environmental microbiology

Environmental Microbiology (2012) **14**(1), 147–161 doi:10.1111/j.1462-2920.2011.02560.x

Photoautotrophic symbiont and geography are major factors affecting highly structured and diverse bacterial communities in the lichen microbiome

Brendan P. Hodkinson,[1]* Neil R. Gottel,[2] Christopher W. Schadt[2,3] and François Lutzoni[4]

[1]*International Plant Science Center, The New York Botanical Garden, Bronx, NY 10458-5126, USA.*
[2]*Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831, USA.*
[3]*Department of Microbiology, University of Tennessee, Knoxville, Tennessee 37796, USA.*
[4]*Biology Department, Duke University, Durham, North Carolina 27708, USA.*

Summary

Although common knowledge dictates that the lichen thallus is formed solely by a fungus (mycobiont) that develops a symbiotic relationship with an alga and/or cyanobacterium (photobiont), the non-photoautotrophic bacteria found in lichen microbiomes are increasingly regarded as integral components of lichen thalli. For this study, comparative analyses were conducted on lichen-associated bacterial communities to test for effects of photobiont-types (i.e. green algal vs. cyanobacterial), mycobiont-types and large-scale spatial distances (from tropical to arctic latitudes). Amplicons of the 16S (SSU) rRNA gene were examined using both Sanger sequencing of cloned fragments and barcoded pyrosequencing. Rhizobiales is typically the most abundant and taxonomically diverse order in lichen microbiomes; however, overall bacterial diversity in lichens is shown to be much higher than previously reported. Members of Acidobacteriaceae, Acetobacteraceae, Brucellaceae and sequence group LAR1 are the most commonly found groups across the phylogenetically and geographically broad array of lichens examined here. Major bacterial community trends are significantly correlated with differences in large-scale geography, photobiont-type and mycobiont-type. The lichen as a microcosm represents a structured, unique microbial habitat with greater ecological complexity and bacterial diversity than previously appreciated and can serve as a model system for studying larger ecological and evolutionary principles.

Received 1 April, 2011; accepted 1 July, 2011. *For correspondence. E-mail brendan.hodkinson@gmail.com; Tel. (+1) 443 340 0917; Fax (+1) 718 817 8648.

© 2011 Society for Applied Microbiology and Blackwell Publishing Ltd

Introduction

In recent decades, a number of classic symbiotic systems that were previously characterized as simple bi-eukaryotic mutualisms have been shown to host a rich diversity of prokaryotes that interact with their hosts in unique ways (Founoune *et al.*, 2002; Artursson *et al.*, 2006; Taylor *et al.*, 2007; Warnecke *et al.*, 2007; Hoffman and Arnold, 2010). Usually, the term 'lichen' refers to a specific structural entity formed by a symbiosis between a fungus (mycobiont) and an alga and/or cyanobacterium (photobiont). However, recent studies have demonstrated that lichen thalli consistently host hyperdiverse fungi (other than the fungal partner forming the lichen thallus), which are more closely related to endophytic fungi found in healthy above ground tissues of plants than to lichen-forming fungi (Arnold *et al.*, 2009). These non-lichen-forming fungi found in lichens are referred to as endolichenic fungi. Lichens have also been found to host communities of non-photoautotrophic bacteria that are substantial in terms of cell density (Grube *et al.*, 2009) and phylogenetic diversity (Bates *et al.*, 2011). The non-photoautotrophic bacteria of the lichen microbiome are increasingly regarded as significant players in the ecology of lichens. Because many lichens are able to grow on extremely nutrient-poor substrates, it has been suggested that these bacteria provide some lichen thalli with a substantial source of crucial nutrients (González *et al.*, 2005; Cardinale *et al.*, 2006; Liba *et al.*, 2006).

Alphaproteobacteria have been shown to dominate bacterial communities associated with lichens having green algal photobionts (i.e. green algal lichens, hereafter referred to as 'chlorolichens') (Cardinale *et al.*, 2008; Grube *et al.*, 2009). One lineage of Alphaproteobacteria ('LAR1' in the order Rhizobiales) has been consistently detected across a diversity of chlorolichens, and has been suggested as a contributor to the lichen symbiosis (Hodkinson and Lutzoni, 2009; Bates *et al.*, 2011). Importantly, the lineage is almost exclusively associated with lichens; however, only a few lichen species (10 by Hodkinson and Lutzoni, 2009; 4 by Bates *et al.*, 2011) from a limited geographic range in the Southern USA have been examined for the presence of LAR1. Therefore, based on this

sampling it was not possible to determine how broadly across lichens and bioclimatic zones LAR1 would be found. Moreover, LAR1 remains a poorly understood group in terms of its physiology and genetics, mostly due to the lack of a cultured representative (Hodkinson and Lutzoni, 2010).

In a pyrosequencing survey, Bates and colleagues (2011) suggested that fine-scale geographic patterns do not exist for lichen-associated bacterial communities. Instead, they found the identity of the lichen-forming fungal partner (i.e. mycobiont-type) to be the best predictor of community structure, as Grube and colleagues (2009) had indicated previously using fingerprinting techniques. Although these two experiments, each conducted at a single geographical site, demonstrate that lichens with different mycobionts may harbour different bacterial communities, no study has examined how these communities might vary across large spatial distances.

Another major factor that could potentially influence bacterial community composition is photobiont-type. While green algae fix carbon and release it in the form of complex sugar alcohols, Cyanobacteria typically release fixed carbon in the form of glucose (Palmqvist, 2002; Adams *et al*., 2006); additionally, Cyanobacteria can fix atmospheric nitrogen, while green algae cannot (Belnap, 2001; Rai, 2002). Therefore, lichen thalli housing these different types of photobionts are likely to have different ecologies based on carbon and nitrogen composition and concentration differences, which could presumably lead different microbial communities. Hodkinson and Lutzoni (2009) presented preliminary evidence that the structure of bacterial communities associated with cyanobacterial lichens (hereafter referred to as 'cyanolichens') differ from those found with chlorolichens; however, studies of overall community differences between these lichen types are lacking.

For this study, community-wide comparative analyses were conducted on lichen-associated bacteria from samples representing a wide geographic/bioclimatic range, the two major types of photobionts, and multiple lineages of lichen-forming fungi, to test for effects that are attributable to spatial distances/bioclimatic zone, photobiont-type and mycobiont-type. Data were obtained through cloning and Sanger sequencing of 16S fragments from specific members of the order Rhizobiales (i.e. LAR1 and *Methylobacterium*) as well as 454 sequencing of 16S amplicons produced with general primers to represent overall bacterial diversity.

Results

Clone libraries

A data set of 1190 cloned, chimera-checked, Rhizobiales 16S sequences (nucleotide sites 533 to 1185, according to *E. coli* numbering; Brosius *et al*., 1978) was produced from the group of lichen samples processed. All sample types produced sequences except for the collections of *Alectoria* from Nome, AK. ANOSIM analyses on the UniFrac and Bray-Curtis distance matrices conducted on data from individual sites with each mycobiont/photobiont/site/cluster combination treated separately indicate that community composition is not significantly correlated with local spatial scale ($0.250 \leq P \leq 0.967$, $r \leq 0.0333$). Therefore, all further analyses were conducted with data from different clusters combined (i.e. each mycobiont/photobiont/site combination treated separately). These analyses indicate that photobiont-type, mycobiont-type and geography are all significantly correlated ($P < 0.05$) with lichen-associated bacterial community composition (Table 1). The analysis of clone library data using an unweighted UniFrac algorithm gave the strongest signifi-

Table 1. Results of ANOSIM tests performed on UniFrac (phylogenetically-based) distance and Bray-Curtis (OTU-based) dissimilarity matrices.

Data set	Metric	Variable of interest	Partial predictor	*r*	*P*
Clones	UniFrac-UW	Photobiont	Geography	0.2747	**0.001**
		Mycobiont	Photobiont	0.3824	**0.012**
			Geography	0.3824	**0.001**
		Geography	Photobiont	0.3323	**0.001**
			Mycobiont	0.3323	**0.001**
	UniFrac-WN	Photobiont	Geography	0.3291	**0.001**
		Mycobiont	Photobiont	0.2187	0.291
			Geography	0.2187	**0.032**
		Geography	Photobiont	0.1353	**0.012**
			Mycobiont	0.1353	**0.048**
	Bray-Curtis	Photobiont	Geography	0.3230	**0.001**
		Mycobiont	Photobiont	0.2723	*0.064*
			Geography	0.2723	**0.008**
		Geography	Photobiont	0.2968	**0.001**
			Mycobiont	0.2968	**0.001**
454	UniFrac-UW	Photobiont	Geography	0.2485	**0.013**
		Mycobiont	Photobiont	0.1007	0.568
			Geography	0.1007	0.225
		Geography	Photobiont	0.1278	*0.075*
			Mycobiont	0.1278	*0.051*
	UniFrac-WN	Photobiont	Geography	0.3001	**0.001**
		Mycobiont	Photobiont	0.4786	**0.028**
			Geography	0.4786	**0.004**
		Geography	Photobiont	0.1003	0.141
			Mycobiont	0.1003	0.133
	Bray-Curtis	Photobiont	Geography	0.2820	**0.001**
		Mycobiont	Photobiont	0.3692	*0.080*
			Geography	0.3692	**0.017**
		Geography	Photobiont	0.0462	0.297
			Mycobiont	0.0462	0.234

Analyses were performed with samples from the three 'clusters' at each site combined. Correlation (*r*) and significance (*P*) values are shown for both unweighted (UW) and weighted + normalized (WN) UniFrac distance matrices, as well as Bray-Curtis dissimilarity matrices based on the relative abundance of each OTU in each sample. Significance values for factors significant at the $P < 0.05$ level are in bold; values considered to be marginally significant ($0.05 \leq P < 0.10$) are italicized.

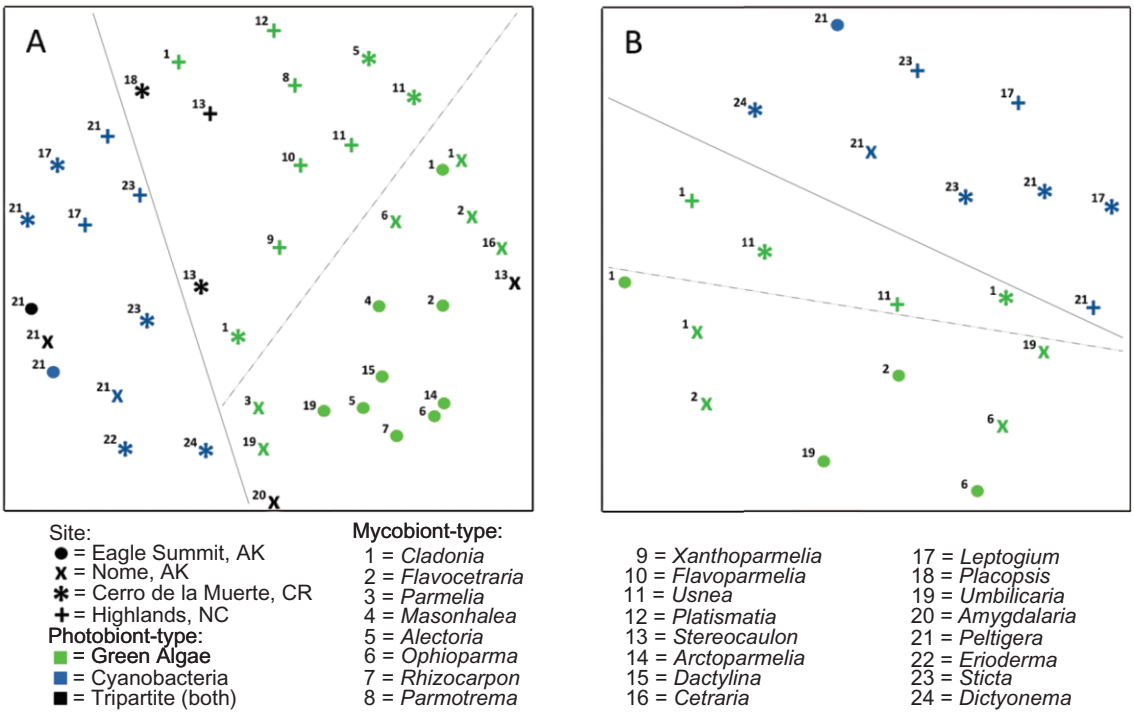

Fig. 1. Non-metric multidimensional scaling plots produced from OTU-based Bray-Curtis dissimilarities. The first plot (A) shows results obtained from clone library data of 16S sequences from the order Rhizobiales, while the second plot (B) was produced from 454 barcoded 16S amplicon data representing approximately half the number of samples but ~100 times as many sequences from a much wider range of bacterial diversity. Continuous lines act as visual aids to delimit communities associated with the two major photobiont-types, whereas dashed lines delimit communities associated with chlorolichens from northern versus southern sites. 166 × 107 mm (600 × 600 DPI).

cance for all factors, even when accounting for each other partial predictor [note that photobiont-type was not analysed with mycobiont-type as a partial predictor because mycobiont-type is almost entirely nested within photobiont-type (i.e. no mycobiont-type can associate exclusively with either green algae or Cyanobacteria, although a small number can have either Cyanobacteria or both photobiont-types simultaneously)]. Other methods showed varying degrees of significance for these three factors. Photobiont and geography were significant factors in all three analysis types ($P = 0.001$ and $P = 0.001$–0.048, respectively). With the clone data set, mycobiont-type appeared to be significant when only accounting for geography as a partial predictor ($P = 0.01$–0.032), but results were mixed when accounting for photobiont-type. In analyses of weighted + normalized UniFrac distances, mycobiont-type was not shown to be significant when accounting for photobiont-type as a partial predictor; however, an analysis of Bray-Curtis dissimilarities showed marginal significance ($P = 0.064$) and an analysis of unweighted UniFrac distances showed mycobiont-type to be a significant factor at the level of $P = 0.012$. In terms of correlation coefficients (r-values), mycobiont was shown to be most strongly correlated with bacterial community composition in the unweighted UniFrac distance-based analyses, while photobiont was most strongly correlated in the weighted + normalized UniFrac distance-based analyses and Bray-Curtis dissimilarity-based analyses (see Table 1 r-values).

A visual representation of the clone library data based on non-metric multidimensional scaling of Bray-Curtis dissimilarities is shown in Fig. 1A. A continuous line is drawn as a visual aid dividing communities associated with chlorolichens from those associated with cyanolichens; communities associated with tripartite lichens (those with both cyanobacterial and green algal associates) are found on both sides of this line. A dashed line is drawn as a visual aid dividing communities associated with chlorolichens sampled in Alaska from those associated with chlorolichens sampled from the more southern sites (Costa Rica and North Carolina). Globally, tendencies based on mycobiont-type can be loosely inferred, but photobiont-type is the most visually striking determinant of bacterial community composition in this plot, with geography (northern versus southern sites) also being shown as a major determinant for bacterial communities found in chlorolichens (Fig. 1A). Non-metric multidimensional scaling plots of UniFrac distances showed similar structure based on these factors.

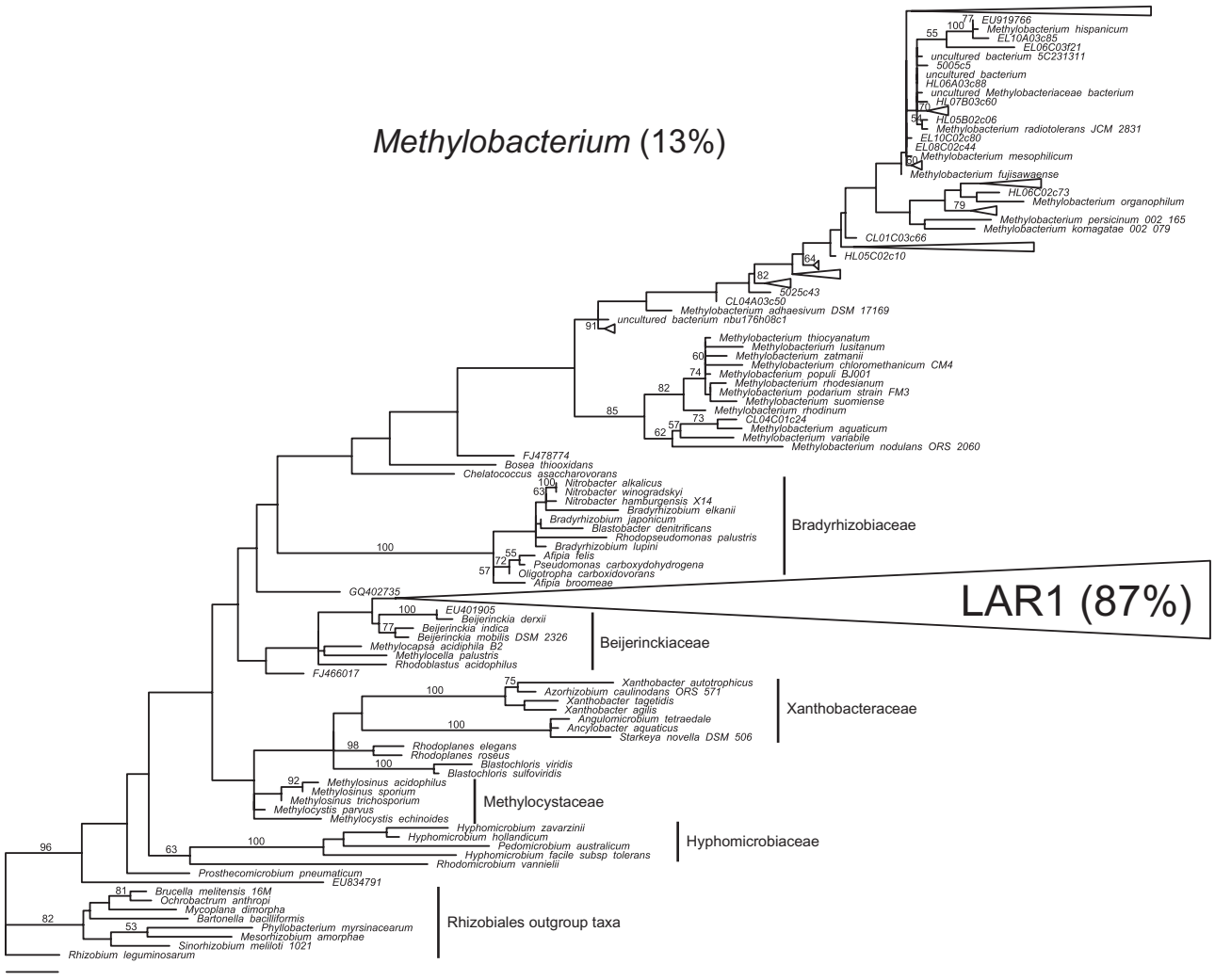

Fig. 2. Maximum likelihood (ML) phylogeny of cloned sequences generated as part of this study (10-digit alpha-numeric codes) with closest BLAST hits (8-digit accession numbers), sequences from reference strains [organism names; GenInfo Identifiers are listed by Hodkinson and Lutzoni (2009)], and lichen-associated cloned sequences generated by Hodkinson and Lutzoni (2009; 6- or 7-digit alpha-numeric codes). Numbers above nodes represent ML bootstrap proportions (BP) > 50%. Clades comprised of cloned sequences only are collapsed (for full tree file and list of removed identical sequences, see Dryad data package). The LAR1 lineage accounts for 87% (1034/1190) of all cloned sequences generated as part of this study, while sequences from bacteria scattered throughout the genus *Methylobacterium* account for the other 13% (156/1190). 201 × 150 mm (600 × 600 DPI).

In the maximum-likelihood phylogeny (summarized in Fig. 2 and shown in full in Fig. S1), LAR1 is reconstructed as a single monophyletic group, although without significant support. All cloned Rhizobiales sequences fall either into the LAR1 clade or are scattered throughout the genus *Methylobacterium*. Of the cloned sequences generated as part of this study, 87% were found to be part of the LAR1 lineage, with the remaining 13% belonging to *Methylobacterium*.

Pyrosequencing

Barcoded 454 sequencing with Titanium chemistry on half of a picotitre plate produced 351 582 sequence reads before analyses with Mothur. After the multiple quality check procedures outlined below, there remained 119 322 bacterial 16S rDNA sequences of ~500 bp each (nucleotide sites 533 to ~1000–1050, according to *E. coli* numbering; Brosius *et al*., 1978).

A taxonomic summary of the diversity found in each 454-sequenced amplicon pool (Fig. 3) shows that lichen-associated bacterial communities are typically dominated by Alphaproteobacteria (60% of the total for the data set as a whole; 71 824 sequences out of 119 322 sequences). However, in contrast, the two samples of *Ophioparma* [the only crustose (saxicolous) lichen sampled using pyrosequencing] are dominated by Acidobacteria. Despite the fact the Acidobacteria contain many

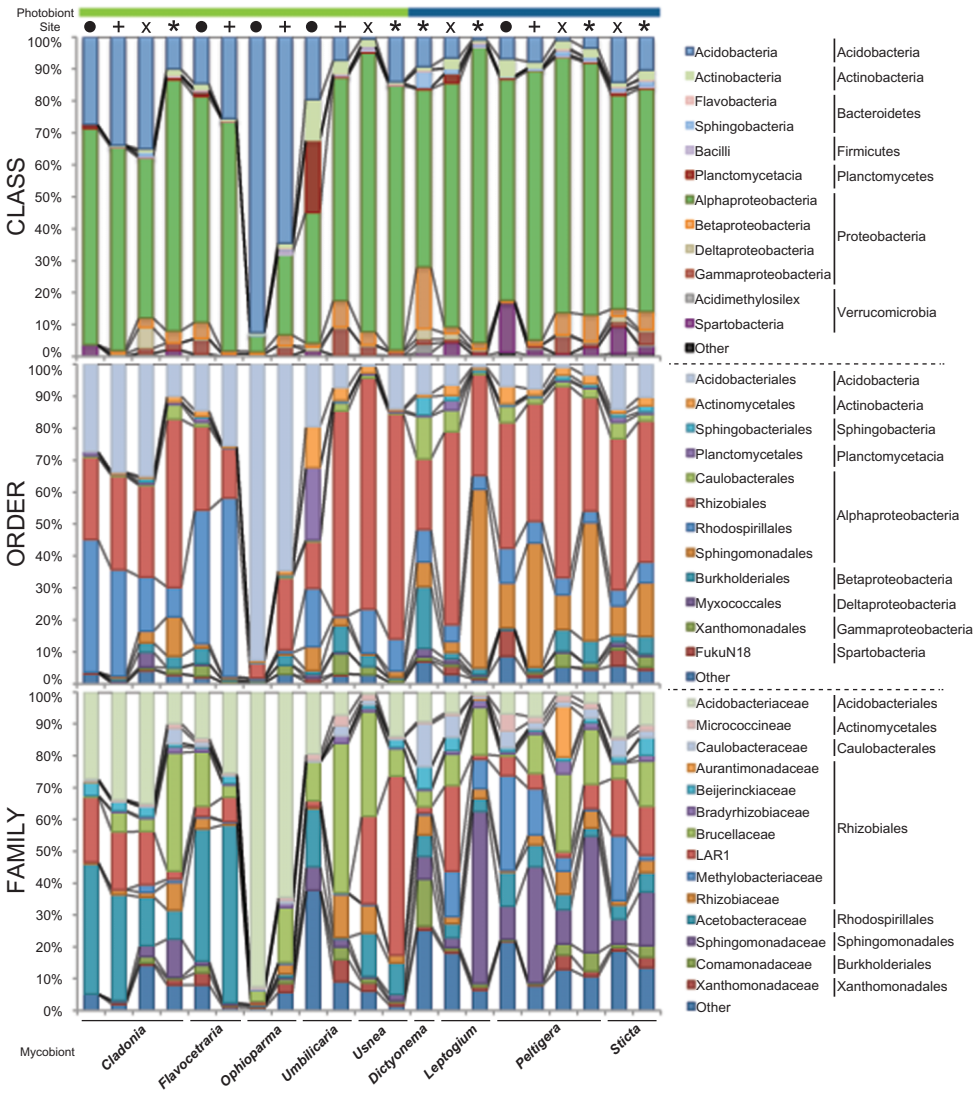

Fig. 3. Relative abundances of bacterial taxa recovered from each lichen sample analysed with 454 sequencing. Taxa included are the classes with > 0.1% of 16S sequences in the full 454 data set, orders with > 0.5% of sequences, and 'families' with > 1% of sequences; at each rank, all sequences that did not fit into one of these categories were classified as 'other'. Photobiont is indicated by the bar across the top of the figure (light green = green algae; dark blue = Cyanobacteria), while the site is indicated with symbols immediately below the bar (• = Eagle Summit, AK; + = Nome, AK; x = Highlands, NC, and * = Cerro de la Muerte, CR). Mycobionts are indicated by name along the base of the figure. 160 × 182 mm (600 × 600 DPI).

uncultured lineages (Barns *et al.*, 2007), the vast majority (> 99%) of Acidobacteria sequences generated as part of this study were assigned to three named genera: *Acidobacterium* (65%; 19 341 of 29 864 sequences), *Edaphobacter* (32%; 9615 of 29 864 sequences) and *Terriglobus* (2%; 622 of 29 864 sequences).

Of the orders within Alphaproteobacteria, Rhizobiales generally dominates, but is often co-dominant with Rhodospirillales (typically in chlorolichens) or Sphingomonadales (typically in cyanolichens). At the familial rank, Rhizobiales is the most diverse order, with seven 'families' in the order each having > 1% of the overall diversity in the sample set (Aurantimonadaceae, Beijerinckiaceae, Bradyrhizobiaceae, Brucellaceae, LAR1, Methylobacteriaceae and Rhizobiaceae). Sequences from Aurantimonadaceae are all assigned to *Aurantimonas* (100%; 3139 of 3139 sequences), those from Beijerinckiaceae mostly were not assignable to a genus (66%; 1478 of 2239 sequences) (although 27% were assigned to *Beijerinckia*; 614 of 2239 sequences), sequences from Bradyrhizobiaceae were variously assigned, with the most common being *Bradyrhizobium* (41%; 735 of 1774 sequences), those from Brucellaceae are almost all assigned to the genus *Ochrobactrum* (nearly 100%; 18 671 of 18 687 sequences), those in Methylobacteriaceae are all (100%; 3760 of 3760 sequences) assigned

to *Methylobacterium*, and members of Rhizobiaceae are almost all assigned to *Rhizobium* (nearly 100%; 5394 of 5396 sequences). Overall, the most dominant taxa from the order Rhizobiales are *Aurantimonas* (7%; 3139 of 43 444 sequences), *Ochrobactrum* (43%; 18 671 of 43 444 sequences), LAR1 (18%; 7860 of 43 444 sequences), *Methylobacterium* (9%; 3760 of 43 444 sequences) and *Rhizobium* (12%; 5394 of 43 444 sequences). Members of the genus *Methylobacterium* are preferentially associated with cyanolichens.

Each of the non-Rhizobiales families having > 1% of the overall diversity in the sample set belongs to its own order. Sequences assigned to the order Rhodospirillales almost entirely represent the family Acetobacteraceae (96%; 15 904 of 16 525 sequences), with most of those being further unclassifiable (84%; 13 353 of 15 904 sequences). The order Sphingomonadales is mostly represented by the family Sphingomonadaceae (96%; 7811 of 8159 sequences), with most of those representing the genus *Sphingomonas* (93%; 7278 of 7811 sequences).

Of the four genera proposed as the most ubiquitous by Grube and colleagues (2009) from culture-based analyses, two were not found to be present at all in this study (*Bacillus* and *Paenibacillus*), one was found extremely rarely (*Acinetobacter*, represented by four sequences across two samples), and one was found to be present in 18 of 21 samples sequenced using 454 (*Burkholderia*), but consistently remained at low levels (< 1% overall; 733 of 119 322 total sequences). The families Acidobacteriaceae, Acetobacteraceae, Brucellaceae, and LAR1 were the most abundant families of Bacteria across the lichens sampled here (forming the top two-percentile of families overall). A full taxonomic summary of sequence assignments for all samples is available in the 'Taxonomy_454' directory of the Dryad data package (http://dx.doi.org/10.5061/dryad.t99b1/1 doi:10.5061/dryad.t99b1/1). Further investigation into the top four families showed that LAR1 was especially high in diversity (rivalled only by Acetobacteraceae) and also the proportion of sequences that were non-redundant; however, the same analyses showed the family Brucellaceae to be low in terms of both diversity and the proportion of non-redundant sequences (Fig. 4).

ANOSIM analyses of UniFrac distances and Bray-Curtis dissimilarities based on 454 data indicate that photobiont-type and mycobiont-type are significantly correlated ($P < 0.05$) with bacterial community composition, while geography shows only marginal significance ($0.05 \leq P < 0.10$) (Table 1). Photobiont-type was a significant factor in all three analyses ($P = 0.001$–0.013), but was only more strongly correlated with community composition than the other factors in the unweighted UniFrac analysis (see Table 1 *r*-values). Mycobiont-type was found to be significant when accounting for partial predictors in

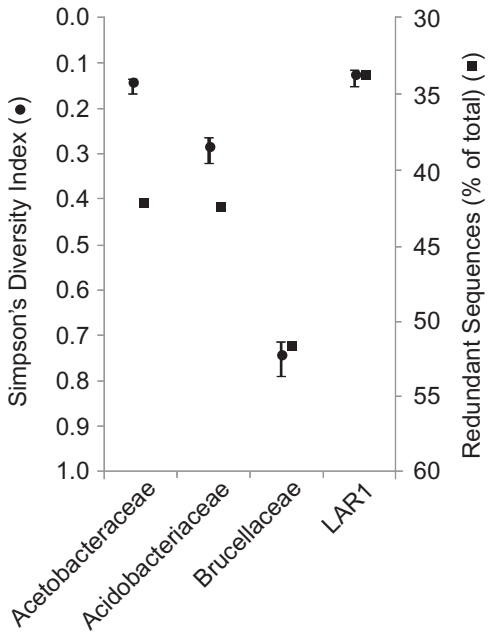

Fig. 4. Simpson's diversity indices (0 = high diversity, 1 = low diversity) and percentage of redundant sequences calculated for each of the four most abundant families of Bacteria across the lichens sampled as part of this study. Brucellaceae shows especially low diversity and a high redundancy in sequence types, while LAR1 has both high diversity and low redundancy. 98 × 128 mm (600 × 600 DPI).

analyses of weighted + normalized UniFrac distances ($P = 0.004$–0.028), but not in analyses of unweighted UniFrac distances ($P = 0.225$–0.568). In the Bray-Curtis dissimilarity-based analyses, mycobiont-type appeared to be significant when accounting for geography as a partial predictor ($P = 0.017$), but was only marginally significant with photobiont as a partial predictor ($P = 0.08$) (this compares favourably with results from analyses of clone library data). In the weighted + normalized UniFrac distance-based and Bray-Curtis dissimilarity-based analyses, mycobiont showed the strongest correlation with overall community composition ($r = 0.4786$ and 0.3692, respectively). In the analyses performed on the 454 data set, geography was not significant at the $P < 0.05$ level ($P = 0.051$–0.297), although its significance could be considered marginal ($P = 0.051$–0.075) when accounting for partial predictors using unweighted UniFrac distances.

A non-metric multidimensional scaling plot of Bray-Curtis dissimilarities is shown in Fig. 1B. Although samples with the same mycobiont-type sometimes group together, clusters are loose and overlapping, making photobiont-type and geographical distribution of chlorolichens the factors that are visually most clearly correlated with bacterial community composition in this plot. These results mirror those from the clone library data. Similar structure based on the various factors was revealed in non-metric multidimensional scaling plots of UniFrac distances.

Discussion

Additional factors come to light

While it has been demonstrated previously that mycobiont-type can be significantly correlated with lichen-associated bacterial community composition at small spatial scales (Grube *et al*., 2009; Bates *et al*., 2011), the present study demonstrates the importance of photobiont-type and large-scale geography in determining the overall structure of these communities. These results indicate not only that communities differ between different types of lichens in different regions, but more generally that the relationships between lichens and their bacterial associates are more complex than previously suggested. Based on the data presented in Table 1, it is clear that the use of any single statistical method may either reveal or obscure significant ecological community patterns; this is in line with recent simulation experiments (Kuczynski *et al*., 2010) and is in fact the reasoning behind using multiple analytical methods on two distinctly different data sets to infer ecological trends in lichen-associated bacterial communities.

Photobiont–bacterial interactions

The present study is the first to examine bacterial community composition of cyanolichens and tripartite associations encompassing both types of photobionts side-by-side with chlorolichens, allowing the elucidation of photobiont-associated patterns that have not been seen before. All statistical tests demonstrate a significant correlation between photobiont-type and bacterial community composition (Table 1). One explanation for the major difference between cyanolichen- and chlorolichen-associated communities is the availability of fixed nitrogen. Because cyanobacteria can fix atmospheric nitrogen, this nutrient is likely to be more readily available to microbes living in or on the surface of cyanolichens (Belnap, 2001; Rai, 2002). Therefore, bacteria with reduced fitnesses in nitrogen-limiting environments may thrive in cyanolichen thalli.

Another potential explanation for the bacterial differences seen between chlorolichens and cyanolichens has to do with the forms of fixed carbon commonly available from the different photobiont-types. Green algae that associate with lichens typically release carbon fixed through photosynthesis in the form of sugar alcohols ('polyols'), while cyanobacteria typically release their fixed carbon in the form of glucose (Palmqvist, 2002; Adams *et al*., 2006). Glucose is a simple sugar that many organisms can process; however, the sugar alcohols of green algae require a special set of enzymes such as polyol dehydrogenases, and therefore, the communities associated with chlorolichens may be biased towards bacteria with these types of enzymes.

Mycobiont–bacterial interactions

The works of Cardinale and colleagues (2008) and Grube and colleagues (2009) suggest that the majority of lichen-associated bacteria live on or near the fungal surfaces of lichen thalli, often in biofilm-like coverings. It has been previously demonstrated that bacterial community composition patterns are correlated with mycobiont-type (Grube and Berg, 2009; Grube *et al*., 2009; Bates *et al*., 2011), and the analyses conducted as part of the present study further corroborate this story (Table 1). One factor that could account for this trend is the presence of secondary compounds (e.g. depsides and depsidones, among others) that are especially abundant in lichens. Many of these compounds have been shown to possess antibiotic properties, suggesting that they could be playing an active role in selecting for specific types of bacteria that are able to withstand them. Further research on varying susceptibility across a diversity of potential bacterial symbionts is needed to specifically test this hypothesis.

Effects of geography

While community composition of multicellular Eukaryotes is clearly geographically and bioclimatically patterned, it has been demonstrated that bacterial communities in many environments have little to no significant biogeographical structure, even across different bioclimatic zones (Humbert *et al*., 2009; Chu *et al*., 2010). Previous studies of lichens have shown no evidence of geography or spatial scale being a factor in bacterial community composition (Grube *et al*., 2009; Bates *et al*., 2011); however, these studies were localized and did not examine trends across a large spatial scale or across bioclimatic zones. Based on the analyses conducted as part of this study, lichen-associated bacteria do seem to be clearly patterned based on geographical/bioclimatic differences. The issue typically cited to determine geographical patterning is dispersal (i.e. a lack of dispersal leads to geographical differentiation of communities, while frequent dispersal leads to a lack of community structure based on geography). Organisms that require a very specific environment, therefore, will less frequently be able to disperse and colonize a new area, leading to distinct geographical patterning. Indeed, it is often bacteria in specific symbiotic systems that are found to have significant geographical patterning (Taylor *et al*., 2005; Najar-Rodríguez *et al*., 2009; Russell *et al*., 2009). Lichen-associated bacterial communities, likewise, fall into this category, suggesting that many groups of lichen-associated bacteria may be obligately associated with lichens and that their dispersal is limited by a reliance on phenotypic traits of their hosts. A dispersal-based expla-

nation of geographical patterning could account for the seemingly contradictory evidence suggesting that geography is not significant on a small spatial scale (where lichen dispersal is a relatively frequent occurrence) but significant on a large spatial scale (where host dispersal can be a limiting factor). It is also possible that the dispersal ability of the host is independent of the dispersal ability of its microbiotic bacteria where, for example, the ecological amplitude of certain lichen-associated bacteria is less than its lichen hosts.

Lichen growth forms and secondary compounds as putative factors affecting bacterial communities

Previous studies have shown a dominance of Alphaproteobacteria in lichen microbiomes (Cardinale *et al*., 2008; Grube *et al*., 2009; Bates *et al*., 2011), a result that is mostly corroborated by the data presented here. However, the two lichen thalli with the mycobiont *Ophioparma* had bacterial communities dominated by Acidobacteria (Fig. 3). These two lichens were the only ones in the 454 data set with a crustose habit, meaning that they are in full contact with the substrate (siliceous rock in this case) and they do not have lower cortices, both of which might impact community structure. The one crustose lichen examined by Grube *et al*., 2009 (*Lecanora polytropa*) was shown to be different from the macrolichens that were part of the same study, in both the spatial distribution of the bacteria across the thallus and in the composition of the bacterial community. However, neither the fingerprinting-based technique nor microscopy could reveal which organisms differed in the crustose community. Further studies of growth habit as a potential determinant of bacterial community composition are needed to fully understand the effects of this factor.

The *Ophioparma* samples examined also produce the naphthaquinone compound haemoventosin, which is extremely abundant in the apothecia (Rycroft *et al*., 1996) (see brownish-red apothecia in Fig. 5C). Naphthaquinone compounds have previously been shown to have antibacterial properties (Riffel *et al*., 2002), and the Acidobacteria found in these thalli may be specially adapted to withstand such compounds. Because lichens typically produce secondary compounds, often in abundance (Culberson *et al*., 1984), the effects of these chemicals on bacterial community composition certainly warrant further study. Additionally, in soils and other environments the ratio of Proteobacteria to Acidobacteria is often considered to be an indicator of copiotrophic versus oligotrophic conditions (Smit *et al*., 2001; Fierer *et al*., 2007; Castro *et al*., 2010). It is possible that the slow growth rate of certain lichens may lead to more oligotrophic conditions and therefore the dominance of Acidobacteria in these environments.

Taxonomic profiles of Bacteria in lichen microbiomes

Within the Alphaproteobacteria, Rhizobiales generally dominate lichen microbiomes, but in chlorolichens Rhodospirillales (mostly unclassifiable Acetobacteraceae) is often co-dominant, while in cyanolichens Sphingomonadales (mostly *Sphingomonas*) typically co-dominate (Fig. 3), accounting for some degree of the community differences seen between lichens with different photoautotrophic symbiont types. The predominance of Rhizobiales bacteria across a broad array of lichen types is demonstrated for the first time in this study. Previous authors concluded that members of the Acetobacteraceae were most abundant (Cardinale *et al*., 2008), with *Sphingomonas* being present at lower levels (Grube *et al*., 2009). However, while all PCR-based methods have their biases, the results obtained previously are likely to be heavily biased, as the sequences were cloned after nested PCR. Aside from the predominance of Rhizobiales bacteria, there is also a great deal of diversity within this order, with seven 'families' (including LAR1) each accounting for $>1\%$ of the diversity in the full 454 sequence data set (Fig. 3). Of the two most abundant Rhizobiales families, Brucellaceae and LAR1, the latter contains a great deal of diversity within lichen microbiomes (Fig. 4). Each of the non-Rhizobiales families having $>1\%$ of the overall diversity in the sample set belongs to its own order, indicating that there is low diversity within each of the other orders. The four genera seen as the most ubiquitous by Grube and colleagues (2009) from culture-based analyses were found rarely, if at all. These data indicate that previous research into the physiological potential of lichen-associated bacteria using culture-based methods is likely to have little, if any, bearing on the physiology of lichen thalli themselves.

The lichen microbiome as a model system for environmental microbiology

Lichen-associated bacterial communities are rich and diverse (Fig. 3), but are simpler in composition than many soil communities (Bates *et al*., 2010). Our data suggest that dispersal of lichen-associated bacteria may be limited to a greater degree than in typical soil communities (Chu *et al*., 2010), making lichens optimal systems in which to study long-term microbial community changes (Mushegian *et al*., 2011). Lichens are extremely slow-growing, often persisting for decades or even centuries in a single location (Lawrey, 2009), and may serve as model systems for extended monitoring to examine microbial succession or change due to factors manipulated through field experiments. Phylogenetic data also support the notion that lichen thalli may act as specialized niches for at least one specific lineage of bacteria from the order Rhizobiales

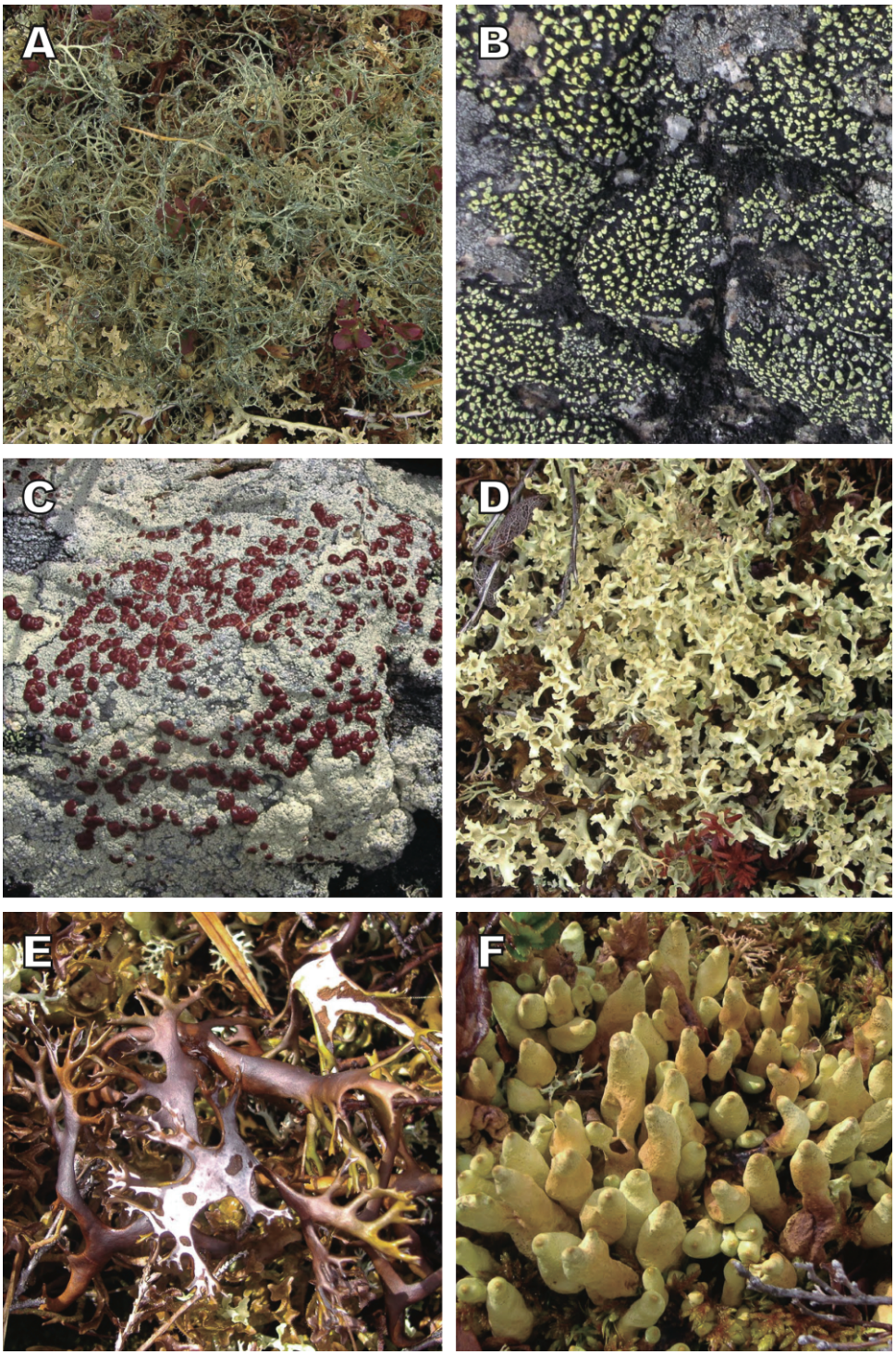

Fig. 5. Photos of selected lichens analysed as part of this study. The mycobiont species shown are (A) *Alectoria ochroleuca* (B) *Rhizocarpon geographicum* (C) *Ophioparma ventosa* (D) *Flavocetraria cucullata* (E) *Masonhalea richardsonii* and (F) *Dactylina arctica*, all collected at the Eagle Summit site near Fairbanks, Alaska. 151 × 227 mm (300 × 300 DPI).

(LAR1, a group that seems to have diversified widely within lichens; Fig. 4), suggesting the possibility of a long co-evolutionary history between lichens and some of their non-photoautotrophic bacterial associates. Interestingly, the lichen-associated clades examined through cloning in this study (LAR1 and *Methylobacterium*) seemingly have an overall rate of evolution that is high in comparison with other closely related families as shown by the overall

branch length differences in Figs 2 and S1. This accelerated rate of evolution could be driven by their adaptation to living within the lichen microbiome and potentially a transition to a mutualistic state (Moran, 1996; Lutzoni and Pagel, 1997; Wernegreen and Moran, 1999). Another possible reason for this pattern could be an accelerated rate of gene exchange, leading to hybrid sequences that are inherited, giving rise to whole new hybrid lineages; in this scenario, gene exchange may be facilitated through a mechanism like quorum-sensing within lichen thalli (Schaefer *et al.*, 2002). When all of these factors are taken together, it becomes clear that the lichen, as a microbiome, has the potential to provide a great number of further insights into the ecology and evolution of environmental microbes.

Experimental procedures

Sampling procedure

Lichens were collected at four sites (see Table 2) in three clusters. A cluster was defined as an area of approximately 10 m in diameter, and each cluster was approximately 100 m from the previous one. Lichens representing 10 different mycobiont taxa (see examples in Fig. 5) were collected at each cluster. Within each site, the same lichen-forming genera were collected at each of the three clusters. Among sites, located in different biomes, the 10 lichen-forming genera were often different. Mycobionts and photobionts were identified based on morphology and chemical spot tests (Brodo *et al.*, 2001). From each collection, three subsamples were taken from the leading edge of the lichen thallus for DNA extraction using sterilized forceps, for a total of 360 subsamples, when including all four sites. Subsamples were stored at −20°C before DNA extraction and all collections are permanently accessioned in the Duke Herbarium (DUKE).

'SQERL' DNA extraction

A Simultaneous QuickExtract/ReadyLyse (SQERL) DNA extraction protocol was developed for the high-throughput processing of samples. First, small fragments of lichen thalli (~1 mg each) were taken for extraction and placed into standard 300 μl PCR tubes/wells. Each well was filled with 100 μl of QuickExtract Plant extraction solution (EPICENTRE Biotechnologies, Madison, WI, USA). A 1/100 mixture of ReadyLyse Lysozyme (EPICENTRE Biotechnologies) and TES Buffer (10 mM Tris-HCl (pH = 7.5), 1 mM EDTA, 100 mM NaCl) was made to produce a bacterial lysis buffer. Two microlitres of freshly mixed lysis buffer were added to each well (creating a fresh QuickExtract/ReadyLyse solution), and tubes were placed on the thermocycler. Extraction conditions consisted of 20 min at 37°C, 6 min at 65°C and 2 min at 98°C, followed by a cool-down at 4°C. 1/10 dilutions of DNA were made for downstream applications; greater dilutions were made for samples with especially strong pigments in order to dilute down PCR inhibitors.

16S clone library construction

Polymerase chain reaction was performed and clone libraries were constructed individually from each DNA extract as described by Hodkinson and Lutzoni (2009) using the primer

Table 2. Summary of sample types collected for this study.

Site	Mycobiont	Photobiont	Identifier
C	*Peltigera*	Cyanobacteria	CL01*
	Alectoria	Green Algae	CL02
	Cladonia	Green Algae	CL03*
	Placopsis	Tripartite	CL04
	Erioderma	Cyanobacteria	CL05
	Sticta	Cyanobacteria	CL06*
	Dictyonema	Cyanobacteria	CL07*
	Leptogium	Cyanobacteria	CL08*
	Stereocaulon	Tripartite	CL09
	Usnea	Green Algae	CL10*
E	*Peltigera*	Cyano/Tri[1]	EL01*
	Alectoria	Green Algae	EL02
	Cladonia	Green Algae	EL03*
	Rhizocarpon	Green Algae	EL04
	Ophioparma	Green Algae	EL05*
	Flavocetraria	Green Algae	EL06*
	Arctoparmelia	Green Algae	EL07
	Umbilicaria	Green Algae	EL08*
	Masonhalea	Green Algae	EL09
	Dactylina	Green Algae	EL10
H	*Peltigera*	Cyanobacteria	HL01*
	Platismatia	Green Algae	HL02
	Cladonia	Green Algae	HL03*
	Parmotrema	Green Algae	HL04
	Flavoparmelia	Green Algae	HL05
	Sticta	Cyanobacteria	HL06*
	Xanthoparmelia	Green Algae	HL07
	Leptogium	Cyanobacteria	HL08*
	Stereocaulon	Tripartite	HL09
	Usnea	Green Algae	HL10*
N	*Peltigera*	Cyano/Tri[1]	NL01*
	Alectoria	Green Algae	NL02
	Cladonia	Green Algae	NL03*
	Amygdalaria	Tripartite	NL04
	Ophioparma	Green Algae	NL05*
	Flavocetraria	Green Algae	NL06
	Parmelia	Green Algae	NL07*
	Umbilicaria	Green Algae	NL08*
	Stereocaulon	Tripartite	NL09
	Cetraria	Green Algae	NL10

Collection sites (C = Cerro de la Muerte, Costa Rica; E = Eagle Summit, near Fairbanks, Alaska; H = Highlands, North Carolina; N = Nome, Alaska), mycobiont-types (generic name of the dominant fungus), photobiont-types (Cyanobacteria vs. Green Algae[1]), and identifiers (used as the first four characters of the 10-digit alphanumeric clone sequence names) are listed for each sample type. Asterisks (*) indicate sample groups used for barcoded 454 amplicon sequencing.

[1]Note: 'Cyano/Tri' indicates that some samples had Cyanobacteria while some were tripartite, with both Cyanobacteria and green algae. Cloned sequences derived from each type of sample are listed in the UniFrac ID mapping file, available in the accompanying Dryad data package (clones from tripartite *Peltigera* lichens are identified as 'EL01t' or 'NL01t' therein); 454 data were generated only from samples with a single photobiont-type.

pair 533F/1185aR (the latter being specific to certain lichen-associated Rhizobiales bacteria in the LAR1 lineage and the genus *Methylobacterium*). Sequences were identified using 10-digit alpha-numeric codes as follows: site identifier (C, E, H or N; see Table 2 caption for site specifics), sample type (L for 'lichen'), two-digit site-specific mycobiont-type identifier (01–10; see Table 2 for mycobiont-types correlated with each of the numbers at each site), cluster within the site (A, B or C), two-digit sample-specific subsample number (01–03), and a subsample-specific 3-digit alpha-numeric clone identifier (e.g. c35). Corrected sequences were grouped based on 97% similarity using Sequencher 4.8 (Gene Codes, Ann Arbor, MI, USA) and representatives of each group were checked using the RDP Classifier tool (Release 10.22; http://rdp.cme.msu.edu/classifier/; Cole *et al*., 2009); sequences that were not classified in the order Rhizobiales were excluded from further analyses along with their 97% similar associates. All sequences were hand-aligned to the published alignment produced by Hodkinson and Lutzoni (2009; available at http://lutzonilab.net/) using both Sequencher 4.8 and MacClade 4.08 (Maddison and Maddison, 2005) (see Dryad data package for PERL scripts, i.e. 'Clon_16S_fasta_renamer.pl' and 'Insert_488_spaces_in_front.pl', and associated instructions for facilitated sequence processing). For the purpose of chimera checking, the full alignment for the present study was split into two aligned FASTA files (one with the taxa from the original Hodkinson and Lutzoni, (2009) alignment and one with the newly generated cloned sequences) and processed using the chimera.slayer function in Mothur 1.15 (http://www.mothur.org/; Schloss *et al*., 2009) with the sequences from Hodkinson and Lutzoni (2009) used as a template for checking the newly generated cloned sequences. All potentially chimeric sequences were excluded from further downstream analyses.

Clone library analyses

Nearest BLAST hits of the chimera-checked cloned sequences, obtained by running a PERL script (Porter *et al*., 2008) that was written using the BioPerl toolkit (Stajich *et al*., 2002; http://bio.perl.org/), were integrated and aligned by hand using MacClade 4.08. Ambiguously aligned regions were defined manually and excluded following Lutzoni and colleagues (2000). A phylogeny was inferred for integration into ecological analyses using the constraint tree published by Hodkinson and Lutzoni (2009) as a backbone for a neighbour-joining analysis in PAUP* 4.0b10 (Swofford, 2002) employing J-C distances (Jukes and Cantor, 1969). Phylogenetically based ecological analyses were run using UniFrac (Lozupone and Knight, 2005), as implemented on the Fast UniFrac web server (Hamady *et al*., 2009), with the tree constructed using PAUP* and manually generated environment and category mapping files. Pairwise distance matrices produced by Fast UniFrac were downloaded for further analyses. An aligned fasta file of sequences cloned as part of this study, with ambiguously aligned regions excluded, was used to produce matrices with Mothur 1.15 showing the relative abundance of each operational taxonomic unit (OTU) in each sample, using 99% sequence similarity for OTU definitions (171 OTUs across all samples). The OTU data matrices were transformed into Bray-Curtis dissimilarity matrices (Bray and Curtis, 1957) using VEGDIST in the Vegan package (Oksanen *et al*., 2009) in R (R Development Core Team, 2010). ANOSIM analyses were run and NMDS plots were made with (i) unweighted UniFrac (ii) weighted + normalized UniFrac and (iii) Bray-Curtis matrices in R using the Vegan and EcoDist (Goslee and Urban, 2007) packages (all R input files and R-scripts for running all analyses are available under the 'R_scripts' directory in the Dryad data package).

In order to infer deeper relationships between the cloned sequences generated, a maximum likelihood phylogenetic tree was constructed. RAxML-HPC-MPI-SSE3 7.2.8α (Stamatakis, 2006) was used to remove duplicate sequences, perform topology searches ($\times$ 1000), and run bootstrap analyses ($\times$ 1000) employing a GTRGAMMAI model across 200 processors on the Duke Shared Cluster Resource (DSCR). For the topology search, the aforementioned Hodkinson and Lutzoni (2009) backbone constraint tree was used, excluding *Methylobacterium aminovorans* and *Methylobacterium extorquens*, because their sequences were identical to the sequence from *Methylobacterium podarium* strain FM3 at all sites included in the analysis. RAxML-HPC 7.0.4 was used to map unconstrained bootstrap proportions to the final constrained topology. The ML tree was visualized with all nodes that did not contain sequences from cultured reference strains collapsed (Fig. 2). A shell script was written and run in order to reinsert the identical sequences that were previously removed from the ML analysis (see 'search_replace_identicals.sh' and associated instructions in the *Supporting information* for scripting details) and was visualized as a fully expanded tree with sequence identifiers colour-coded based on the photobiont-type of the sample of origin (Fig. S1). LAR1 and *Methylobacterium* sequences cloned as part of this study were extracted from the final RAxML topology by isolating the individual clades in MacClade 4.08 and running a custom PERL script ('capture_cloned_seqs.pl'; available in the Dryad data package) to pull out the names matching the 10-digit alpha-numeric code devised for this study; the percentages of LAR1 and *Methylobacterium* sequences in the total cloned sequence data set were then calculated.

454 sequencing of 16S amplicons

DNA extracts derived from samples with the same photobiont, mycobiont and site were pooled for use in 16S PCR amplification reactions. Twenty-one DNA sample pools were selected for 454 16S amplicon sequencing (Table 2). PCR chemistry followed Hodkinson and Lutzoni (2009) and reactions were performed in triplicate. Forward primers consisted of the general 16S primer 533F, along with the LibL 'A' adaptor at the 5′ end and one of the Roche 10-base MID barcodes intervening. Each reaction also contained an equimolar mix of the four constituent primers of the degenerate 1185mR (a primer designed to target a diversity of bacterial lineages while largely excluding Cyanobacteria and photobiont plastids; Hodkinson and Lutzoni, 2009) with LibL 'B' adaptors attached at the 5′ end. Thermocycler settings followed Fierer and colleagues (2008). Identical reactions were performed in triplicate and subsequently pooled. For each pool of products, small fragment removal was

accomplished according to Roche protocols using AMPure XP beads (Beckman Coulter Genomics). Amplicon pools from individual samples were then run separately on a DNA 1000 chip (Agilent Tech.). Based on the DNA concentration, as determined by using a Qubit fluorometer (Invitrogen), the samples were diluted and pooled so that the final pool with all samples would have approximately equivalent amounts of DNA from each of the individual samples. This final pool was run on a High Sensitivity DNA Chip (Agilent Tech.). The samples were then run on two quarter 454 GS-FLX plates with Titanium series reagents (Roche), sequencing from the 'A' adaptor only.

454 sequence processing and community analyses

Fasta and quality files were generated using GS Run Processor 2.5 with default settings (Roche). Subsequently, Mothur 1.15 was run for sequence processing and analyses using Silva reference files (available at http://www.mothur.org/wiki/Silva_reference_files), modified to include LAR1, as outlined in the Dryad data package ('SilvaRefFileMod.doc'). A Mothur batch file was written to perform the following functions on the 454 data set: (i) initially process sequence data (including trimming low-quality sequence data and properly combining data from multiple plate sections in which the same MID barcodes represent different samples); (ii) identify and exclude non-16S/non-bacterial sequences; (iii) identify and exclude cyanobacterial/plastid sequences; (iv) identify and exclude chimeras; (v) identify and exclude singletons (sequences that form single-sequence clusters at 97% identity using furthest-neighbour clustering); (vi) save a fasta file of cleaned-up sequence data with a corresponding group file indicating the sample of origin for each sequence; (vii) cluster sequences into OTUs; (viii) create an OTU-by-sample matrix showing the relative abundance of each OTU in each sample (using 97% sequence similarity for OTU definitions), and (ix) classify each sequence and summarize the taxonomic composition of the bacterial community found in each sample. For the details of each step, see the Mothur batch file used for overall sequence processing (available in the Dryad data package; 'HodkinsonLichen16S.batch'). The OTU-by-sample matrix (2112 OTUs across 21 samples) was transformed into a Bray-Curtis dissimilarity matrix and OTU-based comparative community analyses (ANOSIM and NMDS) were run in R. The four most abundant families of Bacteria across the lichens sampled here (forming the top two-percentile of families overall; Acidobacteriaceae, Acetobacteraceae, Brucellaceae and LAR1) were further examined by calculating a Simpson's diversity index and the percentage of redundant sequences for each family. A second Mothur batch file was written to perform these analyses (using a modified version of Mothur 1.16) and is available in the Dryad data package ('HodkinsonLichen16S_top4fam.batch').

The '.fasta' file containing all sequences in the final cleaned-up sequence data set was edited by adding the name of the sample of origin to the front of each sequence name (using the Mothur-generated '.groups' file for the same sequence set in a procedure described by the 'Fast_UniFrac_Initialization_Instructions.doc' file in the Dryad data package), the sequence set was subjected to BLASTn (against the GreenGenes core set), and the 'create_unifrac_env_file_BLAST.py' python script was run, in accordance with the Fast UniFrac tutorial (http://bmf2.colorado.edu/fastunifrac/tutorial.psp). The python output was used (along with a manually generated category mapping file) to run Fast UniFrac and the pairwise distance matrices were downloaded for further analyses. ANOSIM and NMDS were run on the UniFrac distance matrices in R.

Sequence data archiving

Cloned sequences were deposited in GenBank under accession numbers JF813830-JF815019 (GI: 332648420–332649609; PopSet: 332648420). The cloned sequence data set, the full set of sequence reads generated by 454 sequencing and a cleaned-up 454 data set comprised only of reads used in the final analyses presented here were all permanently deposited in the Dryad data repository (http://datadryad.org/; http://dx.doi.org/10.5061/dryad.t99b1/1 doi: 10.5061/dryad.t99b1/1).

Acknowledgments

We would like to thank Lisa Bukovnik for technical assistance with DNA sequencing, Rytas Vilgalys, Jonathan Olson, Amy Schmid, Bill Morris and Betsy Arnold for valuable guidance, and Bernie Ball, Ester Gaya, Jolanta Miadlikowska, Kathryn Picard, Tami McDonald, Daniele Armaleo, Terri Porter, Greg Bonito, Jason Jackson, Anthony Amend, Alexis Stamatakis, Tom Gihring, Scott Bates, Greg Wray, Jana U'Ren, Kacper Skakuj and Molly McMullen, all of whom have provided important research assistance. Sarah Hodkinson is thanked for support, field assistance and resource retrieval. This work was supported in part by a Doctoral Dissertation Improvement Grant (DEB-1011504) from the National Science Foundation, a James B. Duke Fellowship for Graduate Study, Duke Biology Grants-in-Aid from the Keever Endowment, a Duke Graduate International Research Travel Award, Explorers Club Diversa Awards, a Christiane and Christopher Tyson OTS Research Fellowship, a Lewis and Clark Fund for Exploration and Field Research Grant, a Sally Hughes-Schrader Travel Grant from the Duke Chapter of Sigma Xi, a Sigma Xi Grant-In-Aid of Research, and a Mycological Society of America Graduate Fellowship to BPH. Travel to Alaska was financed in part by a National Science Foundation award (DEB-0640956) to FL. This project was also funded in part by a subcontract (112442) to Daniele Armaleo, Fred Dietrich and FL as part of the Pacific Northwest National Laboratory (PNNL) foundational scientific focus area (FSFA) under DOE-BER's genomic sciences program in collaboration with Scott Baker and Jon Magnuson. The participation of CWS and NG was similarly supported by the DOE-BER scientific focus area on Plant-Microbe Interfaces (http://pmi.ornl.gov) at Oak Ridge National Laboratory (ORNL). This work would not have been possible without the Duke Shared Cluster Resource (DSCR) and outstanding services provided by John Pormann and Tom Milledge.

References

Adams, D.G., Bergman, B., Nierzwicki-Bauer, S.A., Rai, A.N., and Schüßler, A. (2006) Cyanobacterial–plant symbioses. In *The Prokaryotes: A Handbook on the Biology of Bacteria*. 3rd edn. Vol. 1: Symbiotic associations, biotechnology, applied microbiology. Dworkin, M., Falkow, S., Rosenberg, E., Schleifer, K.H., and Stackebrandt, E. (eds). New York, USA: Springer, pp. 331–363.

Arnold, A.E., Miadlikowska, J., Higgins, K.L., Sarvate, S.D., Gugger, P., Way, A., *et al*. (2009) A phylogenetic estimation of trophic transition networks for ascomycetous fungi: are lichens cradles of symbiotrophic fungal diversification? *Syst Biol* **58:** 283–297.

Artursson, V., Finlay, R.D., and Jansson, J.K. (2006) Interactions between arbuscular mycorrhizal fungi and bacteria and their potential for stimulating plant growth. *Environ Microbiol* **8:** 1–10.

Barns, S.M., Cain, E.C., Sommerville, L., and Kuske, C.R. (2007) Acidobactetia phylum sequences in uranium-contaminated subsurface sediments greatly expand the known diversity within the phylum. *Appl Environ Microbiol* **73:** 3113–3116.

Bates, S.T., Cropsey, G.W.G., Caporaso, J.G., Walters, W.A., Knight, R., and Fierer, N. (2010) Bacterial communities of lichens: structure and potential functions. In *Botany 2010*. Stogran, J. (ed.). St. Louis, Missouri: Botanical Society of America, p. 1103.

Bates, S.T., Cropsey, G.W.G., Caporaso, J.G., Knight, R., and Fierer, N. (2011) Bacterial communities associated with the lichen symbiosis. *Appl Environ Microbiol* **77:** 1309–1314.

Belnap, J. (2001) Factors influencing nitrogen fixation and nitrogen release in biological soil crusts. In *Biological Soil Crusts: Structure, Function, and Management*. Belnap, J., and Lange, O.L. (eds). New York, USA: Springer-Verlag, pp. 241–261.

Bray, J.R., and Curtis, J.T. (1957) An ordination of the upland forest communities of southern Wisconsin. *Ecol. Monogr.* **27:** 325–349.

Brodo, I.M., Sharnoff, S.D., and Sharnoff, S. (2001) *Lichens of North America*. New Haven, CT: Yale University Press.

Brosius, J., Palmer, J.L., Kennedy, J.P., and Noller, H.F. (1978) Complete nucleotide sequence of a 16S ribosomal RNA gene from *Escherichia coli*. *Proc Natl Acad Sci USA* **75:** 4801–4805.

Cardinale, M., Puglia, A.M., and Grube, M. (2006) Molecular analysis of lichen-associated bacterial communities. *FEMS Microbiol Ecology* **57:** 484–495.

Cardinale, M., Müller, H., Berg, G., de Castro J., and Grube, M. (2008) In situ analysis of the bacterial community associated with the reindeer lichen *Cladonia arbuscula* reveals predominance of Alphaproteobacteria. *FEMS Microbiol Ecology* **66:** 63–71.

Castro, H.F., Classen, A.T., Austin, E.E., Norby, R.J., and Schadt, C.W. (2010) Soil microbial community responses to multiple experimental climate change drivers. *Appl Environ Microbiol* **76:** 999–1007.

Chu, H., Fierer, N., Lauber, C.L., Caporaso, J.G., Knight, R., and Grogan, P. (2010) Soil bacterial diversity in the Arctic is not fundamentally different from that found in other biomes. *Environ Microbiol* **12:** 2998–3006.

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

Culberson, C.F., Culberson, W.L., and Johnson, A. (1984) Secondary compounds produced exclusively by lichens. In *CRC Handbook of Microbiology* 2nd edn, Vol. 5. Laskin, A.I., and Lechevalier, H.A. (eds). Boca Raton, FL: CRC Press, pp. 793–833.

Fierer, N., Bradford, M.A., and Jackson, R.B. (2007) Toward an ecological classification of soil bacteria. *Ecology* **88:** 1354–1364.

Fierer, N., Hamady, M., Lauber, C.L., and Knight, R. (2008) The influence of sex, handedness, and washing on the diversity of hand surface bacteria. *Proc Natl Acad Sci USA* **105:** 17994–17999.

Founoune, H., Duponnois, R., Bâ, A.M., Sall, S., Branget, I., Lorquin, J., *et al*. (2002) Mycorrhiza helper bacteria stimulate ectomycorrhizal symbiosis of Acacia holosericea with *Pisolithus alba*. *New Phytol* **153:** 81–89.

González, I., Ayuso-Sacido, A., Anderson, A., and Genilloud, O. (2005) Actinomycetes isolated from lichens: evaluation of their diversity and detection of biosynthetic gene sequences. *FEMS Microbiol Ecology* **54:** 401–415.

Goslee, S., and Urban, D. (2007) The ecodist package for dissimilarity-based analysis of ecological data. *J Stat Softw* **22:** 1–19.

Grube, M., and Berg, G. (2009) Microbial consortia of bacteria and fungi with focus on the lichen symbiosis. *Fungal Biol Rev* **23:** 72–85.

Grube, M., Cardinale, M., de Castro, J.V., Jr, Muller, H., and Berg, G. (2009) Species-specific structural and functional diversity of bacterial communities in lichen symbioses. *ISME J* **3:** 1105–1115.

Hamady, M., Lozupone, C., and Knight, R. (2009) Fast UniFrac: facilitating high-throughput phylogenetic analyses of microbial communities including analysis of pyrosequencing and PhyloChip data. *ISME J* **4:** 17–27.

Hodkinson, B.P., and Lutzoni, F. (2009) A microbiotic survey of lichen-associated bacteria reveals a new lineage from the Rhizobiales. *Symbiosis* **49:** 163–180.

Hodkinson, B.P., and Lutzoni, F. (2010) Do lichens harbor their own rhizobia? A large-scale phylogenetic survey of lichen-associated bacteria from the order Rhizobiales. *Inoculum* **61:** 55–56.

Hoffman, M.T., and Arnold, A.E. (2010) Diverse bacteria inhabit living hyphae of phylogenetically diverse fungal endophytes. *Appl Environ Microbiol* **76:** 4063–4075.

Humbert, J.F., Dorigo, U., Cecchi, P., Berre, B.L., Debroas, D., and Bouvy, M. (2009) Comparison of the structure and composition of bacterial communities from temperate and tropical freshwater ecosystems. *Environ Microbiol* **11:** 2339–2350.

Jukes, T.H., and Cantor, C.R. (1969) Evolution of protein molecules. In *Mammalian Protein Metabolism*. Munro, H.N. (ed.). New York, USA: Academic Press, pp. 21–132.

Kuczynski, J., Liu, Z., Lozupone, C., McDonald, D., Fierer, N., and Knight, R. (2010) Microbial community resemblance

methods differ in their ability to detect biologically relevant patterns. *Nat Methods* **7:** 813–819.

Lawrey, J.D. (2009) Chemical defense in lichen symbiosis. In *Defensive Mutualism in Microbial Symbiosis*. White, J.F., and Torres, M.S. (eds). Boca Raton, FL: CRC Press, pp. 167–181.

Liba, C., Ferrara, F., Manfio, G., Fantinatti-Garboggini, F., Albuquerque, R., Pavan, C., *et al*. (2006) Nitrogen-fixing chemo- organotrophic bacteria isolated from cyanobacteria-deprived lichens and their ability to solubilize phosphate and to release amino acids and phytohormones. *J Appl Microbiol* **101:** 1076–1086.

Lozupone, C., and Knight, R. (2005) UniFrac: a new phylogenetic method for comparing microbial communities. *Appl Environ Microbiol* **71:** 8228–8235.

Lutzoni, F., and Pagel, M. (1997) Accelerated evolution as a consequence of transitions to mutualism. *Proc Natl Acad Sci USA* **94:** 11422–11427.

Lutzoni, F., Wagner, P., Reeb, V., and Zoller, S. (2000) Integrating ambiguously aligned regions of DNA sequences in phylogenetic analyses without violating positional homology. *Syst Biol* **49:** 628–651.

Maddison, D., and Maddison, W. (2005) *MacClade 4.08*. Sunderland, MA: Sinauer Associates.

Moran, N.A. (1996) Accelerated evolution and Muller's ratchet in endosymbiotic bacteria. *Proc Natl Acad Sci USA* **93:** 2873–2878.

Mushegian, A.A., Peterson, C.N., Baker, C.C.M., and Pringle, A. (2011) Bacterial diversity across individual lichens. *Appl Environ Microbiol* **77:** 4249–4252.

Najar-Rodríguez, A.J., McGraw, E.A., Mensah, R.K., Pittman, G.W., and Walter, G.H. (2009) The microbial flora of *Aphis gossypii*: patterns across host plants and geographical space. *J Invertebr Pathol* **100:** 123–126.

Oksanen, J., Kindt, R., Legendre, P., O'Hara, B., Simpson, G.L., and Stevens, M.H.H. (2009) vegan: Community Ecology Package. R package version 1.15-4. [WWW document]. URL http://cran.r-project.org/; http://vegan.r-forge.r-project.org/.

Palmqvist, K. (2002) Cyanolichens: Carbon Metabolism. In *Cyanobacteria in Symbiosis*. Rai, A.N., Berman, B., and Rasmussen, U. (eds). Dordrecht, The Netherlands: Kluwer Academic Publishers, pp. 73–96.

Porter, T.M., Skillman, J.E., and Moncalvo, J.M. (2008) Fruiting body and soil rDNA sampling detects complementary assemblage of Agaricomycotina (Basidiomycota, Fungi) in a hemlock-dominated forest plot in southern Ontario. *Mol Ecol* **17:** 3037–3050.

R Development Core Team (2010) *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Rai, A.N. (2002) Cyanolichens: Nitrogen Metabolism. In *Cyanobacteria in Symbiosis*. Rai, A.N., Berman, B., and Rasmussen, U. (eds). Dordrecht, The Netherlands: Kluwer Academic Publishers, pp. 97–115.

Riffel, A., Medina, L.F., Stefani, V., Santos, R.C., Bizani, D., and Brandelli, A. (2002) In vitro antimicrobial activity of a new series of 1,4 naphthoquinones. *Braz J Med Biol Res* **35:** 811–818.

Russell, J.A., Goldman-Huertas, B., Moreau, C.S., Baldo, L., Stahlhut, J.K., Werren, J.H., and Pierce, N.E. (2009) Specialization and geographic isolation among W*olbachia* symbionts from ants and lycaenid butterflies. *Evolution* **63:** 624–640.

Rycroft, D.S., Connolly, J.D., Huneck, S., and Himmelreich, U. (1996) Revised Structure of Haemoventosin. ChemInform 27: 10.1002/chin.199610298.

Schaefer, A.L., Taylor, T.A., Beatty, J.T., and Greenberg, E.P. (2002) Long-Chain Acyl-Homoserine Lactone Quorum-Sensing Regulation of R*hodobacter capsulatus* Gene Transfer Agent Production. *J Bacteriol* **184:** 6515–6521.

Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., *et al*. (2009) Introducing Mothur: open-source, platform- independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol* **75:** 7537–7541.

Smit, E., Leeflang, P., Gommans, S., van den Broek, J., van Mil, S., and Wernars, K. (2001) Diversity and seasonal fluctuations of the dominant members of the bacterial soil community in a wheat field as determined by cultivation and molecular methods. *Appl Environ Microbiol* **67:** 2284–2291.

Stajich, J.E., Block, D., Boulez, K., Brenner, S.E., Chervitz, S.A., Dagdigian, C., *et al*. (2002) The Bioperl toolkit: Perl modules for the life sciences. *Genome Res* **12:** 1611–1618.

Stamatakis, A. (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **22:** 2688–2690.

Swofford, D.L. (2002) *PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods), Version 4.0b10*. Sunderland, Mass: Sinauer Associates.

Taylor, M.W., Schupp, P.J., de Nys, R., Kjelleberg, S., and Steinberg, P.D. (2005) Biogeography of bacteria associated with the marine sponge C*ymbastela concentrica*. *Environ Microbiol* **7:** 419–433.

Taylor, M.W., Radax, R., Steger, D., and Wagner, M. (2007) Sponge-associated microorganisms: evolution, ecology, and biotechnological potential. *Microbiol Mol Biol Rev* **71:** 295–347.

Warnecke, F., Luginbuhl, P., Ivanova, N., Ghassemian, M., Richardson, T.H., Stege, J.T., *et al*. (2007) Metagenomic and functional analysis of hindgut microbiota of a wood-feeding higher termite. *Nature* **450:** 560–565.

Wernegreen, J.J., and Moran, N.A. (1999) Evidence for genetic drift in endosymbionts (B*uchnera)*: analyses of protein-coding genes. *Mol Biol Evol* **16:** 83–97.

Supporting information

Additional Supporting Information may be found in the online version of this article:

Fig. S1. Maximum likelihood (ML) phylogeny of cloned sequences generated as part of this study (10-digit alpha-numeric codes) with closest BLAST hits (8-digit accession numbers), sequences from cultured reference strains [organism names; GenInfo Identifiers are listed by Hodkinson and Lutzoni (2009)], and lichen-associated cloned sequences generated by Hodkinson and Lutzoni (2009; 6- or 7-digit alpha-numeric codes). Numbers above branches represent ML bootstrap proportions (BP) $>50\%$. Identical

sequences that were removed pre-analysis were reinserted using a script described in the Dryad data package. Sequence identifiers in light green are derived from chlorolichens, while those in dark blue are derived from cyanolichens (identifiers for those derived from tripartite lichens remain in black text). The 10-digit alpha-numeric codes used to identify sequences generated as part of this study encode detailed information about the origin of each sequence; this information is conveyed in the following order: site identifier [C (Cerro de la Muerte, Costa Rica), E (Eagle Summit, Alaska), H (Highlands, North Carolina) or N (Nome, Alaska)], sample type [L (lichen)], two-digit site-specific mycobiont-type identifier (01–10; see Table 2 for mycobiont-types correlated with each of the numbers at each site), cluster within the site (A, B or C), two-digit sample-specific subsample number (01–03), and a subsample-specific 3-digit alpha-numeric clone identifier (e.g. c35). The LAR1 lineage accounts for 87% (1034/1190) of all cloned sequences generated as part of this study, while sequences from bacteria scattered throughout the genus *Methylobacterium* account for the other 13% (156/1190).

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.