The ISME Journal (2012) 6, 1812–1822

www.nature.com/ismej

ORIGINAL ARTICLE

Leaf microbiota in an agroecosystem: spatiotemporal variation in bacterial community composition on field-grown lettuce

Gurdeep Rastogi[1], Adrian Sbodio[2], Jan J Tech[1], Trevor V Suslow[2], Gitta L Coaker[1] and Johan HJ Leveau[1]

[1]*Department of Plant Pathology, University of California, Davis, CA, USA and* [2]*Department of Plant Sciences, University of California, Davis, CA, USA*

The presence, size and importance of bacterial communities on plant leaf surfaces are widely appreciated. However, information is scarce regarding their composition and how it changes along geographical and seasonal scales. We collected 106 samples of field-grown Romaine lettuce from commercial production regions in California and Arizona during the 2009–2010 crop cycle. Total bacterial populations averaged between 10^5 and 10^6 per gram of tissue, whereas counts of culturable bacteria were on average one (summer season) or two (winter season) orders of magnitude lower. Pyrosequencing of 16S rRNA gene amplicons from 88 samples revealed that *Proteobacteria, Firmicutes, Bacteroidetes* and *Actinobacteria* were the most abundantly represented phyla. At the genus level, *Pseudomonas*, *Bacillus*, *Massilia*, *Arthrobacter* and *Pantoea* were the most consistently found across samples, suggesting that they form the bacterial 'core' phyllosphere microbiota on lettuce. The foliar presence of *Xanthomonas campestris* pv. *vitians*, which is the causal agent of bacterial leaf spot of lettuce, correlated positively with the relative representation of bacteria from the genus *Alkanindiges*, but negatively with *Bacillus*, *Erwinia* and *Pantoea*. Summer samples showed an overrepresentation of *Enterobacteriaceae* sequences and culturable coliforms compared with winter samples. The distance between fields or the timing of a dust storm, but not Romaine cultivar, explained differences in bacterial community composition between several of the fields sampled. As one of the largest surveys of leaf surface microbiology, this study offers new insights into the extent and underlying causes of variability in bacterial community composition on plant leaves as a function of time, space and environment.

The ISME Journal (2012) **6,** 1812–1822; doi:10.1038/ismej.2012.32; published online 26 April 2012
Subject Category: microbial population and community ecology
Keywords: phyllosphere; bacterial epiphytes; lettuce; bacterial diversity; pyrosequencing; *Lactuca sativa*

Introduction

The phyllosphere or leaf surface (Ruinen, 1956) represents a biome that is inhabited by a variety of bacteria, fungi, archaea and other microorganisms (Lindow and Leveau, 2002; Lindow and Brandl, 2003; Leveau, 2006). Culture-independent methods based on the analysis of ribosomal RNA (rRNA) genes have revealed that phyllosphere microbiota are complex and more diverse than the results from culture-dependent approaches would suggest (Yang *et al.*, 2001; Handschur *et al.*, 2005; Jackson *et al.*, 2006; Lambais *et al.*, 2006; Redford and Fierer, 2009; Yashiro *et al.*, 2011). In recent years, next-generation sequencing of rRNA amplicons has made its way into the toolbox of phyllosphere microbiologists, providing in-depth descriptions of the bacterial and fungal community composition associated with leaves of oak (Jumpponen and Jones, 2009; Jumpponen and Jones, 2010), soybean, clover and mouse-ear cress (Delmotte *et al.*, 2009), various tree species (Redford *et al.*, 2010), salt cedar (Finkel *et al.*, 2011), spinach (Lopez-Velasco *et al.*, 2011) and grape (Leveau and Tech, 2011).

These new technologies and resulting databases provoke great excitement over the prospect to address long-standing questions in the field of phyllosphere microbiology. For example, more complete descriptions of foliar microbial communities are starting to shed light on what constitutes true 'residents' of the phyllosphere as opposed to 'transient' colonizers (Hirano and Upper, 1991). The former are generally regarded as 'core' taxa

Correspondence: JHJ Leveau or GL Coaker, Department of Plant Pathology, University of California, One Shields Avenue, Hutchison Hall, Davis, CA 95616, USA.
E-mail: jleveau@ucdavis.edu (JHJL) or glcoaker@ucdavis.edu (GLC)
Received 19 December 2011; revised 23 February 2012; accepted 23 February 2012; published online 26 April 2012

(Unterseher *et al.*, 2010), which are consistently and abundantly found on the foliage of various plant species and are likely to possess adaptations specific for survival and reproduction under the harsh circumstances that typify the phyllosphere (Leveau, 2006). Another question that is beginning to be addressed relates to the mechanisms and consequences of variability in the composition of microbial communities on leaf surfaces. Such differences have been documented, and the roles of both the local environment and plant species have been discussed (Whipps *et al.*, 2008, Knief *et al.*, 2010). Good evidence exists in support of the hypothesis that plants select for specific microbiota on their leaves and that this correlates with plant genetic components (Balint-Kurti *et al.*, 2010; Hunter *et al.*, 2010). Our understanding of leaf microbiota as providers of specific services, for example, pathogen exclusion (Newton *et al.*, 2010) and nitrogen fixation (Fürnkranz *et al.*, 2008), relies on continued efforts to catalog the microbial communities on plant foliage.

Historically, phyllosphere research has been driven by the desire to understand the biology and ecology of foliar plant pathogens on crops (Leveau, 2006). Subject to intense management of typically monoculture crops across a wide range of seasons and production regions, today's agroecosystems are ideal testing grounds for interpreting differences in the microbial diversity associated with plant leaves in the context of spatiotemporal and environmental factors. One crop of interest to phyllosphere microbiologists is lettuce (*Lactuca sativa* L.), which is economically important worldwide with an estimated annual production value of > $2 billion in the United States (USDA, 2011). Given its positive health benefits, consumption of lettuce and other fresh leafy produce has increased in recent years (Li *et al.*, 2010). At the same time, there is concern about recurrent outbreaks of foodborne illnesses linked to the consumption of leafy greens contaminated with enteric pathogens such as *Escherichia coli* O157:H7 (Erickson *et al.*, 2010) and *Salmonella* (Klerks *et al.*, 2007). Another challenge facing the leafy greens industry is preharvest losses due to plant pathogens (Davis *et al.*, 1997), including *Xanthomonas campestris* pv. *vitians* (*Xcv*) causing bacterial leaf spot (BLS) (Barak *et al.*, 2001).

Several studies have already taken a culture-independent approach to describe the bacterial communities naturally associated with lettuce leaves, either pre- or postharvest (Handschur *et al.*, 2005; Zwielehner *et al.*, 2008; Hunter *et al.*, 2010). They revealed that community composition varied with season (Zwielehner *et al.*, 2008), lettuce cultivar (Hunter *et al.*, 2010) and stage in the farm-to-fork supply chain (Handschur *et al.*, 2005). With the present study, we aimed to expand on these insights by providing a comprehensive database of the bacterial community composition on preharvest lettuce leaves from fields at different geographical locations. More than 90% of the lettuce production in the United States occurs in California and Arizona (Boriss and Brunke, 2005). Most of the summer production (from April to October) occurs in California's Salinas Valley, while winter production (from November to March) takes place in the desert regions of Yuma, Arizona and California's Imperial Valley. For this study, we sampled Romaine lettuce at the time of harvest from over 50 commercial fields throughout an entire production cycle. The scope of this survey allowed us to interpret the observations of variability in bacterial community composition in the context of growing season, field location and environmental conditions.

Materials and methods

Sample collection, bacterial retrieval, colony counting, DNA isolation and quantitative PCR

Samples of field-grown Romaine lettuce were collected from production regions in California and Arizona (Figure 1; Supplementary Table 1). Per field, two samples were taken at the time of harvest ~100 m apart, that is, from the opposing corners of an ~0.5-ha section in the center of the field. Each sample consisted of four lettuce heads from which were picked two outermost and two inner leaves from the fourth leaf circle for a total of 16 leaves per sample. Bacteria were retrieved from leaves by washing as described earlier (Rastogi *et al.*, 2010). Aliquots of leaf washings were plated and incubated at 20 °C for 48 h on 0.1 × Tryptic Soy agar (TSA) or King's B (KB) agar to obtain estimates for the number of colony-forming units (CFUs) per gram of leaf tissue. Aliquots were also plated and incubated at 37 °C for 24 h on CHROMagar ECC to count the total number of coliforms. Bacteria in the remainder of each leaf washing were concentrated by centrifugation and the resulting pellet was used to extract total DNA using a PowerSoil DNA Isolation Kit (MO-BIO Laboratories Inc., Carlsbad, CA, USA). DNA was used in a quantitative PCR (qPCR) to estimate the total bacterial population sizes as described earlier (Rastogi *et al.*, 2010).

Barcoded pyrosequencing, DNA sequence processing and taxonomic analysis

DNA extracted from 88 selected lettuce leaf surface samples was used in a PCR to amplify bacterial 16S rRNA gene sequences encompassing the V5, V6 and V7 hypervariable regions as described earlier (Leveau and Tech, 2011). Pyrosequencing of these barcoded amplicons was performed on two halves of two standard PicoTiter plates using the GS-FLX 454 Titanium platform (Roche, Basel, Switzerland) at the Core for Applied Genomics and Ecology (CAGE) at the University of Nebraska (Lincoln, NE, USA). Reproducibility between the two runs was confirmed by including technical replicates (Supplementary

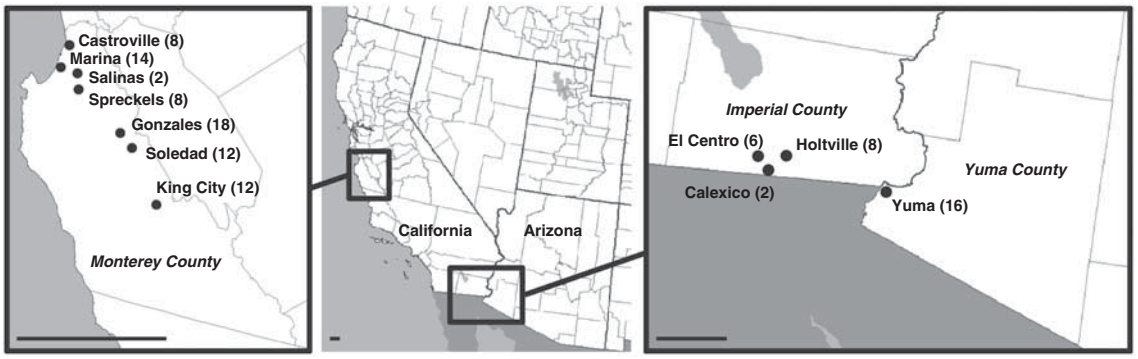

Figure 1 Maps showing the location of cities in California and Arizona near which summer (left) and winter (right) production fields of Romaine lettuce were sampled for this study. Maps were obtained from the website www.nationalatlas.gov. The bar in the left bottom corner of each panel represents about 50 km. Numbers in parentheses indicate the number of samples that were taken from near each city. From June to September 2009, 74 samples were collected on 11 sampling trips to summer production fields near Castroville, Marina, Salinas, Spreckles, Gonzales, Soledad and King City in Monterey County, CA. From December 2009 to January 2010, an additional 32 samples were collected on 4 sampling trips to winter production fields near El Centro, Holtville and Calexico in Imperial County, CA and Yuma in Yuma County, AZ. Meteorological data at date of sampling were retrieved from the nearest weather station installed by the California Irrigation Management Information System and the Arizona Meteorological Network.

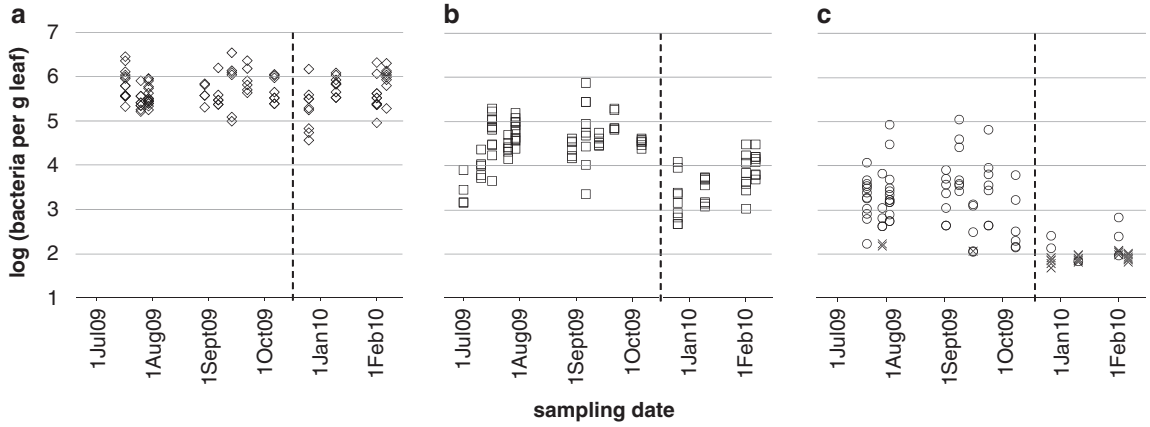

Figure 2 Bacterial population sizes on the foliage of field-grown lettuce. Shown are the sizes of total bacteria (**a**), culturable bacteria (**b**) and culturable coliforms (**c**) on lettuce leaves from fields that were sampled between June 2009 and February 2010. The values in panel a were derived from qPCR-based estimates of 16S rRNA gene copies and assuming one 16S rRNA gene copy per cell. The values shown in panel b represent CFUs on 0.1 × TSA plates. The corresponding values for KB plates were very similar (Supplementary Figure 3). The vertical dashed line separates the summer data from the winter data in each panel. The data points marked × in panel **c** represent the limit of detection for samples from which no colonies were recovered on CHROMagar ECC plates.

Figure 1). Sequences were screened through the CAGE custom length and quality filters, as explained before (Leveau and Tech, 2011), trimmed to 450 bp, and analyzed through the Ribosomal Database Project (RDP) pyrosequencing pipeline (http://pyro.cme.msu.edu). We obtained a total of 818 013 DNA sequences, with an average of 9296 ± 3210 per sample (Supplementary Table 1). From each sequence data set representing a single sample, 2836 sequences were randomly selected using PANGEA (Giongo *et al.*, 2010); this number corresponded to the lowest number of reads found in a single sample (Supplementary Table 1; Supplementary Figure 2). These randomly selected sequences from each sample were pooled into a total data set of 249 568 sequences and the taxonomic identity of individual sequences was assigned using the Bayesian rRNA classifier at an 80% confidence threshold (Cole *et al.*, 2009). Operational taxonomic units (OTUs) were determined at the genus (95%) level. Analysis by weighted Fast UniFrac (Hamady *et al.*, 2010) was used to compare the bacterial communities across different lettuce samples. Only OTUs represented by two or more sequences were included (Behnke *et al.*, 2011).

Results

Bacterial population sizes on field-grown lettuce

Between June 2009 and February 2010, we collected 106 samples of field-grown Romaine lettuce from production regions in California and Arizona at the time of harvest (Figure 1; Supplementary Table 1). Total bacterial population sizes were estimated by qPCR and expressed as log[bacteria per gram leaf]. At an average of 5.70 ± 0.39, these population sizes appeared relatively constant across samples

(Figure 2a), with no significant difference between samples from summer (5.72 ± 0.35) and winter (5.65 ± 0.42) production regions. In contrast, numbers of total culturable bacteria (expressed as log[CFUs per gram leaf]) were significantly lower (Figure 2b), on average 10-fold for the summer samples (4.50 ± 0.46) and 100-fold for the winter samples (3.65 ± 0.48). Population sizes of culturable coliform bacteria varied widely among summer samples around an average of 3.28 ± 0.64, whereas for most winter samples these populations fell below the average limit of detection of 1.92 ± 0.10 (Figure 2c).

Identification of a core bacterial community and estimation of leaf-associated bacterial diversity

From 88 selected lettuce samples, that is, 2 samples per field from 44 fields across summer and winter production regions, we analyzed a total of 249 568 bacterial 16S rRNA sequences. On the basis of RDP Classifier analysis, the majority of these sequences were affiliated with one of four phyla: *Proteobacteria* (74%), *Firmicutes* (13%), *Bacteroidetes* (7%) and *Actinobacteria* (3%) (Figure 3). Of the 478 bacterial genera identified, the 13 most abundantly represented were *Pseudomonas* (17%), *Bacillus* (7%), *Pantoea* (6%), *Massilia* (5%), *Xanthomonas* (4%), *Alkanindiges* (3%), *Erwinia, Duganella* and *Acinetobacter* (2% each), and *Flavobacterium*, *Naxibacter*, *Exiguobacterium* and *Arthrobacter* (1% each) (Supplementary Figure 4). For all other genera, the number of representative sequences was <1%. *Pseudomonas*, *Bacillus*, *Massilia* and *Arthrobacter* were the only genera to have at least one representative sequence in all 88 samples. We tentatively define these genera as making up the core bacterial microbiota of the lettuce phyllosphere, together with *Pantoea*, for which sequences were found in all but one sample. We also analyzed the summer and winter data separately, which added the genera *Skermanella*, *Rhizobium* and *Brevundimonas* to the summer core community, and *Exiguobacterium* and *Planomicrobium* to the winter core community.

Foliar detection of a bacterial pathogen correlates with the presence or absence of specific taxa

The genus most unevenly distributed between samples was *Xanthomonas,* which was identified only in 33 of the 88 samples. In 13 of these samples, the relative abundance of *Xanthomonas* sequences was high, that is, between 2% and 69% of all reads. Most of these *Xanthomonas* sequences revealed close resemblance to the 16S rRNA genes from *X. campestris,* and by using a specific primer set (Barak *et al*., 2001) we were able to confirm (data not shown) that these samples contained high loads of *Xcv*, the causative agent of BLS on lettuce. All high-*Xcv* samples came from summer production regions, specifically from seven fields near Gonzales,

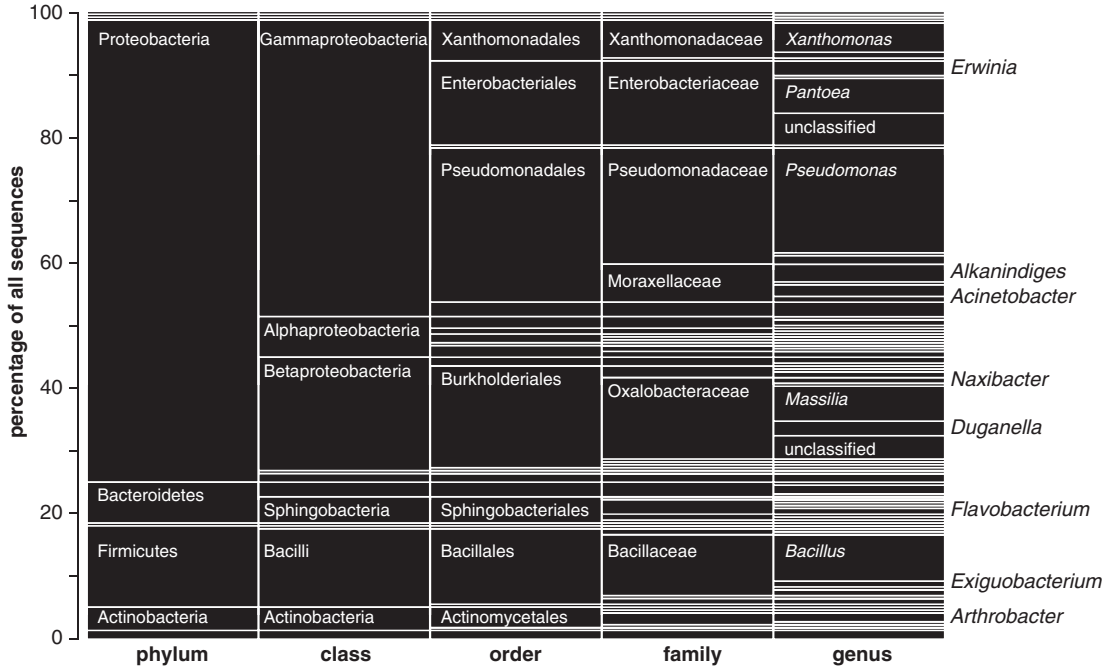

Figure 3 Bacterial community structure associated with Romaine lettuce leaves. Shown is the RDP Classifier analysis of sequences pooled from 88 leaf samples, with 2836 sequences representing each sample. The five columns represent different taxonomic ranks (Phylum, Class, Order, Family and Genus), and each box within a column represents a distinct taxon within that taxonomic rank. The height of each box corresponds to the percentage of all sequences that were assigned to the taxon, which is represented by that box, and the heights of all boxes in a single column add up to 100%. The boxes of abundantly represented taxa are labeled with the name of those taxa. Of all sequences, 97% could be assigned at the Class level, 93% at the Order level, 89% at the Family level and 71% at the Genus level; less than 2% were labeled as 'Unclassified Bacteria'. The largest group of sequences that could be assigned at the Family but not Genus level were 'Unclassified Enterobacteriaceae' (5%).

Soledad and King City. Of all other samples from fields near these cities, eleven showed no *Xanthomonas*. When we compared the relative representation of bacterial genera between high- and no-*Xcv* samples (Figure 4), there was a clear overrepresentation of *Alkanindiges* sequences on leaves with high-*Xcv* loads, whereas no-*Xcv* leaves were on average more abundant for *Bacillus*, *Erwinia* and *Pantoea*. For comparison, the relative abundance of *Massilia* was not different between high- and no-*Xcv* populations (Figure 4).

Comparison of bacterial community composition between samples from the same and different fields

To compare in more detail the bacterial community associated with lettuce leaves from different fields and sampling dates, we used a weighted Fast UniFrac analysis, which placed each of the 88 sequenced samples from 44 fields into one of nine discrete clusters (Figure 5). For 16 fields, the two samples from each field were more similar to one another than to any other sample. For 23 other fields, the two samples from each field were not most similar to each other, but fell within the same cluster. As most of these clusters showed interspersed samples from different geographical regions and different sampling dates, one needs to assume that factors other than field location and date of collection had a role in determining the community composition of the lettuce leaf microbiota. For the remaining five fields, the two samples belonged to different clusters, suggesting substantial in-field variation.

Bacterial communities from summer and winter production regions were different

As shown in Figure 5, samples that were collected in the summer (clusters 1–4) never shared a cluster with samples from the winter (clusters 5–9). This separation did not appear to be a function of the average air temperature, which was very similar between the summer and winter season ($15 \pm 2\,^{\circ}C$ and $12 \pm 1\,^{\circ}C$, respectively). However, summer and winter samplings differed at day of sampling in average solar radiation (522 ± 110 Ly per day in summer versus 281 ± 42 Ly per day in winter), average relative humidity ($75 \pm 10\%$ versus $62 \pm 13\%$) and average soil temperature ($23 \pm 3\,^{\circ}C$ versus $12 \pm 1\,^{\circ}C$). The most conspicuously overrepresented group of bacteria in winter samples were members from the family *Oxalobacteraceae*: on average, 22% of all sequences were classified as such, compared with only 8% in the summer samples. Another striking difference in community composition between samples from the summer and winter production regions involved sequences classified as *Enterobacteriaceae*: these were more ubiquitous in summer than in winter, that is, on average 19% versus 4%. The number of OTUs were different between summer and winter samples at averages of 149 ± 41 and 185 ± 74, respectively (Supplementary Figure 5). However, this difference could be attributed to the six samples in cluster 5. Without these samples, the average number of OTUs in winter was 155 ± 35, that is, not significantly different from summer.

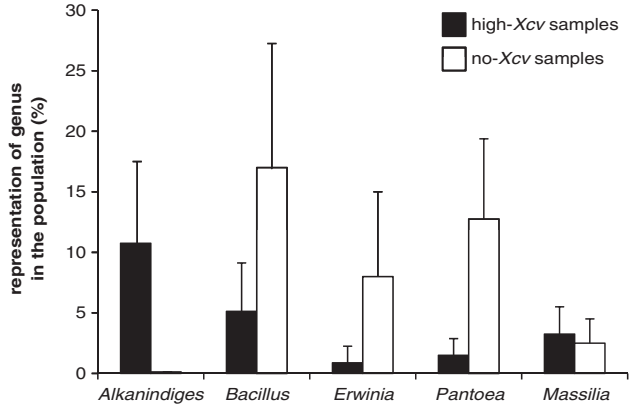

Figure 4 Comparison of the relative abundance of representatives in the genera *Alkanindiges*, *Bacillus*, *Erwinia*, *Pantoea* and *Massilia* in lettuce leaf samples that either harbored high levels of *Xanthomonas* (black bars; averaged over $n = 13$ samples) or no *Xanthomonas* (white bars; averaged over $n = 11$ samples). The error bars indicate standard deviations. Differences between the average abundance of *Alkanindiges*, *Bacillus*, *Erwinia*, *Pantoea*, but not *Massilia*, were significantly different ($P < 0.01$) between high-*Xanthomonas* and no-*Xanthomonas* samples.

Differences among fields from summer production regions

None of the summer clusters 1, 2, 3 and 4 (Figure 5) contained only samples that were collected on the same day or from the same field. Cluster 4 encompassed samples only from fields near Gonzales, but from three different sampling dates. Cluster 1 included samples from fields near Marina early in the season (mid-June to mid-July 2009) and from fields near Salinas, Soledad, Gonzales and King City late in the season (September 2009). All samples from the middle of the summer season (late-July to late-August 2009) fell into clusters 2, 3 and 4. In at least two instances, two fields sampled on the same day in the same general area fell into two separate clusters. For example, Marina samples M2a and M2b were collected on the same day as M3a and M3b, but belonged to cluster 3 instead of 1. Similarly, Gonzales samples G1a and G1b (cluster 2) were taken on the same day as G2a and G2b (cluster 4). These observations require an explanation that is field specific and independent of environmental conditions around the time of sampling, which are presumed to have been similar between M2a/b and M3a/b and between G1a/b and G2a/b. The difference between G1a/b and G2a/b is linked to the fact that, unlike G2a and G2b, samples G1a and G1b were collected from fields with high-*Xcv* loads.

The fields we sampled were planted with different cultivars of Romaine lettuce (Supplementary Table 1). To test for an effect of Romaine cultivar on

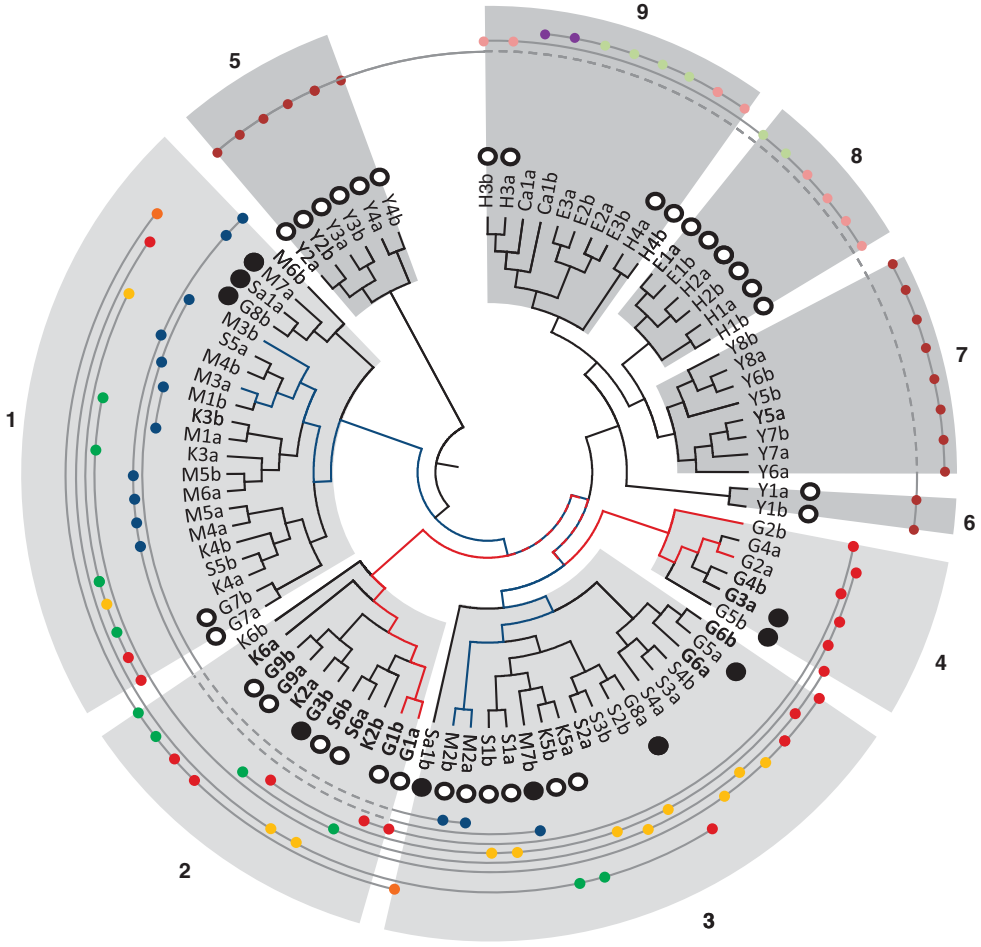

Figure 5 Clustering tree based on weighted Fast Unifrac analysis of the bacterial communities associated with 88 lettuce leaf samples from production fields near Marina (M), Gonzales (G), Soledad (S), King City (K), Salinas (Sa), Yuma (Y), El Centro (E), Holtville (H) and Calexico (Ca). Discrete clusters are highlighted in gray and numbered from 1 to 9. Any sample that was more similar to the corresponding sample from the same field than to any other sample is labeled with a white dot; any sample for which the corresponding sample belonged to a different cluster is labeled with a black dot. The sample names shown in bold represent the 13 samples with high loads of *Xanthomonas*. Colored markings on the outer ring segments indicate the collection dates for each sample; the color of the marking indicates geographical location (e.g. blue represents Marina). For clusters 1 to 4, there are 10 such ring segments, representing the following sampling dates in 2009 (from inside to outside): 6/15, 7/9, 7/15, 7/23, 7/27, 8/25, 9/1, 9/8, 9/16 and 9/30. For clusters 5 to 9, there are 4 ring segments, representing sampling dates (from inside to outside) 12/22 in 2009, and 1/5, 1/26, and 1/31 in 2010. The red line on the cluster tree connects samples G1a/b and G2a/b, whereas the blue line connects samples M2a/b and M3a/b; these are discussed in the text as examples of samples that were taken from different fields in the same general location on the same day, but belonging to different clusters.

bacterial community composition, we performed a principal component analysis on a subset of samples from the summer season for which cultivar information was available (Figure 6). This analysis revealed a tendency of samples from the same field to cluster together, but we did not see a correlation with Romaine cultivar.

Given the availability of data from summer samples that were taken from the same general growing region but on different days, or on the same day but from fields in different growing regions, we were able to address the question of how bacterial community composition on lettuce leaf surfaces changed over distances in time and space. It is clear from Figure 7a that differences in the bacterial community composition on leaves that were collected at different dates from fields near Marina did not correlate with the number of days between sampling events. We reached the same conclusion for other growing regions (data not shown). On the other hand, we saw a positive correlation between the Fast UniFrac distance and the distance between samples from different fields on the same sampling date (Figures 7b and c), suggesting that bacterial communities on lettuce leaves became more dissimilar over short geographical distances.

Differences among fields from winter production regions

For most samples from the winter production region, clustering was correlated with sampling

date. Samples taken on 5 January 2010 belonged for the most part to cluster 8, with the exception of two, which fell into cluster 9. Samples in cluster 7 were all from fields that were wet with recent rain at the time of sampling; their communities showed higher than average abundances of members from the families *Flavobacteriaceae* (phylum Bacteroidetes) and *Methylophilaceae* (class Betaproteobacteria). All samples in clusters 5 and 6 were taken on 22 December 2009 in Yuma County and were characterized by high relative abundances of members from the family *Geodermatophilaceae* (phylum Actinobacteria), that is, 0.5–8.5% of total sequences compared with 0–0.2% in all other samples.

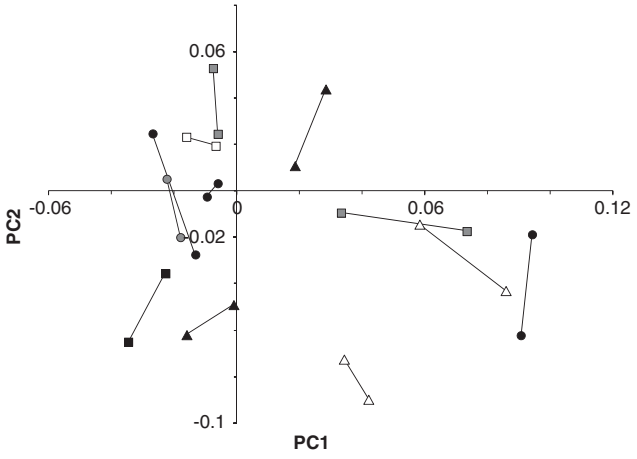

Figure 6 Principal component analysis of the bacterial community composition associated with 24 samples from 12 fields in 3 different regions (Gonzales: squares; Soledad: circles; King City: triangles), planted with 3 different Romaine cultivars (black, gray and white symbols). Plotted are the PC1 and PC2 values that were derived from the same data set that was used to create Figure 5. PC1 and PC2 explained 30.0% (eigenvalue 0.29) and 17.9% (eigenvalue 0.18) of the variation, respectively. Samples that came from the same field are connected with a line.

Discussion

This study is the first comprehensive pyrosequencing analysis of bacterial communities associated with the phyllosphere of a spatiotemporally varying agroecosystem. In terms of the number of samples, sampling locations and times, and the sequencing depth per sample, our culture-independent survey exceeded others dealing specifically with leaf surface microbiota (Delmotte *et al*., 2009; Redford *et al*., 2010; Lopez-Velasco *et al*., 2011; Leveau and Tech, 2011). Our findings not only confirmed previously observed phyllosphere associations of certain bacterial taxa, but also identified new associations, possibly unique to lettuce, as well as unanticipated correlations between the relative abundance of certain bacterial taxa. Given the scope of the survey, we can start appreciating the factors that are likely to shape the bacterial community composition on foliage of field-grown crops, including geographical location and environmental conditions. Of added value to this study was the combination with culture-dependent assessments of bacterial populations, which allowed us to interpret counts of culturable bacteria in the context of bacterial community composition, and *vice versa*.

We identified *Proteobacteria*, *Firmicutes*, *Bacteroidetes* and *Actinobacteria* as the most abundantly represented bacterial phyla on lettuce foliage. This is consistent with what has been reported for many other plants, as deduced from pyrosequencing, such as grape (Leveau and Tech, 2011), various tree species (Redford *et al*., 2010) as well as fresh spinach (Lopez-Velasco *et al*., 2011). At the genus level, we found *Pseudomonas*, *Bacillus*, *Massilia* and *Arthrobacter* in all samples, independent of sampling location and sampling date. *Pseudomonas* was identified previously as dominating the phyllosphere of field-grown lettuce and other plants, based on culture-dependent and -independent analyses

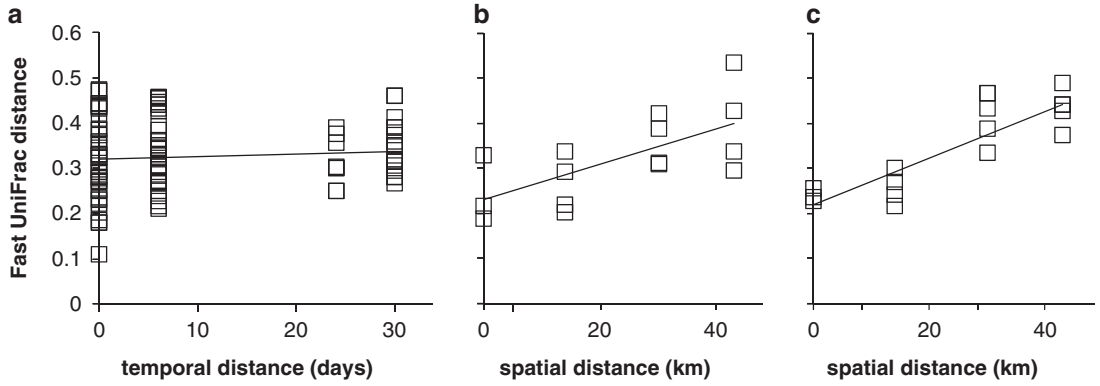

Figure 7 Differences in bacterial community composition on lettuce leaves as a function of time and space. In these graphs, the Fast Unifrac distances for pairs of samples from Marina collected on different dates (**a**) or on the same date from different fields near Gonzales, Soledad and King City (**b** and **c**) were plotted as a function of the time between sampling events (**a**) or of the distance between two sampling events on the same day (**b** and **c**). To create panel a, we used 14 samples collected near Marina from three different dates to plot 91 pairwise comparisons ($F = 0.18$, d.f. = 1,89, ANOVA P-value = 0.67). To create panels b ($F = 7.11$, d.f. = 1,13, ANOVA P-value = 0.02) and c ($F = 57.96$, df = 1,13, ANOVA P-value < 0.001), we used samples collected on 25 August 2009 and 1 September 2009, respectively. The R^2 values for the trendlines in panels **a**, **b** and **c** were 0.0068, 0.4579 and 0.7617, respectively.

(Handschur *et al.*, 2005; Lambais *et al.*, 2006; Zwielehner *et al.*, 2008; Hunter *et al.*, 2010; Lopez-Velasco *et al.*, 2011). *Bacillus* and *Arthrobacter* are also commonly recognized and isolated as leaf surface colonizers (Salerno *et al.*, 1997; Krimm *et al.*, 2005; Enya *et al.*, 2007). Members of the genus *Massilia* have been detected in the phyllosphere earlier, but only through culture-independent approaches (for example, Delmotte *et al.*, 2009; Yashiro *et al.*, 2011). On spinach leaves, *Massilia* represented as much as 7.4% of the total bacterial population (Lopez-Velasco *et al.*, 2011). Interestingly, *Massilia* was identified earlier as a major constituent of agricultural aerosols in central California (Rawa *et al.*, 2011), which suggests a bacterial source/sink relationship between lettuce foliage and the surrounding air. Culturable representatives of *Massilia* have, to our knowledge, never been isolated from leaf surfaces. This finding underscores the important contribution of culture-independent approaches towards a more complete understanding of phyllosphere microbiology. Unlike what has been reported for leaf surfaces of apple (Yashiro *et al.*, 2011), tropical trees (Lambais *et al.*, 2006), grape (Leveau and Tech, 2011), tomato (Enya *et al.*, 2007), soybean, clover and mouse-ear cress (Delmotte *et al.*, 2009), we did not find a dominant representation of *Sphingomonas* in our samples. This observation matches other surveys of lettuce surface microbiota (Zwielehner *et al.*, 2008; Hunter *et al.*, 2010).

The family *Enterobacteriaceae* was prominently represented in the lettuce phyllosphere by *Erwinia* and *Pantoea*, as observed previously (Hunter *et al.*, 2010; Lopez-Velasco *et al.*, 2011). These two genera encompass most of the recognized plant-associated coliforms (Leclerc *et al.*, 2001), which readily explained our detection of coliforms in lettuce leaf washes (Figure 2c). In the summer samples, we found culturable coliforms at levels that were comparable to those reported elsewhere for field-grown lettuce (Mukherjee *et al.*, 2004; Phillips and Harrison, 2005). In the winter samples, coliform levels were much lower and this difference is of potential interest, because it follows the seasonality of historical detections of *Escherichia coli* O157:H7 in the Salinas Valley. However, no correlation has so far been shown between the abundance of coliforms and contamination with *E. coli* O157:H7. On the basis of the data presented here, we contend that coliforms are a natural part of the lettuce microbiota and that the results from tests that use total coliform counts as a proxy for fecal contamination of lettuce should be interpreted accordingly.

We observed a clear separation between summer and winter production regions in terms of the bacterial community composition that characterized the lettuce that was grown in these two regions. To what degree and by what mechanisms seasonal differences such as relative humidity or temperature, or irrigation practices (Supplementary Table 1) were the cause of the observed variation in bacterial community composition remains an open question. We cannot exclude the possibility that location *per se*, rather than local weather conditions, contributed to the observed variation. For example, bacterial communities on lettuce leaves may have been different between summer and winter because they were drawn from metacommunities that differed between northern and southern production regions. This summer/winter (or north/south) incongruity certainly is an important finding that invites an experimental approach to tease out the factors that contribute to the observed differences.

A salient finding was the correlation between the abundance of *Xcv*, the bacterial pathogen causing BLS of lettuce, and the presence or absence of the genera *Alkanindiges*, *Bacillus*, *Erwinia* and *Pantoea*. As for the causes and effects that underlie these positive and negative correlations, three hypotheses can be formulated. First, naturally occurring *Bacillus*, *Erwinia* and *Pantoea* strains in the lettuce phyllosphere are antagonists of *Xcv* establishment, whereas *Alkanindiges* acts as a facilitator. Second, establishment of *Xcv* on lettuce in the field differentially impacts specific groups of bacteria in the phyllosphere bacterial community, particularly *Bacillus*, *Erwinia*, *Pantoea* and *Alkanindiges* species. Third, the changes in abundance of *Xcv*, *Bacillus*, *Erwinia*, *Pantoea* and *Alkanindiges* are not a function of one another, but occur in response to other factors, for example plant genotype. These hypotheses remain to be tested, but there are some observations in the literature that are worth mentioning in this context. A possible antagonistic activity of *Bacillus* against *Xcv* agrees with the fact that a reduction in BLS severity can be achieved by use of the biological pesticide Serenade (Bull and Koike, 2005), which contains *Bacillus* as an active ingredient. Little is known about the genus *Alkanindiges*, although it has been reported to occur on lettuce leaf surfaces (Hunter *et al.*, 2010). One of the only two species described so far, *Alkanindiges illinoisensis*, is an obligate degrader of alkane (Bogan *et al.*, 2003), which is a major component of the waxy cuticle that covers plant leaf surfaces (Müller and Riederer, 2005). Possibly, a compromise in the integrity of the leaf cuticle underlies a facilitative role of *Alkanindiges* in *Xcv* establishment on lettuce leaves. As for the alternative hypothesis that *Xcv* infection was the cause of changes in the bacterial community composition on lettuce, it should be noted that none of the leaves that we analyzed exhibited symptoms typical of BLS, ruling out macroscopic-level lesion formation as an explanation for possible *Xcv*-induced changes that we observed in the bacterial community composition. Finally, recent evidence suggests that different accessions of lettuce harbor different foliar microbiota (Hunter *et al.*, 2010). However, little variation in the bacterial community composition was observed among different cultivars of Romaine (also called Cos), compared with the variation seen

with other lettuce types, such as butterhead, iceberg and curly leaved. This agrees with our finding that there was no cultivar-dependent variation between different Romaine varieties (Figure 6).

We were intrigued by the odd placement of the six samples in cluster 5 in relation to the two samples taken on the same day (but placed in cluster 6) in Figure 5. The overrepresentation of *Geodermatophilaceae* (phylum Actinobacteria) in all eight samples provided an important clue to what might be underlying this placement. Bacteria belonging to this family were first discovered in desert soils (Luedemann, 1968) and have also been detected in airborne dust (Polymenakou *et al.*, 2008). Interestingly, on the day of sampling, a severe dust storm struck Yuma County, which was predicted by the National Weather Service (2009) and confirmed by observations on the ground at the time of sampling. We suspect that the bacterial communities on the leaf samples from this date were a mixture of 'true' lettuce leaf residents and recently immigrated dust-associated bacteria. The peculiar position of cluster 5 in the tree (Figure 5) is explained by the fact that samples Y2a/b, Y3a/b and Y4a/b, which were taken later in the day than Y1a/b, received more dust. Thus, samples Y1a and Y1b still grouped most closely together and with other lettuce samples from the winter season, whereas the other six samples collectively out-grouped every other sample. This observation offers good evidence not only for the temporally dynamic nature of the bacterial community composition on plant foliage, but also for the role of dust and other bioaerosols as a source of many phyllosphere bacteria (Suslow *et al.*, 2003).

The 10- to 100-fold discrepancy that we observed between qPCR- and CFU-based bacterial population sizes on lettuce leaves (Figures 2a and b) has been noted earlier (Rastogi *et al.*, 2010). It may be attributed in part to the fact that our qPCR-based numbers are overestimates because we assumed a single-copy occurrence of the 16S rRNA gene in bacterial cells. On the basis of the pyrosequencing results presented here, we can abandon this assumption and instead, for each taxon identified by Classifier, calculate an adjusted relative abundance for that taxon based on known 16S rRNA gene copy numbers (Lee *et al.*, 2009). This analysis (data not shown) revealed that, on average, we overestimated the total bacterial load on lettuce by a factor 4. This still does not entirely explain the difference between qPCR- and CFU-based estimates, especially for the winter samples, in which even after taking into account multiple copies of 16S rRNA genes per cell, only 4% of the cells made a colony on plate. Such low culturability does not agree with the relatively high representation of genera like *Pseudomonas* in these samples, which we would expect to grow on the plates that we used to count CFUs. Therefore, we suspect that a fraction of the bacteria that were washed off the field-grown lettuce from these regions were either non-viable or viable-but-non-culturable but got included in the qPCR-based total count because their DNA was extracted along with that of other, viable bacteria. *Pseudomonas* cells have indeed been shown to enter a non-culturable state after long periods of incubation on leaves (Wilson and Lindow, 1992).

Bacteria are not the only microorganisms that colonize lettuce leaves, and so this study offers only a partial picture of the lettuce microbiota. Leaves of plants and trees in general are known to harbor rich communities of fungi (Jumponnen and Jones, 2009; Jumpponen and Jones, 2010; Unterseher *et al.*, 2010; Finkel *et al.*, 2011). Within the small group of sequences in our data set, which were classified as not being of bacterial origin, we found several instances of matches with the 18S rRNA gene from the foliar fungal pathogen *Bremia lactucae*, causing downy mildew of lettuce, in addition to several other oomycetes. This preliminary result not only confirms that the protocol we used to isolate DNA from lettuce leaves permits the recovery of fungal DNA; but also opens the window on a yet unexplored part of the lettuce leaf microbiota.

Acknowledgements

This study was supported by grants from the California Leafy Greens Research Program and Center for Produce Safety (contract number LGR-2009–2010) and USDA-NIFA grant 2011-67017-30024 from the United States Department of Agriculture. We sincerely thank the lettuce growers, who kindly allowed us to sample their fields and Carolee Bull for advice on the *Xcv* primers. We would also like to thank members of the CAGE team for their suggestions on matters related to pyrosequencing and RDP staff for providing valuable assistance with the pyrosequencing pipeline.

References

Balint-Kurti P, Simmons SJ, Blum JE, Ballare CL, Stapleton AE. (2010). Maize leaf epiphytic bacteria diversity patterns are genetically correlated with resistance to fungal pathogen infection. *Mol Plant-Microb Interact* **23**: 473–484.

Barak JD, Koike ST, Gilbertson L. (2001). Role of crop debris and weeds in the epidemiology of bacterial leaf spot of lettuce in California**85**: 169–178.

Behnke A, Engel M, Christen R, Nebel M, Klein RR, Stoeck T. (2011). Depicting more accurate pictures of protistan community complexity using pyrosequencing of hypervariable SSU rRNA gene regions. *Environ Microbiol* **13**: 340–349.

Bogan BW, Sullivan WR, Kayser KJ, Derr KD, Aldrich HC, Paterek JR. (2003). *Alkanindiges illinoisensis* gen. nov., sp. nov., an obligately hydrocarbonoclastic, aerobic squalane-degrading bacterium isolated from oilfield soils. *Int J Syst Evol Microbiol* **53**: 1389–1395.

Boriss H, Brunke H. (2005). *Commodity Profile: Lettuce.* Agricultural Issues Center University of California.

Davis. Available at http://aic.ucdavis.edu/profiles/lettuce-2005.pdf.

Bull CT, Koike ST. (2005). Evaluating the efficacy of commercial products for management of bacterial leaf spot on lettuce. *Plant Health Progress*; doi:10.1094/PHP-2005-1121-01-RS. Available at http://www.plantmanagementnetwork.org/pub/php/research/2005/lettuce/.

Cole JR, Wang Q, Cardenas E, Fish J, Chai B, Farris RJ *et al.* (2009). The Ribosomal Database Project: improved alignments and new tools for rRNA analysis. *Nucleic Acids Res* **37**: D141–D145.

Davis RM, Subbarao KV, Raid RN, Kurtz EA. (1997). *Compendium of lettuce diseases*. APS Press: St Paul, MN.

Delmotte N, Knief C, Chaffron S, Innerebner G, Roschitzki B, Schlapbach R *et al.* (2009). Community proteogenomics reveals insights into the physiology of phyllosphere bacteria. *Proc Natl Acad Sci USA* **22**: 16428–16433.

Enya J, Shinohara H, Yoshida S, Tsukiboshi T, Negishi H, Suyama K *et al.* (2007). Culturable leaf-associated bacteria on tomato plants and their potential as biological control agents. *Microb Ecol* **53**: 524–536.

Erickson MC, Webb CC, Diaz-Perez JC, Phatak SC, Silvoy JJ, Davey L *et al.* (2010). Surface and internalized *Escherichia coli O157:H7* on field-grown spinach and lettuce treated with spray-contaminated irrigation water. *J Food Prot* **73**: 1023–1029.

Finkel OM, Burch AY, Lindow SE, Post AF, Belkin S. (2011). Geographical location determines the population structure in phyllosphere microbial communities of a salt-excreting desert tree. *Appl Environ Microbiol* **77**: 7647–7655.

Fürnkranz M, Wanek W, Richter A, Abell G, Rasche F, Sessitsch A. (2008). Nitrogen fixation by phyllosphere bacteria associated with higher plants and their colonizing epiphytes of a tropical lowland rainforest of Costa Rica. *ISME J* **2**: 561–570.

Giongo A, Crabb DB, Davis-Richardson AG, Chauliac D, Mobberley JM, Gano KA *et al.* (2010). PANGEA: pipeline for analysis of next generation amplicons. *ISME J* **4**: 852–861.

Hamady M, Lozupone C, Knight R. (2010). Fast UniFrac: facilitating high-throughput phylogenetic analyses of microbial communities including analysis of pyrosequencing and PhyloChip data. *ISME J* **4**: 17–27.

Handschur M, Pinar G, Gallist B, Lubitz W, Haslberger AG. (2005). Culture free DGGE and cloning based monitoring of changes in bacterial communities of salad due to processing. *Food Chem Toxicol* **43**: 1595–1605.

Hirano SS, Upper CD. (1991). Bacterial community dynamics. In: Andrews JH, Hirano SS (eds) *Microbial Ecology on Leaves*. Springer-Verlag: New York, pp 271–294.

Hunter PJ, Hand P, Pink D, Whipps JM, Bending GD. (2010). Both leaf properties and microbe-microbe interactions influence within-species variation in bacterial population diversity and structure in the lettuce (*Lactuca* species) phyllosphere. *Appl Environ Microbiol* **76**: 8117–8125.

Jackson EF, Echlin HL, Jackson CR. (2006). Changes in the phyllosphere community of the resurrection fern, *Polypodium polypodoides*, associated with rainfall and wetting. *FEMS Microbiol Ecol* **58**: 236–246.

Jumpponen A, Jones KL. (2009). Massively parallel 454 sequencing indicates hyperdiverse fungal communities in temperate *Quercus macrocarpa* phyllosphere. *New Phytologist* **184**: 438–448.

Jumpponen A, Jones KL. (2010). Seasonally dynamic fungal communities in the *Quercus macrocarpa* phyllosphere differ between urban and nonurban environments. *New Phytol* **186**: 496–513.

Klerks MM, Franz E, van Gent-Pelzer M, Zijlstra C, van Bruggen AHC. (2007). Differential interaction of *Salmonella enterica* serovars with lettuce cultivars and plant-microbe factors influencing the colonization efficiency. *ISME J* **1**: 620–631.

Knief C, Ramette A, Frances L, Alonso-Blanco C, Vorholt JA. (2010). Site and plant species are important determinants of the *Methylobacterium* community composition in the plant phyllosphere. *ISME J* **4**: 719–728.

Krimm U, Abanda-Nkpwatt D, Schwab W, Schreiber L. (2005). Epiphytic microorganisms on strawberry plants (*Fragaria ananassa* cv. *Elsanta*): identification of bacterial isolates and analysis of their interaction with leaf surfaces. *FEMS Microbiol Ecol* **53**: 483–492.

Lambais MR, Crowley DE, Cury JC, Bull RC, Rodrigues RR. (2006). Bacterial diversity in the tree canopies of the Atlantic forest. *Science* **312**: 1917.

Leclerc H, Mossel DA, Edberg SC, Struijk CB. (2001). Advances in the bacteriology of the coliform group: their suitability as markers of microbial water safety. *Annu Rev Microbiol* **55**: 201–234.

Lee ZM, Bussema C III, Schmidt TM. (2009). rrnDB: Documenting the number of rRNA and tRNA genes in Bacteria and Archaea. *Nucleic Acids Res* **37**: D489–D493.

Leveau JHJ. (2006). Microbial communities in the phyllosphere. In: Riederer M, Mueller C (eds). *Biology of the Plant Cuticle*. Blackwell Publishing: Oxford, UK, pp 334–367.

Leveau JHJ, Tech JJ. (2011). Grapevine microbiomics: bacterial diversity on grape leaves and berries revealed by high-throughput sequence analysis of 16S rRNA amplicons. *Acta Horticulturae* **905**: 31–42.

Li Z, Zhao X, Sandhu AK, Gu L. (2010). Effects of exogenous abscisic acid on yield, antioxidant capacities, and phytochemical contents of greenhouse grown lettuces. *J Agric Food Chem* **26**: 6503–6509.

Lindow SE, Brandl MT. (2003). Microbiology of the phyllosphere. *Appl Environ Microbiol* **69**: 1875–1883.

Lindow SE, Leveau JH. (2002). Phyllosphere microbiology. *Curr Opin Biotechnol* **13**: 238–243.

Lopez-Velasco G, Welbaum GE, Boyer RR, Mane SP, Ponder MA. (2011). Changes in spinach phylloepiphytic bacteria communities following minimal processing and refrigerated storage described using pyrosequencing of 16S rRNA amplicons. *J Appl Microbiol* **110**: 1203–1214.

Luedemann GM. (1968). *Geodermatophilus*, a new genus of the Dermatophilaceae (Actinomycetales). *J Bacteriol* **96**: 1848–1858.

Müller C, Riederer M. (2005). Plant surface properties in chemical ecology. *Journal of Chemical Ecology* **31**: 2621–2651.

Mukherjee A, Speh D, Dyck E, Diez-Gonzalez F. (2004). Preharvest evaluation of coliforms, *Escherichia coli*, *Salmonella*, and *Escherichia coli O157:H7* in organic and conventional produce grown by Minnesota farmers. *J Food Protection* **67**: 894–900.

National Weather Service. (2009). Available at http://pinalcountyaz.gov/Departments/AirQuality/Documents/Dust/122209_dust_storm_warning.pdf.

Newton AC, Gravouil C, Fountaine JM. (2010). Managing the ecology of foliar pathogens: ecological tolerance in crops. *Ann Appl Biol* **157**: 343–359.

Phillips CA, Harrison MA. (2005). Comparison of the microflora on organically and conventionally grown spring mix from a California processor. *J Food Prot* **68**: 1143–1146.

Polymenakou PN, Mandalakis M, Stephanou EG, Tselepides A. (2008). Particle size distribution of airborne microorganisms and pathogens during an intense african dust event in the Eastern Mediterranean. *Environ Health Perspect* **116**: 292–296.

Rastogi G, Tech JJ, Coaker GL, Leveau JH. (2010). A PCR-based toolbox for the culture-independent quantification of total bacterial abundances in plant environments. *J Microbiol Methods* **83**: 127–132.

Rawa SV, Sarreal CZ, Mandrell RE. (2011). Bacterial communities in aerosols and manure samples from two different dairies in Central and Sonoma valleys of California. *PLoS One* **6**: e17281.

Redford AJ, Bowers RM, Knight R, Linhart Y, Fierer N. (2010). The ecology of the phyllosphere: geographic and phylogenetic variability in the distribution of bacteria on tree leaves. *Environ Microbiol* **12**: 2885–2893.

Redford AJ, Fierer N. (2009). Bacterial succession on the leaf surface: a novel system for studying successional dynamics. *Microb Ecol* **58**: 189–198.

Ruinen J. (1956). Occurrence of *Beijerinckia* species in the 'phyllosphere'. *Nature* **177**: 220–221.

Salerno CM, Montero MC, Sagardoy MA. (1997). Bacterial population in different growth stages of soybean leaves (Glycine max L. Merrill). *Revista Argentina de Microbiologia* **29**: 122–130.

Suslow TV, Oria MP, Beuchat LR, Garrett EH, Parish ME, Harris LJ *et al*. (2003). Production practices as risk factors in microbial food safety of fresh and fresh-cut produce. *Comp Rev Food Sci Food Safety* **2S**: 38–77.

Unterseher M, Jumpponen A, Öpik M, Tedersoo L, Moora M, Dormann CF *et al*. (2010). Species abundance distributions and richness estimations in fungal metagenomics – lessons learned from community ecology. *Mol Ecol* **20**: 275–285.

USDA National Agricultural Statistics Service. (2011) Crop Values 2009 Summary. Available at http://usda.mannlib.cornell.edu/MannUsda/viewDocumentInfo.do?documentID=1050.

Whipps JM, Hand P, Pink D, Bending GD. (2008). Phyllosphere microbiology with special reference to diversity and plant genotype. *J Appl Microbiol* **105**: 1744–1755.

Wilson M, Lindow SE. (1992). Relationship of total viable and culturable cells in epiphytic populations of *Pseudomonas syringae*. *Appl Environ Microbiol* **58**: 3908–3913.

Yang CH, Crowley DE, Borneman J, Keen NT. (2001). Microbial phyllosphere populations are more complex than previously realized. *Proc Natl Acad Sci USA* **98**: 3889–3894.

Yashiro E, Spear RN, McManus PS. (2011). Culture-dependent and culture-independent assessment of bacteria in the apple phyllosphere. *J Appl Microbiol* **110**: 1284–1296.

Zwielehner J, Handschur M, Michaelsen A, Irez S, Demel M, Denner EB *et al*. (2008). DGGE and real-time PCR analysis of lactic acid bacteria in bacterial communities of the phyllosphere of lettuce. *Mol Nutr Food Res* **52**: 614–623.

Supplementary Information accompanies the paper on The ISME Journal website (http://www.nature.com/ismej)