

UC Berkeley

UC Berkeley Previously Published Works

Title

High-Level Culturability of Epiphytic Bacteria and Frequency of Biosurfactant Producers on Leaves

Permalink

<https://escholarship.org/uc/item/5zb3q670>

Journal

Applied and Environmental Microbiology, 82(19)

ISSN

0099-2240

Authors

Burch, Adrien Y
T., Paulina
Sbodio, Adrian
[et al.](#)

Publication Date

2016-10-01

DOI

10.1128/aem.01751-16

Peer reviewed

1 High culturability of epiphytic bacteria and frequency of biosurfactant
2 producers on leaves

3
4 ^aAdrien Y. Burch, ^aPaulina T. Do, ^bAdrian Sbodio, ^bTrevor V. Suslow and ^aSteven E. Lindow[#]

5
6 ^aDepartment of Plant and Microbial Biology, University of California, Berkeley, CA, 94720

7 ^bDepartment of Plant Sciences, University of California, Davis, CA, 95616

8

9

10

11

12 [#]Corresponding author:

13 Department of Plant and Microbial Biology

14 University of California

15 111 Koshland Hall

16 Berkeley, CA 94720-3102

17 Phone: (510) 642-4174; Fax: (510) 642-4995

18 e-mail: icelab@berkeley.edu

19

20

21

22 **RUNNING TITLE:** Biosurfactants in the phyllosphere

23 **ABSTRACT**

24 To better characterize the bacterial community members capable of biosurfactant production on
25 leaves, we distinguished culturable biosurfactant-producing bacteria from non-producers and
26 used community sequencing to compare the composition of these distinct cultured populations
27 with that from DNA directly recovered from leaves. Communities on spinach, romaine, and head
28 lettuce leaves were compared with adjacent samples of soil and irrigation source water. Soil
29 communities were poorly described by culturing, with recovery of cultured representatives from
30 only 21% of the prevalent OTUs (>0.2% reads) identified. The dominant biosurfactant producers
31 cultured from soil included Bacilli and Pseudomonads. In contrast, the cultured communities
32 from leaves are highly representative of the culture-independent communities with over 85% of
33 the prevalent OTUs recovered. The dominant taxa of surfactant producers from leaves were
34 Pseudomonads as well as the infrequently-studied genus *Chryseobacterium*. The proportion of
35 bacteria cultured from head lettuce and romaine leaves that produce biosurfactants was directly
36 correlated with the culture-independent proportion of Pseudomonads in a given sample, whereas
37 spinach harbored a wider diversity of biosurfactant producers. A subset of the culturable bacteria
38 in irrigation water also became enriched on romaine leaves that were overhead irrigated.
39 Although our study was designed to identify surfactant producers on plants, we also provide
40 evidence that most bacteria in some habitats such as agronomic plant surfaces are culturable, and
41 these communities can be readily investigated and described by more classical culturing
42 methods.

43

44

45 **IMPORTANCE**

46 The importance of biosurfactant production to the bacteria that live on waxy leaf surfaces as well
47 as their ability to be accurately assessed using culture-based methodologies was determined by
48 interrogating epiphytic populations by both culture-dependent and culture-independent methods.
49 Biosurfactant production was much more frequently observed in cultured communities on leaves
50 than in other nearby habitats such as soil and water suggesting that this trait is important to life
51 on a leaf by altering either the leaf itself or the interaction of bacteria with water. While
52 Pseudomonads were the most common biosurfactant producers isolated, this habitat also selects
53 for taxa such as *Chryseobacterium* for which this trait was previously unrecognized. The finding
54 that most epiphytic bacterial taxa were culturable validates strategies using more classical
55 culturing methodologies for their study in this habitat.

56

57 **INTRODUCTION**

58 The phyllosphere is recognized to be a selective habitat in which epiphytic bacteria must be able
59 to access limited and spatially heterogeneous nutrient supplies and endure frequent fluctuations
60 in moisture availability on a water-repellent surface (1, 2). Recent culture-independent
61 community analyses have confirmed the apparent selectivity of leaf surfaces, finding that leaf
62 surfaces harbor less diverse communities than environments such as the immediate subtending
63 soil (3, 4). Furthermore, plant species identity also appears to be a factor that can influence the
64 composition of associated phyllosphere bacterial communities (5-9). While epiphytes are
65 apparently distinct in their ability to grow and survive on leaves, their adaptations needed to
66 thrive in this habitat remain largely unknown. Recently, we demonstrated that production of the
67 biosurfactant syringafactin provides fitness benefits to *Pseudomonas syringae* on bean leaves
68 (10). We have also observed a higher incidence of biosurfactant producers in the culturable
69 bacteria communities on plants compared to other habitats (11), suggesting that they contribute
70 to the capacity to colonize the waxy, hydrophobic leaf surface. Biosurfactants are biologically-
71 produced amphiphilic compounds that exhibit surface activity through the actions of their
72 hydrophilic and hydrophobic groups. Epiphytic bacteria could potentially use biosurfactants to
73 increase the wettability of the leaf, to enhance diffusion of nutrients across the waxy cuticle,
74 and/or aid in motility to favorable growth sites (2). Their many possible contributions to
75 epiphytic fitness have led us to focus on the role of biosurfactants in bacterial colonization of
76 leaves.

77

78 Despite the many potential roles of biosurfactants on leaves, only a few studies have examined
79 their production in the phyllosphere (10, 12-17). Furthermore, these studies have focused only on

80 their possible ecological role in specific strains and have not addressed the diversity or relative
81 abundance of surfactant producers on leaf surfaces. A comprehensive examination of
82 phyllosphere inhabitants likely would reveal strains and biosurfactants not normally encountered
83 in other habitats. Although evidence from previous investigations link specific surfactants to
84 particular activities, biosurfactants differ widely in their molecular structures, which could confer
85 a myriad of properties and thus might lend them to contribute to very different functions.
86 Furthermore, this structurally diverse class of compounds has broad industrial applications and
87 considerable economic potential (18). Thus, in addition to better understanding their biological
88 significance, investigation of biosurfactant producers recovered from leaves might reveal new
89 classes of surfactants for structure/functional analysis.

90
91 One inherent shortcoming of studies of the diversity of biosurfactant producers, as well as for
92 any functional trait without a conserved genetic signature, is its reliance on the need to culture
93 the bacteria. It is now widely recognized that cultured bacteria often are not representative of the
94 community from which they are derived (19, 20); with many being unculturable, and some
95 culturable bacteria apparently so rare in the community that they are difficult to detect with
96 culture-independent methods. The cultivable community is clearly dictated to a great degree by
97 the choice of medium and incubation conditions, which often differ from conditions *in situ*.

98 However, agar surfaces likely approximate leaf surface conditions more than environments such
99 as heterogeneous soil particles. While quantitative PCR-based methods suggest that as few as
100 0.1-5% of the bacteria on leaves are cultivable, a proportion similar to most other habitats (20),
101 frequencies of 50% and higher have been recently described when more sophisticated methods
102 are employed (21-23). Therefore, as part of our investigation into the diversity of biosurfactant

103 producers, we also focused on identifying those bacterial taxa that might not be interrogated for
104 lack of their culturability.

105

106 Previous studies have typically assessed biosurfactant production individually in selected
107 cultured organisms. That strategy is quite laborious and thus usually restricts such examinations
108 to relatively few bacterial isolates in any given study. Given the interest in biosurfactants, many
109 studies have described new bacterial strains capable of their production (18). However, the
110 majority of biosurfactant research to date has been focused on Pseudomonads and Bacilli (24), in
111 part because they are frequently the most common biosurfactant-producing bacteria in
112 environmental samples (11). The process by which novel biosurfactant producers can be
113 determined is thus inefficient, requiring phenotyping of a large number of random isolates.
114 Furthermore, unless one determines the phylogenetic identities of all of the strains tested, no
115 information is obtained about the prevalence of a given biosurfactant-producing taxon in the
116 examined environment, nor is it known whether this trait is characteristic of a given species. In
117 this study, we have coupled a high throughput assay that utilizes atomized droplets of oil to
118 screen for biosurfactant production on agar plates (25) with a system in which large numbers of
119 biosurfactant producing and nonproducing bacterial isolates cultured from leaves are separately
120 pooled and community composition determined by high-throughput sequencing of their
121 amplified ribosomal RNA genes. Additionally, we simultaneously characterized the total
122 community recovered from leaves by sequencing ribosomal genes amplified from DNA
123 recovered directly from leaves. Together, this strategy allowed us to identify even relatively
124 uncommon biosurfactant producing bacteria on plants, determine the relative abundance of
125 different taxa of biosurfactant producing organisms, calculate the frequency with which a given

126 taxa could produce biosurfactants, and importantly, address the relative cultivability of various
127 taxa on leaves of highly domesticated horticulture crops, in the respective production soil, and in
128 the irrigation water sources from a major production region. We demonstrate that the genus
129 *Chryseobacterium*, previously undocumented as a producer of biosurfactants, is an important and
130 widespread biosurfactant producer on leaf surfaces and that, unlike many other habitats, most
131 bacterial taxa on the leaves of these plants had culturable representatives.

132

133 **MATERIALS AND METHODS**

134 **Sample collection and processing.**

135 Head lettuce, romaine, spinach, soil and irrigation water samples were collected from
136 commercial fields in the Salinas valley of California on April 25 and June 26, 2012. To collect
137 plant samples, entire lettuce and spinach plants were harvested at the base with a sterile scalpel,
138 placed separately in polypropylene bags and transported in coolers with ice packs.
139 Corresponding soil and water samples were collected into 50 ml Falcon tubes. For each plant
140 variety, four plant samples, one soil and one water sample (when available) were randomly
141 chosen for processing the same day. Water samples were from sources of water being used for
142 irrigation. Plant samples were prepared by combining 1-2 leaves (head lettuce [*Lactuca sativa*
143 var. *capitata*] and romaine [*Lactuca sativa* var. *longifolia*]; outer mature leaves were discarded
144 and only inner edible leaves were used) or 6-8 leaves (spinach [*Spinacia oleracea*]; free of dirt or
145 visible contaminants) from a single plant with 200 ml washing buffer (10 mM KPO₄; pH 7.5) in
146 a 3.8 L Ziplok[®] bag. Soil samples consisted of 0.25 g soil which was added to 30 ml washing
147 buffer. Samples were sonicated (Branson 5510; Branson Ultrasonics Corp, Danbury CT; 135W
148 42KHZ) for 5 min and agitated for 10 min. This method has been shown to remove from 60 to

149 over 80% of all cells of epiphytic bacterial species from leaves, depending on the physical
150 environment in which the leaves were previously exposed (26). Appropriate 10-fold dilutions
151 were plated on 10% Trypticase Soy Agar (TSA, Difco) containing 1.5% agar and natamycin
152 (21.6 µg/ml) and incubated for four days at 28 °C. The remainder of the wash buffer or water
153 sample was filtered on individual 0.2 µm Nalgene® Filter Funnels (Thermo Scientific) and frozen
154 for subsequent DNA isolation. In total, we processed 8 cultures for each plant type, 6 for soil,
155 and 4 for water. One head lettuce sample did not yield cultured bacteria (June) and thus was
156 excluded from analyses that incorporated culture-dependent data.

157

158 **Biosurfactant detection.**

159 The atomized oil assay was performed as previously described (25); 100 randomly-chosen
160 bacterial colonies from dilution plates for each sample were evenly spotted onto TSA agar plates
161 using sterile toothpicks and grown overnight before screening. An artist's airbrush was used to
162 apply a fine mist of mineral oil droplets, and biosurfactant halos were visualized with an oblique
163 source of bright light. Equivalent samples of bacterial colonies that either produced or did not
164 produce were then separately collected from a given source with sterile toothpicks and
165 transferred to a moistened filter paper (Whatman #1).

166

167 **DNA extraction and bar-coded pyrosequencing.**

168 Filter papers (0.2 µm) from whole-leaf washes were cut into strips and added to bead-beating
169 tubes and processed with the PowerSoil® DNA isolation kit (MoBio) according to the
170 manufacturer's protocol. Whatman filter papers containing either biosurfactant producing or
171 non-producing bacteria from the atomized oil assay were added to a Falcon tube with 40 ml

172 washing buffer and 5-10 sterile glass beads and vortexed for 1 min. From this slurry, 4 ml was
173 then centrifuged and the pellet was added to the PowerSoil[®] kit and processed as mentioned
174 above. Triplicate PCR reactions were performed on extracted DNA as described previously (27),
175 amplifying a 16S region encompassing the hypervariable V5-V7 regions with the primers 799f
176 and 1492r which exclude chloroplast amplification (28). PCR products for each sample were
177 pooled and run on an electrophoresis gel. Bands corresponding to approximately 735 bp were
178 excised and extracted from the gel (Qiaquick). Final DNA concentrations were determined with
179 a Qubit[®] kit (Invitrogen). The final multiplex for a given plant species, soil, or water contained
180 eight, six, or four, respectively, total community DNA samples and four pooled collections of
181 biosurfactant producing and non-producing bacterial colonies (one pool each of surfactant
182 producers and non-producers for each collection period). DNA of equal quantity for all samples
183 were pooled to 15 ng/ μ l and sent to the W.M. Keck Center for Comparative and Functional
184 Genomics (University of Illinois at Urbana-Champaign) facility for downstream processing.
185 Pyrosequencing was performed on 1/8th of a PicoTiter plate on the Roche GS-FLX+ system
186 using software version 2.8 with acyclic flow pattern (Roche). This generated 101,951 high
187 quality reads.

188

189 **Processing of pyrosequencing data.**

190 Sequence analysis was performed using the Quantitative Insights Into Microbial Ecology
191 (QIIME 1.8.0) software package (29). Prior to taxonomic and phylogenetic analysis, the
192 following processing and filtering steps were performed using the standard QIIME pipeline (29):
193 Trimming of sequences to 480 bp and excluding shorter sequences, quality filtering (minimum
194 average score of 25), assignment of sequences to samples based on their barcodes and barcode

195 trimming, and clustering of sequences into operational taxonomic units (OTUs) at the 97%
196 similarity level using UCLUST (30). Representative sequences of each OTU were aligned
197 against the Greengenes core dataset (31) with the PyNAST algorithm (32). Chimeras were
198 detected using ChimeraSlayer (33) and excluded from downstream analysis. OTUs that were
199 observed only once (singletons) were removed from the data set. This process narrowed our
200 analysis to a total of 75,924 reads in 1740 total OTUs. Taxonomic classifications were assigned
201 using UCLUST (30) or the RDP classifier (34) at 80% confidence threshold to enable
202 comparison.

203

204 **Statistical analyses.**

205 For alpha- (Table 1) and beta- (Fig. 6) diversity analyses of individual samples, sequences were
206 rarefied by randomly subsampling 200 reads per sample for 10 iterations. One head lettuce and
207 two soil samples yielded fewer than 200 reads and were excluded from these analyses. In order
208 to compare the cultured and uncultured (total) communities, individual samples were combined
209 in the following manner: For culture-independent samples, each culture-independent sample was
210 rarefied to 100 reads (the number of cultured colonies from each sample that were screened and
211 pooled) and reads belonging to each sample type in each collection period were additively
212 combined. For cultured samples, we combined the reads from each producing and non-producing
213 pair for each sample type in each collection period, rarefied to the physical number of colonies
214 added to each pool. As an example, because 400 bacterial isolates were screened from spinach in
215 June (cumulatively from four samples), and 119 of these produced surfactants, we added 119
216 reads from the cultured spinach surfactant producing pool in June with 281 reads from the
217 cultured spinach non-producing June pool, to provide a balanced comparison against the

218 combined 400 rarefied reads from the culture-independent June spinach samples. Because all
219 samples within a comparison (replicates and corresponding culture pools; 4 to 6 samples) were
220 required in order for comparisons to be calculated, we were unable to increase the subsample
221 read-size without reducing the total number of comparisons available for analysis. All
222 rarefaction/pooling procedures were repeated for 10 iterations in order to mitigate random
223 sampling bias. Communities presented in Fig. 2 are an average of the two collections (April and
224 June) for each sample type.

225

226 The combined samples (rarefied to 200 reads each) were subsequently analyzed by calculating
227 distance matrices (weighted and unweighted UniFrac, Bray-Curtis, binary and abundance-based
228 Jaccard) for each of the 10 iterations, and these were averaged and used to generate nonmetric
229 multidimensional scaling plots (NMDS). Nonparametric tests were conducted using the R
230 package Vegan (35), accessed through the Qiime interface, by the functions ‘adonis’
231 (PERMANOVA) and ‘anosim’ (mantel). Student’s *t*-tests were performed for basic comparisons
232 of collected data or beta diversity metrics. Further statistical comparisons of collection and
233 taxonomic data were performed in Excel and Statistica (StatSoft). Table S2 was generated by
234 averaging the 10 iterations of the combined subsamples and only including OTUs that had an
235 average observation frequency in culture-independent samples of at least one (OTUs which
236 represented at least 1% of the community on at least one leaf/soil/water sample had a reasonable
237 chance of being cultured if they were indeed culturable, since 100 colonies were screened per
238 sample). For these OTUs, it was determined what the “culturability” of each OTU was by
239 comparing the relative proportion of that OTU in the culture-independent reads compared to the
240 cultured reads. This table captured 96.8% of the cultured OTUs, and 88.7% of the culture-

241 independent OTUs. Surfactant production data was calculated by first extrapolating the estimated
242 number of colonies of each OTU that were recorded as either producers or non-producers
243 (proportion of reads in a given sorted sample multiplied by the number of colonies placed in the
244 sample) and then determining the overall percent of estimated colonies that were assigned as
245 producers.

246

247 **Nucleotide sequence accession numbers.**

248 The high-throughput sequencing data were submitted to the public Sequence Read Archive
249 (SRA) of NCBI under BioProject PRJNA309173.

250

251 **RESULTS**

252 **Surfactant producers are prevalent on leaf surfaces.**

253 To address surfactant production on leaves we chose agricultural plants from commercial fields
254 in the Salinas, California growing region. On each of two dates, four replicate samples of the
255 interior leaves from head lettuce and romaine lettuce, spinach, as well as three adjacent soil and
256 two irrigation water samples were examined. Bacterial cells were recovered by washing of leaves
257 and a portion of each sample was applied to a low-nutrient medium to culture as many cells as
258 possible, while DNA recovered from cells directly washed from leaves or recovered from other
259 samples was used to determine the total community composition by high-throughput sequencing
260 of amplified small subunit ribosomal genes. The fraction of biosurfactant-producing bacteria was
261 assessed by screening 100 randomly selected bacterial colonies using an atomized oil assay for
262 each sample (Table 1). A higher proportion of the culturable bacteria on leaf surfaces exhibited
263 biosurfactant production (30%) than those in soil (15%), while a significantly lower proportion

264 (2%) of bacteria from water samples produced surfactants when tested on agar plates (Table 1).
265 Screened colonies were then manually sorted to create pools of strains that either produced or did
266 not produce biosurfactants from each habitat during each collection period. DNA was
267 subsequently isolated from the pooled, cultured strains to ascertain the composition of the
268 surfactant producing and non-producing members of the cultured community by high-throughput
269 sequencing of amplified small subunit ribosomal genes.

270

271 A positive correlation was observed between the frequency of surfactant production in the
272 bacterial community and the total population size of culturable bacteria on a given leaf. This
273 relationship was seen for each plant species as well as when leaves of all species were considered
274 as a group ($r^2 = 0.24$, $p = 0.01$) (Fig. 1). When soil and water samples were included in the
275 correlation analysis, no such relationship was observed.

276

277 **Comparison of the culture-independent and culture-dependent communities of edible**
278 **plants, soil and water.**

279 To determine the extent to which the epiphytic bacterial communities on these various plant
280 species were culturable, community composition was assessed from DNA recovered directly
281 from leaves (total bacteria) as well as DNA from the pooled cultured communities (see materials
282 and methods). Chao and Faith's phylodiversity (PD) estimates calculated from the culture-
283 independent approach revealed that the spinach leaves harbored more complex communities than
284 interior romaine leaves, and that both were more complex than interior head lettuce leaves (Table
285 1). Similar to other studies (3, 36), we observed much higher diversity in soil samples than on
286 leaf surfaces. The diversity of irrigation water as measured by Chao was similar to that of leaves

287 while estimates made using Faith's PD suggested that water communities were more
288 phylogenetically diverse than those on the edible plants, although still substantially less than soil
289 communities.

290

291 Not surprisingly, the compositions of the abundant taxonomic groups of bacteria identified using
292 culture-independent analyses were much more similar on the various plant species studied than
293 when compared to nearby soil and water (Fig. 2). Similar to other studies of edible leaf

294 communities (27, 37), Pseudomonadales (primarily *Pseudomonas*) was the most dominant taxon
295 observed on leaf surfaces. While on average they represented 46% of the bacteria on leaves, their
296 proportion on individual leaf samples varied widely, ranging from 4 to 86%. Burkholderiales

297 (14%) and Enterobacteriales (20%) were also commonly found on the leaves of all three plant

298 species. In agreement with corresponding estimates of bacterial diversity, the fewest orders of
299 epiphytic bacteria were observed on head lettuce while the most were observed on spinach

300 leaves. Furthermore, the plant species significantly influenced the leaf community as measured

301 by an anosim test based on unweighted UniFrac distance matrices ($r = 0.16$, $p = 0.02$). Seasonal
302 dynamics have been reported to impact the phyllosphere community (27) and we also observed a

303 significant influence of the sampling date on the leaf community by anosim ($r = 0.24$, $p = 0.005$).

304 In contrast to leaf samples, we observed a much larger variety of bacterial taxa in agricultural

305 soils, consistent with the relatively high Chao and PD values measured. These taxa were also

306 more evenly distributed, lacking a dominant taxon, a common observation made in other

307 unsaturated surface soils (38). Furthermore, we observed a low density of archaea

308 (Nitrososphaerales, 3%) in soil samples, but chose to exclude these from our analysis due to their

309 inability to be cultured with our laboratory conditions.

310

311 A prior study by Rastogi *et al.* (27) provided a comprehensive culture-independent analysis of
312 romaine epiphytes collected from the same growing region as our study but in a different year.

313 The results of that study are compared to the current study in Figure 2. Whereas we sampled only
314 internal leaves from romaine plants, samples in the earlier study also included the most exterior
315 leaves. All orders that were observed at a higher frequency in soil than on romaine leaves in this
316 study also constituted a higher proportion of the community on spinach in this study as well as
317 on romaine leaves in the earlier study. These orders also constituted a lower relative proportion
318 of the bacteria recovered from head lettuce. Thus both the outer leaves of romaine lettuce and
319 spinach appear to harbor bacteria of soil origin to a greater degree than that of the inner leaves of
320 either romaine or head lettuce. When compared at the genus level of phylogenetic classification,
321 *Massilia* was observed at a relatively high relative abundance (5%) in the earlier study but was
322 not found in our analysis. However, Rastogi *et al.* used the RDP pipeline for phylogenetic
323 analysis while we used UCLUST. When the phylogenetic placement of our sequences were re-
324 assigned using the RDP classifier, some of our unassigned OTUs previously grouped within
325 Burkholderiales were re-classified as *Massilia* (data not shown). The only genera observed in the
326 earlier study of romaine lettuce samples at a frequency of 1% prevalence or higher that we did
327 not also find in this study, even when both data sets were analyzed with the RDP classifier, were
328 *Duganella* (2%) and *Naxibacter* (1%), neither of which were considered part of the “core”
329 romaine community.

330

331 It was possible to address the extent to which cultured bacterial communities in different habitats
332 reflected those determined by culture-independent methods. An accurate estimate of the

333 composition of cultured bacterial communities was obtained by sequencing small subunit
334 ribosomal amplicons in pools of large numbers of randomly collected biosurfactant producing
335 and non-producing bacterial strains. Because the number of colonies placed into each cultured
336 pool was known and similar volumes of each colony were introduced, we used weighted values
337 to reconstruct the composition of overall cultured communities in a given sample. In this manner,
338 we could compare the abundance of different culturable bacterial orders (representing both
339 biosurfactant producers and non-producers) with that determined from sequencing of amplicons
340 from cells directly recovered from samples (Fig. 2). Overall, the relative abundance of a given
341 bacterial order in cultured communities closely matched that of total (culture-independent)
342 communities on the leaves of a given plant species (Fig. 2). All orders of bacteria detected on
343 plants at a frequency greater than 0.2% in culture-independent analyses were also present within
344 the cultured communities. In contrast, the relative portion of different bacterial orders of cultured
345 strains from soil often was very different from that estimated by direct DNA sequencing (Fig. 2).
346 For example, Bacillales dominated the cultured community (72%), but represented only 13% of
347 the sequences from DNA isolated directly from soil. Orders that comprised more than 1% of the
348 uncultured soil population but which were not found in any cultures included Myxococcales,
349 Rhodospirillales, Rubrobacterales and an unidentified order belonging to Gemmatimonadetes.
350 For agricultural water samples, the relative proportions of several dominant taxa differed
351 substantially between cultured and total bacterial populations; several bacterial orders dominant
352 in the culture-independent communities such as the Aeromonadales were much less abundant in
353 the cultured bacterial community and vice a versa (Fig. 2).
354

355 While at the level of order bacterial cells on leaves appeared to be highly culturable, we
356 examined culturability at a higher taxonomic resolution to determine if there were taxa that were
357 less likely to be culturable. In order to estimate the culturability of the bacteria in our study, we
358 calculated the ratio of the proportional incidence of occurrence of a given genus with that found
359 in culture-independent analyses (Table 2) taking into context the habitat in which they were
360 identified and the class in which they were found. For the genera examined (those with a
361 prevalence of at least 0.2% in a given habitat), 60% of the genera from plants were cultured, 40%
362 from the water samples and 18% from soil. Additionally, more than half of the genera that were
363 cultured from plants had very similar proportional representations in the cultured and total reads
364 (incidence ratios close to one). We also distinguished these strains at the OTU-level of
365 classification for a given habitat and for their overall prevalence (Table S1 and S2, respectively).
366 Most of the common OTUs of culturable bacteria on leaves were also the most dominant in total
367 bacterial communities as determined by culture-independent assessment. For instance, the OTU
368 corresponding to surfactant-producing *Pseudomonas viridiflava* accounted for 6% of culturable
369 bacteria on leaves, and also represented 13% of sequences of DNA directly isolated from leaves.
370 Likewise, an OTU that includes *Pseudomonas fluorescens* was very dominant on leaves,
371 accounting for 49% of the cultured bacteria and 30% of all sequences of directly isolated DNA.
372 Overall, 40 OTUs were detected at a frequency of at least 0.2% of the sequences determined
373 from direct DNA isolation from leaves. Only six of these 40 OTU's were not represented in the
374 cultured bacteria from leaves (Table S1). In contrast, 60 of the 76 OTUs detected at a frequency
375 of at least 0.2% of the sequences determined from direct DNA isolation from soil were not
376 cultured. Some genera from Table 2 were eliminated at this resolution if their component OTUs
377 cumulatively accounted for >0.2% of sequences, but no OTU individually reached this

378 prevalence. Four of the six OTUs on leaves without culturable representatives were present at
379 very low frequency in the culture-independent community (less than 0.35%) and their lack of
380 culturable representatives might reflect inadequate sampling for such relatively rare members.
381 Interestingly however, an OTU representing the genus *Alkanindiges* (Pseudomonadales), an
382 alkane degrader, was regularly observed in sequences from direct DNA isolated from leaves (1%
383 of total sequences) but was not encountered in the cultured community. Likewise, an OTU
384 representing *Buchnera* (Enterobacteriales), an obligate and unculturable aphid endosymbiont,
385 was found only in one collection of spinach leaves assessed by the culture independent method
386 (at a high density up to 14% of sample sequences). Thus, while most bacterial taxa on leaves
387 appear to be highly culturable, at lower phylogenetic levels a few culture-resistant community
388 members, presumably having specialized habitat requirements, can be identified. The relative
389 cultivability of all OTUs that were observed at a frequency of at least 1% in at least one sample
390 of DNA directly isolated from various habitats, calculated as a ratio of its frequency of
391 occurrence in cultured cells of that same sample is shown (Table S2).

392

393 **Comparison of biosurfactant-producing and non-producing bacteria**

394 The richness and diversity of surfactant-producing and non-producing bacterial communities was
395 estimated by high throughput sequencing of amplicons derived from pooled cultures sorted for
396 this phenotype. A minimum of 80 and 190 isolates, respectively, in pools of surfactant producers
397 and non-producers, were used to generate Chao estimates of the total number of different taxa in
398 each pool (Table 1); if insufficient isolates were obtained during culturing (such as surfactant
399 producers from water), Chao estimates were excluded from this analysis. Spinach had a tendency
400 to host more diverse populations of surfactant-producing bacteria than the other plant species

401 examined, although there were not enough replications to determine statistical significance. In
402 general, more phylogenetic diversity was observed in pools of bacteria that did not produce
403 biosurfactants compared to those that did (Table 1).
404
405 The taxonomic placement of those bacteria capable of producing biosurfactants differed
406 somewhat between the different habitats (Fig. 3). The dominant surfactant-producing bacteria
407 recovered from plants were Pseudomonadales (89% of all biosurfactant producers recovered
408 from this habitat). In addition, Flavobacteriales, primarily *Chryseobacterium*, which has not
409 previously been demonstrated to produce biosurfactant, was often isolated from plants (7% of
410 isolates). While not frequently isolated, several strains of Xanthomonadales (1% of all
411 biosurfactant producing isolates from plants) and Enterobacteriales, specifically *Erwinia* sp., (1%
412 of isolates) that produced biosurfactant were recovered from plants. Curiously, unlike other
413 orders that were consistently isolated from a given habitat over time, Exiguobacteriales
414 (*Exiguobacterium*) accounted for 5% of the sequences recovered from biosurfactant-producing
415 bacteria cultured from romaine lettuce in June (and DNA signatures of this taxon were also
416 observed on each romaine sample subjected to direct DNA sequencing in June, with an average
417 frequency of 13%), but cultured representatives capable of biosurfactant production were not
418 observed in any other leaf samples, and thus it may not be a common biosurfactant producing
419 taxon on leaves. In addition, very small numbers of Bacillales (*Bacillus* and *Paenibacillus*) were
420 identified as cultured biosurfactant producers in a single surfactant pool from romaine leaves, yet
421 were commonly recovered in relatively high numbers from soil. A few additional biosurfactant-
422 producing taxa were found rarely in leaf samples (*Agrobacterium* and *Sphingobacterium*),
423 primarily from spinach. Due to our sampling process, it was impossible to determine if any taxa

424 were erroneously introduced into our pools of surfactant-producing bacteria due to the
425 occurrence of mixed colonies. Therefore, we have included in Table S2 a column that includes
426 the frequency at which a given cultured OTU was observed to be a surfactant producer (as
427 opposed to a non-producer) for all OTUs that occurred at a frequency of 1% or higher in at least
428 one sample, with the presumption that a taxon that was frequently observed to produce
429 surfactants is most likely a true surfactant producer. Additionally, we have created a large library
430 of purified biosurfactant producing strains such as *Agrobacterium*, *Chryseobacterium* and
431 *Sphingobacterium*, many of which were taxa that were frequently identified as producers in this
432 survey. For the most part, OTUs that were observed to produce surfactants at low frequencies
433 (Table S2) belonged to families and genera that also contained one or more OTUs that produced
434 with high frequencies (or were observed in our strain library of biosurfactant producers).
435 Furthermore, many cultured families and genera were never observed in the biosurfactant-
436 producing pools, indicative of very little contamination from mixed colonies. However, one
437 OTU belonging to the family Oxalobacteraceae was observed in our surfactant pools with only
438 1% frequency of being a surfactant producer, whereas the 8 other Oxalobacteraceae OTUs (5 of
439 which were cultured) were not observed as surfactant producers. Therefore, this OTU was likely
440 introduced as part of a mixed colony and we make no claims on its ability to produce
441 biosurfactants.

442
443 The incidence of culturable biosurfactant-producing bacteria in soil was much less than that of
444 plants. These strains were primarily composed of Bacillales and Pseudomonadales. The
445 incidence of biosurfactant-producing bacteria in irrigation water was also very low, with
446 representatives of only the orders Exiguobacterales (60% of the reads recovered), Rhizobiales

447 (*Agrobacterium*, 26%) and Burkholderiales (13%) recovered; only 7 colonies were pooled to
448 obtain these values for water samples.

449

450 The predominant culturable plant-associated taxa incapable of producing biosurfactants (non-
451 producers) differed substantially from those capable of producing biosurfactants. The dominant
452 cultured non-producers recovered from plants were Pseudomonadales (47% overall),
453 Enterobacteriales (29%) and Burkholderiales (13%). It was noteworthy that Pseudomonadales
454 were a predominant taxa of both cultured biosurfactant-producing and non-producing bacteria. It
455 seemed possible that at a higher level of genetic resolution biosurfactant producing and non-
456 producing Pseudomonadales might be distinguished from each other. When we examined the
457 frequency with which each Pseudomonad OTU was associated with biosurfactant production
458 (Table S2), we observed many more OTUs that included both surfactant producing and non-
459 producing representatives than OTUs that consisted solely of surfactant producers. Two
460 predominant surfactant-producing *Pseudomonas* OTUs were identified as examples at either end
461 of the spectrum (primarily producer or mixture). One, characterized as a *Pseudomonas*
462 *viridiflava*, accounted for 9% of all culture-independent reads, and had a 98% chance of being
463 identified in culture as a surfactant producer. This OTU was observed on leaves of each plant
464 species but was not observed in any soil or water samples. In contrast, another *Pseudomonas*
465 OTU accounted for a staggering 23% of culture-independent reads and included strains that both
466 produced and did not produce biosurfactants (44% chance of being identified in culture as a
467 surfactant producer). While definition of OTUs at a higher level of similarity (99%) identified
468 more OTUs, each of which contained fewer representatives, cultured biosurfactant producers and
469 non-producers were still not unambiguously associated with a given OTU (data not shown).

470 Thus, whereas biosurfactant production can be associated relatively unambiguously with a few
471 OTUs (such as *Pseudomonas viridiflava*), it is clear that 16S rRNA signatures of the V5-V7
472 hypervariable region are not predictive of this phenotype in most cases.

473

474 **Examining underlying microbial structures.**

475 The relative structure and beta diversity of cultured and total bacterial communities from various
476 habitats was addressed using nonmetric multidimensional scaling (NMDS) plots. This
477 abundance-based analysis, generated from weighted UniFrac distances, revealed that both
478 cultured and uncultured bacterial communities from plants were distinct from those in either
479 water or soil (Fig. 4A, Fig. S1). Whereas the separate collections of cultured soil bacteria
480 clustered closer to each other than to total soil bacteria, cultured and total bacteria from head
481 lettuce and romaine leaves were nearly coincident at a given sampling time (particularly in
482 April), highlighting the effectiveness of our culturing method in preserving epiphytic community
483 structure. Other analyses using Bray-Curtis and abundance-based Jaccard metrics revealed
484 similar patterns.

485

486 Because the majority of the leaf-associated bacterial communities were comprised of relatively
487 few OTUs (the most abundant 40 OTUs accounted for 88.8% of all sequences), we also
488 addressed the community structure with binary Jaccard matrices in order to incorporate less
489 common bacteria in our analysis (Fig. 4B). When each OTU was considered independent of its
490 relative abundance, the communities in different habitats were slightly less distinguishable from
491 one another. Overall, cultured bacteria from leaves all clustered together and differed most from
492 those of total soil and water communities. Total bacteria from head lettuce leaves were most

493 similar to cultured bacteria from leaf samples, while those from spinach were much less similar
494 and trending towards soil samples (Fig. 4B). Interestingly, the total bacterial community on one
495 romaine lettuce collection (June) was positioned surprisingly close to that of the cultured
496 bacterial community from water (June). We had already observed that a surfactant-producing
497 *Exiguobacterium* strain was common both on these romaine lettuce leaves and in the nearby
498 irrigation water. NMDS analysis however suggested that these samples shared other bacteria as
499 well. When we expanded our binary Jaccard analysis to include individual leaf samples (instead
500 of only the combined communities), we observed that the communities on romaine samples
501 collected in June were significantly more similar to the cultured communities in water samples
502 from June than that of total (culture-independent) communities in the June water samples
503 ($p < 0.05$). Additionally, the communities on romaine leaves in June are significantly more similar
504 to the cultured communities in water collected in June than are those on romaine leaves collected
505 in April ($p < 0.005$). Although one might speculate that the bacteria on these romaine lettuce
506 leaves are derived from irrigation water that might have been deposited onto leaves, it is
507 important to note that the total bacteria on romaine clustered more with the cultured bacteria
508 from water than from total water bacteria, suggesting that not all waterborne bacteria were
509 equally present on leaves. Rather, the data suggests that specific immigrants (the ones that are
510 easily cultured) from irrigation water subsequently flourished on these leaves. A more detailed
511 analysis of the culturable and total bacterial communities on water and romaine lettuce provides
512 support for the conjecture of selective enrichment of some taxa from water on leaves. As an
513 example, bacteria belonging to orders such as Alteromonadales were present in the total bacterial
514 community of water, and were enriched in both the culturable water communities and the total
515 bacterial community on sprinkle irrigated romaine (Fig. 5). On the other hand, Rhodobacterales

516 were 2% of the total water community in June, but the corresponding OTUs belonging to this
517 order were never observed in cultured water samples or on leaves wetted with that water. Thus,
518 romaine leaves might be considered “green petri dishes” in that they appeared to enable the
519 proliferation of culturable bacteria in irrigation water.

520

521 **Correlating underlying community structure with observable traits.**

522 While we have demonstrated that, when considered over many plant species and samples,
523 cultured bacterial communities on leaves closely resemble that of total bacterial communities, we
524 questioned whether ascertaining particular phenotypes of the cultured community would be
525 predictive of the structure of the total bacterial communities on leaves. Since data for the
526 proportional representation of biosurfactant producers on multiple independent samples of each
527 of three plant species as well as total community composition was available, we determined the
528 relationship of these factors using UniFrac distance-based NMDS plots and hierarchical
529 clustering (Fig. 6A, Fig. S2). This analysis revealed that the individual leaf samples exhibiting
530 the highest proportions of culturable surfactant-producing bacteria grouped together irrespective
531 of the plants on which they were found, whereas those samples having the lowest incidence of
532 surfactant-producing bacteria did not cluster together (Fig. 6A), apparently driven by the higher
533 diversity of non-producing than surfactant-producing bacteria. Probably because of the high
534 diversity of non-producers, there was not a statistically significant relationship between the
535 incidence of biosurfactant production and overall community structure as determined using an
536 adonis analysis.

537

538 Because the majority of the cultured surfactant producers from leaves are Pseudomonads, we
539 specifically tested for the predictive nature of biosurfactant producing bacteria as an indicator for
540 the prevalence of these organisms. A high correlation between surfactant production and the
541 proportion of Pseudomonadales in a given sample was observed ($r^2 = 0.57$, $p < 0.005$) (Fig. 6B).
542 This relationship was very strong for head lettuce ($r^2 = 0.81$, $p < 0.03$) and romaine leaves ($r^2 =$
543 0.83 , $p = 0.01$), but not for spinach leaves ($r^2 = 0.02$, $p > 0.95$). It should also be noted that
544 spinach harbored the most diverse surfactant producers, as well as the largest proportion of
545 Pseudomonadales categorized as non-producers (Fig. 3).

546

547 **DISCUSSION**

548 The majority of biosurfactant research to date has focused on the diverse products synthesized by
549 Pseudomonads and Bacilli (24), and the identification of a broader diversity of biosurfactant-
550 producing bacteria should help identify novel biosurfactants with potential new industrial
551 applications. Our research has previously suggested that the phyllosphere is an environment that
552 hosts particularly high frequencies of biosurfactant-producing bacteria (11). Our current study
553 similarly observed high frequencies of biosurfactant producing bacteria on leaf surfaces (~30%)
554 and provides comprehensive information on the identity of those bacteria that produce
555 surfactants in the phyllosphere. As suggested from limited studies of other plant species (11, 16),
556 we found Pseudomonads to be the predominant biosurfactant-producing taxa on leaves. It was
557 unexpected however that *Chryseobacterium* would be found to be a relatively common genus
558 capable of biosurfactant production on all plant species, as it is a poorly-studied taxon with no
559 prior strains demonstrated to produce surfactant. Given the large number of studies that have
560 attempted to isolate biosurfactant producers from other habitats such as soil and water,

561 *Chryseobacterium* strains with this phenotype might be common only on plants. Other genera
562 from plants that were found to produce surfactants included *Agrobacterium*, *Bacillus*, *Erwinia*,
563 *Exiguobacterium*, *Paenibacillus*, *Rhodanobacter* and *Sphingobacterium* (Table S2). While
564 limited reports have appeared that note surfactant production in some of these taxa, more work
565 will be needed to understand the significance of their biosurfactant production in the natural
566 habitats in which they have been described. In order to address this issue in the future, we have
567 established a large library of biosurfactant-producing isolates including many representative
568 isolates such as *Chryseobacterium* for further studies. Our current report uses one assay to
569 broadly assess biosurfactant production, and future purification, tensiometric measurements and
570 chemical characterization will help elucidate how these surfactants differ in structure and
571 function.

572

573 Our approach to also identify the bacteria incapable of biosurfactant production provided insight
574 into the frequency with which biosurfactant production is exhibited by a given OTU (Table S2).
575 It is clear that biosurfactant production is not an invariant phenotype associated with a given
576 taxon and that strains with similar ribosomal sequences are physiologically diverse. However,
577 some taxa are more frequently observed to produce surfactants than others; this type of
578 information could allow us to infer levels of endogenous biosurfactant production on different
579 plant species in future culture-independent studies of the phyllosphere.

580

581 Previously, we have observed a positive correlation between the incidence of surfactant
582 producing bacteria and the overall population size of *Pseudomonas syringae* on a given leaf of
583 inoculated plants using mixed inoculum experiments (10), suggesting that biosurfactant

584 production increases the fitness of *P. syringae* on leaf surfaces. Similarly in this study, we
585 observed a positive correlation between the relative abundance of surfactant producing bacteria
586 and total culturable bacterial population size. It thus appears that biosurfactants might be
587 generally beneficial to bacteria on leaves. This hypothesis is also supported by the overall high
588 frequency of surfactant producers encountered in the phyllosphere compared to other
589 environments.

590

591 Considerable evidence suggests that most bacteria in most habitats remain uncultured (19).
592 These observations have prompted many in the microbial ecology community to disregard
593 community analyses that only examine those community members that are culturable.
594 Unfortunately only a few other studies have addressed the culturability of phyllosphere bacteria
595 (22, 23, 39). The low incidence of culturability is particularly noteworthy in soil where the
596 apparent presence of a heterogeneous mixture of microenvironments with distinct chemical and
597 physical restrictions presumably select for bacteria with distinct features capable of exploiting
598 only localized regions of the soil (19). Indeed, in this study we were only able to culture
599 representatives of 16 of the 76 relatively abundant bacterial OTUs in the soil community.

600 However, this study supports the findings of a few recent studies that reveal that not all habitats
601 harbor bacteria that are similarly difficult to culture since cultured representatives of 34 out of 40
602 relatively-abundant (>0.2%) OTUs on leaves could be recovered. Although this was surprising, it
603 was not unprecedented; a recent study of phyllosphere bacteria using fluidic force microscopy
604 demonstrated that 69 of 100 randomly-chosen cells of clover epiphytes recovered from leaves
605 were capable of growth in a low-nutrient medium (22), and high-throughput sequencing efforts
606 in *Arabidopsis* similarly found nearly 50% of the relatively-abundant (>0.1%) OTUs determined

607 by direct DNA sequencing to also be represented in a library of cultured strains from these same
608 samples (23). Furthermore, a previous study on the apple phyllosphere sequenced 300 clones and
609 found 85% of them corresponded to known cultured OTUs, even though DAPI staining would
610 have indicated that less than 1% of the cells were enumerated by culturing (39). Thus while a
611 relatively high proportion of the bacterial diversity on leaves apparently is culturable, it also
612 suggests that a high proportion of the bacteria on leaves are dead at a given time in some
613 settings, a finding consistent with the harsh environmental conditions and heterogeneous
614 nutritional distribution expected on leaves and supported by some studies of the viability of
615 particular taxa on leaves (40).

616

617 Although the leaf surface probably offers epiphytes a diversity of habitats differing in water
618 availability and nutrient abundance, etc., it appears that much of the growth of epiphytes on
619 leaves occurs at relatively few sites where nutrients are locally more abundant (41). It seems
620 likely that these most favorable sites for bacterial growth provide primarily plant-produced
621 sugars and other simple nutrients (42, 43). Furthermore, those bacteria that could thrive on the
622 simple sugars, organic acids and amino acids that might predominate on leaves (44) probably
623 also can grow well on laboratory media consisting of low concentrations of similar nutrients such
624 as used in this study. The high similarity in the community composition seen between those on
625 romaine leaves and the bacteria that could be cultured from the irrigation water suggests that the
626 converse may also be true; bacteria capable of growth on simple nutrients in culture might
627 readily colonize leaves given other appropriate environmental conditions that would otherwise
628 limit their growth. It is generally agreed that the leaf surface is a harsh environment with rapid
629 fluctuations in bacterial population densities; it also would be expected that the high flux of UV

630 irradiation and processes that would lead to cell lysis would quickly destroy genetic signatures of
631 the bacteria upon death. Such rapid genetic degradation would cause the analysis of epiphytic
632 communities by sequencing of total extractable DNA to resemble that of the culturable/viable
633 bacteria, assuming that most of the bacteria on leaves could be cultured.

634

635 Similar to studies of other edible green vegetables, we found that the majority of bacteria on our
636 leaf samples could be assigned to Pseudomonadales and Enterobacteriales (27, 45, 46). Although
637 *Methylobacterium* have been noted as important members of soybean, Arabidopsis, and clover
638 leaves (3), they were only observed at low densities on a few of our leaf samples (primarily in
639 April). Likewise, *Sphingomonas* species were also commonly found on other plants (3), but were
640 quite uncommon on the plants studied here. While the most abundant bacterial taxa on leaves
641 had a very high probability of being cultured, several taxa with apparently unique features that
642 would make them unlikely to be easily cultured were also detected. *Alkanindiges* were never
643 cultured but were commonly observed as a small proportion of the total bacterial community
644 from leaves, similar to other studies of lettuce (6, 27). It seems likely that this taxon would be
645 common on leaves due to its apparent ability to catabolize the alkanes found in leaf surface
646 waxes. As such, they may coexist with other leaf colonists because of nutritional niche
647 partitioning, and thus would not be cultured with our standard media that lacks the apparently
648 unique carbon sources that they consume. In contrast, the obligate aphid gut symbiont that was
649 detected by culture independent methods on the June spinach samples probably represents a
650 contamination that was likely associated with an insect infestation; our sampling methods would
651 not have excluded small aphid body parts or gut fluid.

652

653 Other studies have used culture independent methods to show that different plant species harbor
654 somewhat distinct bacterial communities, including previous comparisons of head lettuce and
655 romaine, as well as different ecotypes of a single plant species (*Arabidopsis*); this is perhaps due
656 to varying leaf surface morphology, differential availability of soluble carbohydrates, and/or
657 other compounds present on the leaf surface (6, 47). Similarly, we observed significant
658 differences between the types of bacteria present on the different species of plants examined in
659 this study. Head lettuce leaves harbored the least diverse bacterial community while also
660 exhibiting the lowest dissimilarity between its total and culturable community. Conversely,
661 spinach harbored the most diverse communities, and compared to the other plants, its total
662 bacterial communities were the most similar to soil communities (Fig. 4B). It is notable that the
663 structure of leaves of these plant species vary, with head lettuce presumably providing a more
664 protected environment to bacteria due to the closed nature of its canopy. Due to the overlapping
665 leaves, epiphytes likely would be less influenced by external factors and sources of immigrant
666 bacteria such as from soil. In contrast, romaine and spinach both have a more open canopy
667 structure, and the short stature of spinach places its leaves in close proximity to soil, thereby
668 facilitating immigration from this source. Comparison of the results of community analysis of
669 romaine lettuce in this study and that of Rastogi *et al.* (27) provide further support for the impact
670 of the proximity of soil on leaf bacterial communities. Whereas internal leaves of romaine lettuce
671 were sampled in this study Rastogi *et al.* included the most exterior romaine lettuce leaves.
672 While taxa common in soil were infrequent colonists of romaine lettuce in our study, they were
673 relatively much more abundant on the leaves in the Rastogi study, presumably reflecting
674 immigration of soil or soil-borne bacteria to the outermost leaves of their plants. The inner leaves
675 studied here presumably would be resistant to such immigration. The nature of the external

676 environment probably also strongly influences bacterial community composition on leaves, as
677 factors such as temperature, moisture, and light intensity and duration would select for particular
678 bacterial fitness traits or species, but also influence host plants. Other studies have documented
679 leaf community variability due to seasonal dynamics (27, 45, 48, 49). Likewise, we also
680 observed significant differences in leaf community compositions on our agricultural samples
681 collected from the same region but on different dates in April and June. Studies to isolate the
682 particular environmental factors that can lead to seasonal changes in community composition
683 should be fruitful.

684

685 Even if most dominant epiphytes are culturable, we were still surprised at the quantitative
686 similarity of culturable epiphytic communities and those determined by culture independent
687 methodologies. This finding is in disagreement with many other phyllosphere studies that have
688 suggested that the uncultured community is very different than that obtained by culturing.
689 However, these discrepancies are probably due in large part to the different methodologies used
690 in these studies. Some previous studies have used other methods such as broth culturing in which
691 competition *in vitro* could quickly alter the composition of communities from those originally
692 recovered from plants (50). Furthermore, some studies of epiphytic communities have been made
693 on plants where epiphytic populations were likely very low (50, 51) and thus particularly subject
694 to the effects of emigration from outside sources of bacterial inoculum, or from contaminating
695 sequences present in the reagents and equipment used, a problem which has recently been shown
696 to be acute in such a setting (52-54). In our study, we attempted to minimize the over-
697 representation of bacteria capable of rapid growth by culturing the cells on solid media having
698 low nutrient concentrations that presumably would not inhibit more oligotrophic colonists. Using

699 this procedure, the only OTUs that would not be represented in the cultured community either
700 would be unable to grow on plates, or would require more than 4 days to produce a visible
701 colony. (Very few such slow-growing colonies were observed however). Variation in cell size
702 and the presence of extra-cellular compounds could have impacted the biomass of a given
703 bacterium harvested and thus the proportion of its 16S rRNA gene signature in a pool in our
704 process of determining culturable communities. Despite such possible sources of variation the
705 proportional representation of many taxa by culturing closely matched that determined by a
706 culture-independent methodology (ratio close to 1, Tables 2 and S2). Although these conditions
707 appear optimal for epiphytic bacteria, it is likely that different media and/or incubation periods
708 would likely increase the apparent culturability of soil and water bacteria (55).

709

710 Unlike popular belief, this study supports other recent studies that suggest that culture-based
711 assays can be highly practical indicators of the true underlying community structure in certain
712 natural habitats such as leaves. Therefore, surveys of surfactant production or other traits of
713 phyllosphere communities can provide a comprehensive assessment of the diversity of bacteria
714 in the leaf habitat. These results also suggest that the large earlier body of work on phyllosphere
715 microbiology that used culture-based methodologies is robust, and can safely guide new
716 experiments using different technologies.

717

718 **ACKNOWLEDGMENTS**

719 We would like to express our gratitude to the CA Leafy Greens Research Board and cooperating
720 growers for facilitating farm access. We are grateful to Renee Koutsoukis for experimental
721 assistance and Chris Wright at the W.M. Keck Center for Comparative and Functional Genomics

722 for technical assistance. We also would like to thank Despoina Lympelopoulou for critical
723 reading of the manuscript.

724

725 **FUNDING INFORMATION**

726 This work was supported by a USDA-NIFA grant (2011-67017-30097).

727 **REFERENCES**

728

- 729 1. **Hirano SS, Upper CD.** 2000. Bacteria in the leaf ecosystem with emphasis on
730 *Pseudomonas syringae* - a pathogen, ice nucleus, and epiphyte. *Microbiol Mol Biol Rev* **64**:624-
731 653.
- 732 2. **Lindow SE, Brandl MT.** 2003. Microbiology of the phyllosphere. *Appl Environ*
733 *Microbiol* **69**:1875-1883.
- 734 3. **Delmotte N, Knief C, Chaffron S, Innerebner G, Roschitzki B, Schlapbach R, von**
735 **Mering C, Vorholt J.** 2009. Community proteogenomics reveals insights into the physiology of
736 phyllosphere bacteria. *Proc Natl Acad Sci USA* **106**:16428-16433.
- 737 4. **Knief C, Delmotte N, Chaffron S, Stark M, Innerebner G, Wassmann R, von**
738 **Mering C, Vorholt J.** 2012. Metaproteogenomic analysis of microbial communities in the
739 phyllosphere and rhizosphere of rice. *ISME J* **6**:1378-1390.
- 740 5. **Lambais MR, Crowley DE, Cury J, Büll R, Rodrigues RR.** 2006. Bacterial diversity
741 in tree canopies of the Atlantic forest. *Science* **312**:1917.
- 742 6. **Hunter PJ, Hand P, Pink D, Whipps JM, Bending GD.** 2010. Both leaf properties and
743 microbe-microbe interactions influence within-species variation in bacterial population diversity
744 and structure in the lettuce (*Lactuca* species) phyllosphere. *Appl Environ Microbiol* **76**:8117-
745 8125.
- 746 7. **Redford AJ, Bowers RM, Knight R, Linhart Y, Fierer N.** 2010. The ecology of the
747 phyllosphere: geographic and phylogenetic variability in the distribution of bacteria on tree
748 leaves. *Environ Microbiol* **12**:2885-2893.

- 749 8. **Finkel OM, Burch AY, Lindow SE, Post AF, Belkin S.** 2011. Geographical location
750 determines the population structure in phyllosphere microbial communities of a salt-excreting
751 desert tree. *Appl Environ Microbiol* **77**:7647-7655.
- 752 9. **Vokou D, Vareli K, Zarali E, Karamanoli K, Constantinidou H-IA, Monokrousos N,**
753 **Halley J, Sainis I.** 2012. Exploring biodiversity in the bacterial community of the Mediterranean
754 phyllosphere and its relationship with airborne bacteria. *Microb Ecol* **64**:714-724.
- 755 10. **Burch AY, Zeisler V, Yokota K, Schreiber L, Lindow SE.** 2014. The hygroscopic
756 biosurfactant syringafactin produced by *Pseudomonas syringae* enhances fitness on leaf surfaces
757 during fluctuating humidity. *Environ Microbiol* **16**:2086-2098.
- 758 11. **Burch A, Browne P, Dunlap C, Price N, Lindow S.** 2011. Comparison of biosurfactant
759 detection methods reveals hydrophobic surfactants and contact-regulated production. *Environ*
760 *Microbiol* **13**:2681-2691.
- 761 12. **Bunster L, Fokkema NJ, Schippers B.** 1989. Effect of surface-active *Pseudomonas* spp.
762 on leaf wettability. *Appl Environ Microbiol* **55**:1340-1345.
- 763 13. **Hildebrand P, Braun P, McRae K, Lu X.** 1998. Role of the biosurfactant viscosin in
764 broccoli head rot caused by a pectolytic strain of *Pseudomonas fluorescens*. *Can J Plant Pathol*
765 **20**:296-303.
- 766 14. **Hernandez-Anguiano AM, Suslow TV, Leloup L, Kado CI.** 2004. Biosurfactants
767 produced by *Pseudomonas fluorescens* and soft-rotting of harvested florets of broccoli and
768 cauliflower. *Plant Pathol* **53**:596-601.
- 769 15. **Kruijt M, Tran H, Raaijmakers JM.** 2009. Functional, genetic and chemical
770 characterization of biosurfactants produced by plant growth-promoting *Pseudomonas putida* 267.
771 *J Appl Microbiol* **107**:546-556.

- 772 16. **D'aes J, De Maeyer K, Pauwelyn E, Höfte M.** 2010. Biosurfactants in plant–
773 *Pseudomonas* interactions and their importance to biocontrol. *Environ Microbiol Rep* **2**:359-372.
- 774 17. **Pauwelyn E, Huang C-J, Ongena M, Leclère V, Jacques P, Bleyaert P, Budzikiewicz**
775 **H, Schafer M, Hofte M.** 2013. New linear lipopeptides produced by *Pseudomonas cichorii* SF1-
776 54 are involved in virulence, swarming motility, and biofilm formation. *Mol Plant Microbe*
777 *Interact* **26**:585-598.
- 778 18. **Banat IM, Franzetti A, Gandolfi I, Bestetti G, Martinotti MG, Fracchia L, Smyth T,**
779 **Marchant R.** 2010. Microbial biosurfactants production, applications and future potential. *Appl*
780 *Microbiol Biotechnol* **87**:427-444.
- 781 19. **Ritz K.** 2007. The plate debate: cultivable communities have no utility in contemporary
782 environmental microbial ecology. *FEMS Microbiol Ecol* **60**:358-362.
- 783 20. **Müller T, Ruppel S.** 2014. Progress in cultivation-independent phyllosphere
784 microbiology. *FEMS Microbiol Ecol* **87**:2-17
- 785 21. **Niwa R, Yoshida S, Furuya N, Tsuchiya K, Tsushima S.** 2011. Method for simple and
786 rapid enumeration of total epiphytic bacteria in the washing solution of rice plants. *Can J*
787 *Microbiol* **57**:62-67.
- 788 22. **Stiefel P, Zambelli T, Vorholt JA.** 2013. Isolation of optically targeted single bacteria
789 by application of fluidic force microscopy to aerobic anoxygenic phototrophs from the
790 phyllosphere. *Appl Environ Microbiol* **79**:4895-4905.
- 791 23. **Bai Y, Müller DB, Srinivas G, Garrido-Oter R, Potthoff E, Rott M, Dombrowski N,**
792 **Münch P, Spaepen S, Remus-Emsermann M, Hüttel B, McHardy A, Vorholt J, Schulze-**
793 **Lefert P.** 2015. Functional overlap of the Arabidopsis leaf and root microbiota. *Nature* **528**:364-
794 369.

- 795 24. **Raaijmakers J, de Bruijn I, Nybroe O, Ongena M.** 2010. Natural functions of
796 lipopeptides from *Bacillus* and *Pseudomonas*: more than surfactants and antibiotics. FEMS
797 Microbiol Rev **34**:1037-1062.
- 798 25. **Burch A, Shimada B, Browne P, Lindow S.** 2010. Novel high-throughput detection
799 method to assess bacterial surfactant production. Appl Environ Microbiol **76**:5363-5372.
- 800 26. **O'Brien R, Lindow S.** 1989. Effect of plant-species and environmental-conditions on
801 epiphytic population sizes of *Pseudomonas syringae* and other bacteria. Phytopathol **79**:619-627.
- 802 27. **Rastogi G, Sbodio A, Tech JJ, Suslow TV, Coaker GL, Leveau JH.** 2012. Leaf
803 microbiota in an agroecosystem: spatiotemporal variation in bacterial community composition on
804 field-grown lettuce. ISME J **6**:1812-1822.
- 805 28. **Chelius MK, Triplett EW.** 2001. The diversity of archaea and bacteria in association
806 with the roots of *Zea mays* L. Microb Ecol **41**:252-263.
- 807 29. **Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK,**
808 **Fierer N, Gonzalez A, Goodrich J, Gordon J, Huttley G, Kelley S, Knights D, Koenig J, Ley**
809 **R, Lozupone C, McDonald D, Muegge B, Pirrung M, Reeder J, Sevinsky J, Turnbaugh P,**
810 **Walters W, Widmann J, Yatsunenko T, Zaneveld J, Knight R.** 2010. QIIME allows analysis
811 of high-throughput community sequencing data. Nat Methods. **7**:335-336.
- 812 30. **Edgar RC.** 2010. Search and clustering orders of magnitude faster than BLAST.
813 Bioinform **26**:2460-2461.
- 814 31. **DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, Huber T,**
815 **Dalevi D, Hu P, Andersen G.** 2006. Greengenes, a chimera-checked 16S rRNA gene database
816 and workbench compatible with ARB. Appl Environ Microbiol **72**:5069-5072.

- 817 32. **Caporaso JG, Bittinger K, Bushman FD, DeSantis TZ, Andersen GL, Knight R.**
818 2010. PyNAST: a flexible tool for aligning sequences to a template alignment. *Bioinform*
819 **26**:266-267.
- 820 33. **Haas BJ, Gevers D, Earl AM, Feldgarden M, Ward DV, Giannoukos G, Ciulla D,**
821 **Tabbaa D, Highlander S, Sodergren E, Methe B, DeSantis TZ, The Human Microbiome**
822 **Consortium, Petrosino JF, Knight R, Birren BW.** 2011. Chimeric 16S rRNA sequence
823 formation and detection in Sanger and 454-pyrosequenced PCR amplicons. *Genome Res* **21**:494-
824 504.
- 825 34. **Wang Q, Garrity GM, Tiedje JM, Cole JR.** 2007. Naive Bayesian classifier for rapid
826 assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microbiol*
827 **73**:5261-5267.
- 828 35. **Oksanen J, Blanchet G, Kindt R, Legendre P, Minchin P, O'Hara RB, Simpson GL,**
829 **Solyomos P, Stevens MH, Wagner H.** 2015. *vegan*: Community Ecology Package; R package
830 version 2.2-1. <http://CRAN.R-project.org/package=vegan>.
- 831 36. **Kim M, Singh D, Lai-Hoe A, Go R, Rahim RA, Ainuddin A, Chun J, Adams JM.**
832 2012. Distinctive phyllosphere bacterial communities in tropical trees. *Microb Ecol* **63**:674-681.
- 833 37. **Lopez-Velasco G, Welbaum GE, Boyer RR, Mane SP, Ponder MA.** 2011. Changes in
834 spinach phylloepiphytic bacteria communities following minimal processing and refrigerated
835 storage described using pyrosequencing of 16S rRNA amplicons. *J Appl Microbiol* **110**:1203-
836 1214.
- 837 38. **Zhou J, Xia B, Treves DS, Wu L-Y, Marsh TL, O'Neill RV, Palumbo AV, Tiedje**
838 **JM.** 2002. Spatial and resource factors influencing high microbial diversity in soil. *Appl Environ*
839 *Microbiol* **68**:326-334.

- 840 39. **Yashiro E, Spear R, McManus P.** 2011. Culture-dependent and culture-independent
841 assessment of bacteria in the apple phyllosphere. *J Appl Microbiol* **110**:1284-1296.
- 842 40. **Monier J, Lindow S.** 2004. Frequency, size, and localization of bacterial aggregates on
843 bean leaf surfaces. *Appl Environ Microbiol* **70**:346-355.
- 844 41. **Remus-Emsermann MN, Tecon R, Kowalchuk GA, Leveau JH.** 2012. Variation in
845 local carrying capacity and the individual fate of bacterial colonizers in the phyllosphere. *ISME J*
846 **6**:756-765.
- 847 42. **Leveau J, Lindow S.** 2001. Appetite of an epiphyte: quantitative monitoring of bacterial
848 sugar consumption in the phyllosphere. *Proc Natl Acad Sci USA* **98**:3446-3453.
- 849 43. **Vorholt JA.** 2012. Microbial life in the phyllosphere. *Nat Rev Microbiol* **10**:828-840.
- 850 44. **Mercier J, Lindow S.** 2000. Role of leaf surface sugars in colonization of plants by
851 bacterial epiphytes. *Appl Environ Microbiol* **66**:369-374.
- 852 45. **Thompson I, Bailey M, Fenlon J, Fermor T, Lilley A, Lynch J, McCormack PJ,**
853 **McQuilken M, Purdy K, Rainey P, Whipps J.** 1993. Quantitative and qualitative seasonal
854 changes in the microbial community from the phyllosphere of sugar beet (*Beta vulgaris*). *Plant*
855 *Soil* **150**:177-191.
- 856 46. **Lopez-Velasco G, Carder PA, Welbaum GE, Ponder MA.** 2013. Diversity of the
857 spinach (*Spinacia oleracea*) spermosphere and phyllosphere bacterial communities. *FEMS*
858 *Microbiol Lett* **346**:146-154.
- 859 47. **Reisberg EE, Hildebrandt U, Riederer M, Hentschel U.** 2013. Distinct phyllosphere
860 bacterial communities on *Arabidopsis* wax mutant leaves. *PLoS ONE* **8**(11):e78613.
861 doi:10.1371/journal.pone.0078613.

- 862 48. **Redford AJ, Fierer N.** 2009. Bacterial succession on the leaf surface: a novel system for
863 studying successional dynamics. *Microb Ecol* **58**:189-198.
- 864 49. **Jackson CR, Denney WC.** 2011. Annual and seasonal variation in the phyllosphere
865 bacterial community associated with leaves of the southern Magnolia (*Magnolia grandiflora*).
866 *Microb Ecol* **61**:113-122.
- 867 50. **Yang C-H, Crowley DE, Borneman J, Keen NT.** 2001. Microbial phyllosphere
868 populations are more complex than previously realized. *Proc Natl Acad Sci USA* **98**:3889-3894.
- 869 51. **Shade A, McManus PS, Handelsman J.** 2013. Unexpected diversity during community
870 succession in the apple flower microbiome. *MBio* **4**(2):e00602-12. doi:10.1128/mBio.00602-12.
- 871 52. **Laurence M, Hatzis C, Brash DE.** 2014. Common contaminants in next-generation
872 sequencing that hinder discovery of low-abundance microbes. *PLoS ONE* **9**(5):e97876.
873 doi:10.1371/journal.pone.0097876.
- 874 53. **Salter SJ, Cox MJ, Turek EM, Calus ST, Cookson WO, Moffatt MF, Turner P,**
875 **Parkhill J, Loman NJ, Walker AW.** 2014. Reagent and laboratory contamination can critically
876 impact sequence-based microbiome analyses. *BMC Biol* **12**:87. doi:10.1186/s12915-014-0087-z.
- 877 54. **Tremblay J, Singh K, Fern A, Kirton ES, He S, Woyke T, Lee J, Chen F, Dangl J,**
878 **Tringe S.** 2015. Primer and platform effects on 16S rRNA tag sequencing. *Front Microbiol*
879 **6**:771. doi:10.3389/fmicb.2015.00771
- 880 55. **Davis KE, Joseph SJ, Janssen PH.** 2005. Effects of growth medium, inoculum size, and
881 incubation time on culturability and isolation of soil bacteria. *Appl Environ Microbiol* **71**:826-
882 834.
- 883
- 884
- 885

886 **FIGURE LEGENDS**

887 **Figure 1.** Relationship between culturable bacterial population sizes in samples of leaves (filled
888 symbols) and other habitats (open symbols) and the proportion of bacteria capable of
889 biosurfactant production. A significant positive correlation was observed for phyllosphere
890 samples ($r^2 = 0.24$, $p = 0.01$). The line drawn represents this relationship: $y = 0.15x - 0.47$.

891

892 **Figure 2.** Composition of total and cultured bacterial communities on leaves and in associated
893 water and soil samples. Culture-independent (total) communities are averages of 4-8 individual
894 samples. Pooled cultured communities (cultured) reflect averages from two separate sampling
895 times. Shown are orders that comprised at least 2% of the sequences in at least one sample of a
896 given habitat. The culture-independent community of romaine as described previously by
897 Rastogi *et al.* (27) is also included (Rastogi*).

898

899 **Figure 3.** Composition of cultured bacteria characterized as surfactant producers and non-
900 producers from various habitats. The bars shown are averages of the proportional incidence of
901 isolation from two separate times of collection. The orders shown were observed in at least 1%
902 of the amplicon sequences in at least one of the pools. Water samples are excluded due to the
903 low number of surfactant producers recovered.

904

905 **Figure 4.** Nonmetric multidimensional scaling (NMDS) plots based on weighted UniFrac (A)
906 and binary Jaccard (B) distance matrices comparing sequences of total bacteria (filled symbols)
907 and cultured bacteria (open symbols) recovered from various habitats (color-coded) sampled in
908 April (circles) or June (triangles).

909

910 **Figure 5.** Comparison of the cultured and total bacterial communities in irrigation water in June,
911 as well as the total bacterial communities on romaine leaves in June and April.

912

913 **Figure 6.** Correlations between frequency of biosurfactant production in cultured bacteria on
914 leaves and culture-independent community structure. Nonmetric multidimensional scaling
915 (NMDS) plot based on weighted UniFrac distances comparing sequences of the total (culture-
916 independent) bacterial communities on individual leaf samples differing in the proportion of
917 surfactant producing bacteria observed in the cultured community (A). Relationship of the
918 proportion of Pseudomonads in the total population of bacteria on leaves (abscissa) to the
919 proportion of surfactant producers in the cultured community (B).

920

921

922

923 **TABLES**924 **Table 1.** Culture-based and culture-independent characterization of different agronomic habitats.

925

	Head lettuce	Romaine	Spinach	Soil	Water
Culture-based:					
% surfactant producers	29 a	28 ab	34 a	15 ab	2 b
Population (log cfu/g)	4.99 a	5.19 a	5.39 a	6.78 b	4.19 a
Culture-independent:					
Chao estimate	57.6 a	76.7 a	91.0 a	294.4 b	86.5 a
Faith's Phylodiversity	1.76 a	2.20 a	2.52 a	6.91 b	3.42 a
Diversity (Chao estimate) of pooled cultured isolates					
Surfactant producers	9.0*	13 a	21.1 a	1.6 *	n/a*
Non-producers	37.5 a	69.4*	38.4 a	41.9 a	68.0*

Values are averages; letters denote values that are significantly different as determined by a student's *t*-test ($p < 0.05$) within row comparisons.

(*) denotes when at least one of the two pools were excluded due to insufficient isolates (surfactant producers < 80 or non-producers < 190) or low 16S rRNA gene reads, prohibiting statistical tests.

926

927 **Table 2.** Cultivability of different genera determined as the ratio of the proportional incidence of a given genus in the cultured bacteria
 928 from a given habitat with that found in culture-independent analyses. Values reflect the number of genera observed in each Class
 929 present in at least 0.2% of the culture-independent reads (by habitat) in columns corresponding to their cultivability ratios.

(Cultivability ratio)	Plants				Water				Soil			
	0	<0.5	~1	>1.5	0	<0.5	~1	>1.5	0	<0.5	~1	>1.5
Acidimicrobiia									2			
Actinobacteria	2								4	1		
Alphaproteobacteria		1			2				9	2		
Anaerolineae									1			
Bacilli	2									1	1	2
Betaproteobacteria	1	2	1		1	1			8	2	1	
Deltaproteobacteria					2				8			
Fimbriimonadetes									1			
Flavobacteriia			1					1	1			
Gammaproteobacteria	1		5		1		1	1	3		1	1
Gemm-2									2			
Gemmatimonadetes									4			
JL-ETNP-Z39									1			
Nitrospira									1			
Pedospaerae									2			
Proteobacteria n/a		1								1		
Rubrobacteria									1			
Solibacteres									1			
Sphingobacteriia	2		1		3	1	1		5			
Sva0725									1			
Thermoleophilia									4			
TK17									2			

930











