

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

⁴ **Sébastien Renaud^{1,2}, Jacynthe Masse^{1,2}, Mohamed Hijri^{1,2}** ¹Département de Sciences Biologiques,
⁵ Institut de Recherche en Biologie Végétale, Université de Montréal, 4101 Sherbrooke Est, Montreal, H1X 2B2, Que-
⁶ bec, 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

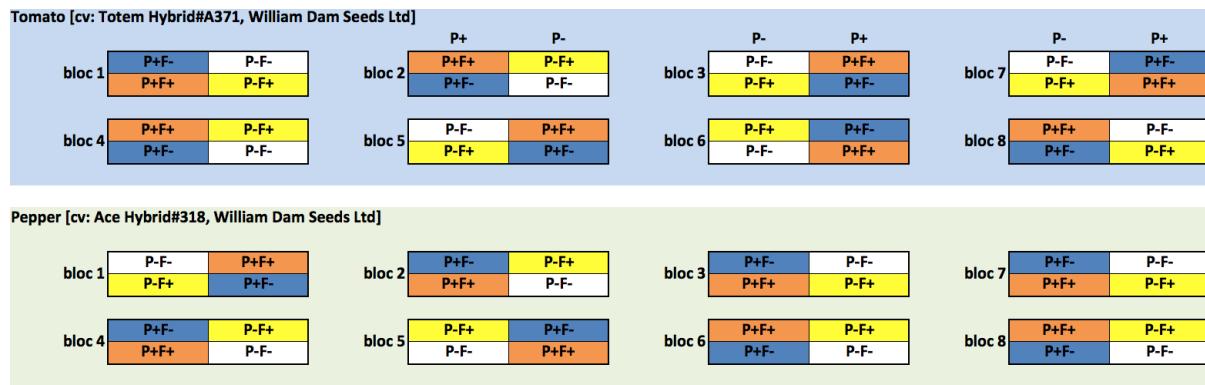
21 MATERIAL AND METHOD

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 specifically, we used fungal primers ITS3_KYO2 (5'-ACACTGACGA CATG-
71 GTTCT ACAGATGAAGAAC GYAGYRAA-3') and ITS4_KYO3 (5'-TACGGT AGCAGAGACTT
72 GGTCTCTBTTV CCKCTTCACTCG-3') to produce a final amplicon size of ~430bp. This primer
73 pair should target the Internal transcribed spacer and inhibit the amplification of plant sequences
74 and enable the selective amplification of fungal communities from soil, mycorrhizal and other en-
75 vironmental 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. This primer pair has been used extensively in high-
80 throughput sequencing studies in a range of environments and was shown to be the least biased
81 among 512 primer pairs evaluated in silico for bacterial amplification Klindworth et al. (2013).

82

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

87

88

89 *Bioinformatics*

90 All bioinformatics, statistical, and graphical analyses further described were performed in R 3.5.1
91 Team & others (2018) and detailed scripts are available here (https://github.com/seb951/Acadian_SeaPlants).
92

93

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

101

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

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 unique to a block or replicate, but not found
140 in the majority of a treatment and keep only *abundant* ones.

141

142 We then conducted community-based analyses looking at the effect of the fertilization treatment
143 on the abundant 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 the grouping of samples, ASVs
161 and their association with productivity variables (*species* scaling based on ASV matrix). 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*

¹⁸³ In Table 1, we present the characteristics of the soil which was collected at the IRDA research
¹⁸⁴ station in St-Bruno (Qc, Canada) and used in the current experimental design.

¹⁸⁵

Table 1: Soil characteristics

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

¹⁸⁶

¹⁸⁷ *productivity*

¹⁸⁸ We tested the effect of the fertilization treatment on six measures of overall plant growth and pro-
¹⁸⁹ ductivity (fruit number, average fruit weight, shoots fresh weight, shoots dry weight, roots fresh

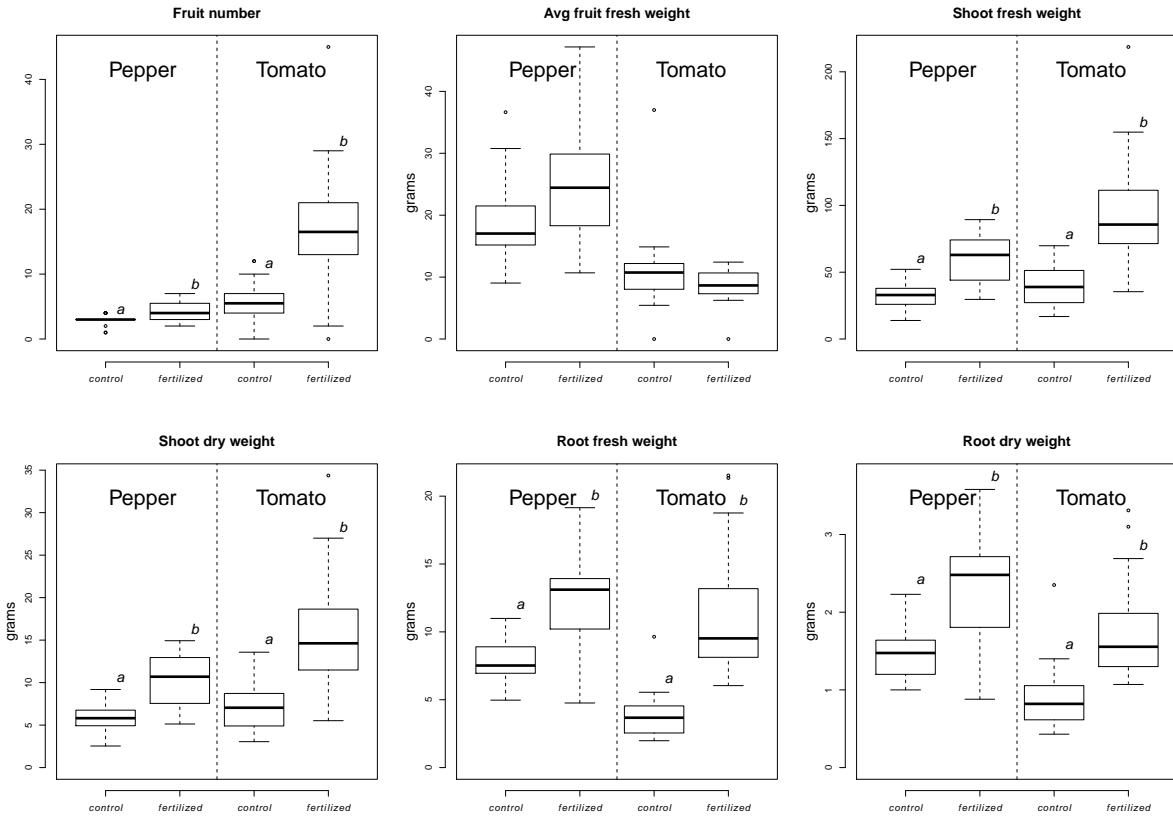
190 weight, roots dry weight) for both tomato and peppers. Visually, both above ground and below
191 ground plant structure grew larger in fertilized plants, in addition to producing more fruits (see
192 Figure 2 for some examples of the striking difference between fertilized and unfertilized plants).



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

197
198 Statistically, all six productivity measures significantly differed according to species, and five of
199 those were significantly different according to the fertilization treatment. The only exception was
200 the average fruit weight which did not differ between fertilized and control plants (LMM, $F_{(1,69)}$
201 = 1.27, p -value=0.26). However the model did reveal a significant interaction between treatment
202 and plant ($F_{(1,69)} = 9.6$, p -value=0.0028). In fact, when testing only the pepper plants, the effect of
203 fertilization on average fruit weight was significantly higher in the fertilized pepper plants ($F_{(1,23)}$
204 = 10.84, p -value=0.0032).

205



206

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

208

209

210 *Sequencing*

211 A total of 2.7 million paired-end raw reads were obtained for all samples combined (976,000 for
 212 fungi-soil, 920,000 for fungi-root, 309,000 for bacteria-soil and 535,000 for bacteria-root, Table 2).
 213 Note that sequencing samples were analysed separately for fungal-soil, fungal-root, bacteria-soil
 214 and bacteria-root conditions. On average, 47,664 paired-end reads were obtained per sample. Af-
 215 ter quality filters were applied, including removing chimeras, and paired-end reads were merged,
 216 an average of 19,690 sequences remained per sample. While 192 soil samples for fungi and bac-
 217 teria, and 96 root samples for fungi and bacteria were sequenced, seven fungi-soil samples, 15
 218 fungi-root samples and one bacteria-root samples were removed because they had too few reads
 219 based on our strict quality thresholds.

220

221 The dada2 pipeline inferred, on average, 170 Amplicon Sequence Variants per sample (average
222 of 176 fungal-soil ASV, 37 fungal-root ASVs, 269 bacterial-soil ASVs and 92 bacterial-root ASVs).
223 Many of those were unique to one of a few samples (total number of 6,112 fungal-soil, 845 fungal-
224 root, 9,352 bacterial-soil and 2,023 bacterial-roots ASVs). After quality filtering ASVs that were
225 found in fewer than 10% of the samples, we retained 413, 106, 811 and 325 ASVs and which
226 comprised 94%, 95%, 89% and 98% of all reads in the fungal-soil, fungal-root, bacterial-soil and
227 bacterial-root samples, respectively.

228

229

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	50,847	32,208	47,907	56,365
Nb_seq_mean_filtered	32,626	12,714	29,662	37,642
Nb_seq_mean_filt_merged	29,300	12,094	14,060	30,706
Nb_seq_mean_filt_merg_non_chimeras	25,476	9,849	13,521	30,408
Nb_samples	192	96	192	96
Nb_samples_trimmed	189	81	192	95
ASV_sum	6,112	845	9,352	2,023
ASV_sum_trimmed	413	106	811	325
ASV_persample	176	37	269	92

230

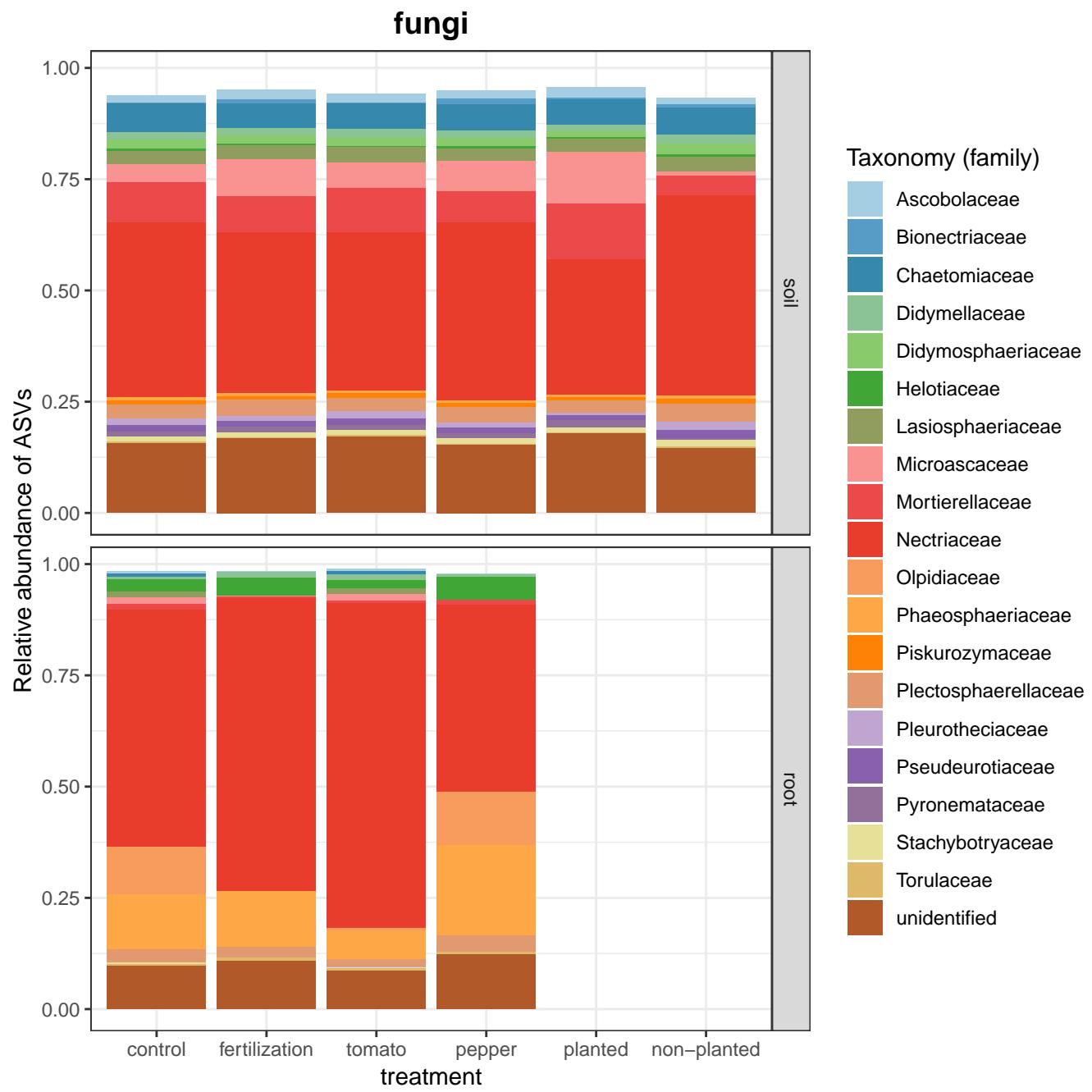
231

232 *Root, soil, microbial and bacterial diversity*

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

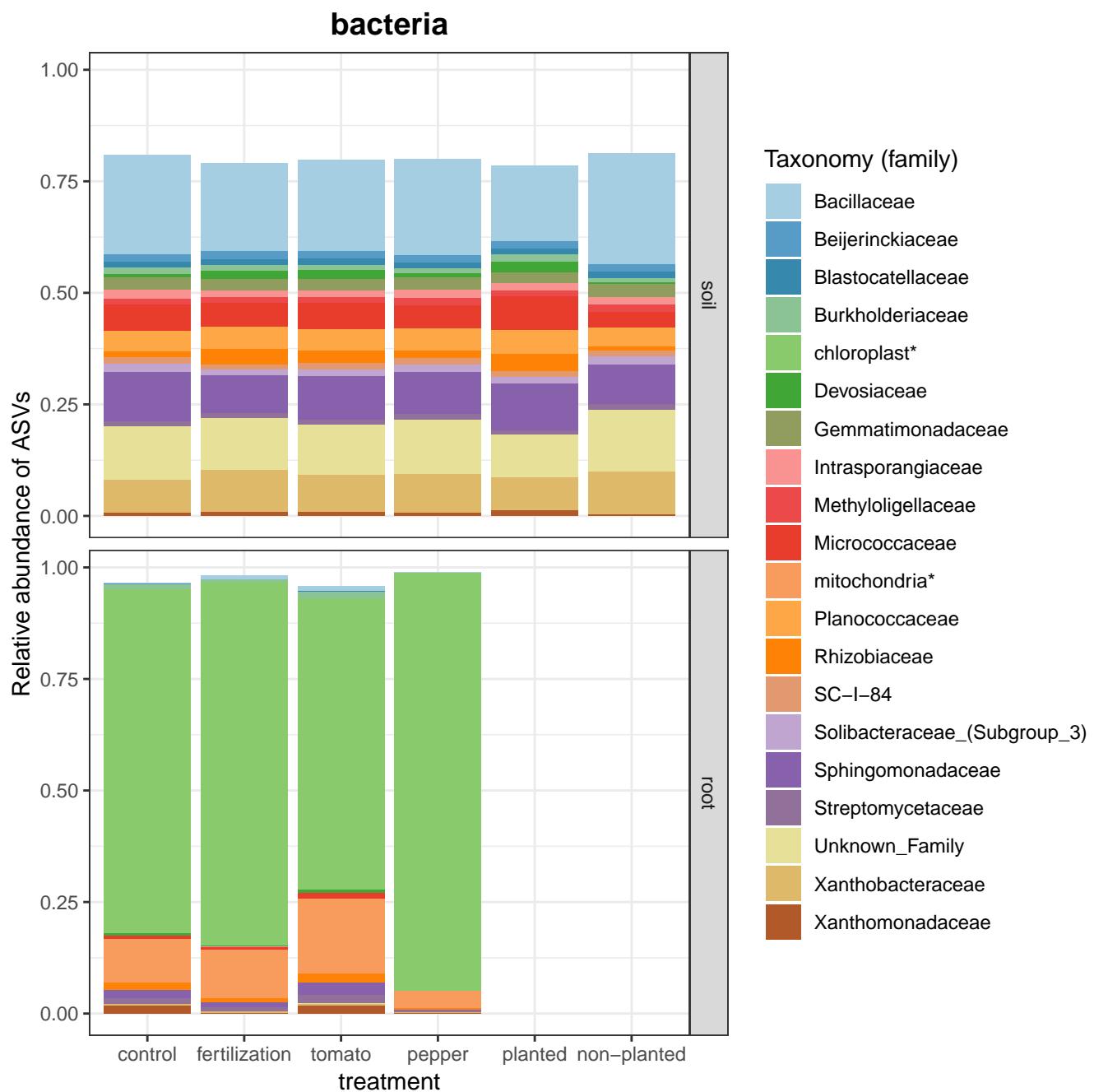
236 family Bacilaceae dominated to a lesser extent the soil communities. Bacterial root communities
 237 were largely dominated by the Cyanobacteria phylum (identified as *chloroplast* according to the
 238 Ribosomal Database Project Naive Bayesian Classifier and the silva database). In fact, these ASVs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

241



242

243 **Figure 4a: Barplots fo the relative abundance of fungal ASVs for**

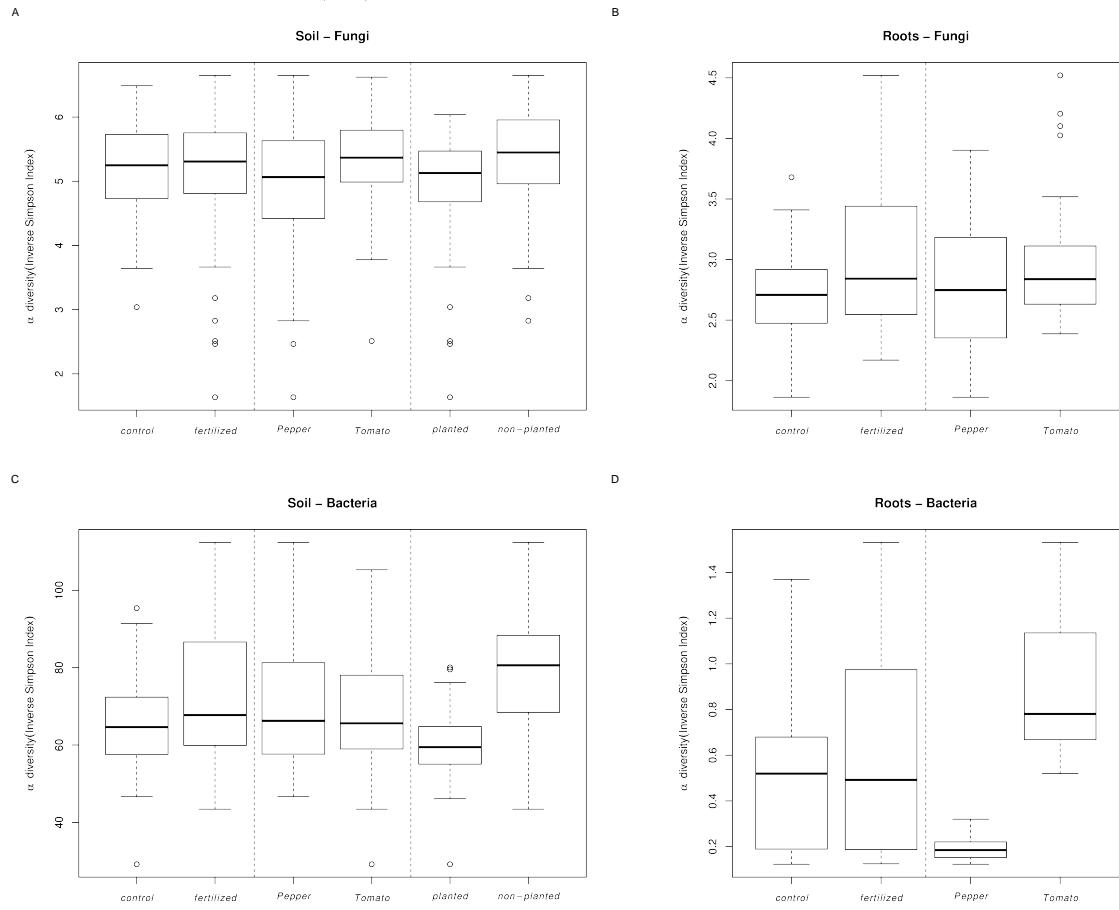


246 **Figure 4b: Barplots fo the relative abundance of bacterial ASVs for**

248 *Local (a-diversity)*

249 The diversity of each site (*a*-diversity) was calculated seperately for each sample and under each
 250 experimental conditions (fungi-soil, fungi-root, bacteria-soil and bacteria-root, Figure 5). Linear
 251 mixed effects models were used to assess significance. In soil samples, fungal diversity did not

252 differ with respect to the fertilization ($F_{(1,161)}=0.17$, $p\text{-value}=0.6853$), but did so with respect to
 253 planting ($F_{(1,161)}=9.00$, $p\text{-value}<0.0032$) treatment and species ($F_{(1,161)}=13.03$, $p\text{-value}=0.0003$). In
 254 root samples, fungal diversity differed with respect to the fertilization treatment ($F_{(1,56)}=10.1$, $p\text{-}
 255 value}=0.003$), and the species tested ($F_{(1,56)}=4.5$, $p\text{-value}=0.04$). In soil samples, bacterial diver-
 256 sity differed with respect to the fertilization treatment ($F_{(1,165)}=17.13$, $p\text{-value}<0.0001$), planting
 257 ($F_{(1,165)}=139.0$, $p\text{-value}<0.0001$) but not species ($F_{(1,165)}=1.89$, $p\text{-value}=0.17$). In root samples, bac-
 258 terial diversity differed with respect to the fertilization treatment ($F_{(1,67)}=17.27$, $p\text{-value}=0.0001$),
 259 and the species tested ($F_{(1,67)}=359.69$, $p\text{-value}<0.0001$).



260
 261 **Figure 5: Boxplot of alpha diversity according to the treatment, species and planting effect for**
 262 **fungal-root, fungal-soil, bacteria-soil and bacteria-root.**

263

264

265 *Differences in species composition among sites*

266 Using a PERMANOVA statistical framework, we identified that for all conditions, communities
 267 differed with respect to the fertilization treatment (Table 3). Soil fungal and bacterial communities
 268 differed the most according to whether the tray was planted (greatest % of variance explained,
 269 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)
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

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

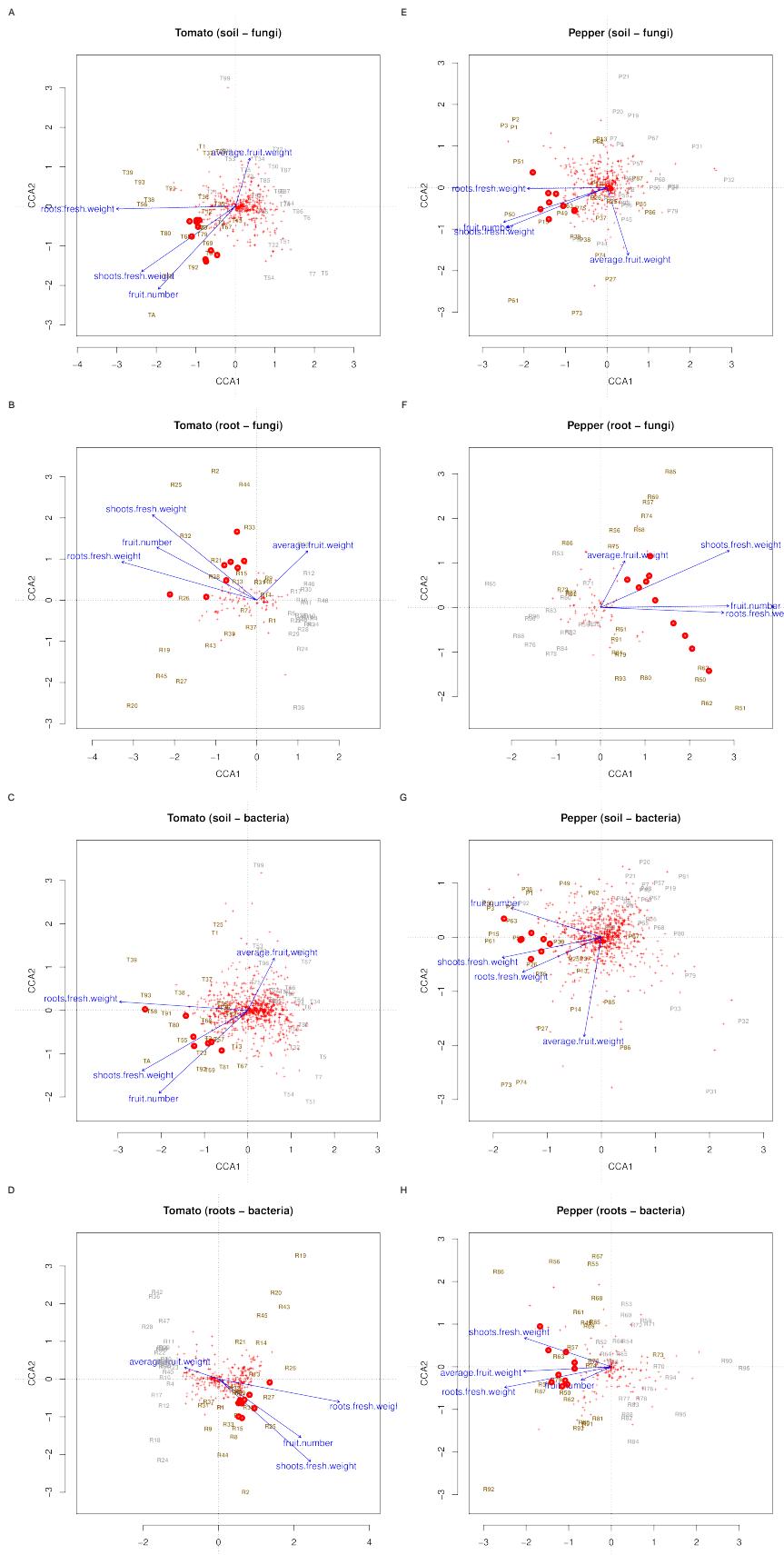
272

273 *Canonical correspondence analyses and candidate ASVs*

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

279

280



²⁸² **Figure 6:** Canonical correspondence analyses for tomato (A-D) and peppers (E-H) for soil-fungi,
²⁸³ root-fungi, soil-bacteria and root-bacteria. Samples were labelled and colored in gray (unfer-
²⁸⁴ tilized) or dark yellow (fertilized). Red crosses represent individual ASVs, while red points
²⁸⁵ represent the ten ASVs most closely associated wih the three productivity measures of root
²⁸⁶ fresh weight, shoots fresh weight and fruit number. Blue arrows are the four productivity mea-
²⁸⁷ sures used as constrains in the ordinations.

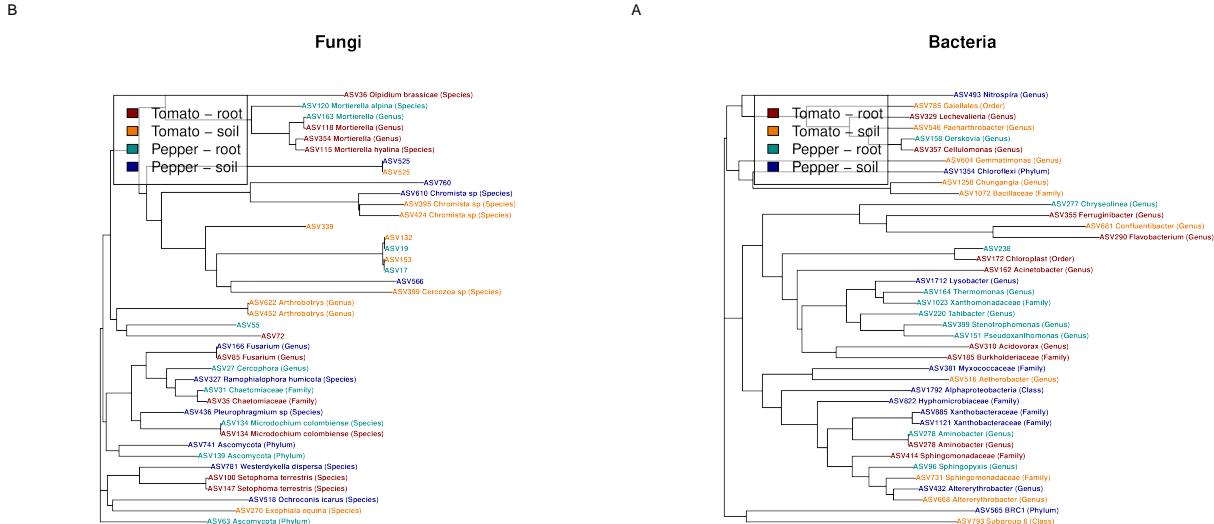
²⁸⁸

Next, we identified, for each ordination, the ten ASVs most closely related to the three constraints which behaved in a similar fashion (productivity measures of root fresh weight, shoots fresh weight and fruit number). These ASVs were considered as putative candidates sequences most positively impacted (increase presence of the ASV) by fertilization. We further analysed the corresponding sequences for these eighty candidates (ten candidates * eight ordinations) ASVs in two separate alignments (one for fungi and one for bacterial ASVs) and their accompanying phylogenetic trees. In fungi, we identified one cluster of ASVs taxonomically assigned to *Mortierella* (soil saprotrophs in the phylum Zygomycota) positively associated to productivity in both tomato and pepper roots. In addition, we identified a cluster of four different ASVs in tomato soil (ASV132, ASV153) and pepper-root (ASV19 & AV17) closely related phylogenetically. Given that no taxonomy was assigned to these sequences through the dada2 RDP bootstrap approach, we used a BLASTn Altschul et al. (1997) approach (against NCBI nr) to identify the most closely related sequences. We identified this cluster of ASVs as *Rhogostoma schuessleri* (BLASTn, *e*-value=4e-76), a protist in the phylum Cercozoa, which are known to be present in the soil and phyllosphere Dumack et al. (2017).

304

In bacteria-roots, we identified a number of different ASVs most positively impacted (increase presence of the ASV) by fertilization. For example, we 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). Also mention that most ASVs are unique AND very rare, such that when
322 we remove the rare ones (found in <10% of the samples), we discard ~90% of ASV, but keep >95%
323 of reads. So this is still a very valuable approach. • Also, talk about advantages / disadvantages
324 of the RDP bayesian classifier approach to taxonomy rather than a BLAST approach (a BLAST
325 approach will always give you a result even if this result is false, an RDP will only give you a tax-
326 onomy (from Kingdom to species) if there is enough power/confidence in it). Also any taxonomy
327 classification largely depends on the database you have.

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

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

332 • Follow-up work: ?

333 **ACKNOWLEDGMENTS**

334 We thank Mengxuan Kong for technical work measuring productivity.

335 REFERENCES

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

- 363 based diversity studies. *Nucleic acids research* 41:e1–e1.
- 364 Oksanen J., Blanchet FG., Kindt R., Legendre P., Minchin PR., O'hara R., Simpson GL., Solymos
365 P., Stevens MHH., Wagner H., others. 2013. Package “vegan”. *Community ecology package, version*
366 2.
- 367 Paradis E., Claude J., Strimmer K. 2004. APE: Analyses of phylogenetics and evolution in r
368 language. *Bioinformatics* 20:289–290.
- 369 Pinheiro J., Bates D., DebRoy S., Sarkar D., Team RC. 2017. Nlme: Linear and nonlinear mixed-
370 effects models. r package version 3.1-128. 2016. *R software*.
- 371 Schliep KP. 2010. Phangorn: Phylogenetic analysis in r. *Bioinformatics* 27:592–593.
- 372 Schloss PD., Westcott SL., Ryabin T., Hall JR., Hartmann M., Hollister EB., Lesniewski RA.,
373 Oakley BB., Parks DH., Robinson CJ., others. 2009. Introducing mothur: Open-source, platform-
374 independent, community-supported software for describing and comparing microbial communi-
375 ties. *Applied and environmental microbiology* 75:7537–7541.
- 376 Team RC., others. 2018. R: A language and environment for statistical computing.
- 377 Toju H., Tanabe AS., Yamamoto S., Sato H. 2012. High-coverage its primers for the dna-based
378 identification of ascomycetes and basidiomycetes in environmental samples. *PloS one* 7:e40863.
- 379 Wang Q., Garrity GM., Tiedje JM., Cole JR. 2007. Naive bayesian classifier for rapid assign-
380 ment of rRNA sequences into the new bacterial taxonomy. *Applied and environmental microbiology*
381 73:5261–5267.
- 382 Wickham H. 2016. *Ggplot2: Elegant graphics for data analysis*. Springer.
- 383 Wickham H., Francois R., Henry L., Müller K. 2015. Dplyr: A grammar of data manipulation.
384 R package version 0.4.3.
- 385 Wright ES. 2016. Using decipher v2. 0 to analyze big biological sequence data in r. *R Journal* 8.