

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

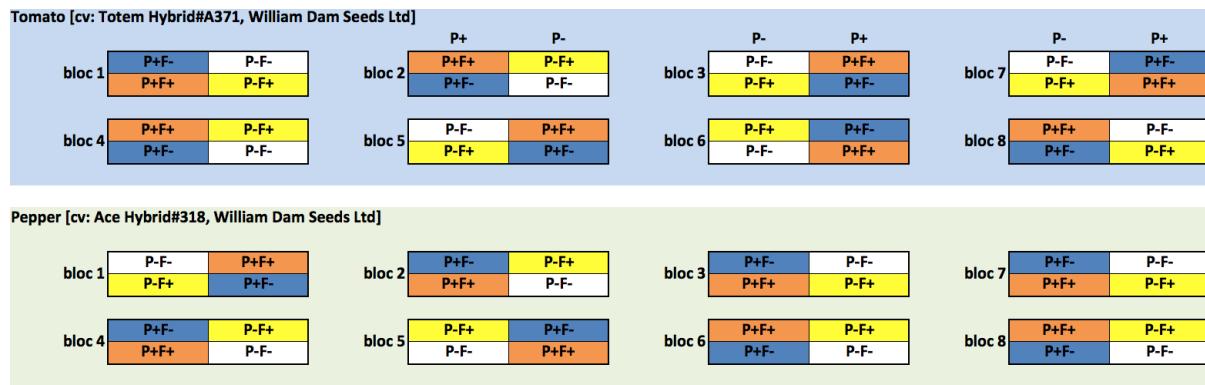
21 MATERIAL AND METHOD

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 ma-
 40 nure, 18 g per tray repeated every 4 weeks, 5-3-2) from Acti-sol (Notre-Dame-du-Bon-Conseil,
 41 Qc, Canada) in addition to Stella Maris® (3.5 ml per 1L, each tray received 250 ml, repeated every
 42 2 weeks) for the duration of the experiment. The other half were unfertilized. Stella Maris® is a
 43 registered trademark from Acadian Seaplants Ltd. (Darmouth, NS, Canada). It is primarily com-
 44 posed of *Ascophyllum nodosum* seaweed and is advertised as a natural activator of the crops' own
 45 growth and defense mechanisms to improve root growth and resist temperature, drought, and
 46 salinity stress in order to maximize yield and crop qualities (Acadian Seaplants Ltd. 2018). Half
 47 of the pepper plants were treated using solely Stella Maris (3.5 ml per 1L, each tray received 250
 48 ml, repeated every 2 weeks) for the duration of the experiment. The other half were untreated.
 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, thus allowing sequences and abundance counts to
100 be compared among studies Callahan et al. (2016).

101

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

117

118 *Statistical analyses - plant productivity*

119 We tested for the effect of species (tomato vs pepper), fertilization and their interaction on six plant

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

129

130

131 *Statistical analyses - microbial and fungal diversity*

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

140

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

150 statistical framework.

151

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

157

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

168

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

179

180 **RESULTS**

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

182

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

183

184 *productivity*

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

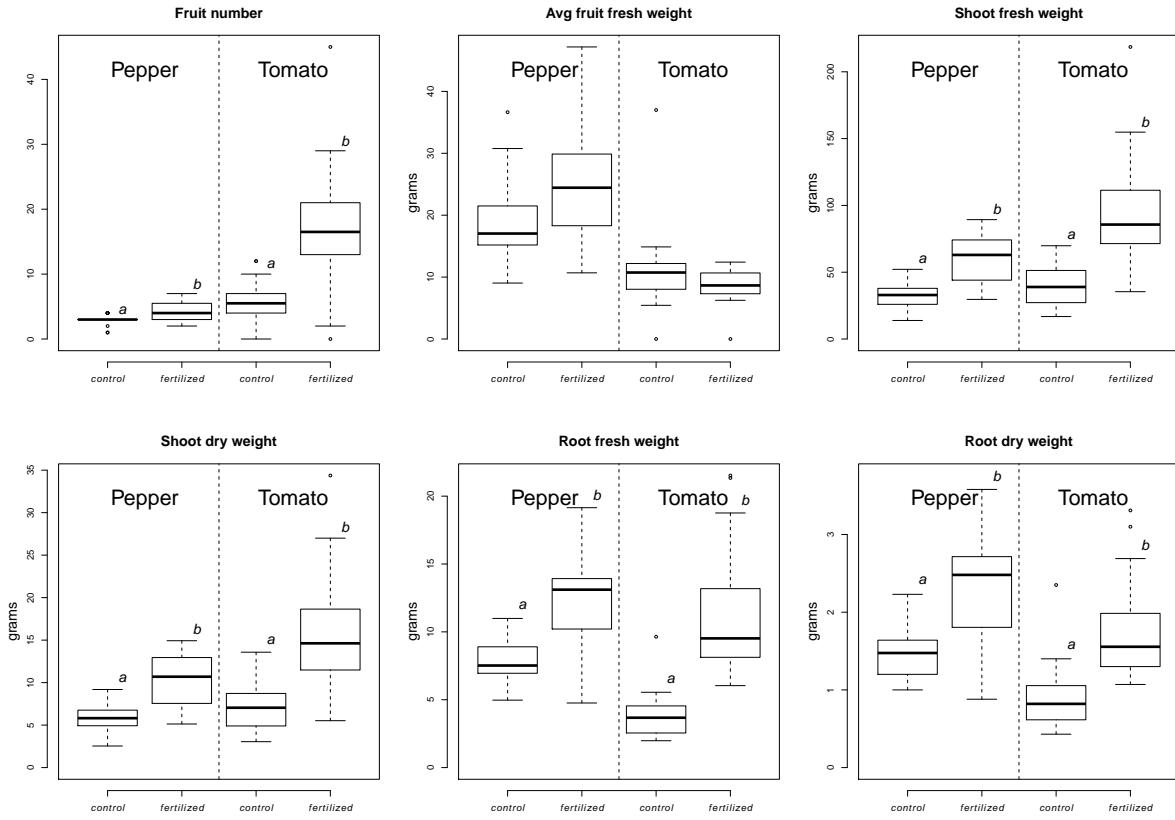
188 ground plant structure grew larger in fertilized plants, in addition to producing more fruits (see
189 Figure 2 for some examples of the striking difference between fertilized and unfertilized plants).



190
191 **Figure 2: Plant productivity.** Photos were taken at the end of the experimental treatment. In
192 each photo, fertilized plants are on the left. **A:** pepper plants, **B:** pepper roots, **C:** pepper fruits
193 and **D:** tomato fruits.

194
195 Statistically, all six productivity measures significantly differed according to species, and five of
196 those were significantly different according to the fertilization treatment. The only exception was
197 the average fruit weight which did not differ between fertilized and control plants (LMM, $F_{(1,69)}$
198 = 1.27, p -value=0.26). However the model did reveal a significant interaction between treatment
199 and plant ($F_{(1,69)} = 9.6$, p -value=0.0028). In fact, when testing only the pepper plants, the effect of
200 fertilization on average fruit weight was significantly higher in the fertilized pepper plants ($F_{(1,23)}$
201 = 10.84, p -value=0.0032).

202



203

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

205

206

207 *Sequencing*

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

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

217

218 The dada2 pipeline inferred, on average, 112 Amplicon Sequence Variants per sample (average of
219 163 fungal-soil ASV, 49 fungal-root ASVs, 112 bacterial-soil ASVs and 122 bacterial-root ASVs).
220 Many of those were unique to one of a few samples (total number of 6,178 fungal-soil, 930 fungal-
221 root, 10,120 bacterial-soil and 3,143 bacterial-roots ASVs). After quality filtering ASVs that were
222 found in fewer than 10% of the samples, we retained 418, 169, 206 and 250 ASVs and which
223 comprised 91%, 88%, 50% and 85% of all reads in the fungal-soil, fungal-root, bacterial-soil and
224 bacterial-root samples, respectively.

225

226

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

227

228

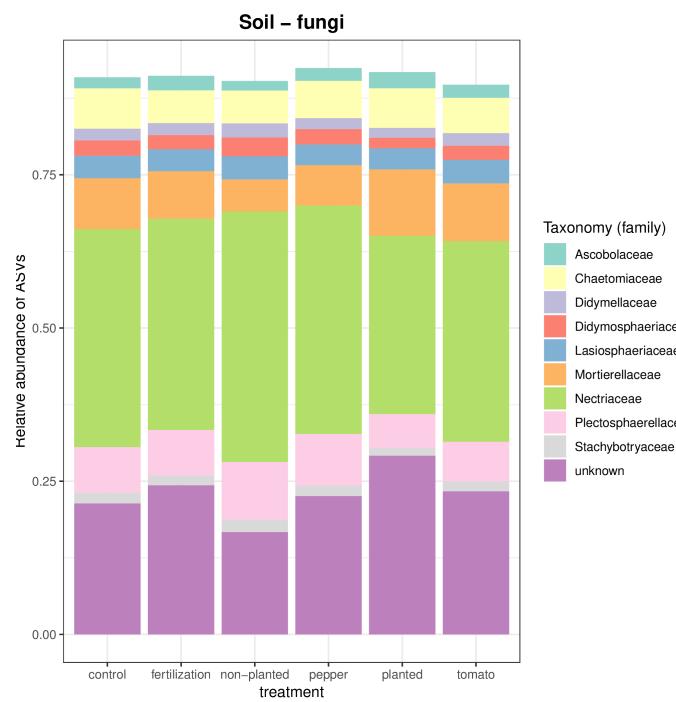
229 *Root, soil, microbial and bacterial diversity*

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

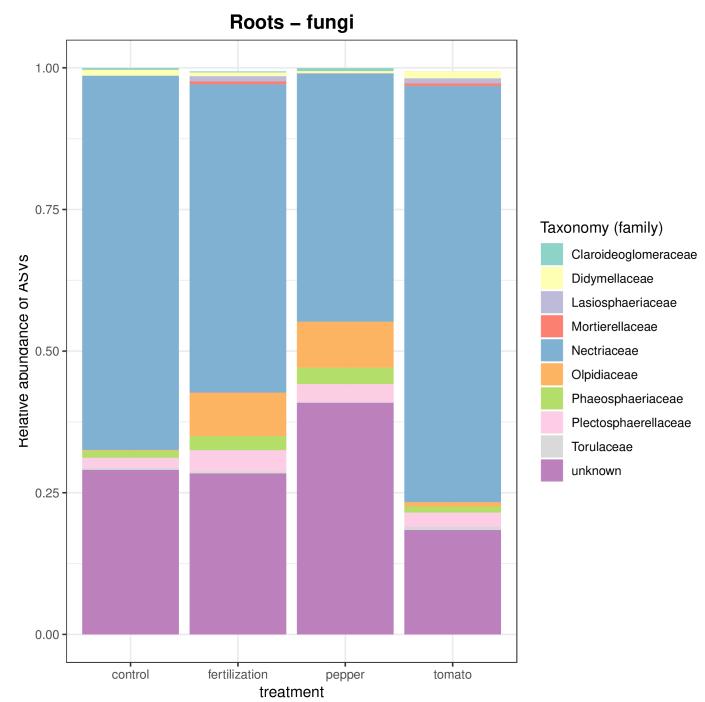
233 communities were largely dominated by the Cyanobacteria phylum (identified as *chloroplast* ac-
234 cording to the Ribosomal Database Project Naive Bayesian Classifier and the silva database). In
235 fact, these ASVs are likely sequenced chloroplasts from the plants themselves, despite the fact that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

238

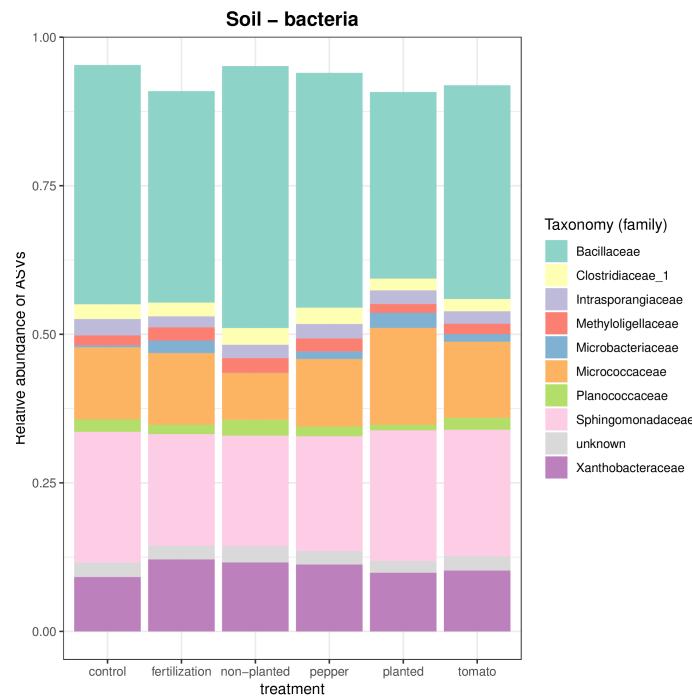
A



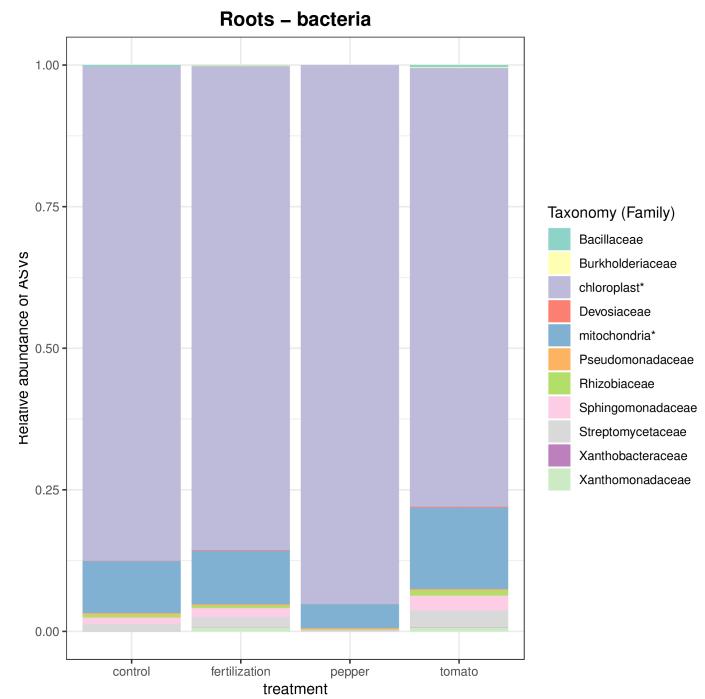
B



C



D



239

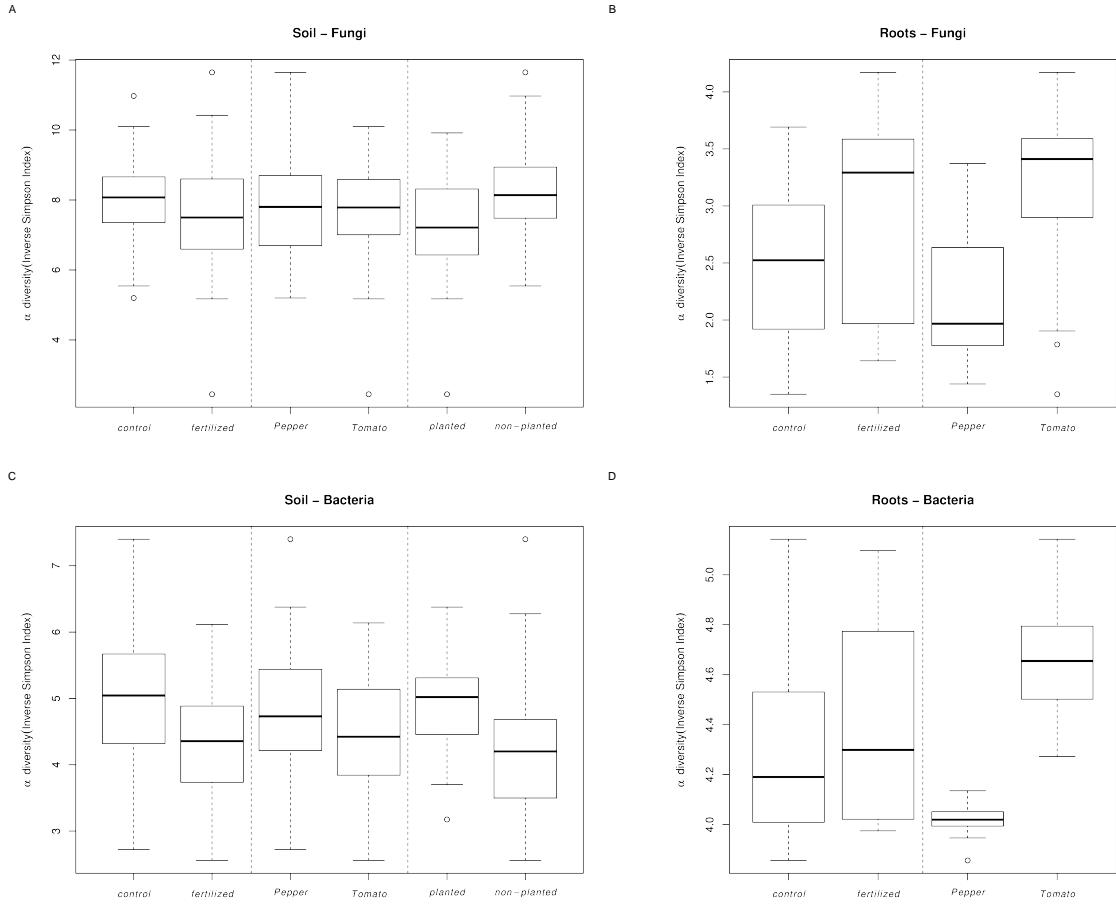
240 **Figure 4: Barplots fo the relative abundance of ASVs for fungal-root, fungal-soil, bacteria-soil**241 **and bacteria-root**

242

243

244 *Local (a-diversity)*

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



256

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

259

260

261 *Differences in species composition among sites*

262 Using a PERMANOVA statistical framework, we identified that for all conditions, communities
 263 differed with respect to the fertilization treatment (Table 3). Soil fungal and bacterial communities
 264 differed the most according to whether the tray was planted (greatest % of variance explained,
 265 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 (4e-04)	0.04 (0.0013)	0.03 (1e-04)	0.01 (0.0705)

	fungi-soil	fungi-root	bacteria-soil	bacteria-root
planted	0.15 (1e-04)	NA	0.06 (1e-04)	NA
species	0.02 (2e-04)	0.2 (1e-04)	0.01 (0.0032)	0.42 (1e-04)
fertilization:planted	0.01 (0.008)	NA	0.02 (1e-04)	NA
fertilization:species	0.01 (0.0705)	0.03 (0.0094)	0.01 (0.002)	0.01 (0.0973)
planted:species	0.01 (0.1597)	NA	0.01 (0.1767)	NA
fertilization:planted:species	0 (0.7956)	NA	0.01 (0.1179)	NA

²⁶⁶ * r^2 [percentage of variance explained by the term in the model] and associated *p*-values in
²⁶⁷ parentheses.

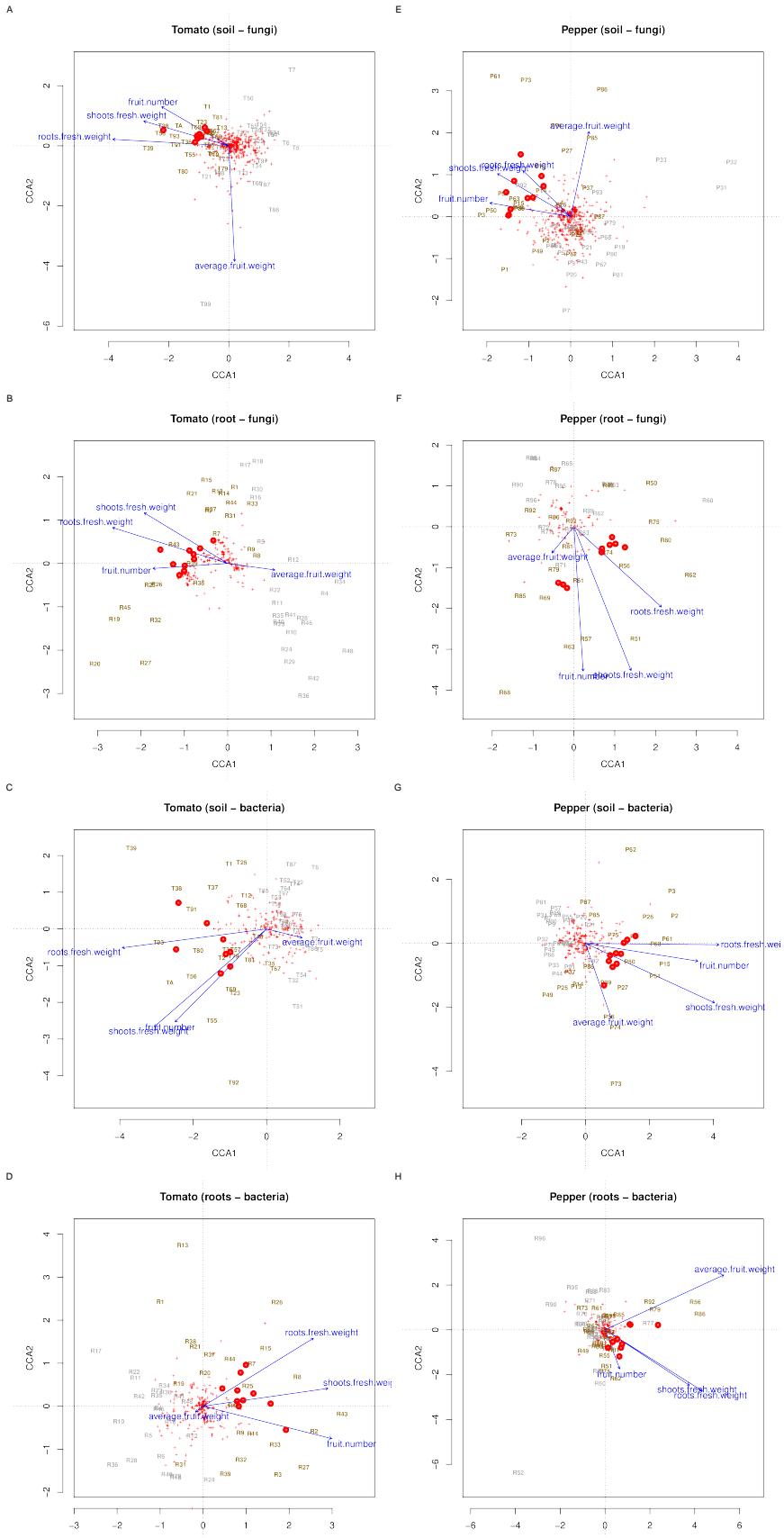
²⁶⁸

²⁶⁹ *Canonical correspondence analyses and candidate ASVs*

²⁷⁰ Canonical correspondence analyses indicated how fertilized samples clustered together according
²⁷¹ to their fungal or bacterial communities (Figure 6). It also shows a similar association of three of
²⁷² the constrain variables (productivity measures of root fresh weight, shoots fresh weight and fruit
²⁷³ number), while average fruit weight behave differentially (in fact nearly orthogonally to the other
²⁷⁴ three constrains in most ordinations).

²⁷⁵

²⁷⁶



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

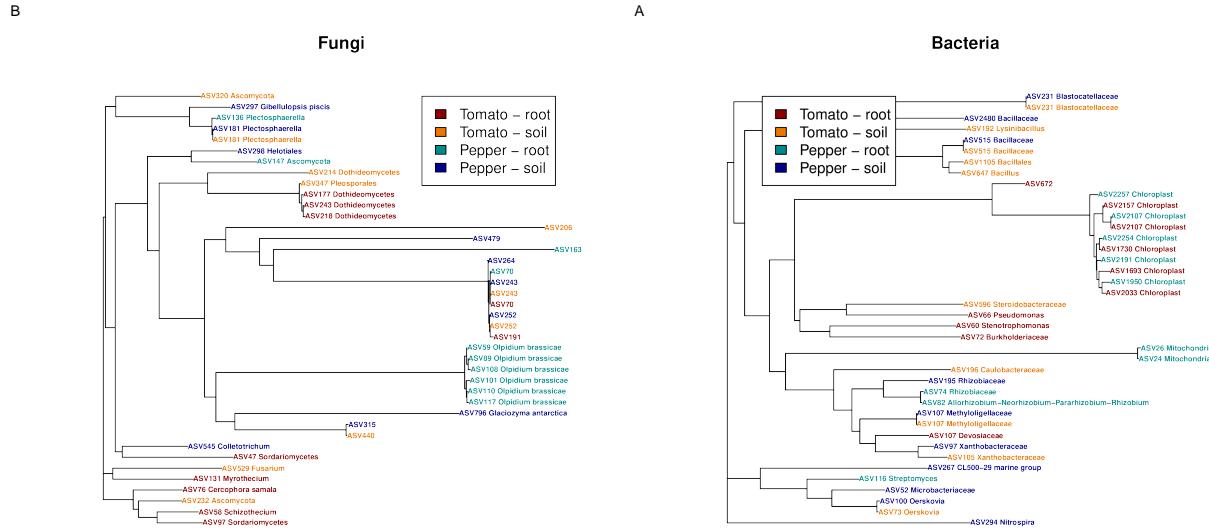
284

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

300

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

306



307

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

310

311 **DISCUSSION**

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

314 • A few words about the goal of the sequencing: very broad approach looking at lots of factors
315 (fertilization, species, planting root, soil, fungal, bacterial). • A few words about the *dada2* ap-
316 proach and its advantages (it's reproducible, open & R based. It identifies unique (non-clustered)
317 ASVs which can be compared among species) & disadvantages (essentially, the main disadvan-
318 tage is that it identifies non-clustered ASVs, which means that most ASVs are unique, and not
319 found anywhere else, which makes it hard to compare: see *Sequencing* paragraph in results).

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

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

328 • Discuss some of the candidate ASVs identified based on the ordinations and Figure 7. • Follow-
329 up work: ?

330 REFERENCES

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

- 358 based diversity studies. *Nucleic acids research* 41:e1–e1.
- 359 Oksanen J., Blanchet FG., Kindt R., Legendre P., Minchin PR., O'hara R., Simpson GL., Solymos
360 P., Stevens MHH., Wagner H., others. 2013. Package “vegan”. *Community ecology package, version*
361 2.
- 362 Paradis E., Claude J., Strimmer K. 2004. APE: Analyses of phylogenetics and evolution in r
363 language. *Bioinformatics* 20:289–290.
- 364 Pinheiro J., Bates D., DebRoy S., Sarkar D., Team RC. 2017. Nlme: Linear and nonlinear mixed-
365 effects models. r package version 3.1-128. 2016. *R software*.
- 366 Schliep KP. 2010. Phangorn: Phylogenetic analysis in r. *Bioinformatics* 27:592–593.
- 367 Schloss PD., Westcott SL., Ryabin T., Hall JR., Hartmann M., Hollister EB., Lesniewski RA.,
368 Oakley BB., Parks DH., Robinson CJ., others. 2009. Introducing mothur: Open-source, platform-
369 independent, community-supported software for describing and comparing microbial communi-
370 ties. *Applied and environmental microbiology* 75:7537–7541.
- 371 Toju H., Tanabe AS., Yamamoto S., Sato H. 2012. High-coverage its primers for the dna-based
372 identification of ascomycetes and basidiomycetes in environmental samples. *PloS one* 7:e40863.
- 373 Wang Q., Garrity GM., Tiedje JM., Cole JR. 2007. Naive bayesian classifier for rapid assign-
374 ment of rRNA sequences into the new bacterial taxonomy. *Applied and environmental microbiology*
375 73:5261–5267.
- 376 Wickham H. 2016. *Ggplot2: Elegant graphics for data analysis*. Springer.
- 377 Wickham H., Francois R., Henry L., Müller K. 2015. Dplyr: A grammar of data manipulation.
378 R package version 0.4.3.
- 379 Wright ES. 2016. Using decipher v2. 0 to analyze big biological sequence data in r. *R Journal* 8.