Phytobiomes are compositionally nested from the ground up

Anthony S Amend [Corresp., 1] , Gerald M Cobian [2] , Aki J Laruson [3] , Kristina Remple [4] , Sarah J Tucker [5, 6] , Kirsten E Poff [7] , Carmen Antaky [8] , Andre Boraks [2] , Casey A Jones [2] , Donna Kuehu [9] , Becca R Lensing [5] , Mersedeh Pejhanmehr [2] , Daniel T Richardson [8] , Paul P Riley [8]

[1] Department of Botany, University of Hawaii at Manoa, Honolulu, Hawaii, United States

[2] Department of Botany, University of Hawaii at Manoa, Honolulu, HI, United States

[3] Department of Biology, University of Hawaii at Manoa, Honolulu, HI, United States

[4] Department of Oceanography, University of Hawaii at Manoa, Honolulu, HI, United States

[5] Marine Biology, University of Hawaii at Manoa, Honolulu, HI, United States

[6] Hawaii Institute of Marine Biology, University of Hawaii at Manoa, Honolulu, HI, United States

[7] Plant and Environmental Protection Services, University of Hawaii at Manoa, Honolulu, HI, United States

[8] Department of Natural Resources and Environmental Management, University of Hawaii at Manoa, Honolulu, HI, United States

[9] Department Molecular Biosciences and Bioengineering, University of Hawaii at Manoa, Honolulu, HI, United States

Corresponding Author: Anthony S Amend
Email address: amend@hawaii.edu

Plant-associated microbes are critical players in host health, fitness and productivity. Despite microbes' importance in plants, seeds are mostly sterile, and most plant microbes are recruited from an environmental pool. Surprisingly little is known about the processes that govern how environmental microbes assemble on plants in nature. In this study we examine how bacteria are distributed across plant parts, and how these distributions interact with spatial gradients. We sequenced amplicons of bacteria from six plant parts and adjacent soil of *Scaevola taccada*, a common beach shrub, along a 60 km transect spanning Oʻahu island's windward coast, as well as within a single intensively-sampled site. Bacteria are more strongly partitioned by plant part as compared with location. Within *S. taccada* plants, microbial communities are highly nested: soil and rhizosphere communities contain much of the diversity found elsewhere, whereas reproductive parts fall at the bottom of the nestedness hierarchy. Nestedness patterns suggest either that microbes follow a source/sink gradient from the ground up, or else that assembly processes correlate with other traits, such as tissue persistence, that are vertically stratified. Our work shines light on the origins and determinants of plant-associated microbes across plant and landscape scales.

1

Phytobiomes are compositionally nested from the ground up

4

5   Anthony S. Amend[1], Gerald M. Cobian[1], Áki J. Láruson[2], Kristina Remple[3], Sarah J.
6   Tucker[4,8], Kirsten E. Poff[5], Carmen Antaky[7], Andre Boraks[1], Casey A. Jones[1], Donna
7   Kuehu[9], Becca R. Lensing[4,6], Mersedeh Pejhanmehr[1], Daniel Richardson[7], Paul P.
8   Riley[7],

9
10   [1]Department of Botany, University of Hawai'i at Mānoa, Honolulu, HI, USA
11   [2]Department of Biology, University of Hawai'i at Mānoa, Honolulu, HI, USA
12   [3]Department of Oceanography, University of Hawai'i at Mānoa, Honolulu, HI, USA
13   [4]Marine Biology Program, University of Hawai'i at Mānoa, Honolulu, HI, USA
14   [5]Plant and Environmental Protection Sciences, University of Hawai'i at Mānoa,
15   Honolulu, HI, USA
16   [6]Microbiology, University of Hawai'i at Mānoa, Honolulu, HI, USA
17   [7]Department of Natural Resources and Environmental Management, University of
18   Hawai'i at Mānoa, Honolulu, HI, USA
19   [8]Hawai'i Institute of Marine Biology, Kaneohe, HI, USA
20   [9]Molecular Biosciences and Bioengineering

21
22   Corresponding Author:
23   Anthony S Amend, Department of Botany, University of Hawaii Manoa, 3190 Maile
24   Way, Honolulu, HI 96822 U.S.A.
25   Email: Amend@hawaii.edu

26

27

Abstract

29   Plant-associated microbes are critical players in host health, fitness and productivity.

30   Despite microbes' importance in plants, seeds are mostly sterile, and most plant microbes are

31   recruited from an environmental pool. Surprisingly little is known about the processes that

32   govern how environmental microbes assemble on plants in nature. In this study we examine how

33   bacteria are distributed across plant parts, and how these distributions interact with spatial

34   gradients. We sequenced amplicons of bacteria from six plant parts and adjacent soil of *Scaevola*

35   *taccada*, a common beach shrub, along a 60 km transect spanning O'ahu island's windward

36   coast, as well as within a single intensively-sampled site. Bacteria are more strongly partitioned

37   by plant part as compared with location. Within *S. taccada* plants, microbial communities are

38   highly nested: soil and rhizosphere communities contain much of the diversity found elsewhere,

39   whereas reproductive parts fall at the bottom of the nestedness hierarchy. Nestedness patterns

40   suggest either that microbes follow a source/sink gradient from the ground up, or else that

41   assembly processes correlate with other traits, such as tissue persistence, that are vertically

42   stratified. Our work shines light on the origins and determinants of plant-associated microbes

43   across plant and landscape scales.

44

Introduction
45

46

47   Many of what we formerly considered "plant" traits we now know to be the direct

48   or indirect result of a consortium of microbial species that colonize the inside and

49   outside of plant tissues. Known as the phytobiome, these microscopic organisms from

50   throughout the tree of life play critical roles in plant chemistry, health, fitness and

51   phenology (Glick, 2014; Bulgarelli *et al.*, 2015; Hacquard *et al.*, 2015; Panke-Buisse *et*

52   *al.*, 2015). Given microbes' central role in plant health, it is remarkable that most

53   microbes are not inherited by birthright (like chloroplasts or mitochondria). Instead, most

54   hosts emerge physically decoupled from their microbiomes, which are accumulated

55   from the surrounding environment (Hodgson *et al.*, 2014). This process may convey

56   some advantages, as it enables a population to adapt to a local habitat more quickly

57   than would be possible relying on the comparatively slow process of evolution (Lau and

58   Lennon, 2012). For this reason, a plant's location can be a strong determinant of

59   phytobiome composition (O'Rorke *et al.*, 2015). However, microbial composition can be

60   conserved *within* plant parts across multiple host species and environments (Lundberg

61   *et al.*, 2012; Ottesen *et al.*, 2013; Bodenhausen *et al.*, 2013; Lambais *et al.*, 2014; Leff

62   *et al.*, 2015; Müller *et al.*, 2016; Laforest-Lapointe *et al.*, 2016). Two leaves, though

63   oceans apart, might nevertheless recruit a similar consortia of microbes.

64   While researchers have expended considerable effort to study the assembly

65   dynamics driving phytobiome structure and diversity, researchers have yet to uncover

66   from where these microbes derive and which processes govern their assembly on

67   plants. Here, we examined how two factors: plant part identity and geographic distance,

68   partition microbial communities. We chose these variables because they represent

69   strongly deterministic and stochastic processes respectively, and because they are

70   reasonably easy to isolate on a volcanic island when aspect and elevation are held

71   constant.

72   Factors contributing to differences among plant parts are highly deterministic and

73   form microhabitats that vary considerably at small spatial scales. Rhizospheres, for

74   example, combine physiochemical properties of soils with nutrient and moisture inputs

75   from plants, whereas leaf surfaces can be hydroscopic and nutrient poor. Nectar

76   producing flowers, on the other hand, often provide a sugar-rich, low water potential,

77   acidic habitat. Tissue longevity further differentiates plant microbes. Whereas stems

78   may persist for the entire life of a plant, most other tissues have more limited lifespans.

79   In some species (i.e. *Hibiscus trionum*, the so-called flower-of-an-hour), flowers are

80   open for less than a single day, a seemingly short timeframe in which to recruit and

81   establish a resident microbiome. Distance, in contrast, serves as proxy for dispersal

82   limitation (Peay *et al.*, 2012) or priority affects (Kennedy *et al.*, 2009; Cadotte *et al.*,

83   2012), which are stochastic environmental drivers assuming that other environmental

84   conditions are held constant.

85   Further, we sought to determine whether within-plant distributions of microbes

86   are consistent with microbial source sink dynamics. We hypothesized that soil and

87   rhizosphere microbial communities serve as a reservoir for above-ground plant parts.

88   By conceptualizing plant parts as nodes in a bipartite network we can determine

89   whether microbial communities in ephermeal and late-emerging plant parts are

90   compositionally nested within early-developing and long-lived plant parts, consistent

91   with the hypothesis that a plant effectively inoculates itself from the ground up.

92   To address these questions, we examined how distance and plant part identity

93   structure the surface bacteria of beach Naupaka, *Scaevola taccada*, a widespread

94   littoral shrub native to Hawaiʻi and much of the tropics. Sampling at three scales (within

95   plant, within site, and across island) we examine how distance and plant part shape

96   phytobiomes, and make comparisons with adjacent soil communities that assemble

97   independent of a biological host. Using a multivariate modeling framework, we

98   determine the extent to which community variance is partitioned by site and by plant

99   part. Finally, to gain insight into phytobiome source sink dynamics, we determine

100   whether microbial communities are compositionally nested with respect to plant part.

101   We find that communities appear to be more sensitive to within-plant location,

102   than to site differences. We find, unsurprisingly, a split between below-ground and

103   above-ground samples, and distinct community structure on vegetative and

104   reproductive plant parts, suggesting that plant part function and longevity may

105   contribute to these differences. Phytobiomes within plants are compositionally nested,

106   and demonstrated hierarchical structuring consistent with expectations of soils as plant

107   microbial reservoirs.

108
109 **Materials & Methods**
110
111 *Sampling*

112   We selected our plant species and sites to minimize environmental

113 heterogeneity. *S. taccada* is an evergreen shrub, native to Hawaiʻi, which produces

114 fruits and flowers across seasons. It is also one of the most common littoral plants on

115 Oahu's East (windward) shore, enabling sampling at regular spatial intervals. The

116 coastal habitat was selected because local climates at this elevation are minimally

117 influenced by land topography, and because all coastlines are accessible to the public

118 by virtue of state law. All samples were collected on January 26, 2017. A single *S.*

119 *taccada* individual was sampled from each of 10 locations along the windward shore of

120 Oʻahu at 6 km intervals (Fig. 1). Generally, there were few individuals containing both

121 reproductive parts (fruits and flowers) at a given site, and if more than one was present

122 the sampled individual was selected haphazardly. Within the Kailua site, nine additional

123 individuals were selected haphazardly within a 50 m^2 plot. Plant locations were

124 recorded with a Garmin Rhino GPS. A single mature, disease-symptom free leaf, flower,

125 fruit, stem and axil was sampled by swabbing with a sterile cotton swab that had been

126 moistened with an alkaline Tris extraction buffer (containing EDTA and KCL;

127 comparable to Extract-N-Amp solution Sigma Aldrich, St. Louis). At two locations we

128 were unable to locate and sample flowers. Root (rhizosphere) samples were collected

129 by selecting an area of the root with diameter 0.5-1.7cm, and buried 2-10 cm below the

130 surface. Swabs of soil were taken within 10cm of the plant base, at a depth of 2-5 cm.

131 Because soil was likely the most heterogeneous component of our study, bulk soil

132 samples were also collected for soil description; including soil taxonomy and pH. Finally,

133 sterile swabs dipped in extraction solution were exposed to the air for approximately 20

134 seconds as extraction negative controls. After sampling, swabs were immersed in 200

135 μL extraction buffer and stored in a cooler until DNA extraction later in the day.

136 Although soil is not technically a plant part, we refer to it as such when differentiating

137 sample types above.

138

139 *Nucleic Acid Extraction and Library Preparation*

140 Swabs were incubated in 50 μL of the extraction buffer at 65ºC for 10 minutes

141 and at 95°C for an additional 10 minutes. After brief centrifugation, swabs were

142 removed from the buffer, and 200 μL of 3% BSA was added to each sample. DNA

143 extractions were purified and concentrated using a carboxylated paramagnetic bead

144 solution. Briefly, 100 μL of gDNA was mixed with 80 μL of bead solution, immobilized

145 with magnets, washed twice with 70% ethanol, and suspended in 30 μL of 0.1x TE

146 buffer.

147

148 *DNA library preparation and sequencing*

149 PCRs targeted the v4 region of the 16S rRNA gene using the primers 515f and

150 816r with overhangs complementary to a construct containing dual 8bp indices and the

151 Illumina i5 and i7 adaptors following the standard Illumina 16S Metagenomic

152 Sequencing Library Preparation protocol. The first and second PCRs contained twenty-

153 five and eight cycles respectively.

154 Negative PCR controls were added to the library preparations. PCR products

155 were purified and normalized to equimolar concentrations using the just-a-plate 96 PCR

156   Purification and Normalization Kit (Charm Biotech, San Diego, California, USA).

157   Normalized PCR products were pooled and sequenced at the Hawai'i Institute of Marine

158   Biology CORE Laboratory using the Illumina MiSeq V3, 600 cycle, paired-end protocol.

159   Sequence data were deposited to the SRA as PRJNA385181.

160

161   *Data processing*

162   Average reverse sequence quality scores declined to <25 after *ca* 150 bp and

163   were, therefore, not further considered. Chimeras were detected and removed from 16S

164   sequences using VSEARCH (Rognes *et al.*, 2016). Sequences were filtered by length

165   (75 bp min) and quality score (mean score 25) and demultiplexed within the QIIME

166   environment (Caporaso *et al.*, 2010). OTU binning at 97% identity and taxonomic

167   assignments were conducted in QIIME's "open reference" workflow using default

168   parameters (Caporaso et al., 2010). Taxa that were not assigned to bacteria (mostly

169   chloroplasts) were removed from datasets, as were OTUs represented by ten or fewer

170   sequences. None of the extraction negative controls or the PCR negative controls

171   produced sequence data.

172   Because of uneven sequencing depth across our dataset, data were down-

173   sampled to 1,000 reads per sample, a cutoff that included the majority of samples.

174   Although there is no consensus about how best to treat uneven sampling depth, down-

175   sampling to a common depth is generally robust for multivariate comparisons of

176   community compositions (Weiss *et al.*, 2017). Eighty-five samples containing greater

177   than 1,000 reads were retained for subsequent analysis.

178

179 *Data Analyses*

180   To determine which factors (site and/or plant part) predicted community

181 composition, binary Jaccard distances were used in a single PERMANOVA model

182 (function "adonis" in the vegan package (Oksanen *et al.*, 2013)) considering marginal

183 values of both factors and their interactions. In order to balance the number of samples

184 among sites, a single individual plant was selected for the Kailua site. This analysis was

185 repeated, including all samples from the Kailua site with qualitatively similar results.

186 Because the analysis can be sensitive to dispersion of beta-diversity values, these were

187 evaluated separately for plant parts. Following a globally significant RDA analysis

188 showing that community dissimilarity was predicted by site, habitat and their interaction,

189 a variance partitioning approach was used to determine the proportion of variance

190 explained by each.

191   To address whether soil or plant parts might serve as a microbial reservoir for

192 self-inoculation, we used two measures of nestedness, both of which indicate the extent

193 to which OTUs present in species-poorer plant parts are contained in those with higher

194 species richness. A bipartite matrix was calculated by summing all reads within each

195 plant part and then downsampling to the lowest sample sum. We calculate the

196 nestedness temperature (Atmar and Patterson, 1993), a popular, though older, method

197 calculating nestedness as a range from 0 (perfectly nested) to 100 (perfectly random)

198 based on hierarchical entropy order and presence absence data. We also calculated

199 nestedness based on overlap and decreasing fill (NODF; Almeida-Neto *et al.*, 2008)

200 scaled from 0 to 1. Both were compared with a distribution of randomized null

201 communities generated using the *quasiswap* method in vegan. Differences in

202 community membership among habitats were decomposed into nestedness and

203 turnover components using the *nestedbetajac* function in vegan.

204   The extent to which certain plant parts "select for" microbes, was calculated

205 using the network-wide H2 index (Blüthgen *et al.*, 2006), which returns a value from 0

206 (generalized) to 1 (specialized) based on potential associations given OTU abundance

207 totals. Significant deviations from null expectations were quantified using a distribution

208 of null community matrixes calculated using the quantitative *r2dtable* method in vegan in

209 which marginal totals are maintained. We calculated the related d' statistic (Blüthgen *et*

210 *al.*, 2006) to assess selectivity of individual plant parts.

211   To test for dissimilarity by distance patterns, binary Jaccard community

212 dissimilarities were compared with geographic distances between samples using Mantel

213 correlations (tests using Bray Curtis indices were qualitatively identical). So that the

214 intensively sampled Kailua site did not skew detected spatial structure, we randomly

215 selected a single Kailua individual for analysis (Mantel correlations including all samples

216 showed similar effect sizes, except for above ground plant parts, which no longer

217 showed a significant pattern; Table S2).

218   Indicator species analysis was calculated using the *indicspecies* package (De

219 Caceres and Jansen, 2016) on single groups (multi-group combinations were

220 suppressed due to the high number of OTUs).

221
222 **Results**
223
224 *Sequencing results*

225   Illumina sequencing of the V4 region of bacteria 16S gene generated 15,354,523

226 reads. Following demultimplexing, quality score filtering and chimera detection

227   8,901,385 bacteria reads remained. Removal of OTUs identified as chloroplasts, and

228   OTUs present in read abundances <10 resulted in 4,643,809 reads and 2,192 bacteria

229   OTUs retained for subsequent analysis.

230   *Specialization and nestedness of habitats*

231   Bacteria were significantly more nested than null expectations (Table 1; Fig. 1)

232   and nestedness contributed 44.7% of total between-sample dissimilarity. Nestedness

233   hierarchy followed the vertical structure of the plant: below-ground samples contained

234   the highest species richness, above-ground vegetative structures were intermediate,

235   and reproductive structures were lowest in the hierarchy (Fig. 2).

236   Microbes significantly specialized on particular plant parts. We examine the

237   degree and nature of specialization across plant parts using a bipartite network

238   architecture index, and also within plant part using an indicator species analysis.

239   Network-wide degrees of specialization were high and 762 "indicator species" were

240   statistically associated with plant parts (SI). The d' statistic, which measures degree of

241   specialization for individual habitats, showed that soil and rhizosphere were the most

242   specialized parts.

243

244   *Plant part and abiotic drivers of community composition*

245   A PERMANOVA model demonstrated that both plant part and location shape *S.*

246   *taccada* phytobiomes (Table 2), but that there was no interaction between those terms.

247   Further examination of community partitioning demonstrates that plant part is a stronger

248   determinant of bacteria community than location. Partitioning of variance, based on

249   RDA analysis, demonstrated differences in the relative importance contributed by these

250   two variables. Plant part accounted for 88% of the explained variance (RDA, *DF*=50,

251   *F*=1.33, *P*=0.043), although residual variance accounted for 81% of the total, suggesting

252   that other unmeasured variables may be more predictive of community composition.

253   NMDS ordination (Fig. 3) demonstrates that bacteria communities assorted into

254   three principle groups associated with plant parts: belowground (soil and rhizosphere),

255   mid-plant vegetative (stem and axil), and reproductive (fruit and flower). Leaf samples

256   are spread among the other above-ground plant parts. Dispersion of bacteria beta

257   diversity (bacteria betadispersion ANOVA, *DF*=6, *P*=0.08; Fig. S1) did not differ

258   significantly between any sample types, suggesting that results of the PERMANOVA

259   analysis reflect differences in multidimensional community "location" rather than spread.

260   Bacteria were strongly partitioned at the order level by plant parts (Fig. 4). Bacillales

261   were ubiquitous and abundant in soil and leaf samples, whereas Rhizobiales and

262   Actinomycetales were abundant across root and vegetative plant parts.

263   While site location was a strong predictor of compositional differences, tests of

264   dissimilarity by distance for microbial community compositions showed mixed results.

265   There was a strong distance-decay relationship for below-ground samples and the

266   relationship for above-ground plant parts and all samples was weaker, though

267   significant (Table 3). Although geographic distance was generally a poor predictor of

268   microbial composition, site location strongly impacted community composition (Table 2).

269   This may be due to a deterministic influence of environmental variables that were not

270   accounted for in our study, including the influence of the surrounding vegetation.

271
272   **Discussion**

273   *Distribution of OTUs across the plant*

274   We show that the phytobiome composition of *S. taccada* is strongly shaped by

275   plant part. Both community similarity and richness patterns are structured by plant parts:

276   stem, axil, and reproductive parts were most similar to each other, rhizosphere samples

277   were most similar to soil samples, and leaves were allied with above-ground parts more

278   generally. Segregation of microbial communities by plant part (Junker *et al.*, 2011;

279   Ottesen *et al.*, 2013; Junker and Keller, 2015) and even within-individual part location

280   (Leff *et al.*, 2015) has been reported for bacteria previously. Notably, Leff and

281   colleagues demonstrated that both fungi and bacteria are partitioned by plant part

282   (seed, rhizosphere and root) among sunflower cultivars (2016). By examining

283   differences among plant parts within a network context, however, we gain insight into

284   the potential sources and sinks of plant microbes.

285   We show that soil may be an important source for plant-surface microbiomes,

286   including above-ground parts. Our analysis demonstrates that plant microbes are

287   nested with respect to plant parts, that nestedness accounts for a large proportion of the

288   among-sample diversity, and that below-ground parts (soil and rhizosphere) are at the

289   top of the hierarchy. Even putative "soil" dwelling microbes, such as Bacillales, are

290   found throughout the plant. These results suggest that soils may serve as an important

291   source for plant-surface microbes, either through initial inoculation as the plant emerges

292   from the ground, or repeated introduction via wind and water dispersal. An alternative

293   hypothesis, that soils and plants could both be sinks from an alternative source (such as

294   airborne bacteria) is also plausible could be tested via experimental manipulations.

295   Despite nestedness patterns overall, each plant part contained a distinct subset

296   of indicator bacteria that were not shared. These patterns suggest that distinct assembly

297   processes may govern different plant parts. Pollinators, for example, are known to be

298   effective vectors for microbes found in nectar and other plant parts (Herrera *et al.*, 2010;

299   Aleklett *et al.*, 2014), although nectar yeasts are the best known examples. Less is

300   known about the origins or mechanisms of inoculation of other plant parts.

301   Factors other than physiochemical plant traits also likely play a role in shaping

302   phytobiomes. The strong division between below- and above-ground communities, for

303   example, may result from dispersal limitation in addition to climatic differences.

304   Furthermore, plant part longevity may play a role in compositional differences. Flowers

305   and fruits, the plant parts with lowest species richness are also the most ephemeral,

306   present on a plant for days to weeks. *S. taccada* flowers studied on the French

307   Polynesian Island of Moorea, for example, remain open for an average of fewer than

308   four days (Liao, 2008), seemingly scant time to recruit microbes from the environment.

309

310   *Plant microbial distributions over space*

311   In microbial biogeography studies, distance between samples is generally a

312   reliable predictor of microbial community dissimilarity, particularly among communities

313   that are not attached to a macro-organism (Martiny *et al.*, 2011; Tedersoo *et al.*, 2014;

314   Zinger *et al.*, 2014), even at the scale of centimeters (Bell, 2010). The distance decay of

315   plant-associated microbial similarity is comparatively less resolved compared with

316   studies of "host-independent" substrates such as ocean, water, or soil. A positive

317   pattern could be attributed to at least two factors and or their cumulative effects. First,

318   dispersal limitation could lead to clinal dissimilarity among communities by enabling

319   "ecological drift" over relatively short timescales (Martiny *et al.*, 2006). Second,

320   geographic distance could correlate with some environmental cline, which selects for a

321   distinct adapted community. Factors such as rainfall (O'Rorke *et al.*, 2017), soil pH

322   (Fierer and Jackson, 2006), soil nutrients, or even host genotypic clines (Bálint *et al.*,

323   2013) all contribute to community dissimilarity.

324   Among plant-associated microbes, evidence for the dissimilarity by distance

325   pattern is mixed. Redford and colleagues (2010) found little evidence of geographic

326   structure in pine bacteria phyllosphere communities spanning continents, and Meaden

327   and others (2016) found no evidence for dispersal limitation among Oak phyllosphere

328   bacteria within a 20 ha plot. In contrast, Stone et al. (2000) found strong isolation by

329   distance patterns among Magnolia phyllosphere bacteria located within 400 m of each

330   other. Oono and colleagues (2017) found that low abundance pine needle fungal

331   endophytes were spatially structured over ~100 km, but that high abundance

332   communities were not.

333   In our study, both above-ground and belowground bacteria were correlated with

334   geographic distance (Table 3), although the latter relationship was stronger. Differing

335   community turnover rates among plant parts could be attributable to different dispersal

336   rates since airborne bacteria would presumably travel farther and more quickly due to

337   prevalent trade winds compared to subsurface soil bound communities. Although there

338   were no obvious environmental clines along our transect, microbes may have

339   responded differently to additional unmeasured variables. Multiple studies show that

340   bacteria are particularly sensitive to soil properties such as pH. Substrates change

341   rapidly over short distances from the high tide, and although pH did not covary in a

342   linear fashion along the transect, pH was variable among sample sites. At a single site

343   for this study, we sampled soils ranging from ph 5.28 to 8.65 (Table S1). Finally,

344   differences in the distance decay relationship could be driven by interactions with the

345   host plant. Obligacy or strong host selection of some community members, for example,

346   could depress turnover if the microbe was required for growth and survival.

347   Plant-associated microbes are critical players in plant fitness and health. Our

348   work shows that while both stochastic and deterministic factors play a role in shaping

349   surface phytobiomes, plant part is the most predictive, with nestedness patterns relating

350   strongly to vertical stratification. Bacteria are highly differentiated across a plant, with

351   distance playing a weak, though significant, role in community composition at the scale

352   examined. Despite these patterns, we were unable to describe the majority of

353   community variance in this system. Identifying environmental reservoirs for

354   phytobiomes, particularly those sources for ephemeral reproductive parts, will help in

355   understanding plant associated microbial distributions and the factors that shape them.

356
357
358
359   **Conclusions**

360   Plant-associated microbes are critical players in plant fitness and health. Our

361   work shows that while both stochastic and deterministic factors play a role in shaping

362   surface phytobiomes, plant part is the most predictive, with nestedness patterns relating

363   strongly to vertical stratification. Bacteria are highly differentiated across a plant, with

364   distance playing a weak, though significant, role in community composition at the scale

365   examined. Despite these patterns, we were unable to describe the majority of

366   community variance in this system. Identifying environmental reservoirs for

367   phytobiomes, particularly those sources for ephemeral reproductive parts, will help in

368   understanding plant associated microbial distributions and the factors that shape them.

369

370   **Acknowledgements**

371

372   The authors wish to thank Michele Langer and Illumina Corporation for their generous

373   support. We also appreciate the assistance and support of Amy Eggers and the Hawai'i

374   Institute of Marine Biology Genetics Core Facility. This paper was the result of a class

375   project: High Throughput Sequencing Approaches to Ecology and Evolution at the

376   University of Hawaii at Manoa.

377

378   **References**

379

380   Aleklett, K., Hart, M., and Shade, A. (2014) The microbial ecology of flowers: an

381   emerging frontier in phyllosphere research. *Botany* **92**: 253–266.

382   Almeida-Neto, M., Guimaraes, P., Guimaraes, P.R., Loyola, R.D., and Ulrich, W. (2008)

383   A consistent metric for nestedness analysis in ecological systems: reconciling

384   concept and measurement. *Oikos* **117**: 1227–1239.

385   Atmar, W. and Patterson, B.D. (1993) The measure of order and disorder in the

386   distribution of species in fragmented habitat. *Oecologia* **96**: 373–382.

387   Bálint, M., Tiffin, P., Hallström, B., O'Hara, R.B., Olson, M.S., Fankhauser, J.D., et al.

388   (2013) Host genotype shapes the foliar fungal microbiome of balsam poplar

389   (*Populus balsamifera*). *PLoS ONE* **8**: e53987.

390   Bell, T. (2010) Experimental tests of the bacterial distance–decay relationship. *ISME J*

391   **4**: 1357–1365.

392   Blüthgen, N., Menzel, F., and Blüthgen, N. (2006) Measuring specialization in species

393   interaction networks. *BMC Ecol.* **6**: 9.

394   Bodenhausen, N., Horton, M.W., and Bergelson, J. (2013) Bacterial communities

395   associated with the leaves and the roots of *Arabidopsis thaliana*. *PLoS ONE* **8**:

396   e56329.

397   Bulgarelli, D., Garrido-Oter, R., Münch, P.C., Weiman, A., Dröge, J., Pan, Y., et al.

398   (2015) Structure and Function of the Bacterial Root Microbiota in Wild and

399   Domesticated Barley. *CHOM* **17**: 392–403.

400   Cadotte, M.W., Cardinale, B.J., and Oakley, T.H. (2012) Phylogenetic relatedness

401   predicts priority effects in nectar yeast communities. *Proc Natl Acad Sci USA* **105**:

402   17012–17017.

403   Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello,

404   E.K., et al. (2010) QIIME allows analysis of high-throughput community sequencing

405   data. *Nat Meth* **7**: 335–336.

406   De Caceres, M. and Jansen, F. (2016) Indicspecies: Relationship between Species and

407   Groups of Sites. R package, Version 1.7. 4. 2014.

408   Fierer, N. and Jackson, R.B. (2006) The diversity and biogeography of soil bacterial

409   communities. *Proc. Natl. Acad. Sci. U.S.A.* **103**: 626–631.

410   Glick, B.R. (2014) Bacteria with ACC deaminase can promote plant growth and help to

411   feed the world. *Microb Res* **169**: 30–39.

412   Hacquard, S., Garrido-Oter, R., Gonzalez, A., Spaepen, S., Ackermann, G., Lebeis, S.,

413   et al. (2015) Microbiota and Host Nutrition across Plant and Animal Kingdoms. *Cell*

414   *Host Microbe* **17**: 603–616.

415 Herrera, C.M., Canto, A., Pozo, M.I., and Bazaga, P. (2010) Inhospitable sweetness:

416   nectar filtering of pollinator-borne inocula leads to impoverished, phylogenetically

417   clustered yeast communities. *Proc R Soc B* **277**: 747–754.

418 Hodgson, S., de Cates, C., Hodgson, J., Morley, N.J., Sutton, B.C., and Gange, A.C.

419   (2014) Vertical transmission of fungal endophytes is widespread in forbs. *Ecol Evol*

420   **4**: 1199–1208.

421 Junker, R.R. and Keller, A. (2015) Microhabitat heterogeneity across leaves and flower

422   organs promotes bacterial diversity. *Fems Microbiol Ecol* **91**: 1–9.

423 Junker, R.R., Loewel, C., Gross, R., Dötterl, S., Keller, A., and Blüthgen, N. (2011)

424   Composition of epiphytic bacterial communities differs on petals and leaves. *Plant*

425   *Biology* **13**: 918–924.

426 Kennedy, P.G., Peay, K.G., and Bruns, T.D. (2009) Root tip competition among

427   ectomycorrhizal fungi: Are priority effects a rule or an exception? *Ecology* **90**: 2098–

428   2107.

429 Laforest-Lapointe, I., Messier, C., and Kembel, S.W. (2016) Tree phyllosphere bacterial

430   communities: exploring the magnitude of intra-and inter-individual variation among

431   host species. *PeerJ*.

432 Lambais, M.R., Lucheta, A.R., and Crowley, D.E. (2014) Bacterial Community

433   Assemblages Associated with the Phyllosphere, Dermosphere, and Rhizosphere of

434   Tree Species of the Atlantic Forest are Host Taxon Dependent. *Microb Ecol* **68**:

435   567–574.

436 Lau, J.A. and Lennon, J.T. (2012) Rapid responses of soil microorganisms improve

437   plant fitness in novel environments. *Proc Natl Acad Sci USA* **109**: 14058–14062.

438 Leff, J.W., Del Tredici, P., Friedman, W.E., and Fierer, N. (2015) Spatial structuring of

439   bacterial communities within individual *Ginkgo biloba* trees. *Environ Microbiol* **17**:

440   2352–2361.

441 Leff, J.W., Lynch, R.C., Kane, N.C., and Fierer, N. (2016) Plant domestication and the

442   assembly of bacterial and fungal communities associated with strains of the

443   common sunflower, *Helianthus annuus*. *New Phytol.* **214**: 412–423.

444 Liao, I.T. (2008) Pollination biology and reproductive ecology of Scaevola taccada

445   (Goodeniaceae) on Mo'orea, French Polynesia. *Biology and Geomorphology of*

446   *Tropical Islands Student Research Papers*.

447 Lundberg, D.S., Lebeis, S.L., Paredes, S.H., Yourstone, S., Gehring, J., Malfatti, S., et

448   al. (2012) Defining the core *Arabidopsis thaliana* root microbiome. *Nature* **488**: 86–

449   90.

450 Martiny, J.B.H., Bohannan, B.J.M., Brown, J.H., Colwell, R.K., Fuhrman, J.A., Green,

451   J.L., et al. (2006) Microbial biogeography: putting microorganisms on the map. *Nat*

452   *Rev Microbiol* **4**: 102–112.

453 Martiny, J.B.H., Eisen, J.A., Penn, K., Allison, S.D., and Horner-Devine, M.C. (2011)

454   Drivers of bacterial beta-diversity depend on spatial scale. *Proc Natl Acad Sci USA*

455   **108**: 7850–7854.

456 Meaden, S., Metcalf, C.J.E., and Koskella, B. (2016) The effects of host age and spatial

457   location on bacterial community composition in the English Oak tree (Quercus

458   robur). *Environ Microbiol Rep* **8**: 649–658.

459 Müller, D.B., Vogel, C., Bai, Y., and Vorholt, J.A. (2016) The Plant Microbiota: Systems-

460   Level Insights and Perspectives. *Annu. Rev. Genet.* **50**: 211–234.

461   O'Rorke, R., Cobian, G.M., Holland, B.S., Price, M.R., Costello, V., and Amend, A.S.

462   (2015) Dining local: the microbial diet of a snail that grazes microbial communities is

463   geographically structured. *Environ Microbiol* **17**: 1753–1764.

464   O'Rorke, R., Tooman, L., Gaughen, K., Holland, B.S., and Amend, A.S. (2017) Not just

465   browsing: an animal that grazes phyllosphere microbes facilitates community

466   heterogeneity. *Proc Natl Acad Sci USA* **109**: 13022–13027.

467   Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O'Hara, R.B., et al.

468   (2013) vegan: community Ecology Package. R package version 2.0-7. *Community*

469   *ecology package Version, 2, 0-7*.

470   Oono, R., Rasmussen, A., and Lefèvre, E. (2017) Distance decay relationships in foliar

471   fungal endophytes are driven by rare taxa. *Environ Microbiol* **19**: 2794–2805.

472   Ottesen, A.R., Peña, A.G., White, J.R., Pettengill, J.B., Li, C., Allard, S., et al. (2013)

473   Baseline survey of the anatomical microbial ecology of an important food plant:

474   *Solanum lycopersicum* (tomato). *BMC Microbiology 2013 13:1* **13**: 114.

475   Panke-Buisse, K., Poole, A.C., Goodrich, J.K., Ley, R.E., and Kao-Kniffin, J. (2015)

476   Selection on soil microbiomes reveals reproducible impacts on plant function. *ISME*

477   *J* **9**: 980–989.

478   Peay, K.G., Schubert, M.G., Nguyen, N.H., and Bruns, T.D. (2012) Measuring

479   ectomycorrhizal fungal dispersal: macroecological patterns driven by microscopic

480   propagules. *Molecular Ecol* **21**: 4122–4136.

481   Redford, A.J., Bowers, R.M., Knight, R., Linhart, Y., and Fierer, N. (2010) The ecology

482   of the phyllosphere: geographic and phylogenetic variability in the distribution of

483   bacteria on tree leaves. *Environ Microbiol* **12**: 2885–2893.

484   Rognes, T., Flouri, T., Nichols, B., Quince, C., and Mahé, F. (2016) VSEARCH: a

485   versatile open source tool for metagenomics. *PeerJ* **4**: e2584.

486   Stone, J.K., Bacon, C.W., and White, J.F. (2000) An overview of endophytic microbes:

487   endophytism defined. *Microbial endophytes* **3**: 29–33.

488   Tedersoo, L., Bahram, M., PÕLME, S., Kõljalg, U., Yorou, N.S., Wijesundera, R., et al.

489   (2014) Fungal biogeography. Global diversity and geography of soil fungi. *Science*

490   **346**: 1256688.

491   Weiss, S., Xu, Z.Z., Peddada, S., Amir, A., Bittinger, K., Gonzalez, A., et al. (2017)

492   Normalization and microbial differential abundance strategies depend upon data

493   characteristics. **5**: 27.

494   Zinger, L., Boetius, A., and Ramette, A. (2014) Bacterial taxa-area and distance-decay

495   relationships in marine environments. *Molecular Ecol* **23**: 954–964.

496

Figure 1(on next page)

Each sampling location indicates where seven biological samples were collected (six plant parts and a soil sample).

Sites were spaced ca. 6 km along the windward coast. Ten sites were sampled in Kailua (inset). Axes represent decimal degrees. Topography is show to indicate that all samples were collected at sea level and with a similar aspect.
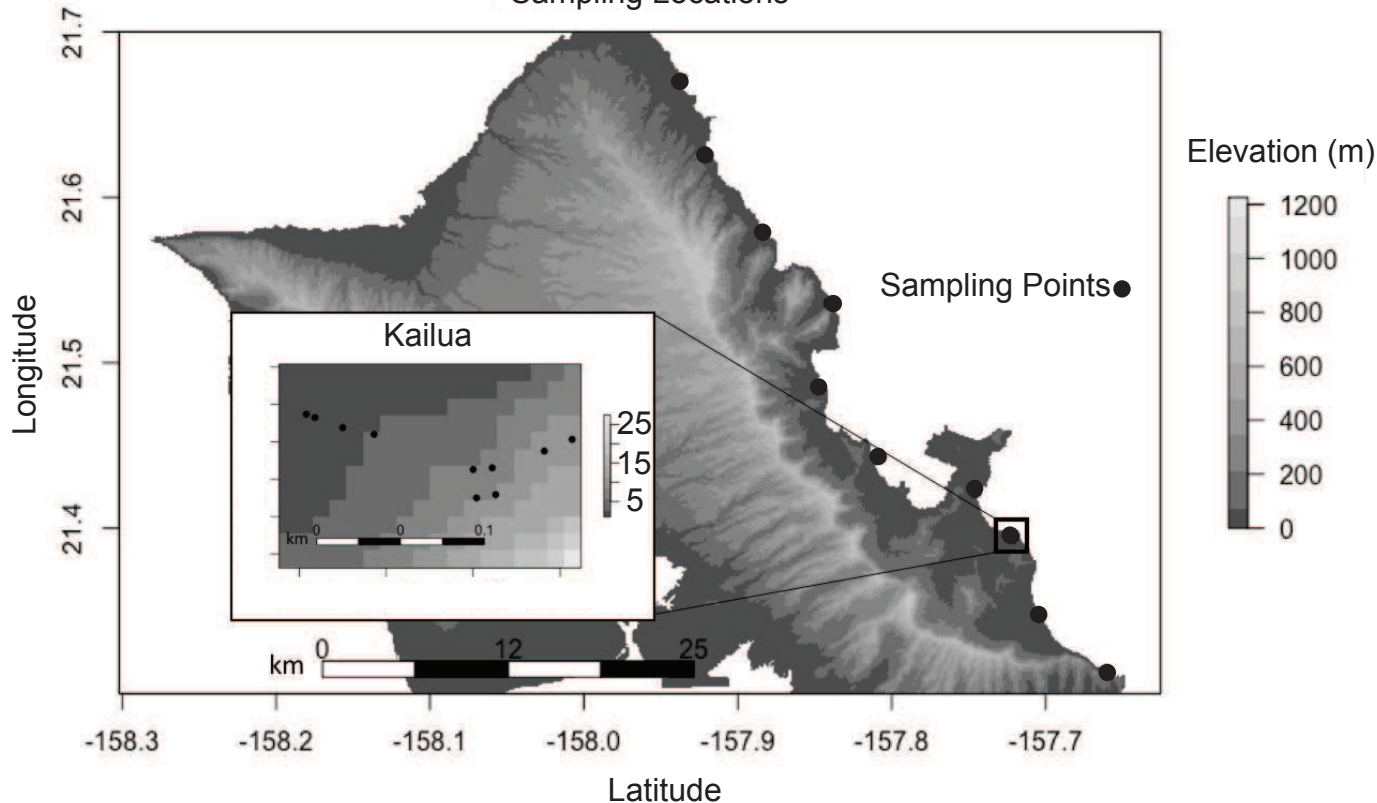
Sampling Locations

Figure 2(on next page)

Nestedness plot of bacteria aggregated by plant part.

Presence of an OTU in a plant part is represented as a rectangle. OTUs are ordered by occupancy (left to right) across plant parts, and rows are ordered by highest OTU richness (top to bottom). If all OTUs occurred above the "Fill line" (curved line), the network would be perfectly nested.

Figure 3(on next page)

Non-metric multidimensional scaling plots of microbial communities colored by habitat.

Ellipses indicate 95% confidence intervals. Above and below-ground parts are differentiated along the first axis, with leaves intermediate.
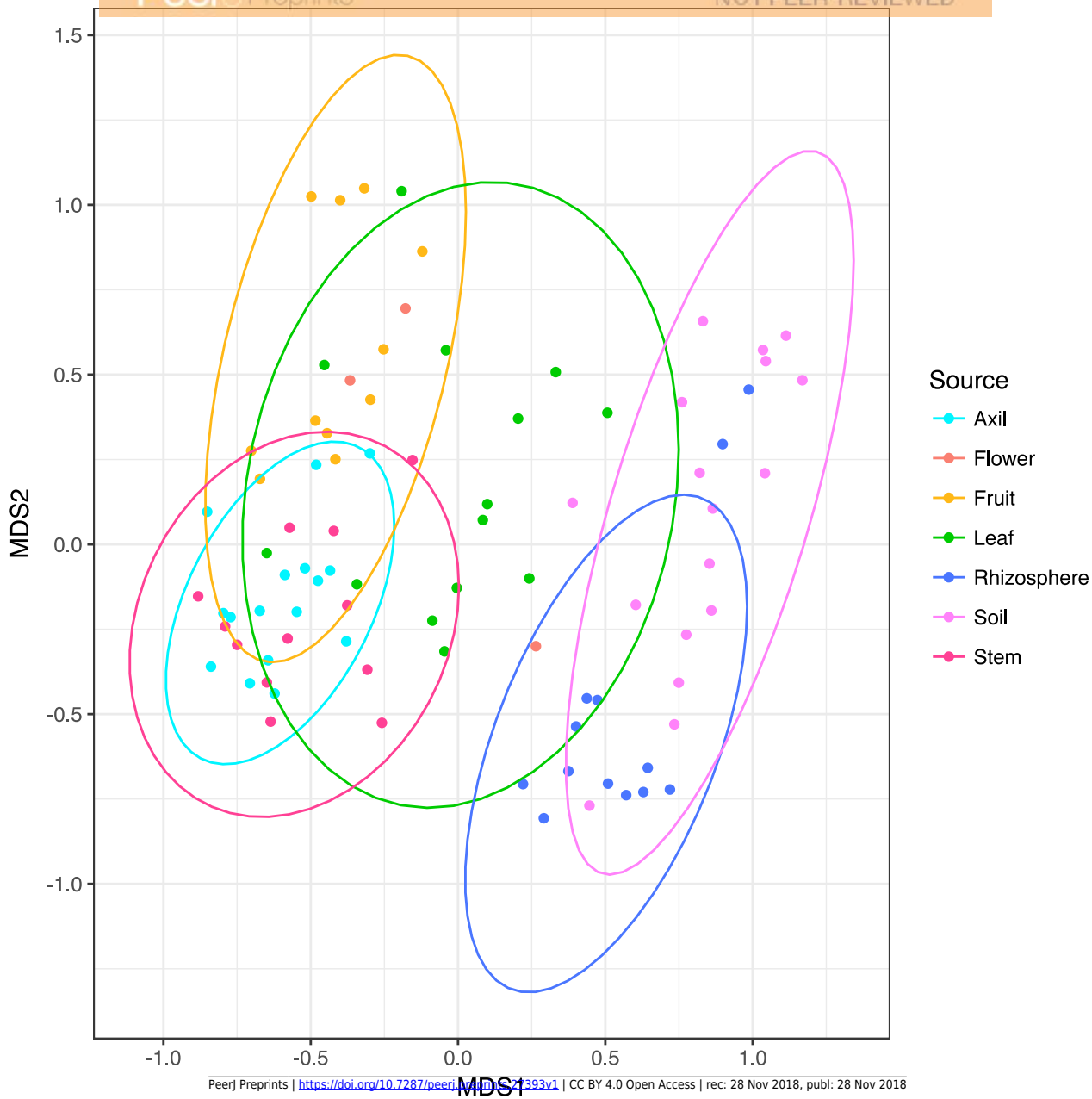
Figure 4(on next page)

Heatmap of order-level taxa as distributed across plant parts.

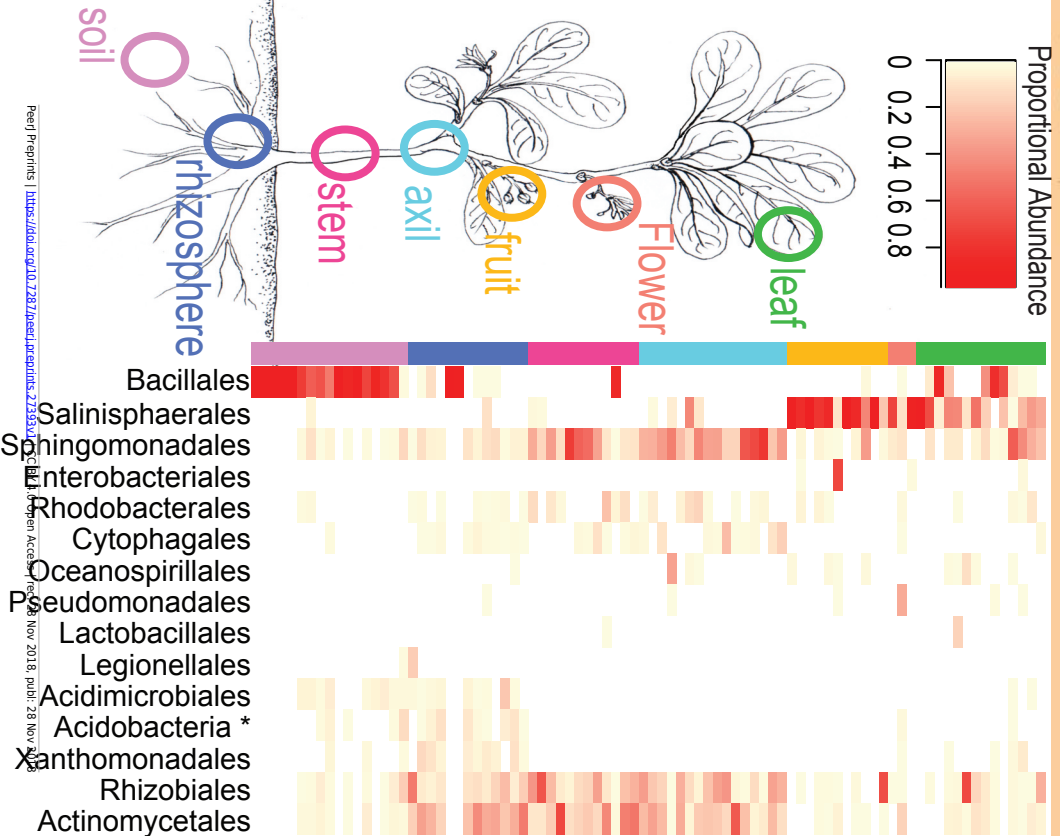Cell values are calculated proportionally across samples (rows).

Table 1(on next page)

Network stucture of microbial communities.

Table 1 Network stucture of microbial communities. The d' statistic is a measure of specialization by habitat ranging from 0 (not specialized) to 1 (completely specialized). H2 is an index of specialization across all taxa within the network and is measured on the same scale. Both networks were significantly more specialized than randomized null simulations would predict. Plant parts are ordered by nestedness structure, with the most species-rich communities on the top. Turnover and nestedness proportions describe community dissimilarity among plant parts explained by that process.

Sample	d'	P
Rhizosphere	0.47	0.001
Soil	0.55	0.001
Leaf	0.29	0.001
Axil	0.31	0.001
Stem	0.34	0.001
Flower	0.25	0.001
Fruit	0.30	0.001
Nestedness Temp.	28.21	0.001
NODF	0.24	0.001
H2	0.64	0.001
Turnover	0.11	
Nestedness	0.09	

1

Table 2(on next page)

PERMANOVA examining community compositional variance explained by plant part, site, or their interaction.

Table 2 PERMANOVA examining community compositional variance
explained by plant part, site, or their interaction.

Component	Degrees Freedom	Sum of Squares	F	P
Part	6	5.1887	5.8391	**0.001**
Site	8	2.894	2.4426	**0.001**
Part:Site	29	6.6346	1.5447	**0.012**
Residual	4	0.5924		

1

Table 3(on next page)

Mantel tests measure the correlation between geographic distance and microbial community dissimilarity for *Scaevola taccada* surface microbes.

Table 3 Mantel testsmeasure the correlation between geographic distance and microbial community dissimilarity for *Scaevola taccada* surface microbes. Sample sites occur as 10 locations spanning 60 km on the island of Oahu, HI. Mantel's statistic measure Pearson's product-moment correlation with 999 permutations. Significant values indicating a relationship between spatial distance and community dissimilarity are bolded. In this analysis a single plant individual was selected from the Kailua site. Above-ground refers to all plant parts, whereas belowground refers to rhizospere and soil.

Plant part	r-value	p-value	*n*
All plant parts	0.06538	**0.046**	48
Above-ground	0.1319	**0.017**	32
Below-ground	0.3781	**0.008**	16
Soil	0.1334	0.231	9

1