root, Figure 5). Linear
248 mixed effects models used to assess significance. In soils samples, fungal diversity differed with
249 respect to the fertilization ($F_{(1,161)}=14.35, p\text{-value}<0.0001$) and planting ($F_{(1,161)}=41.00, p\text{-value}<0.0001$)
250 treatment, but not the species ($F_{(1,161)}=0.13, p\text{-value}=0.72$). In root samples, fungal diversity dif-
251 fered with respect to the fertilization treatment ($F_{(1,56)}=13.56, p\text{-value}=0.001$), and the species tested
252 ($F_{(1,56)}=74.31, p\text{-value}=0.003$). In soil samples, bacterial diversity differed with respect to the fer-
253 tilization treatment ($F_{(1,165)}=46.25, p\text{-value}<0.0001$), planting ($F_{(1,165)}=48.77, p\text{-value}<0.0001$) and
254 species ($F_{(1,165)}=10.22, p\text{-value}=0.002$). In root samples, bacterial diversity differed with respect
255 to the fertilization treatment ($F_{(1,67)}=16.48, p\text{-value}=0.0001$), and the species tested ($F_{(1,67)}=523.42,$
256 $p\text{-value}<0.0001$).



257

258 **Figure 5: Boxplot of alpha diversity according to the treatment, species and planting effect for**
 259 **fungal-root, fungal-soil, bacteria-soil and bacteria-root.**

260

261

262 *Differences in species composition among sites*

263 Using a PERMANOVA statistical framework, we identified that for all conditions, communities
 264 differed with respect to the fertilization treatment (Table 3). Soil fungal and bacterial communities
 265 differed the most according to whether the tray was planted (greatest % of variance explained,
 266 Table 3), while root communities differed the most between tomato and pepper plants.

Table 3: summary of PERMANOVAs*

	fungi-soil	fungi-root	bacteria-soil	bacteria-root
fertilization	0.02 (2e-04)	0.08 (1e-04)	0.04 (1e-04)	0.07 (1e-04)

	fungi-soil	fungi-root	bacteria-soil	bacteria-root
planted	0.21 (1e-04)	NA	0.13 (1e-04)	NA
species	0.02 (1e-04)	0.26 (1e-04)	0.02 (3e-04)	0.52 (1e-04)
fertilization:planted	0.01 (0.003)	NA	0.02 (1e-04)	NA
fertilization:species	0.01 (0.006)	0.04 (0.002)	0.03 (1e-04)	0.05 (2e-04)
planted:species	0.01 (0.09)	NA	0.01 (0.004)	NA
fertilization:planted:species	0.01 (0.16)	NA	0.01 (0.04)	NA

* r^2 [percentage of variance explained by the term in the model] and associated *p*-values in parentheses.

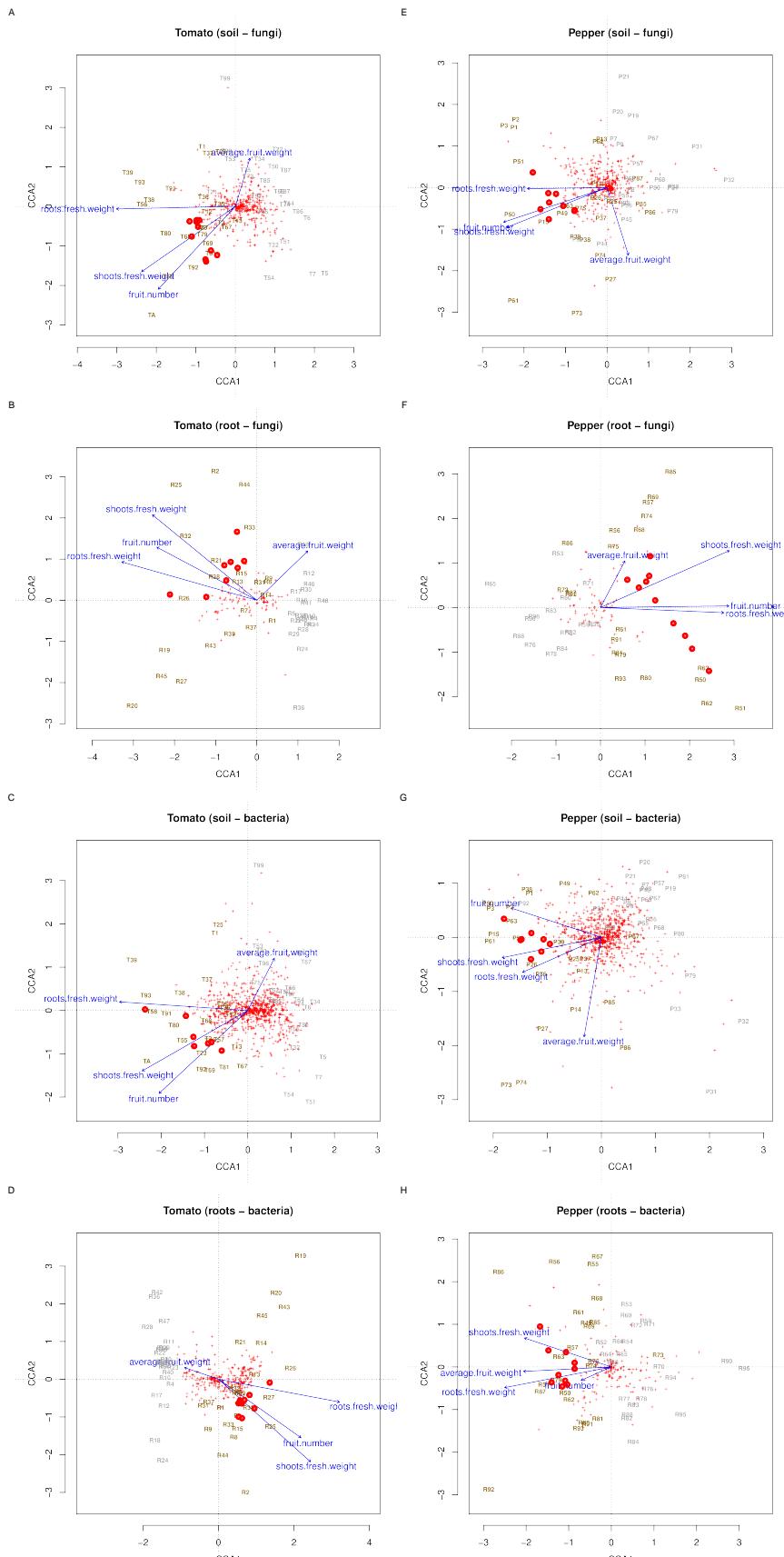
269

270 *Canonical correspondence analyses and candidate ASVs*

271 Canonical correspondence analyses indicated how fertilized samples clustered together according
 272 to their fungal or bacterial communities (Figure 6). It also shows a similar association of three of
 273 the constrain variables (productivity measures of root fresh weight, shoots fresh weight and fruit
 274 number), while average fruit weight behave differentially (in fact nearly orthogonally to the other
 275 three constrains in most ordinations).

276

277



279 **Figure 6: Canonical correspondence analyses according for tomato (A-D) and peppers (E-H) for**
280 **soil-fungi, root-fungi, soil-bacteria and root-bacteria. Samples are labelled and colored in gray**
281 **(unfertilized) or dark yellow (fertilized). Red crosses represent individual ASVs, while red**
282 **points represent the ten ASVs most closely associated wih the three productivity measures of**
283 **root fresh weight, shoots fresh weight and fruit number. Blue arrows are the four productivity**
284 **measures used as constrains in the ordinations.**

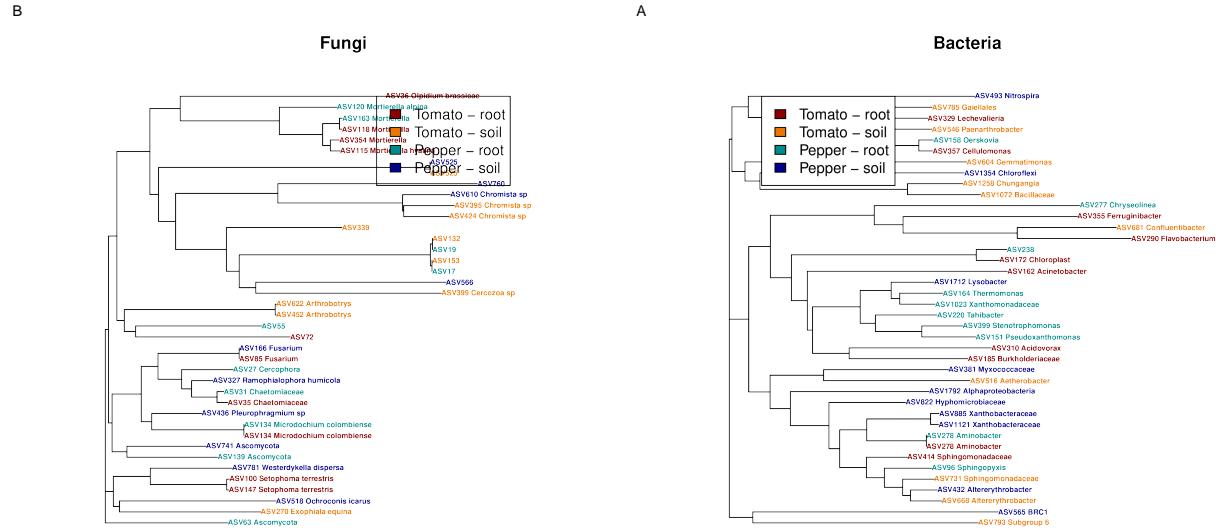
285

286 Next, we identified, for each ordination, the ten ASVs most closely related to the three con-
287 strains which behaved in a similar fashion (productivity measures of root fresh weight, shoots
288 fresh weight and fruit number). These ASVs were considered as putative candidates sequences
289 most positively impacted (increase presence of the ASV) by fertilization. We further analysed the
290 corresponding sequences for these eighty candidates (ten candidates * eight ordinations) ASVs
291 in two separate alignments (one for fungi and one for bacterial ASVs) and their accompanying
292 neighbouring joining trees. In fungi, we identified one cluster of ASVs taxonomically assigned to
293 *Olpidium brassicae* (fungal obligate parasite in the phylum Chytridiomycota) that forms the ma-
294 jority of ASVs most closely related to productivity. In addition, we identified five different ASVs
295 in both species and both root and soil closely related phylogenetically. Given that no taxonomy
296 was assigned to these sequences through the dada2 RDP bootstrap approach, we used a BLASTn
297 Altschul et al. (1997) approach (against NCBI nr) to identify the most closely related sequences.
298 We identified this cluster of ASVs as *Rhogostoma schuessleri* (BLASTn, e-value=2e-74), a protist in
299 the phylum Cercozoa, which are known to be present in the soil and phyllosphere Dumack et al.
300 (2017) .

301

302 In bacteria-roots, we identified a cluster of ten closely related sequences taxonomically assigned
303 to *Chloroplast*, and which likely originate from the plants themselves. We also identified a number
304 of ASVs associated with productivity in the soil of both the pepper and tomato plants. Notably,
305 ASV100 & ASV73 (*Oerskovia* spp.), ASV231 (*Blastocatellaceae*), ASV515, ASV1105 & ASV647 (*Bacil-*
306 *laceae*), ASV107 (*Methyloligellaceae*) and ASV95 & ASV107 (*Santhobacteraceae*) were identified.

307



308

309 Figure 7: Neighbor-Joining trees of candidates ASVs (fungi & roots) associated with produc-
310 tivity measures

311

312 **DISCUSSION**

313 • Overall increases in productivity in both species (but mention that tomato were fertilized with
314 hen manure as well).

315 • A few words about the goal of the sequencing: very broad approach looking at lots of factors
316 (fertilization, species, planting root, soil, fungal, bacterial).

317 • A few words about the *dada2* approach and its advantages (it's reproducible, open & R based. It
318 identifies unique (non-clustered) ASVs which can be compared among species) & disadvantages
319 (essentially, the main disadvantage is that it identifies non-clustered ASVs, which means that most
320 ASVs are unique, and not found anywhere else, which makes it hard to compare: see *Sequencing*
321 paragraph in results).

322 • Also, talk about advantages / disadvantages of the RDP bayesian classifier approach to taxon-
323 omy rather than a BLAST approach (a BLAST approach will always give you a result even if this
324 result is false, an RDP will only give you a taxonomy (from Kingdom to species) if there is enough
325 power/confidence in it). Also any taxonomy classification largely depends on the database you
326 have.

327 • Talk about effect of treatment on root + soil on overall (alpha) diversity. • Talk about effect of
328 treatment on root + soil on differences in species composition among sites (beta diversity) • Talk
329 about the fact in the roots, we most likely sequenced the plant itself, rather than the bacteria.

330 • Discuss some of the candidate ASVs identified based on the ordinations and Figure 7.

331 • Follow-up work: ?

332 REFERENCES

- 333 Altschul SF., Madden TL., Schäffer AA., Zhang J., Zhang Z., Miller W., Lipman DJ. 1997. Gapped
334 blast and psi-blast: A new generation of protein database search programs. *Nucleic acids research*
335 25:3389–3402.
- 336 Anderson MJ. 2001. A new method for non-parametric multivariate analysis of variance. *Aus-
337 tral ecology* 26:32–46.
- 338 Anderson MJ., Legendre P. 1999. An empirical comparison of permutation methods for tests
339 of partial regression coefficients in a linear model. *Journal of statistical computation and simulation*
340 62:271–303.
- 341 Callahan B. 2018. Silva for dada2: Silva taxonomic training data formatted for dada2 (silva
342 version 132). *Zenodo*. DOI: [10.5281/zenodo.1172783](https://doi.org/10.5281/zenodo.1172783).
- 343 Callahan BJ., McMurdie PJ., Rosen MJ., Han AW., Johnson AJA., Holmes SP. 2016. DADA2:
344 High-resolution sample inference from illumina amplicon data. *Nature methods* 13:581.
- 345 Caporaso JG., Kuczynski J., Stombaugh J., Bittinger K., Bushman FD., Costello EK., Fierer N.,
346 Pena AG., Goodrich JK., Gordon JI., others. 2010. QIIME allows analysis of high-throughput
347 community sequencing data. *Nature methods* 7:335.
- 348 Community U. 2018.UNITE general fasta release. version 01.12.2017. Available at <https://files.pluto.ut.ee/doi/C8/E4/C8E4A8E6A7C4C00EACE3499C51E550744A259A98F8FE25993B1C7B9E7D2170B2.zip>
- 351 Craigie JS. 2011. Seaweed extract stimuli in plant science and agriculture. *Journal of Applied
352 Phycology* 23:371–393.
- 353 Dumack K., Flues S., Hermanns K., Bonkowski M. 2017. Rhogostomidae (cercozoa) from soils,
354 roots and plant leaves (*arabidopsis thaliana*): Description of rhogostoma epiphylla sp. nov. and r.
355 cylindrica sp. nov. *European journal of protistology* 60:76–86.
- 356 Jukes T., Cantor C. 1969. Evolution of protein molecules, pp. 21–132 in mammalian protein
357 metabolism, edited by munro hn.
- 358 Klindworth A., Pruesse E., Schweer T., Peplies J., Quast C., Horn M., Glöckner FO. 2013. Evalu-
359 ation of general 16S ribosomal rna gene pcr primers for classical and next-generation sequencing-

360 based diversity studies. *Nucleic acids research* 41:e1–e1.

361 Oksanen J., Blanchet FG., Kindt R., Legendre P., Minchin PR., O'hara R., Simpson GL., Solymos
362 P., Stevens MHH., Wagner H., others. 2013. Package “vegan”. *Community ecology package, version*
363 2.

364 Paradis E., Claude J., Strimmer K. 2004. APE: Analyses of phylogenetics and evolution in r
365 language. *Bioinformatics* 20:289–290.

366 Pinheiro J., Bates D., DebRoy S., Sarkar D., Team RC. 2017. Nlme: Linear and nonlinear mixed-
367 effects models. r package version 3.1-128. 2016. *R software*.

368 Schliep KP. 2010. Phangorn: Phylogenetic analysis in r. *Bioinformatics* 27:592–593.

369 Schloss PD., Westcott SL., Ryabin T., Hall JR., Hartmann M., Hollister EB., Lesniewski RA.,
370 Oakley BB., Parks DH., Robinson CJ., others. 2009. Introducing mothur: Open-source, platform-
371 independent, community-supported software for describing and comparing microbial communi-
372 ties. *Applied and environmental microbiology* 75:7537–7541.

373 Toju H., Tanabe AS., Yamamoto S., Sato H. 2012. High-coverage its primers for the dna-based
374 identification of ascomycetes and basidiomycetes in environmental samples. *PloS one* 7:e40863.

375 Wang Q., Garrity GM., Tiedje JM., Cole JR. 2007. Naive bayesian classifier for rapid assign-
376 ment of rRNA sequences into the new bacterial taxonomy. *Applied and environmental microbiology*
377 73:5261–5267.

378 Wickham H. 2016. *Ggplot2: Elegant graphics for data analysis*. Springer.

379 Wickham H., Francois R., Henry L., Müller K. 2015. Dplyr: A grammar of data manipulation.
380 *R package version 0.4.3.*

381 Wright ES. 2016. Using decipher v2. 0 to analyze big biological sequence data in r. *R Journal* 8.