

# Biogeographic Patterns Between Bacterial Phyllosphere Communities of the Southern Magnolia (*Magnolia grandiflora*) in a Small Forest

Bram W. G. Stone<sup>1</sup> · Colin R. Jackson<sup>1</sup>

Received: 11 September 2015 / Accepted: 8 February 2016 / Published online: 16 February 2016  
© Springer Science+Business Media New York 2016

**Abstract** The phyllosphere presents a unique system of discrete and easily replicable surfaces colonized primarily by bacteria. However, the biogeography of bacteria in the phyllosphere is little understood, especially at small to intermediate scales. Bacterial communities on the leaves of 91 southern magnolia (*Magnolia grandiflora*) trees 1–452 m apart in a small forest plot were analyzed and fragments of the 16S ribosomal RNA (rRNA) gene sequenced using the Illumina platform. Assemblages were dominated by members of the Alphaproteobacteria, Bacteroidetes, and Acidobacteria. Patterns in community composition were measured by both relative abundance (theta) and presence-absence (Jaccard) dissimilarity metrics. Distance-based Moran's eigenvector map analyses of the distance-decay relationship found a significant, positive relationship between each dissimilarity metric and significant eigenfunctions derived from geographic distance between trees, indicating trees that were closer together had more similar bacterial phyllosphere communities. Indirect gradient analyses revealed that several environmental parameters (canopy cover, tree elevation, and the slope and aspect of the ground beneath trees) were significantly related to multivariate ordination scores based on relative bacterial sequence abundances; however, these relationships were not significant when looking at the incidence of bacterial taxa. This suggests

that bacterial growth and abundance in the phyllosphere is shaped by different assembly mechanisms than bacterial presence or absence. More broadly, this study demonstrates that the distance-decay relationship applies to phyllosphere communities at local scales, and that environmental parameters as well as neutral forces may both influence spatial patterns in the phyllosphere.

**Keywords** Phyllosphere · Distance-decay relationship · Biogeography · Bacterial communities · 16S rRNA

## Introduction

The phyllosphere is the aerial portion of plants on which potentially diverse communities composed of Bacteria, Archaea, and microbial eukaryotes exist [1]. The surface area of the phyllosphere is vast, estimated as being several orders of magnitude greater than the land area of the planet [2]. Despite this large area, few studies on the biogeographic distribution of phyllosphere communities have been reported. Rather, the focus has more commonly been on the distribution and population dynamics of single microbial species, typically from a pathological perspective. Most of these spatial studies of the phyllosphere have utilized culture-dependent techniques and focused on the distribution of agricultural pathogens, thus tending to analyze plants that occur in a regular and repeated pattern (i.e., rows of crops) [3]. As such, autocorrelation and autoregressive-moving-average (ARMA) models, which are well suited for quantifying disease, are common in the literature [3]. However, ARMA models operate best when applied unidirectionally, while distance methods are probably more appropriate to compare plants that are not regularly distributed. The tendency for communities to become dissimilar at

---

**Electronic supplementary material** The online version of this article (doi:10.1007/s00248-016-0738-4) contains supplementary material, which is available to authorized users.

---

✉ Bram W. G. Stone  
bstone2@go.olemiss.edu

<sup>1</sup> Department of Biology, University of Mississippi, Shoemaker Hall, University MS, P.O. Box 1848, Oxford, MS 38677, USA

increasing distances is recognized in broader studies of biogeography, and is known as the distance-decay relationship [4].

While there has been increased use of next-generation sequencing (NGS) methods to describe phyllosphere bacterial communities, especially comparing community composition between species [5, 6], few studies have addressed the issue of biogeographic distribution, despite a call for studies at different spatial scales [7]. The application of the distance-decay relationship to microbial ecology in general has yielded conflicting results [8], and a similar disagreement has emerged from the limited number of studies of phyllosphere assemblages. Compositional differences in the bacterial phyllosphere between different plant species have generally been found to be much greater than intraspecific differences between individuals of the same plant species, even with different geographic origins of the plant host [9]. Conversely, however, geographic location has been found to be a greater predictor of phyllosphere community composition than plant host species and environmental factors [10, 11].

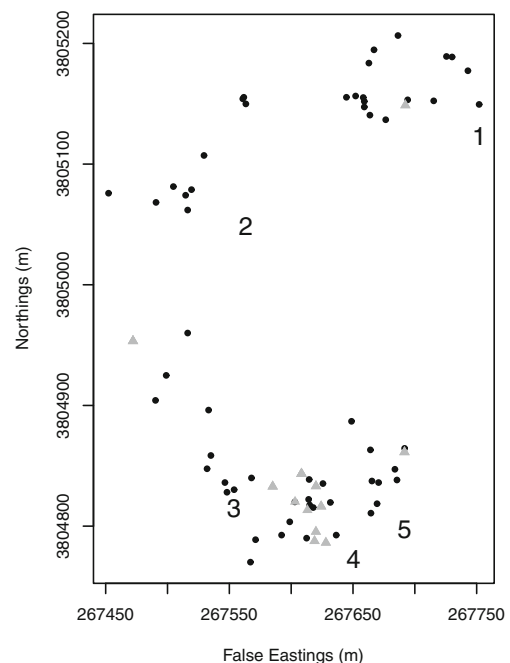
The distance-decay relationship is unique in that it may occur by neutral processes if the pattern is caused by dispersal limitation or by niche-based processes if driven by environmental gradients. Previous culture-independent studies of phyllosphere biogeography have utilized spatial scales in the hundreds of kilometers or more, so that little is known of patterns at smaller scales. The goal of this study was to supplement current understanding of spatial patterns of the phyllosphere by investigating whether phenomena such as the distance-decay relationship might apply within a single wooded ecosystem. This study examined the effect of distance on the composition of the bacterial community in the phyllosphere of a single tree species at a small spatial scale (<1 km) using NGS methods. We sampled the phyllosphere community of almost 100 *Magnolia grandiflora* (southern magnolia) trees at distances of <1 m to 450 m apart. By sampling broadleaf evergreens in winter, variation in community composition was tested in a natural setting with limited interference from other tree species. Distances (<1 km) were hypothesized to affect the bacterial community composition of *M. grandiflora* leaves such that dissimilarity would be higher between communities on trees that were further apart.

## Methods

Leaves were initially sampled from 100 *M. grandiflora* trees in Bailey Woods, a 20-ha tract of mature woods adjacent to the University of Mississippi campus in Oxford, Mississippi, USA, in February 2014. The woods are a remnant of old growth forests in northern Mississippi and are mainly deciduous with *M. grandiflora* accounting for approximately 1 % of the total number of trees [12]. The 100 *M. grandiflora* trees sampled account for the majority of *M. grandiflora* trees that

grow in this system. GPS coordinates were determined for each tree at the time of sampling. To aid in multivariate plotting, trees were mapped a posteriori to five general clusters based on location within the woods (Fig. 1). Unrelated to these five spatial clusters, some smaller trees occurred within close proximity to larger central trees and may have represented offspring from fallen seeds. These were recorded as separate trees in the same location, and their relationship was noted and used in subsequent analyses.

Sampling consisted of collecting two leaves (200 leaves total) from each tree, from different branches at approximately 1.5–1.8 m height. Only leaves that displayed no signs of disease or decay (such as browning or spotting) were selected. Leaves were placed in individual sterile sample bags and stored at  $-20^{\circ}\text{C}$  from as soon as possible after sampling (<2 h) until processed. The circumference of each tree was recorded with measuring tape and then converted to diameter at breast height (DBH). Additionally, at each tree, percent canopy cover was measured using a spherical densiometer. Using the raster package in R [13], GPS coordinates were combined with a digital elevation model to obtain the relative elevation, slope, and slope aspect of each tree.



**Fig. 1** Location of 91 *Magnolia grandiflora* trees within a 20-ha woodland in Oxford, MS, USA, that were sampled for phyllosphere communities. *Black dots* represent single trees while *gray triangles* represent locations of one parent tree surrounded by one or more smaller offspring. *Black diamonds* represent a group of trees (14–18) that showed distinct phyllosphere communities from other trees. *Numbers* indicate hypothetical spatial clusters of trees assigned by geographic or topographic isolation and are placed to the lower right of the indicated area. Location coordinates are expressed using the Universal Transverse Mercator (UTM) coordinate system (zone 6) and with the NAD83 Datum

The phyllosphere community was recovered from each leaf by brushing the leaf for 2 min with a sterile toothbrush in sterile TE (pH 8.0) buffer. The resulting suspension was centrifuged (10,000×*g*, 2 min) and the pellet frozen (−20 °C) for subsequent DNA extraction. DNA was extracted using a PowerSoil extraction kit (MoBio, Carlsbad, CA) following standard procedures; other than that, the final volume of DNA eluted was reduced from 100 to 50 μL. DNA samples were analyzed using a dual index barcoding approach targeting the V4 region of the 16S ribosomal RNA (rRNA) gene [14]. Briefly, DNA (standardized by volume, 1 μL) was combined with 1 μL of each primer (at 10 μM) and 17 μL of AccuPrime Pfx Supermix (Invitrogen, Grand Island, NY). Amplification was conducted through 30 cycles of 95 °C (20 s), 55 °C (15 s), and 72 °C (2 min) after an initial denaturation step of 95 °C (2 min), and followed by a final elongation step of 72 °C (10 min). Amplification products were standardized with SequalPrep Normalization Plates (Life Technologies, Grand Island, NY) and pooled prior to sequencing. The multiplexed library was sequenced using the Illumina MiSeq platform at the Molecular and Genomics Core Facility at the University of Mississippi Medical Center (UMMC).

Raw data files (FASTQ) were processed using the Mothur bioinformatics pipeline [15] following the procedures recommended by Schloss et al. [16] and Kozich et al. [14]. Rare sequences that were similar to abundant sequences (two base differences), and likely the result of sequencing errors, were merged so that remaining singletons would more legitimately reflect bacterial diversity (see Online Resource 1 for an examination of singleton removal). After removal of other potentially erroneous data, chimeras, and mitochondrial and chloroplast sequences, nine trees (unrelated spatially or by lab procedures) had leaves with <10,000 valid bacterial sequences and were removed from the dataset, leaving 91 trees in the final analysis, or 182 leaves. All diversity analyses were conducted using operational taxonomic units (OTUs) defined by 97 % sequence similarity and by subsampling (1000 iterations) the number of reads to that in the lowest remaining sample (27,719 sequences). After assignment of OTUs, several prominent OTUs classified as “unclassified Cyanobacteria” lineages, were subsequently (BLAST searches) determined to likely be of chloroplast origin and these were removed prior to further analyses. Beta diversity was assessed using the abundance-based theta index [17] and non-abundance (i.e., presence-absence) derived Jaccard index, which were both used to calculate non-metric multidimensional scaling (NMDS) axes and specific OTUs correlated to axes scores. Monte Carlo tests were employed to determine if leaves from the same trees were more similar than between different trees. On average, and by both metrics, bacterial communities from leaves of the same tree demonstrated significantly greater similarity than leaves from other trees (Supplementary Fig. S1). Sequences of the remaining leaves

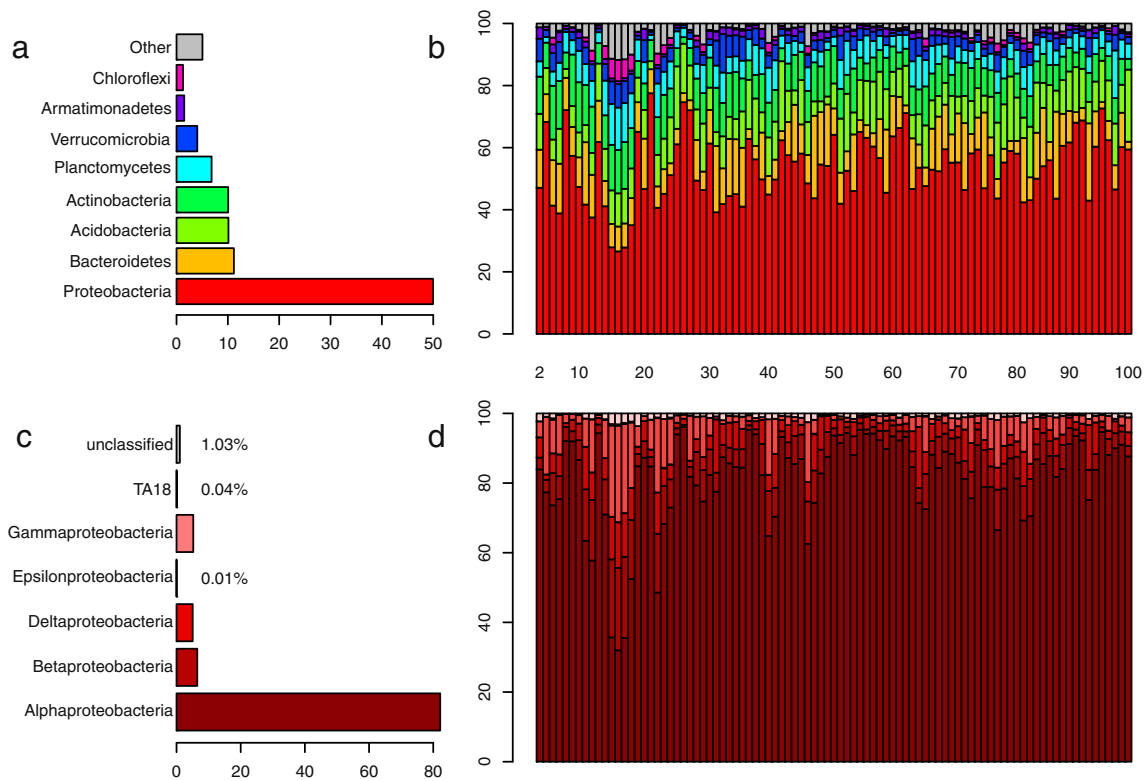
were thus merged by tree to give a final dataset of 91 trees, with at least 27,719 sequences per tree. As the data were analyzed, a small cluster of trees (trees 14–18; located in the western part of cluster 1, Fig. 1) appeared to harbor different phyllosphere communities to the other samples. These trees had high richness of distinctly identifiable sequences (Supplementary Fig. S2), and were reprocessed separately to ensure that sequences were not incorrectly attributed to them during the processing steps.

Further analyses and figure creation were done using R version 3.0.2 [18]. To determine the effect of environmental gradients on bacterial alpha diversity, analysis of covariance (ANCOVA) was performed using DBH, canopy cover, relative elevation, slope, and slope aspect as covariates, and relatedness as a categorical factor. Indirect gradient analyses were performed with those same environmental parameters on multivariate NMDS axes scores using the *envfit* function in the *vegan* package [19]. The significance of the distance-decay relationship was assessed with distance-based Moran’s eigenvector maps (MEMs). Distance-based MEM analysis was carried out using the *PCNM* function in the *PCNM* package [20] and the relationship between all significant eigenfunctions and community dissimilarity matrices was assessed with distance-based redundancy analysis (RDA; using the *dbRDA.D* function [21]). Results of distance-based MEMs are reported along with results from the more commonly used permutation-based Mantel tests using a spatial Euclidean distance matrix between every tree and each community dissimilarity matrix separately. Mothur parameters used for sequence processing and R code used for analyses can be found in Online Resource 2.

## Results

### Bacterial Alpha Diversity and Community Composition

After removal of sequences attributed to archaea, eukarya, mitochondria, and chloroplasts, there were 8,704,825 bacterial sequence reads across all samples, consisting of 20,416 unique sequences. Eight bacterial phyla accounted for >95 % of sequence reads: Proteobacteria (49.9 % of valid bacterial 16S reads, 82.2 % of which were Alphaproteobacteria, Fig. 2c, d); Bacteroidetes (12.1 %); Acidobacteria (10.1 %); Actinobacteria (10.1 %); Planctomycetes (6.8 %); Verrucomicrobia (4.0 %); Armatimonadetes (1.5 %); and Chloroflexi (1.3 %, Fig. 2a, b). At a finer taxonomic level, there were 16 OTUs that were represented by >100,000 reads each, and together accounted for 4,047,874 reads (46.5 % of the reads in the dataset). Of these, a member of the Alphaproteobacteria in the order Rhizobiales was the most abundant and accounted for 385,742 sequences. In total, six of these prominent OTUs belonged to the Alphaproteobacteria and included members of the



**Fig. 2** Bar charts of relative abundances of major bacterial lineages in phyllosphere communities of *Magnolia grandiflora* based off 16S rRNA gene sequence reads. All values are expressed as percentages of 8,704,825 bacterial sequences and 4,343,707 Proteobacteria sequences.

Rhizobiales, Acetobacteraceae, and a *Methylobacterium* species. Two OTUs represented lineages of Acidobacteria and accounted for 486,194 reads, and could both be attributed to the Acidobacteriaceae. Three OTUs (accounting for 592,344 reads) were lineages of Bacteroidetes. Two OTUs belonging to the Actinomycetales (Actinobacteria) were also abundant (202,250 reads). OTUs representing a lineage in the Planctomycetes and an unidentified member of the Verrucomicrobia were also among the most abundant taxa detected.

The alpha diversity of phyllosphere bacterial epiphyte communities (measured by the inverse Simpson metric; found in Online Resource 3) was not related to DBH, relatedness, elevation, slope, or slope aspect of the trees. Stepwise model fitting using AIC scores suggested the model that explained the most variation in inverse Simpson scores included only canopy cover as a significant variable ( $F_{1,89} = 4.499$ ,  $p = 0.037$ ).

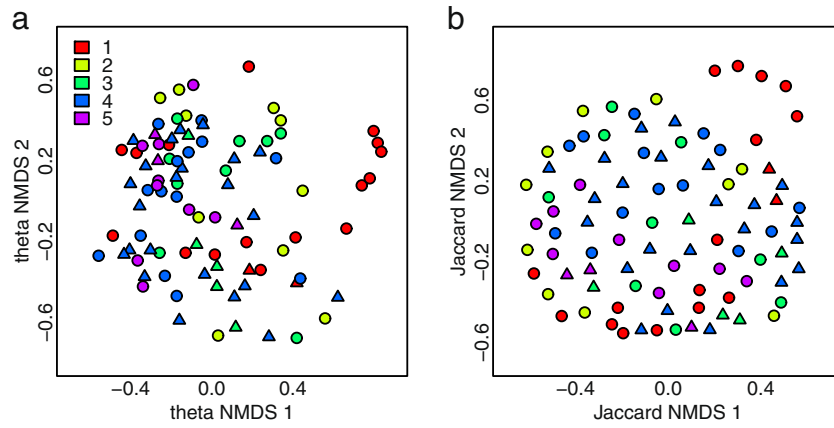
### Patterns in Beta Diversity

Indirect gradient analysis indicated that canopy cover, elevation, slope, and slope aspect were significantly related to relative abundance-based NMDS scores; however, no additional variables were significantly related to presence-absence-based NMDS scores. AMOVAs of community dissimilarity grouped

by the relatedness of the trees showed that theta dissimilarities were not significantly affected by the parent-offspring relationship while Jaccard dissimilarities were ( $F_{1,89} = 2.00$ ,  $p = 0.054$  and  $F_{1,89} = 1.46$ ,  $p = 0.018$ ). However, related trees did not group close together, or distinctly away from, non-related trees on either NMDS plots (Fig. 3).

### Distance-Decay and Similarity

By both relative abundance and presence-absence metrics, there was a positive relationship between tree distance and phyllosphere bacterial community dissimilarity (Fig. 4a, b). Distance-based MEM (dbMEM) analyses found significant relationships between community dissimilarity and spatial predictor eigenfunctions ( $p < 0.001$  in both cases). Because indirect gradient analyses showed significant relationships between several environmental parameters and relative abundance-based community metrics,  $z$ -transformed environmental values were included in the predictor matrix of eigenfunctions when assessing the spatial relationship with relative abundance-based dissimilarities. Theta dissimilarities demonstrated higher adjusted  $R^2$  and  $F$  statistic values than Jaccard dissimilarities ( $R^2_{\theta} = 0.203$ ,  $F_{\theta} = 1.500$ ;  $R^2_J = 0.087$ ,  $F_J = 1.199$ ). Similarly to dbMEM analyses, a partial Mantel test was constructed using dissimilarities of environmental



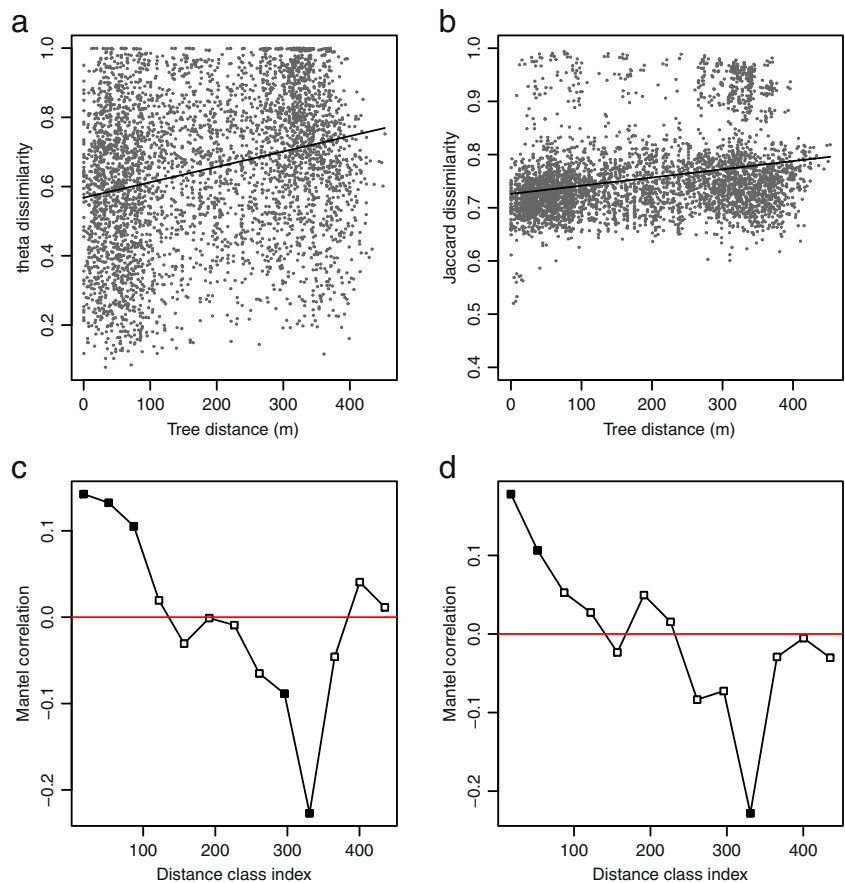
**Fig. 3** Non-metric multidimensional scaling (NMDS) ordinations of community dissimilarity between phyllosphere bacterial communities on leaves of *Magnolia grandiflora* trees as determined from relative abundance-based metrics (theta, **a**, stress=0.246) or presence-absence based metrics (Jaccard, **b**, stress=0.298). Colors correspond to different

spatial clusters of trees (as identified in Fig. 1) and triangular points represent parent and offspring trees which occur in close physical proximity while circular points represent non-related trees. Five trees (14–18) in cluster 1 were separated from others by high values of the first theta NMDS axis and high values of both Jaccard NMDS axes

factor scores (created using factor analysis) as the controlling distance matrix, and showed a significant, positive relationship between the relative abundances of bacterial community members and geographic distance between samples ( $r_M=0.244, p<0.001$ , Fig. 4a). Phyllosphere community dissimilarity based on the presence or absence of bacterial OTUs

(Jaccard index) showed a clearer pattern over geographic distance (Fig. 4b) and a slightly higher Mantel correlation coefficient ( $r_M=0.268, p<0.001$ ). Mantel correlograms indicated that by both dissimilarity metrics, bacterial phyllosphere communities were spatially autocorrelated until 87 m apart (Figs. 4c, d). Additionally, correlograms showed that negative

**Fig. 4** Distance-decay relationship in bacterial phyllosphere communities as visualized by regression of theta and Jaccard dissimilarities against geographic distances between *Magnolia grandiflora* trees (**a, b**). Both theta and Jaccard dissimilarities demonstrated positive changes over geographic distance which were confirmed by Mantel tests showing positive coefficients (theta  $r_M=0.213$ , Jaccard  $r_M=0.268$ ). Mantel correlograms (**c, d**) depict Mantel correlation coefficients plotted over distance classes for theta (**c**) and Jaccard dissimilarities (**d**). Black dots indicate correlation coefficients significantly different than 0 ( $\alpha=0.05$  after applying a Holmes correction)





spatial autocorrelation (i.e., distance-decay, indicated by negative Mantel correlation coefficients) of bacterial leaf communities was significant between trees that were 330 m apart.

## Discussion

The predominant bacterial phyla from the trees sampled in this study match closely with those reported by Jackson and Denney who sequenced clone libraries of *M. grandiflora* bacterial leaf communities from a single tree in the same forest, but repeatedly sampled over different seasons from 2007–2009 [22]. In both cases, Proteobacteria was the most prominent phylum ( $\bar{x} = 53.1$ ,  $SE_{\bar{x}} = 10.7$  from these data, 53–80 % from Jackson and Denney), with Alphaproteobacteria being the most prevalent subphylum. The other proportionally abundant phyla detected in this study (Bacteroidetes, Acidobacteria, and Actinobacteria) were also found at similar frequencies by Jackson and Denney [22], while phyla that were less abundant from this dataset were only transiently present in the clone library data. This suggests that, at least at a broad taxonomic level, there is some consistency to the *M. grandiflora* phyllosphere, and also that comparisons of community composition derived from NGS data to that derived from older cloning-sequencing approaches may be valid.

At a finer taxonomic resolution, specific lineages that were detected by NGS methods in this study were also previously reported in clone libraries. The most prevalent lineages of Alphaproteobacteria reported from *M. grandiflora* clone libraries were the Beijerinckiaceae, Methylobacteriaceae, and the Sphingomonadales [22]; lineages that were also abundant in the larger Illumina dataset. Some of the Acidobacteria sequences derived from the clone libraries were related to *Terriglobus roseus* and *Edaphobacter* [22], and there were numerous sequences identified as a *Terriglobus* species and *Edaphobacter modestum* in the Illumina data. The same patterns were seen for other taxa, and finding such similarities when comparing samples from multiple trees sampled in 2014 to those taken from a single tree from 2007 to 2009, analyzed by different approaches, does suggest some consistency to the *M. grandiflora* phyllosphere microbiome.

Both dbMEM and the Mantel test demonstrated a significant relationship between the spatial structure of the trees and the community dissimilarities. However, a recent critique of the Mantel test in this application argues that dbMEM is more powerful for spatial analyses [21]. Although Mantel tests found similar results as dbMEM methods in this current study, this may be because of the large sample size of the dataset, and Mantel tests may not be appropriate in this context since spatial data do not meet several numerical assumptions [21]. While similar results to Mantel tests may be obtained, future studies on spatial patterns in bacterial communities should

utilize the dbMEM framework because of these considerations.

The theta index was used because its values reflect proportional changes in bacterial abundance that are less dependent on the amplification success of any one sample. In contrast, Bray-Curtis dissimilarities are standardized by shared abundance using the denominator in the equation of the metric, but differences in sequencing depth between samples could exaggerate such pairwise differences in community structure. However, despite these contrasting methodologies, when analyses were repeated using Bray-Curtis dissimilarities, the same results were achieved as with the theta metric.

In this study, the diameter of trees (DBH) was used as a rough estimate of overall tree size which can affect the internal and external leaf structures in several ways (e.g., nitrogen content, lignification, and stomatal density) [23]. However, DBH was not a significant predictor of bacterial community composition on the leaf surface. Other more direct and intensive measurements of specific leaf traits may be necessary to elucidate relationships between individual tree physiology and phyllosphere composition. Canopy cover was used as a proxy for radiation and light that could affect the microbial community of the leaf. Canopy cover differences between trees was significantly correlated with community similarity based on abundances of bacterial OTUs (i.e., using the theta index) but not with that based on presence or absence of individual taxa. This suggests that canopy cover may be important in shaping existing bacterial communities, influencing the proportional abundance of certain populations, rather than affecting bacterial immigration and emigration (i.e., dispersal).

Only NMDS axis scores of phyllosphere community dissimilarity expressed as the theta index were significantly associated to environmental variables, suggesting that relative abundances of bacterial OTUs were sensitive to environmental gradients while the presence or absence of specific OTUs was not. This relationship was further supported by the significance of theta dissimilarities with spatial and environmental factors using dbMEM and constrained RDA ordination. This distinction may reflect different mechanisms of microbial assembly. Dispersal and colonization of the phyllosphere from the atmosphere (which in large part determines the bacteria that occur on a leaf) is a purely stochastic process [24–26] that is contrasted by the presence of predictable, selective forces encountered by colonizers after deposition because of conditions on the leaf surface [1, 24, 25]. Relatedness was significantly associated with Jaccard dissimilarity scores, and likely reflects the very close physical proximity of offspring trees to one another and to their parent tree so that the same OTUs were present.

These findings suggest that environmental gradients may be more important in shaping the relative abundances of existing community members. Stochastic, or neutral, forces may be drivers of the particular bacterial species on plant

leaves. However, leaf age, structure, and morphology were not considered in this study, and these factors may influence bacterial species presence by altering which species may successfully colonize a leaf surface. Further work must be done with the colonization and growth of bacterial communities on leaves, with respect to variation in leaf morphology, in order to determine the true role of neutral forces on phyllosphere community structure.

Previous phyllosphere distance-decay studies have found that dispersal limitations are present along with environmental heterogeneity [10, 11]. The distribution of abundant taxa, in particular, has suggested that environmental heterogeneity is the likely cause of differences in leaf bacterial communities, while distribution patterns in rare bacterial taxa seemed to support dispersal-driven differences [10]. At the spatial scales used in this study, dispersal limitation per se may not occur, but small population sizes may prevent rare or uncommon taxa from dispersing evenly throughout a habitat. Cosmopolitan distribution depends, in part, on large population sizes [8], which rare taxa necessarily lack.

The issue of dispersal limitation in microbial ecology has largely been explored with communities from continuous habitats (e.g., air, soil, or large water bodies) [8, 27–30]. The global biogeographic distribution of microbes in isolated hot springs has been examined with both archaeal (e.g., *Sulfolobus*) and bacterial (e.g., *Synechococcus*) lineages and suggests divergent evolutionary histories, which has been attributed to dispersal barriers [31–33]. At similar scales to those examined in this study, distance-decay pattern of bacterial communities in pools formed by buttressed tree roots was thought to be caused by (unmeasured) environmental variables, not dispersal limitation [34]. However, the method of bacterial description (terminal restriction fragment length polymorphisms) used in that study has biases towards abundant or well-amplified taxa, and analyses were limited to those based on an abundance-based measure of community dissimilarity. If dispersal-based mechanisms are more evident among less common or non-abundant bacterial taxa, the above-mentioned methods may have prevented these patterns from being observed. In this study, we found that the use of both abundance-based and presence-absence dissimilarity metrics demonstrated patterns of community assemblage that would have been difficult to infer otherwise.

The potential influence of both stochastic and deterministic forces make the phyllosphere an intriguing system in which to explore issues of microbial biogeography. The leaf surface offers a discrete, heterogeneous, and repeatable unit that may be used to attain levels of replication that would be prohibitively difficult with communities of larger organisms [35]. Using such isolated leaf communities, this study found that while the abundances of bacterial taxa seemed to depend on conditions around the leaf, community membership was not sensitive to the environment and may rather have been due to

atmospheric dispersal, deposition, and leaf surface traits. Studies of the phyllosphere can also utilize higher sample sizes, such as the 100 trees (and 200 individual leaves) initially sampled in this study, allowing more confidence than limited sampling of communities of macroorganisms. In addition to spatial patterns such as distance-decay and biogeography, exploration of the phyllosphere bacterial communities in other contexts (such as examining temporal variation) may be applied to ongoing ecological questions relating to community stability and resistance, metapopulations and patch dynamics, and mechanisms contributing to diversity.

**Acknowledgments** We thank Heather Valdin and Jennifer Bell for help with the collection and processing of samples.

**Sequence data** The nucleotide sequence data reported are available in the SRA database under the project accession number PRJNA293874.

**Compliance with ethical standards**

**Conflict of Interest** The authors declare that they have no conflict of interests.

## References

1. Vorholt JA (2012) Microbial life in the phyllosphere. *Nat Rev Microbiol* 10:828–840
2. Morris CE, Kinkel LL (2002) Fifty years of phyllosphere microbiology: significant contributions to research in related fields. In: Lindow SE, Hecht-Poinar EJ, Elliott V (eds) *Phyllosphere microbiology*. APS Press, St. Paul, MN, pp 365–375
3. Clayton MK, Hudelson BD (1995) Analysis of spatial patterns in the phyllosphere. In: Andrews JH, Hirano SS (eds) *Microbial ecology of leaves*. Springer, New York, pp 111–131
4. Hubbell SP (2001) *The unified neutral theory of biodiversity and biogeography*. Princeton University Press, Princeton, NJ
5. Kim M, Singh D, Lai-Hoe A, Go R, Abdul Rahim R, Ainuddin AN et al (2012) Distinctive phyllosphere bacterial communities in tropical trees. *Microb Ecol* 63:674–681
6. Kembel SW, O'Connor TK, Arnold HK, Hubbell SP, Wright SJ, Green J (2014) Relationships between phyllosphere bacterial communities and plant functional traits in a neotropical forest. *Proc Natl Acad Sci U S A* 111:13715–13720
7. Andrews JH, Harris RF (2000) The ecology and biogeography of microorganisms on plant surfaces. *Annu Rev Phytopathol* 38:145–180
8. Green J, Bohannan BJM (2006) Spatial scaling of microbial biodiversity. *Trends Ecol Evol* 21:501–507
9. 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
10. 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 Microb* 77:7647–7655
11. Finkel OM, Burch AY, Elad T, Huse SM, Lindow SE, Post AF, Belkin S (2012) Distance-decay relationships partially determine

- diversity patterns of phyllosphere bacteria on *Tamrix* trees across the sonoran desert. *Appl Environ Microb* 78:6187–6193
12. Brewer JS (2001) Current and presettlement tree species composition of some upland forests in northern Mississippi. *J Torrey Bot Soc* 128:332–349
  13. Hijmans RJ (2015) Raster: geographic data analysis and modeling. R package version 2.3-24., <http://CRAN.R-project.org/package=raster>
  14. Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD (2009) Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Appl Environ Microb* 79: 5112–5120
  15. Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB et al (2009) Introducing mothur: open-source, platform-independent, community-supported, software for describing and comparing microbial communities. *Appl Environ Microb* 75:7537–7541
  16. Schloss PD, Gevers D, Westcott SL (2011) Reducing the effects of PCR amplification and sequencing artifacts on 16S rRNA-based studies. *PLoS One* 6:1–14
  17. Yue JC, Clayton MK (2005) A similarity measure based on species proportions. *Commun Stat-Theor M* 34:2123–2131
  18. R Core Team (2013) R: a language and environment for statistical computing. R Foundation for statistical Computing, Vienna, Austria, <http://www.R-project.org/>
  19. Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'Hara RB et al (2015) Vegan: community ecology package. R package version 2.2-1., <http://CRAN.R-project.org/package=vegan>
  20. Legendre P, Borcard D, Blanchet FG, Dray S (2013) PCNM: MEM spatial eigenfunction and principal coordinate analyses. R package version 2.1-2/r109., <http://R-Forge.R-project.org/projects/sedar/>
  21. Legendre P, Fortin MJ, Borcard D (2015) Should the Mantel test be used in spatial analysis? *Methods Ecol Evol*: 1239–1247. Supplementary Information, Appendix S4, dbRDA *F*-test for response data in dissimilarity matrix form: 26–30
  22. Jackson CR, Denney WC (2011) Annual and seasonal variation in the phyllosphere bacterial community associated with leaves of the southern magnolia (*Magnolia grandiflora*). *Microb Ecol* 61:113–122
  23. Steppe K, Niinemets Ü, Teskey RO (2011) Tree size- and age-related changes in leaf physiology and their influence on carbon gain. In: Meinzer FC, Lachenbruch B, Dawson TE (eds) Size- and age-related changes in tree structure and function. Springer, New York, NY, pp 235–254
  24. Kinkel LE (1997) Microbial population dynamics on leaves. *Annu Rev Phytopathol* 35:327–347
  25. 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
  26. Remus-Emsermann MNP, Tecon R, Kowalchuk GA, Leveau JHJ (2012) Variation in local carrying capacity and the individual fate of bacterial colonizers in the phyllosphere. *ISME J* 6:756–765
  27. Lighthart B (1997) The ecology of bacteria in the alfresco atmosphere. *FEMS Microbiol Ecol* 23:263–274
  28. Brandao PFB, Clapp JP, Bull AT (2002) Discrimination and taxonomy of geographically diverse strains of nitrile-metabolizing actinomycetes using chemometric and molecular sequencing techniques. *Environ Microbiol* 4:262–276
  29. Horner-Devine MC, Lage M, Hughes JB, Bohannon BJM (2004) A taxa-area relationship for bacteria. *Nature* 432:750–753
  30. Galand PE, Casamayor EO, Kirchman DL, Lovejoy C (2009) Ecology of the rare microbial biosphere of the Arctic Ocean. *Proc Natl Acad Sci U S A* 106:22427–22432
  31. Whitaker RJ, Grogan DW, Taylor JW (2003) Geographic barriers isolate endemic populations of hyperthermophilic archaea. *Science* 301:976–978
  32. Papke RT, Ramsing NB, Bateson MM, Ward DM (2003) Geographical isolation in hot spring cyanobacteria. *Environ Microbiol* 5:650–659
  33. Martiny JBH, Bohannon BJM, Brown JH, Colwell RK, Fuhman JA, Green JL et al (2006) Microbial biogeography: putting microorganisms on the map. *Nat Rev Microbiol* 4:102–112
  34. Bell T (2010) Experimental tests of the bacterial distance-decay relationship. *ISME J* 4:1357–1365
  35. Meyer KM, Leveau JHJ (2012) Microbiology of the phyllosphere: a playground for testing ecological concepts. *Oecologia* 168:621–629