Wild plant species growing closely connected in a subalpine meadow host distinct root-associated bacterial communities.

Kristin Aleklett, Jonathan W Leff, Noah Fierer, Miranda Hart

Plant roots are known to harbor large and diverse communities of bacteria. It has been suggested that plant identity can structure these root-associated communities, but few studies have specifically assessed how the composition of root microbiota varies within and between plant species growing under natural conditions. We assessed the community composition of endophytic and epiphytic bacteria through high throughput sequencing using 16S rDNA derived from root tissues collected from a population of a wild, clonal plant (Orange hawkweed – *Pilosella aurantiaca*) as well as two neighboring plant species (Oxeye daisy – *Leucanthemum vulgare* and Alsike clover – *Trifolium hybridum*).

Our first goal was to determine if plant species growing in close proximity, under similar environmental conditions, still hosted unique root microbiota. Our results showed that plants of different species host distinct bacterial communities in their roots. In terms of community composition, *Betaproteobacteria* (especially the family *Oxalobacteraceae*) were found to dominate in the root microbiota of *L. vulgare* and *T. hybridum* samples, whereas the root microbiota of *P. aurantiaca* had a more heterogeneous distribution of bacterial abundances where *Gammaproteobacteria* and *Acidobacteria* occupied a larger portion of the community.

We also explored the extent of individual variance within each plant species investigated, and found that in the plant species thought to have the least genetic variance among individuals (*P.aurantiaca*) still hosted just as diverse microbial communities. Whether all plant species host their own distinct root microbiota and plants more closely related to each other share more similar bacterial communities still remains to be fully explored, but among the plants examined in this experiment there was no trend that the two species belonging to the same family shared more similarities in terms of bacterial community composition.

1   Kristin Aleklett[1], Jonathan W. Leff[2], Noah Fierer[2], Miranda Hart[1*]

2   *[1]Department of Biology, University of British Columbia – Okanagan CANADA, [2]Cooperative*

3   *Institute for Research in Environmental Sciences and the Department of Ecology and*

4   *Evolutionary Biology, University of Colorado, Boulder, Colorado, USA*

5   *Corresponding author: Miranda Hart, miranda.hart@ubc.ca, +1-250-807-9398*

6   **Introduction**

7   Plant roots function as distinct habitats within the soil and bacterial communities in root systems

8   have repeatedly been shown to differ from those of the surrounding bulk soil (Smalla et al. 2001;

9   Haichar et al. 2008; Gottel et al. 2011; Lundberg et al. 2012). Even though root associated

10   bacterial communities (both *rhizospheric* – in the soil surrounding the roots, *epiphytic* – living at

11   the surface of roots and *endophytic* – living inside root tissues) have been under investigation for

12   many years, there is still little consensus in how these communities are formed and what

13   determines their composition (Berg and Smalla 2009; Aleklett and Hart 2013; Bulgarelli et al.

14   2013).

15   Traditionally, the composition of bacterial communities living in association with plants has

16   been attributed to environmental factors. For example, soil type has been suggested as the

17   strongest determinant of community structure in root associated microbial communities (De

18   Ridder-Duine et al. 2005; Singh et al. 2007; Lundberg et al. 2012; Bulgarelli et al. 2013). At the

19   same time, it has also been argued that the host plant may play an equally large role in

20   determining the composition of its root microbiota (Marschner et al. 2005; Costa et al. 2006;

21   Hartmann et al. 2008; Doornbos et al. 2011), especially endophytic bacterial communities

22   (Haichar et al. 2008).

23   Recent work has demonstrated that hosts can alter their root microbiota by regulating soil

24   conditions in the vicinity of the root system through root exudation of sugars, phenolics and

25   amino acids that could also function as signaling molecules with the microbes in the surrounding

26   soil (Chaparro et al. 2013). Since root exudation patterns and composition can be associated with

27   plant gene expression, variation in host genetics has the potential to create large differences in the

28   chemical profile of plants and consequently the composition of microbes able to inhabit the root

29   system. Several studies have found that different plant species or genotypes of the same species

30   host distinct microbial communities (Bailey et al. 2005; Marschner et al. 2005; van Overbeek and

31   van Elsas 2008; Schweitzer et al. 2008; Micallef et al. 2009a; Manter et al. 2010; Becklin et al.

32 2012; Peiffer et al. 2013). Even studies where soil type was considered to have the strongest

33 effect on structuring the root microbiota, differences in bacterial community composition

34 between genotypes was still detected (Bulgarelli et al. 2012; Lundberg et al. 2012).

35   The root environment varies greatly among plant species (Bardgett et al. 2014); these

36 differences may lead to the selection of distinct bacterial communities. Plants can differ in terms

37 of both root lifespan (Roumet et al. 2006), root architecture (Hodge et al. 2009), root surface

38 structure and components and patterns of root exudation (Bais et al. 2006). Root exudates are

39 known to provide a food source for the microbes (Farrar et al. 2003), instigators of symbiotic

40 associations (such as mycorrhizal infection or nodule formation) (Bais et al. 2004), and defend

41 the plant against pathogens (Doornbos et al. 2011). All these plant characteristics could contribute

42 to shaping root systems of different plants into local habitats and potentially distinct niches for

43 microbial colonizers.

44   The role of intra-species variation among root associated microbial communities has been

45 overlooked, but might represent a significant proportion of variation in natural systems (Bell et

46 al. 2013). Since we know that natural populations exhibit variation in root exudation patterns and

47 root morphology, one would expect there to be variation among individual plants in their root

48 microbiota as well (Micallef et al. 2009). But variation among plants might also be driven by

49 environmental heterogeneity because we know that small-scale environmental heterogeneity

50 exists in soil systems. Is this variation static across plant taxa, or do different taxa exhibit more

51 variation than others? If plant genetics are determining bacterial community composition, then

52 certainly, populations with low genetic diversity (i.e. asexually reproducing, or metapopulations),

53 would be expected to have less variation than sexually reproducing populations with high levels

54 of gene flow. Because we sampled a plant species known to reproduce clonally through stolons

55 and apomixis (*P. aurantiaca*), we also examined whether individuals within that species had less

56 dispersion in their microbial community composition than individuals within the other two plant

57 species.

58   The majority of studies characterizing bacterial communities in the root microbiota have

59 been conducted with model plants in artificial greenhouse settings or agricultural contexts

60 (Marschner and Yang 2001; Garbeva et al. 2004; Micallef et al. 2009; Manter et al. 2010;

61 Doornbos et al. 2011; Lundberg et al. 2012) where the study of genetically modified plants have

62 been especially informing when it comes to understanding slight differences between plant

63 genotypes (van Overbeek and van Elsas 2008; Weinert et al. 2009; İnceoǧlu et al. 2010). While

64 these studies are crucial for understanding the mechanistic basis of plant:microbe interactions,

65 they do not reflect how natural environmental conditions contribute to variation in bacterial

66 community composition across individual plants, particularly in complex environments where a

67 wide diversity of plants and biota are interacting.

68   In this study, we explored variation in bacterial community composition between individual

69 root systems of neighboring plants in a common field in order to determine how much variation

70 exists within and between plant taxa. We sampled the root microbiota of three plant species

71 growing within 10 meters from each other in a field and asked – are bacterial root communities

72 distinct among plant species growing in a common location, close enough to be exposed to the

73 same environmental conditions? And – do certain plant species contain more intra-species

74 variance in bacterial communities than others?

75 **Methods**

76 Field site and target plant

77 *Field site description*

78 Samples were collected in August, 2011, from a subalpine meadow near Chute Lake, British

79 Columbia, Canada (49.698859N,-119.533133W). The sampling area has not been used for

80   agriculture or forestry but is in proximity to a forestry road as well as a camp site. Since it

81   contains a high number of invasive plant species, it could also be considered a disturbed site. The

82   soil at the site was determined to be a sandy loam, and the site is classified under the

83   biogeoclimatic zone Interior Douglas Fir, dry warm (IDFdw) (Biogeoclimatic Ecosystem

84   Classification (BEC) and Ecology Research program of the British Columbia). This zone is

85   characterized by cold winters and dry, warm summers with low amounts of precipitation due to

86   the rain shadow created by the Coast, Cascade and Columbia mountains. Because the climate is

87   very dry in the summers, forest fires are a common feature in the landscape, and the site is

88   located close to the area that was badly burnt during a large fire in Okanagan Mountain Park in

89   2003. The plant community sampled for this experiment was growing up on a rock plateau, away

90   from the tree coverage, completely exposed to the elements (see Fig.1 in additional material). The

91   dominant vegetation at the field site consisted of Orange hawkweed (*Pilosella aurantiaca* (L.)

92   F.W. Schultz and Schultz-Bip), Hairy vetch (*Vicia villosa,* Roth), Oxeye daisy (*Leucanthemum*

93   *vulgare,* Lam.), Wild strawberry (*Fragaria virginiana,* Duchesne) Timothy (*Phleum pratense,* L.)

94   and Alsike clover (*Trifolium hybridum,* L.) growing homogenously, but at different abundances

95   across the field.

96   *Target plant*

97   Our target plant was *P. aurantiaca* (formerly known as *Hieracium aurantiacum*)*,* which is native

98   to Europe and invasive in North America. Because of *P.aurantiaca*'s stoloniferous reproduction,

99   the species is able to form dense mats in the landscape, making it hard for other plant species to

100   compete for space through seed germination in the field (Giroday and Baker 2006). This is one of

101   the features that have made the species a successful invasive plant within British Columbia

102   (Giroday and Baker 2006). Genetic diversity within *P. aurantiaca* has previously been examined

103   across 48 locations in North America, and results showed that there were only three genotypes, of

104   which two were found only in isolated locations (one in Alaska and one in Oregon) (Loomis and

105   Fishman 2009). By choosing to work with a plant expressing this low diversity in wild

106   populations, we hoped to minimize genetic variance within the population that we sampled.

107   To clarify the role of host identity and intra-species variance in bacterial root microbiota, we

108   additionally sampled two of the co-occurring plant species, *L. vulgare* and *T. hybridum* that were

109   in the same developmental stage (flowering) as *P. aurantiaca*.

110   Experimental design

111   Root systems of *P. aurantiaca* were collected one meter apart along two 10 m transects (n=20) in

112   order to explore within species variance of a clonal plant. To be able to compare bacterial

113   community composition between species, additional samples of *T. hybridum* (n=10) and *L.*

114   *vulgare* (n=10) were also collected where present along the transects, several of which were

115   growing within centimeters of *P. aurantiaca* samples. Because *T. hybridum* and *L. vulgare* were

116   not as abundant within the field, their original sample size was lower. For statistical comparisons

117   between species however, an equal sample size of 8 plants/species was used throughout analyses.

118   Each root system was rinsed from surrounding rhizospheric soil in de-ionized water in

119   order to separate it from roots of neighboring plants. Root systems were then cut up in pieces and

120   a subsample of root tissue, representative of the whole root system, including young fresh roots

121   as well as older root tissues (with no exclusion of nodules in *T.hybridum*), was collected and

122   further used for classification of bacterial community composition. Since no further treatment

123   was performed in order to remove rhizoplane microbes, we assume that the communities

124   extracted could be of either endophytic or epiphytic origin.

125   Bacterial community analysis

126   *Amplification and sequencing of target gene*

127   DNA from all collected plant tissues (0.25g/sample) was extracted using a PowerSoil DNA

128   Isolation Kit (MoBio Laboratories Inc., USA) according to the manufacturer's protocol.

129   Microbial diversity and the relative abundances of individual taxa were assessed by barcoded

130   pyrosequencing of a portion of the 16S rRNA gene. Each DNA sample was amplified in triplicate

131   through PCR reactions using the protocol described in Fierer et al. (2008) except with a different

132   primer pair. The forward primer contained the 454 Life Sciences primer B sequence, the bacterial

133   primer 799f (Chelius and Triplett 2001) and a two-base linker sequence ('AG'). The reverse

134   primer contained the 454 Life Sciences primer A sequence, a unique 12 bp error-correcting Golay

135   barcode (Fierer et al. 2008), a 'GT' linker sequence, and the 'universal' bacterial primer 1115r

136   (Reysenbach and Pace 1995). The targeted gene region has shown to be appropriate for accurate

137   taxonomic classification of bacterial sequences and the primers are designed to exclude

138   chloroplasts from plant tissues in the samples (Redford et al. 2010). Amplicons were visualized

139   via gel electrophoresis purified and quantified. Amplicons from all samples were then combined

140   in equimolar ratios into a single tube. Samples were sequenced at Engencore (University of South

141   Carolina) on a Roche GS-FLX sequencer running the Titanium chemistry.

142   *Processing raw sequence data*

143   All sequences were de-multiplexed and further analyzed using the Quantitative Insights Into

144   Microbial Ecology (QIIME) toolkit (Caporaso et al. 2010). Operational Taxonomic Units (OTUs)

145   were defined at the level of ≥97% similarity and the taxonomy assigned by comparing a

146   representative sequence from each OTU to the Greengenes database (February 14th 2012 version)

147   (DeSantis et al. 2006) using the Basic Local Alignment Search Tool (BLAST) classifier (Altschul

148   et al. 1990). In order to correct for differences in the number of sequences analyzed per sample, a

149   randomly selected subset of 400 sequences per root sample was used to compare relative

150 differences in taxonomic diversity. Only samples from which we obtained a minimum of 400

151 bacterial sequences per sample or more were considered in the study, eliminating 3 samples from

152 the study (one *P. aurantiaca* and two *L. vulgare*). Though 400 sequences cannot fully capture the

153 rare biosphere, it allowed us to compare samples while still maintaining as many samples as

154 possible. It has previously been shown that studies of bacterial communities show similar results

155 even at a lower rarefaction (Hamady and Knight 2009; Kuczynski et al. 2010). In fact, re-

156 analyzing our data set with a higher rarefaction limit showed the same general trends but

157 drastically lowered our number of samples available to analyze.

158

159 Statistics

160 Differences in community composition between samples were calculated using phylogenetic

161 metric (UniFrac) where weighted UniFrac shows an emphasis on the more abundant taxa in

162 samples and un-weighted UniFrac treats all taxa the same (Lozupone et al. 2007; Hamady et al.

163 2010). As a comparison, we also included a taxonomic metric (Bray-Curtis distance) to explore

164 whether dissimilarity patterns were the same in terms of presence/absence of taxa. Before

165 calculating Bray-Curtis distances, all relative abundances were log-transformed. 2-D scatterplots

166 of Principal Coordinates Analysis (PCoA) generated in PRIMER-E (Clarke and Gorley, 2006)

167 were used to visualize the greatest amount of variability in the pair-wise distances between

168 samples.

169   We tested for variance among host plants in their root microbiota using a 2-way

170 PerMANOVA (Anderson 2005) with host species and transect as factors and weighted and un-

171 weighted UniFrac as well as Bray Curtis as our dissimilarity metrics. All analyses were

172 permutated 9999 times. Since it has been shown that PerMANOVA is not robust when sampling

173 efforts are un-equal (Anderson and Walsh 2013), we subsampled 8 samples from each species

174 which were used for both PerMANOVA and PermDISP analyses.

175   Variability in community composition within each of the three species was analyzed

176   through PermDISP (Anderson 2004) (9999 permutations), creating a centroid for each species

177   and measuring the average spread of samples belonging to that species from the centroid. A large

178   spread (high average) would indicate a high variability in community composition among

179   individuals within the species (Anderson 2004).

180   **Results**

181   *Variation between host species*

182   When comparing the phylogenetic overlap between bacterial root microbiota (UniFrac) across

183   three different species of plant hosts growing in a common field, bacterial communities from

184   samples of the same plant species were significantly more similar to each other than to bacterial

185   communities sampled from plants of the two other species (Table 1). This was true for both

186   weighted (Pseudo-F=8.54 p=0.0001) and un-weighted UniFrac (Pseudo-F=1.66 p=0.0001) as

187   well as Bray Curtis dissimilarities (Pseudo-F=2.27 p=0.0001) (Table 1). These patterns were also

188   evident from the principal coordinates analyses which showed little overlap between samples of

189   different plant species (Fig.1).

190   *Variation within host species*

191   Plant species differed in how much variance there was among bacterial communities of

192   individual root samples (Fig.1). There was a significant difference between plant species in the

193   amount of compositional dissimilarity of bacterial taxa between individual plants (Bray Curtis:

194   F=9.56 p=0.02). That is, the amount of dispersion of individual plants from the centroid differed

195   among plant species. In this case, *P. aurantiaca* exhibited the most variance among plant

196   individuals, while *T. hybridum* showed the least (Table 2). This was not the case when the same

197   data was analyzed using phylogenetic measures, since we could not detect any significant

198   difference in dispersion among plant taxa (UniFrac: weighted F=2.57 p=0.15; un-weighted

199   F=2.56 p=0.63) (Table 2).

200   *Relative abundance of taxa across hosts*

201   A total number of 4384 unique OTUs were analyzed within the rarefied data set. A taxonomic

202   summary, showing the average abundance of bacterial phyla in *P. aurantiaca, T. hybridum* and *L.*

203   *vulgare* samples, illustrates the compositional differences between root systems of different plant

204   species (Fig.2). In *P. aurantiaca,* the most abundant phylum was *Betaproteobacteria* which made

205   up, on average, 29% out of all sequences found in *P. aurantiaca* samples, followed by

206   *Bacteroidetes* (19%), *Alphaproteobacteria* (16%) and *Actinobacteria* (12%). In *T. hybridum*,

207   *Betaproteobacteria* made up, on average, 51% of the all bacterial sequences found in the species,

208   followed by *Alphaproteobacteria* (21%) and *Bacteroidetes* (16%). Bacterial communities in *L.*

209   *vulgare* samples were, similarly to *T. hybridum*, dominated by *Betaproteobacteria* (50%),

210   followed by *Bacteroidetes* (18%) and *Alphaproteobacteria* (12%).   A table showing the relative

211   distribution of the 14 most abundant OTUs across all samples as well as their relative abundance

212   within samples of the different species is given in Table 3.

213   A closer examination of the relative abundances of all *Betaproteobacteria* found in samples

214   showed that while *Burkholderiales* was the predominant order across all three species, *T.*

215   *hybridum* and *L. vulgare* samples were heavily dominated by bacteria of the family

216   *Oxalobacteriaceae* – especially bacteria of the genus *Herbaspirillum* (11% of the total bacterial

217   community in *T. hybridum* and 18% in *L. vulgare*) (Fig.3).

218   **Discussion**

219   *Host specificity*

220   Our study shows that root bacterial communities vary significantly between plants belonging to

221   three different species, growing in close proximity to each other in natural plant communities.

222   Because the plant roots collected in our study were growing in close proximity to each other it is

223   unlikely that variation in soil conditions (usually thought to be one of the main drivers of

224   microbial community structure) is a significant source of variation among our samples. While

225   these results need to be further confirmed across multiple field sites in order to draw general

226   conclusions about whether this is just a local trend or a general pattern across plant species and

227   soil types, our results support previous work showing bacterial host plant specificity in roots of

228   agricultural crops (Marschner and Yang 2001; Wieland et al. 2001; Haichar et al. 2008) and wild

229   grass species (Kuske and Ticknor 2002; Osanai et al. 2012).

230   Although all plant species investigated in this study (*P. aurantiaca, T. hybridum* and *L.*

231   *vulgare*), are perennial, there are significant differences in root morphology between the species.

232   For example, *P. aurantiaca* and *L. vulgare* (both belonging to the family *Asteraceae*) have

233   creeping root stocks and produce fibrous root systems whereas *T. hybridum* (family *Fabaceae*)

234   grows a branching tap root system that is known to form nodules with nitrogen fixing bacteria.

235   This variation in root morphology could contribute to the differences in abundance and

236   composition of bacteria in our results (Bardgett et al. 2014). For example, roots that penetrate

237   deeper soil may encounter different microbes than those in shallow layers (Fierer et al. 2003).

238   Similarly, the thickness and/or texture of the root surface (i.e. woody, fibrous) may be more or

239   less penetrable to colonizing bacteria.

240   Part of the variance seen in bacterial community composition between the three plant

241   species could also be caused by species-specific root exudation patterns. For example, several

242   members of the *Asteraceae* family are known to produce allelochemicals that could affect the

243   bacterial community as well as surrounding plants (Alford et al. 2009). However, these

244   differences are difficult to assess in wild plant communities, especially when roots of different

245   plant species grow in close proximity to each other with entangled root systems. In our study, the

246   roots grew so intimately that exudation from one plant species could have influenced root

247   systems of neighboring plants.

248   *Individual variation in root microbiota*

249   We know that genetic differences between plants, even at the genotype level, can affect the

250   composition of the root microbiota (Bailey et al. 2005; Schweitzer et al. 2008; Peiffer et al.

251   2013). Thus, we would expect variation in bacterial community composition among individuals

252   within a population of plants, even when they are growing in a common environment, due to

253   genetic variation in the population in terms of root traits and exudation chemistry, among other

254   factors. Though there is a potential for clonality in *T. hybridum*, we still predicted less individual

255   variance in the root microbiota among individuals from *P. aurantiaca* (thought to consist of

256   mainly one genotype across all of North America (Loomis and Fishman 2009)) than within the

257   two co-occurring out-crossing plant species with presumed higher genetic diversity (*L. vulgare*

258   *and T. hybridum)*.

259   Though our data show a significant difference in compositional turnover within different

260   plant species, it rejects the hypothesis that *P. aurantiaca* had the most similar root communities

261   across individuals. Comparing the average dispersion of bacterial community composition for the

262   three plant species, there was no indication that *P. aurantiaca* had a smaller dispersion than the

263   two other plant species (Table 2). Instead, it shows that *P. aurantiaca* had the highest variation

264   within a species comparing dispersion based on taxonomic differences (Table 2). The fact that we

265   could not detect any differences in dispersion when using phylogenetic metrics suggests that

266   individual root systems differ more in terms of which taxa are present or absent than how related

267   they are, or that there is little phylogenetic conservatism at the individual level.

268   Overall, this study shows that the extent of individual variation seen in root microbiota

269   varies between species, but that a plant species thought to be more genetically homogenous does

270   not necessarily host more homogeneous root communities. It also indicates that individual

271   variation in bacterial community composition in root systems is determined, not only by plant

272   genetics, but also by small scale variance in the surrounding environment (soil chemistry,

273   humidity, biotic interactions etc.) and potentially, events throughout the plant's life that could

274   affect root colonization (Aleklett andand Hart 2013).

275   *Bacterial community composition*

276   Similar patterns of bacterial community composition to what we found in our plants, growing in a

277   subalpine meadow in Canada, have been reported in rhizosphere samples of other studies. For

278   example, roots tissues of the plant species that we sampled were mainly dominated by

279   *Betaproteobacteria*, (Fig. 2), especially members of the order *Burkholderiales* and the family

280   *Oxalobacteriaceae*, which represented as much as 32% of the total bacterial community in *L.*

281   *vulgare* (Fig. 3). Seed- and root-colonizing populations of *Oxalobacteriaceae* have previously

282   shown to be responsive to plant species (Green et al. 2007), supporting our data of plant species

283   hosting distinct bacterial communities. Dominance by these taxa in root systems has also been

284   reported in other studies. For example, roots of *Arabidopsis thaliana*, examined at the same

285   taxonomic resolution by Lundberg et al. (2012), were dominated by *Betaproteobacteria* and

286   *Oxalobacteriaceae* both in samples of rhizosphere soil as well as in the endophytic root

287   compartment. In sphagnum mosses, *Burkholderiales* has also been documented as one of the

288   dominant bacterial groups, thought to be behind the production of anti-fungal compounds and

289   anti-microbial properties of the sphagnum mosses (Opelt et al. 2007).

290   Other studies have found *Actinobacteria* to dominate in root tissues of plants (e.g. Ottesen

291   et al. 2013), especially in communities of the endophytic compartment (Bodenhausen et al.

292   2013). In our study, *Actinobacteria* represented at most 12% out of the total bacterial community

293   in the plant species that we sampled (Fig. 2) and was mainly found in samples of *P. aurantiaca*

294   that, in general, were less dominated by beta-proteobacteria.

295   The dominance of sequences belonging to the genus *Herbaspirillum* was further

296   emphasized when we examined the fourteen most abundant OTUs across all samples (Table 3).

297   *Herbaspirillum* spp. are known to colonize apoplastic or intracellular spaces of plant tissues and

298   several species have shown the ablility to fix nitrogen (Schmid et al. 2006). While it is believed

299   that this nitrogen fixing ability could be beneficial to their plant host, it has also been documented

300   that certain *Herbaspirillum* strains are mild pathogens and a causative agent of "mottled stripe

301   disease" in crops such as sugar-cane (Schmid et al. 2006). Besides *Herbaspirillum*, we also saw

302   high abundances of sequences belonging to *Limnohabitans* and *Cytophaga* (Table 3), two genera

303   more commonly associated with bacterial communities in fresh water (Kirchman 2002; Simek et

304   al. 2010) as well as the species *Methylibium petroleiphilum*, a recognized methylotroph (Kane et

305   al. 2007) and *Janthinobacterium lividum*, known to thrive in soils (Shivaji et al. 1991) and

306   produce antibiotics (Johnson et al. 1990). The high presence of these groups in our samples could

307   be due to the inclusion of epiphytic members of the root microbiota, where bacteria associated

308   with water films and soil particles of the root surface would be expected.

309   In comparison, Bodenhausen and colleagues (2013) found that a *Flavobacterium* of the

310   phylum *Bacteroidetes* stood out as the single most abundant OTU in endophytic root samples,

311   making up 10.15% of the total community. Though bacteria of the phylum *Bacteroidetes*

312   represented a significant part of the community in root samples of the three plant species sampled

313   in our study (Fig.2), they were by no means the most dominant taxonomic group in any of the

314   species (Table 3).

315   As the genus *Trifolium* are known to be hosts of nitrogen fixing bacteria that form nodules

316   in their roots, *T.hybridum* samples were expected to host larger populations of

317   *Alphaproteobacteria*, specifically belonging to the order *Rhizobiales* which is a common

318   symbiont of legumes (Masson-Boivin et al. 2009).This expected pattern was not evident in our

319   results though. The only OTU belonging to the order *Rhizobiales* detected in notable abudances

320   in our study was a *Bradyrhizobium* taxa which made up 1% of the collective community of *T.*

321   *hybridum* samples and 1.5% in *P. aurantiaca* samples (Table 3). Instead, it was evident that the *T.*

322   *hybridum* community was dominated by the family *Oxalobacteraceae* (23.48%) (Fig.3) and

323   specifically one OTU of the genus *Herbaspirillum* (10.75%) (Table 3), which is mainly known to

324   colonize roots of non-leguminous plants, and have nitrogen fixing properties (Baldani et al.1997).

325   What stands out though is that this group of bacteria was even more predominant in *L. vulgare*

326   samples, where *Oxalobacteraceae* made up 32.47% of the community and the same

327   *Herbaspirillum* OTU represented 18.26% of the total community.

328   *Variance in relative abundances of bacterial taxa across plant species*

329   We observed differences in the evenness of bacterial taxa across host plants. While the roots of *L.*

330   *vulgare* and *T. hybridum* seemed dominated by a few select groups of microbes, samples of *P.*

331   *aurantiaca* supported communities with abundances more evenly distributed among bacterial

332   taxa (Fig. 2; Table 3). Though few studies have looked specifically at variance in bacterial

333   evenness between plant species, it could be an important source of variation. For example,

334   dominance of single taxon may indicate specialized plant/bacterial associations whereas high

335   evenness in community composition could reflect generalist associations among plants and

336   bacteria. Alternatively, differences in evenness may result from microbial interactions within the

337   plant, not driven by the plant but microbial competition for plant resources.

338   *Conclusions*

339 In this study, we showed that plant identity plays a major role in explaining the variation seen in

340 root microbiota both between and within plant species growing under natural conditions. Further

341 studies across a larger set of wild plant species and natural sites as well as more detailed

342 investigations of the effect of plant genetics versus plant phenotypic traits on bacterial

343 community assembly could help resolve the relative contribution of host identity at an individual

344 level in shaping the root microbiota. It would also allow us to draw further conclusions as to

345 whether plats more related to each other actually host more similar bacterial communities across

346 plant species and families.

347   The results of our study speak of how intimately related bacterial communities are with

348 their host plants. Root systems of wild plants are never alone; they are constantly surrounded by

349 the roots of other plants, entangled in the soil, competing for resources and space. Yet, our results

350 show that bacterial communities associated with roots of plants growing in a common field are

351 distinct between plant species.

352   Ultimately, we are not able to tell exactly why these three plant species have such distinct

353 bacterial root communities, but further studies linking metabolomics of wild plants with bacterial

354 community composition would be useful for better understanding how plants affect bacterial

355 community assembly.

356 **Acknowledgements**

357 We would like to thank Monika Gorzelak for helping out with sampling efforts in the field.

358 **References**

359 Aleklett K, Hart M. 2013. The root microbiota—a fingerprint in the soil? *Plant and Soil* 370:
360   671–686. doi:10.1007/s11104-013-1647-7.

361   Alford ÉR, Vivanco JM, Paschke MW. 2009. The effects of flavonoid allelochemicals from
362   knapweeds on legume-rhizobia candidates for restoration. *Restoration Ecology* 17: 506–514.
363   doi:10.1111/j.1526-100X.2008.00405.x.

364   Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool.
365   *Journal of Molecular Biology* 215: 403–10. doi:10.1016/S0022-2836(05)80360-2.

366   Anderson MJ. 2004. PERMDISP: a FORTRAN computer program for permutational analysis of
367   multivariate dispersions (for any two-factor ANOVA design) using permutation tests.
368   Department of Statistics, University of Auckland, New Zeeland

369   Anderson MJ. 2005. PERMANOVA: a FORTRAN computer program for permutational
370   multivariate analysis of variance. Department of Statistics, University of Auckland, New
371   Zealand.

372   Anderson MJ, Walsh DCI. 2013. PERMANOVA, ANOSIM, and the Mantel test in the face of
373   heterogeneous dispersions: What null hypothesis are you testing? *Ecological Monographs*
374   83: 557–574.

375   Bailey JK, Deckert R, Schweitzer JA, Rehill BJ, Lindroth RL, Gehring C, Whitham TG. 2005.
376   Host plant genetics affect hidden ecological players: links among Populus , condensed
377   tannins, and fungal endophyte infection. *Canadian Journal of Botany* 83: 356–361.
378   doi:10.1139/b05-008.Bais HP, Park S-W, Weir TL, Callaway RM, Vivanco JM. 2004. How
379   plants communicate using the underground information superhighway. *Trends in Plant*
380   *Science* 9: 26–32. doi:10.1016/j.tplants.2003.11.008.

381   Bais HP, Weir TL, Perry LG, Gilroy S, Vivanco JM. 2006. The role of root exudates in
382   rhizosphere interactions with plants and other organisms. *Annual Review of Plant Biology*
383   57: 233–66. doi:10.1146/annurev.arplant.57.032905.105159.

384   Baldani J, Caruso L, Baldani V, Goi S, Döbereiner J. 1997. Recent advances in BNF with non-
385   legume plants. *Soil Biology and Biochemistry* 29: 911–922.

386   Bardgett RD, Mommer L, De Vries FT. 2014. Going underground: root traits as drivers of
387   ecosystem processes. *Trends in Ecology and Evolution* 29: 692-699.

388   Becklin KM, Hertweck KL, Jumpponen A. 2012. Host identity impacts rhizosphere fungal
389   communities associated with three alpine plant species. *Microbial Ecology* 63: 682–93.
390   doi:10.1007/s00248-011-9968-7.

391   Bell C, Carrillo Y, Boot CM, Rocca JD, Pendall E, Wallenstein MD. 2013. Rhizosphere
392   stoichiometry: are C : N : P ratios of plants, soils, and enzymes conserved at the plant
393   species-level? *New Phytologist*: 1–13. doi:10.1111/nph.12531.

394   Berg G, Smalla K. 2009. Plant species and soil type cooperatively shape the structure and
395   function of microbial communities in the rhizosphere. *FEMS Microbiology Ecology* 68: 1–
396   13. doi:10.1111/j.1574-6941.2009.00654.x.

397 Bodenhausen N, Horton MW, Bergelson J. 2013. Bacterial communities associated with the
398   leaves and the roots of Arabidopsis thaliana. *PloS one* 8: e56329.
399   doi:10.1371/journal.pone.0056329.

400 Bulgarelli D, Rott M, Schlaeppi K, Ver Loren van Themaat E, Ahmadinejad N, Assenza F, Rauf
401   P, Huettel B, Reinhardt R, Schmelzer E, Peplies J, Gloeckner FO, Amann R, Eickhorst T,
402   Schulze-Lefert P. 2012. Revealing structure and assembly cues for Arabidopsis root-
403   inhabiting bacterial microbiota. *Nature* 488: 91–5. doi:10.1038/nature11336.

404 Bulgarelli D, Schlaeppi K, Spaepen S, Ver Loren van Themaat E, Schulze-Lefert P. 2013.
405   Structure and functions of the bacterial microbiota of plants. *Annual Review of Plant*
406   *Biology* 64: 807–38. doi:10.1146/annurev-arplant-050312-120106.

407 Caporaso J, Kuczynski GJ, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Peña
408   AG, Goodrich JK, Gordon JI, Huttley GA, Kelley ST, Knights D, Koenig JE, Ley RE,
409   Lozupone CA, McDonald D, Muegge BD, Pirrung M, Reeder J, Sevinsky JR, Turnbaugh PJ,
410   Walters WA, Widmann J, Yatsunenko T, Zaneveld J, Knight R. 2010. QIIME allows analysis
411   of high-throughput community sequencing data. *Nature Methods* 7: 335–6.
412   doi:10.1038/nmeth.f.303.

413 Chaparro JM, Badri DV, Bakker MG, Sugiyama A, Manter DK, Vivanco JM. 2013. Root
414   exudation of phytochemicals in Arabidopsis follows specific patterns that are
415   developmentally programmed and correlate with soil microbial functions. *PloS one* 8:
416   e55731. doi:10.1371/journal.pone.0055731.

417 Chelius MK, Triplett EW. 2001. The diversity of archaea and bacteria in association with the
418   roots of Zea mays L. *Microbial Ecology* 41: 252–263. doi:10.1007/s002480000087.

419 Clarke KR, Gorley RN. 2006. PRIMER v6: User Manual/Tutorial. Plymouth Marine Laboratory,
420   Plymouth, UK.

421 Costa R, Götz M, Mrotzek N, Lottmann J, Berg G and Smalla K. 2006. Effects of site and plant
422   species on rhizosphere community structure as revealed by molecular analysis of microbial
423   guilds. *FEMS Microbiology Ecology* 56: 236–49. doi:10.1111/j.1574-6941.2005.00026.x.

424 DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, Huber T, Dalevi D, Hu P,
425   Andersen GL. 2006. Greengenes, a chimera-checked 16S rRNA gene database and
426   workbench compatible with ARB. *Applied and Environmental Microbiology* 72: 5069–72.
427   doi:10.1128/AEM.03006-05.

428 Doornbos RF, van Loon LC, Bakker PAHM. 2011. Impact of root exudates and plant defense
429   signaling on bacterial communities in the rhizosphere. A review. A*gronomy for Sustainable*
430   *Development* 32: 227–243. doi:10.1007/s13593-011-0028-y.

431 Farrar J, Hawes M, Jones D, Lindow S. 2003. How roots control the flux of carbon to the
432   rhizosphere. *Ecology* 84: 827–837.

433   Fierer N, Hamady M, Lauber CL, Knight R. 2008. The influence of sex, handedness, and
434   washing on the diversity of hand surface bacteria. *Proceedings of the Natural Academy of*
435   *Science USA* 105: 17994–9. doi:10.1073/pnas.0807920105.

436   Fierer N, Schimel JP, Holden PA. 2003. Variations in microbial community composition through
437   two soil depth profiles. *Soil Biology and Biochemistry* 35: 167–176. doi:10.1016/S0038-
438   0717(02)00251-1.

439   Garbeva P, van Veen JA and van Elsas JD. 2004. Microbial diversity in soil: selection microbial
440   populations by plant and soil type and implications for disease suppressiveness. *Annual*
441   *review of phytopathology* 42: 243–70. doi:10.1146/annurev.phyto.42.012604.135455.

442   Giroday H, Baker V. 2006. Invasive hawkweeds (Hieracium spp.) in Northeastern British
443   Columbia - Invasive plant risk assessment and literature search. *British Coulmbia Ministry*
444   *of Forests and Range*, Prince George, Canada.

445   Gottel, NR, Castro HF, Kerley M, Yang Z, Pelletier DA, Podar M, Karpinets T, Uberbacher E,
446   Tuskan GA, Vilgalys R, Doktycz MJ, Schadt CW. 2011. Distinct microbial communities
447   within the endosphere and rhizosphere of Populus deltoides roots across contrasting soil
448   types. *Applied and Environmental Microbiology* 77: 5934–44. doi:10.1128/AEM.05255-11.

449   Green SJ, Michel FC, Hadar Y, Minz D. 2007. Contrasting patterns of seed and root colonization
450   by bacteria from the genus Chryseobacterium and from the family Oxalobacteraceae. *The*
451   *ISME Journal* 1: 291–9. doi:10.1038/ismej.2007.33.

452   Haichar FEZ, Marol C, Berge O, Rangel-Castro JI, Prosser JI, Balesdent J, Heulin T, Achouak W.
453   2008. Plant host habitat and root exudates shape soil bacterial community structure. *The*
454   *ISME journal* 2: 1221–30. doi:10.1038/ismej.2008.80.

455   Hamady M, Knight R. 2009. Microbial community profiling for human microbiome projects:
456   Tools, techniques, and challenges. *Genome research*, 19, 1141–52.

457   Hamady M, Lozupone C, R. Knight. 2010. Fast UniFrac: facilitating high-throughput
458   phylogenetic analyses of microbial communities including analysis of pyrosequencing and
459   PhyloChip data. *The ISME Journal* 4: 17–27. doi:10.1038/ismej.2009.97.

460   Hartmann A, Schmid M, van Tuinen D, Berg G. 2008. Plant-driven selection of microbes. *Plant*
461   *and Soil* 321: 235–257. doi:10.1007/s11104-008-9814-y.

462   Hodge A, Berta G, Doussan C, Merchan F, Crespi M. 2009. Plant root growth, architecture and
463   function. *Plant and Soil* 321: 153–187. doi:10.1007/s11104-009-9929-9.

464   Inceoğlu O, Salles JF, van Overbeek L, van Elsas JD. 2010. Effects of plant genotype and growth
465   stage on the betaproteobacterial communities associated with different potato cultivars in
466   two fields. *Applied and Environmental Microbiology*, 76, 3675–84.

467   Kane SR, Chakicherla AY, Chain PSG, Schmidt R, Shin MW, Legler TC, Scow KM, Larimer
468   FW, Lucas SM, Richardson PM, Hristova KR. 2007. Whole-genome analysis of the methyl

469   tert-butyl ether-degrading beta-proteobacterium Methylibium petroleiphilum PM1. *Journal*
470   *of Bacteriology* 189: 1931–45. doi:10.1128/JB.01259-06.

471   Kirchman DL. 2002. The ecology of Cytophaga-Flavobacteria in aquatic environments. *FEMS*
472   *Microbiology Ecology* 39: 91–100. doi:10.1111/j.1574-6941.2002.tb00910.x.

473   Kuczynski J, Liu Z, Lozupone C, McDonald D, Fierer N, Knight R. 2010. Microbial community
474   resemblance methods differ in their ability to detect biologically relevant patterns. *Nature*
475   *methods* 7: 813–9.

476   Kuske C, Ticknor L. 2002. Comparison of soil bacterial communities in rhizospheres of three
477   plant species and the interspaces in an arid grassland. *Applied and Environmental*
478   *Microbiology* 68: 1854–1863. doi:10.1128/AEM.68.4.1854–1863.2002.

479   Loomis ES, Fishman L. 2009. A continent wide clone: population genetic variation of the
480   invasive plant *Hieracium aurantiacum* (Orange Hawkweed; Asteraceae) in North America.
481   *International Journal of Plant Science* 170: 759–765. doi:10.1086/599241.

482   Lozupone CA, Hamady M, Kelley ST, Knight R. 2007. Quantitative and qualitative beta diversity
483   measures lead to different insights into factors that structure microbial communities. *Applied*
484   *and Environmental Microbiology* 73: 1576–85. doi:10.1128/AEM.01996-06.

485   Lundberg DS, Lebeis SL, Paredes SH, Yourstone S, Gehring J, Malfatti S, Tremblay J,
486   Engelbrektson A, Kunin V, del Rio TG, Edgar RC, Eickhorst T, Ley RE, Hugenholtz P,
487   Tringe SG, Dangl JL. 2012. Defining the core Arabidopsis thaliana root microbiome. *Nature*
488   488: 86–90. doi:10.1038/nature11237.

489   Manter DK, Delgado JA, Holm DG, Stong RA. 2010. Pyrosequencing reveals a highly diverse
490   and cultivar-specific bacterial endophyte community in potato roots. *Microbial Ecology* 60:
491   157–66. doi:10.1007/s00248-010-9658-x.

492   Marschner P, Grierson PF, Rengel Z. 2005. Microbial community composition and functioning in
493   the rhizosphere of three Banksia species in native woodland in Western Australia. *Applied*
494   *Soil Ecology* 28: 191–201. doi:10.1016/j.apsoil.2004.09.001.

495   Marschner P, Yang C. 2001. Soil and plant specific effects on bacterial community composition
496   in the rhizosphere. Soil Biology and Biochemistry 33: 1437–1445.

497   Masson-Boivin C, Giraud E, Perret X, Batut J. 2009. Establishing nitrogen-fixing symbiosis with
498   legumes: how many rhizobium recipes? *Trends in Microbiology* 17: 458–66.

499   Micallef SA, Channer S, Shiaris MP, Colón-Carmona A. 2009. Plant age and genotype impact the
500   progression of bacterial community succession in the Arabidopsis rhizosphere. *Plant*
501   *signaling and behavior* 4: 777–80. doi:10.1093/jxb/erp053.

502   Osanai Y, Bougoure DS, Hayden HL, Hovenden MJ. 2012. Co-occurring grass species differ in
503   their associated microbial community composition in a temperate native grassland. *Plant*
504   *and Soil* 368: 419–431. doi:10.1007/s11104-012-1529-4.

505 Ottesen AR, González Peña A, White JR, Pettengill JB, Li C, Allard S, Rideout S, Allard M, Hill
506   T, Evans P, Strain E, Musser S, Knight R, Brown E. 2013. Baseline survey of the anatomical
507   microbial ecology of an important food plant: Solanum lycopersicum (tomato). *BMC*
508   *Microbiology* 13: 114. doi:10.1186/1471-2180-13-114.

509 Opelt K, Chobot V, Hadacek F, Schönmann S, Eberl L and Berg G. 2007. Investigations of the
510   structure and function of bacterial communities associated with Sphagnum mosses.
511   Environmental microbiology 9: 2795–809. doi:10.1111/j.1462-2920.2007.01391.x.

512 van Overbeek L, van Elsas JD. 2008. Effects of plant genotype and growth stage on the structure
513   of bacterial communities associated with potato (Solanum tuberosum L.). *FEMS*
514   *Microbiology Ecology* 64: 283–96. doi:10.1111/j.1574-6941.2008.00469.x.

515 Peiffer JA, Spor A, Koren O, Jin Z, Tringe SG, Dangl JL, Buckler ES, Ley RE. 2013. Diversity
516   and heritability of the maize rhizosphere microbiome under field conditions. *Proceedings of*
517   *the Natural Acadademy of Science USA* 110: 6548–53. doi:10.1073/pnas.1302837110.

518 Redford, A. J., Bowers RM, Knight R, Y. Linhart and N. Fierer. 2010. The ecology of the
519   phyllosphere: geographic and phylogenetic variability in the distribution of bacteria on tree
520   leaves. *Environmental microbiology* 12: 2885–93. doi:10.1111/j.1462-2920.2010.02258.x.

521 de Ridder-Duine AS, Kowalchuk GA, Klein Gunnewiek PJA, Smant W, van Veen JA, de Boer W.
522   2005. Rhizosphere bacterial community composition in natural stands of Carex arenaria
523   (sand sedge) is determined by bulk soil community composition. *Soil Biology and*
524   *Biochemistry* 37: 349–357. doi:10.1016/j.soilbio.2004.08.005.

525 Roumet C, Urcelay C., Díaz S. 2006. Suites of root traits differ between annual and perennial
526   species growing in the field. *The New Phytologist* 170: 357–68. doi:10.1111/j.1469-
527   8137.2006.01667.x.

528 Schmid M, Baldani JIVO, Hartmann A. 2006. The Prokaryotes. Pages 141–150 (M. Dworkin, S.
529   Falkow, E. Rosenberg, K.-H. Schleifer, and E. Stackebrandt, Eds.). 141–150, New York, NY.

530 Schweitzer JA, Bailey JK, Fischer DG, LeRoy CJ, Lonsdorf EV, Whitham TG, Hart SC. 2008.
531   Plant–soil–microorganism interactions: heritable relationship between plant genotype and
532   associated soil microorganisms. *Ecology* 89: 773–781. doi:10.1890/07-0337.1.

533 Shivaji S, Ray M, Kumar G, Reddy G, Saisree L, Wynn-Williams D. 1991. Identification of
534   Janthinobacterium lividum from the soils of the islands of Scotia Ridge and from Antarctic
535   peninsula. *Polar biology* 11: 267–271.

536 Simek K, Kasalicky V, Jezbera J, Jezberová J, Hejzlar J, Hahn MW. 2010. Broad habitat range of
537   the phylogenetically narrow R-BT065 cluster, representing a core group of the
538   Betaproteobacterial genus *Limnohabitans*. *Applied and Environmental Microbiology* 76:
539   631–9. doi:10.1128/AEM.02203-09.

540 Singh, B. K., S. Munro, J. M. Potts and P. Millard. 2007. Influence of grass species and soil type
541   on rhizosphere microbial community structure in grassland soils. *Applied Soil Ecology* 36:
542   147–155. doi:10.1016/j.apsoil.2007.01.004.

543 Smalla K, Wieland G, Buchner A, Zock A, Parzy J, Kaiser S, Roskot N. 2001. Bulk and
544   rhizosphere soil bacterial communities studied by denaturing gradient gel electrophoresis :
545   plant-dependent enrichment and seasonal shifts revealed. *Applied and Environmental*
546   *Microbiology* 67: 4742–4751. doi:10.1128/AEM.67.10.4742.

547 Weinert N, Meincke R, Gottwald C, Heuer H, Gomes NCM, Schloter M, Berg G, Smalla K.
548   2009. Rhizosphere communities of genetically modified zeaxanthin-accumulating potato
549   plants and their parent cultivar differ less than those of different potato cultivars. *Applied*
550   *and Environmental Microbiology* 75: 3859-3865.

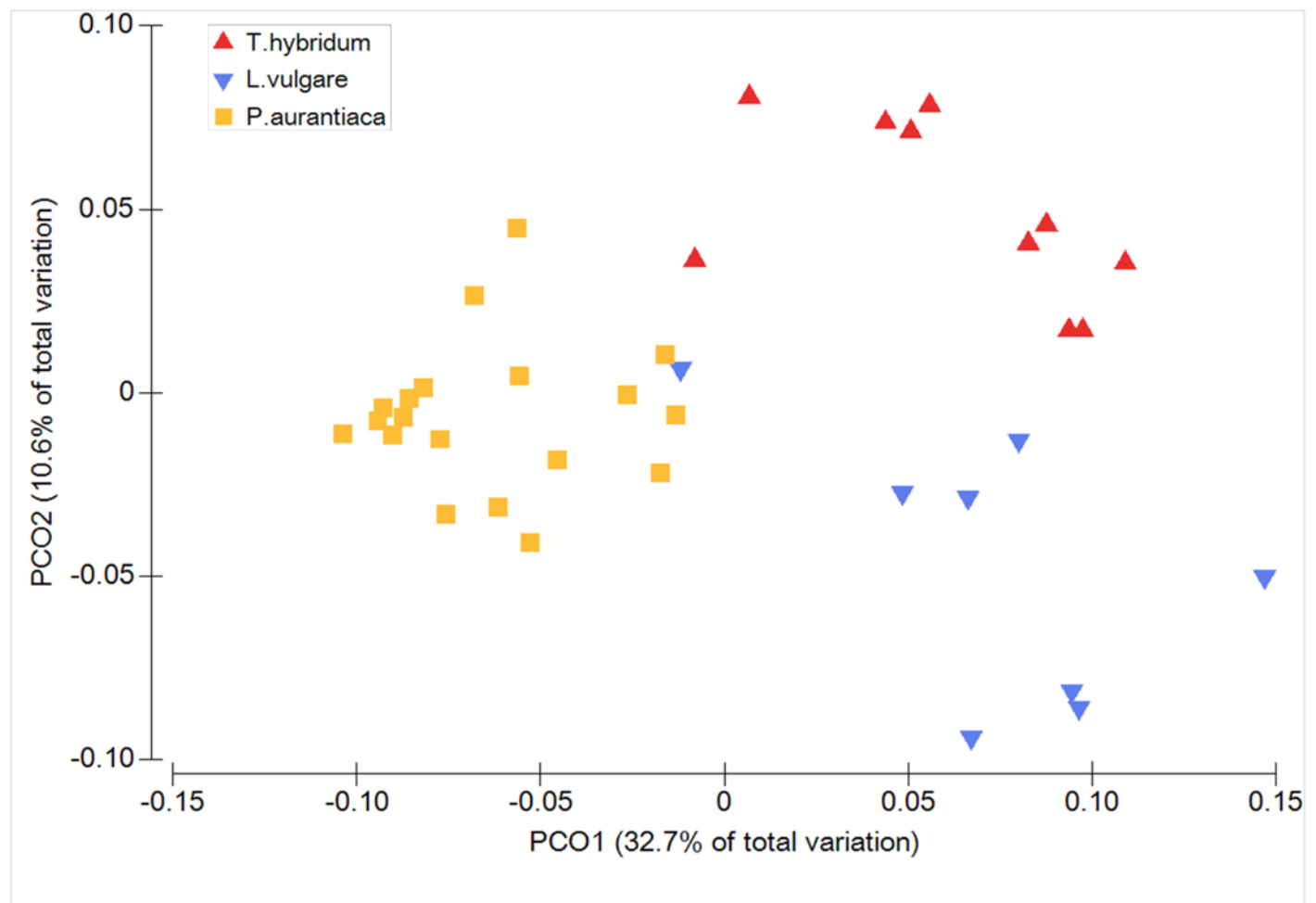
551 Wieland G, Neumann R, Backhaus H. 2001. Variation of microbial communities in soil,
552   rhizosphere, and rhizoplane in response to crop species, soil type, and crop development.
553   *Applied and Environmental Microbiology* 67: 5849–5854. doi:10.1128/AEM.67.12.5849–
554   5854.2001.

PerMANOVA results, comparing bacterial community resemblance between plant species and transects using different diversity metrics.

Factor	Diversity metric	Pseudo-F	P (perm)
Species	Weighted UniFrac	6.15	0.0001
	Un-weighted UniFrac	1.43	0.0001
	Bray Curtis	1.88	0.0001
Transect	Weighted UniFrac	1.22	0.23
	Un-weighted UniFrac	1.06	0.22
	Bray Curtis	1.07	0.24
SpXTr	Weighted UniFrac	1.54	0.09
	Un-weighted UniFrac	1.03	0.28
	Bray Curtis	1.06	0.20

1

Principal coordinates analysis plot illustrating the phylogenetic overlap in root prokaryotic community composition among samples from three different plant species.

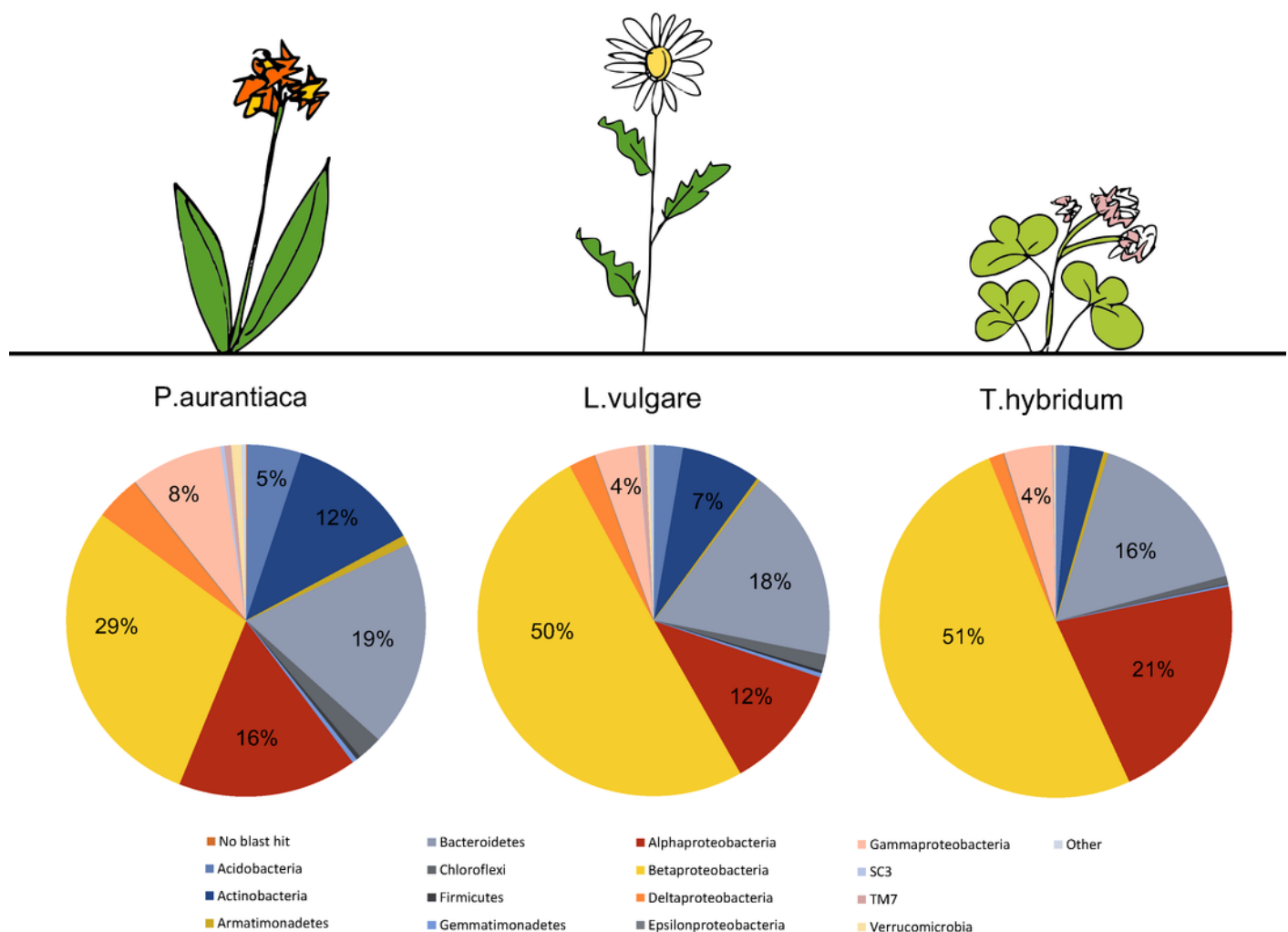
Phylogenetic overlap between communities was assessed using weighted UniFrac. Community composition was significantly different among plant species ($P<0.001$; PerMANOVA).

PermDISP results showing the average spread from centroid and standard error (SE) for samples of each species.

The PERMANOVA (P(perm)) values are assessing whether there is a significant difference between species in sample dispersion, using different diversity metrics.

Diversity metric	F	P (perm)	Species	Average	SE
Weighted UniFrac	3.10	0.11	*P. aurantiaca*	9.96 E^{-2}	4.28 E^{-3}
			T. hybridum	8.57 E^{-2}	5.70 E^{-3}
			L. vulgare	0.12	1.29 E^{-2}
Un-weighted UniFrac	1.04	0.37	*P. aurantiaca*	0.47	3.39 E^{-3}
			T. hybridum	0.46	5.42 E^{-3}
			L. vulgare	0.47	7.25 E^{-3}
Bray Curtis	9.30	0.005	*P. aurantiaca*	51.43	0.73
			T. hybridum	47.13	0.44
			L. vulgare	51.05	1.05

2

Comparison of the average bacterial community composition and relative abundances, at the phylum level (*Proteobacteria* divided into class) in root samples from three different plant species.

Results show a strong dominance of sequences belonging to *Betaproteobacteria* in all three plant species, but especially in *T. hybridum* (51%) and *L. vulgare* (50%). The phyla representing less than 1% out of the total community were grouped as "Other" and consisted of: NKB19, *Nitrospirae*, PAUC34f, *Cyanobacteria*, *Elusimicrobia*, *Fibrobacteres*, *Chlamydiae*, SC4, *Spirochaetes* and *Thermi*. Sequences not matching the database were recorded as "No blast hit".

The core root microbiota represented by the fourteen OTUs with the highest abundances across all samples

Values are calculated as the average percent out of the 400 sequences recorded for each sample, across all species (Total %). As a comparison, data is also included for what percentage (on average) the fourteen OTUs make up within the bacterial communities of the three plant species examined (*P. aurantiaca, T. hybridum* and *L. vulgare*).

# OTU ID	Total %	*P. aurantiaca* %	*T. hybridum* %	*L. vulgare* %	Phylum	Class	Order	Family	Genus	Species
1537	7.6	2.8	9.4	16.8	*Proteobacteria*	*Betaproteobacteria*	*Burkholderiales*	*Oxalobacteraceae*	*Herbaspirillum*	
19032	3.1	3.2	3.9	2.2	*Proteobacteria*	*Betaproteobacteria*	*Burkholderiales*			
22328	2.7	1.0	5.1	3.8	*Proteobacteria*	*Betaproteobacteria*	*Burkholderiales*	*Comamonadaceae*	*Limnohabitans*	
30435	1.5	1.7	1.3	1.2	*Proteobacteria*	*Betaproteobacteria*	*Burkholderiales*		*Methylibium*	*Methylibium petroleiphilum*
4453	1.3	0.8	1.7	2.1	*Proteobacteria*	*Betaproteobacteria*	*Burkholderiales*	*Oxalobacteraceae*	*Janthinobacterium*	*Janthinobacterium lividum*
1231	1.2	1.9	0.5	0.4	*Bacteroidetes*	*Sphingobacteria*	*Sphingobacteriales*	*Flexibacteraceae*	*Cytophaga*	
22285	1.0	1.5	1.0	0.4	*Proteobacteria*	*Alphaproteobacteria*	*Rhizobiales*	*Bradyrhizobiaceae*	*Bradyrhizobium*	
20009	1.0	1.5	0.2	0.6	*Actinobacteria*	*Actinobacteria*	*Actinomycetales*	*Thermomonosporaceae*	*Actinocorallia*	*Actinocorallia longicatena*
25072	0.9	1.0	1.0	0.7	*Proteobacteria*	*Gammaproteobacteria*	*Xanthomonadales*	*Xanthomonadaceae*	*Rhodanobacter*	*Rhodanobacter lindaniclasticus*
1340	0.9	1.0	1.0	0.6	*Bacteroidetes*	*Sphingobacteria*	*Sphingobacteriales*	*Sphingobacteriaceae*	*Sphingobacterium*	*Sphingobacterium faecium*
5213	0.8	0.6	0.9	1.3	*Bacteroidetes*	*Sphingobacteria*	*Sphingobacteriales*	*Sphingobacteriaceae*		
29492	0.8	1.3	0.1	0.6	*Proteobacteria*	*Gammaproteobacteria*	*Chromatiales*	*Sinobacteraceae*		
26917	0.8	1.1	0.3	0.7	*Chloroflexi*	*Chloroflexi*	*Roseiflexales*	*Kouleothrixaceae*	*Kouleothrix*	
1184	0.8	0.4	1.3	1.1	*Bacteroidetes*	*Flavobacteria*	*Flavobacteriales*	*Flavobacteriaceae*	*Chryseobacterium*	

3

Average relative abundances of *Betaproteobacteria* families and *Oxalobacteraceae* genera found in root samples of the three plant species.

Values are given as the percentage of sequences belonging to a certain taxa out of the total average bacterial community for each of the three plant species (rarefied at 400 sequences/sample). The heat map is colour coded from blue (low abundance) to red (high abundance).

Taxon	P.aurantiaca	T.hybridum	L.vulgare
Unclassified Betaproteobacteria	2.33	1.43	1.50
Unclassified Burkholderiales	9.92	13.65	7.19
Burkholderiales_Alcaligenaceae	0.32	0.53	0.31
Burkholderiales_Burkholderiaceae	0.57	0.63	0.31
Burkholderiales_Comamonadaceae	2.67	9.78	5.88
Burkholderiales_Oxalobacteraceae	9.28	23.48	32.47
Gallionellales_Gallionellaceae	0.08	0.00	0.09
Hydrogenophilales_Hydrogenophilaceae	0.01	0.05	0.00
Unclassified Methylophilales	0.05	0.03	0.03
Methylophilales_Methylophilaceae	1.37	0.70	0.88
Unclassified Rhodocyclales	0.76	0.08	0.44
Rhodocyclales_Rhodocyclaceae	1.87	0.35	1.22
Total percentage Betaproteobacteria	29%	51%	50%

	P.aurantiaca	T.hybridum	L.vulgare
Unclassified Oxalobacteraceae	3.39	6.93	8.94
Oxalobacteraceae_Collimonas	0.38	0.88	0.72
Oxalobacteraceae_Herbaspirillum	3.21	10.75	18.28
Oxalobacteraceae_Janthinobacterium	1.17	2.13	2.66
Oxalobacteraceae_Massilia	0.61	1.50	1.13
Oxalobacteraceae_Oxalobacter	0.50	1.30	0.75