MOLECULAR ECOLOGY

Molecular Ecology (2012) doi: 10.1111/j.1365-294X.2012.05496.x

Environment or kin: whence do bees obtain acidophilic bacteria?

QUINN S. McFREDERICK,* WILLIAM T. WCISLO,† DOUGLAS R. TAYLOR,‡ HEATHER D. ISHAK,* SCOT E. DOWD§ and ULRICH G. MUELLER*

*Section of Integrative Biology, University of Texas at Austin, 2401 Speedway Drive #C0930. Austin, TX 78712, USA, †Smithsonian Tropical Research Institute, Box 0843-03092, Balboa, Ancon, Republic of Panama, ‡Department of Biology, University of Virginia, Charlottesville, VA 22904-4328, USA, §MR DNA (Molecular Research), 503 Clovis Rd., Shallowater, TX 79363, USA

Abstract

As honey bee populations decline, interest in pathogenic and mutualistic relationships between bees and microorganisms has increased. Honey bees and bumble bees appear to have a simple intestinal bacterial fauna that includes acidophilic bacteria. Here, we explore the hypothesis that sweat bees can acquire acidophilic bacteria from the environment. To quantify bacterial communities associated with two species of North American and one species of Neotropical sweat bees, we conducted 16S rDNA amplicon 454 pyrosequencing of bacteria associated with the bees, their brood cells and their nests. *Lactobacillus* spp. were the most abundant bacteria in many, but not all, of the samples. To determine whether bee-associated lactobacilli can also be found in the environment, we reconstructed the phylogenetic relationships of the genus *Lactobacillus*. Previously described groups that associate with *Bombus* and *Apis* appeared relatively specific to these genera. Close relatives of several bacteria that have been isolated from flowers, however, were isolated from bees. Additionally, all three sweat bee species associated with lactobacilli related to flower-associated lactobacilli. These data suggest that there may be at least two different means by which bees acquire putative probiotics. Some lactobacilli appear specific to corbiculate apids, possibly because they are largely maternally inherited (vertically transmitted). Other lactobacilli, however, may be regularly acquired from environmental sources such as flowers. Sweat bee–associated lactobacilli were found to be abundant in the pollen and frass inside the nests of halictids, suggesting that they could play a role in suppressing the growth of moulds and other spoilage organisms.

Keywords: 16S rDNA amplicon 454 pyrosequencing, acidophilic bacteria, insect diseases, insect symbiosis, *Lactobacillus*, phylogeny

Received 8 July 2011; revision received 6 January 2012; accepted 13 January 2012

Introduction

As *Apis mellifera* populations continue to decline, interest in the role of microbes in the health of pollinators continues to increase. For example, several recent studies have highlighted the possible roles of pathogens in pollinator declines (Cox-Foster *et al.* 2007; Johnson *et al.* 2009; Bromenshenk *et al.* 2010; Cameron *et al.* 2011). On the opposite end of the spectrum, there is accumulating evidence for a worldwide association between *Ap. mellifera* and a core set of bacterial phylotypes (Jeyaprakash *et al.* 2003; Mohr & Tebbe 2006, 2007; Babendreier *et al.* 2007; Cox-Foster *et al.* 2007; Martinson *et al.* 2011). Similar associations have been reported between other corbiculate apids and bacteria, especially *Bombus* spp. (Mohr & Tebbe 2006; Koch & Schmid-Hempel 2011a; Martinson *et al.* 2011).

Because specific bacteria are consistently associated with *Ap. mellifera*, it has been suggested that these bacteria

Correspondence: Quinn S. McFrederick, Fax: +1 (512) 471-3878; E-mail: quinnmcfrederick@gmail.com

are beneficial mutualists (Martinson *et al.* 2011). Experimental evidence of the benefits of core *Apis* bacteria on bee health, however, is lacking, with the exception of the lactic acid bacteria (LAB). A mixture of *Lactobacillus* and *Bifidobacterium* phylotypes has been shown to decrease the infection rate of *Ap. mellifera* larvae that were exposed to the pathogen *Paenibacillus larvae* (Forsgren *et al.* 2010). In addition, LAB have been isolated from the crop of *Ap. mellifera*, freshly collected pollen and freshly fermented bee bread, which is the protein source for larval *Ap. mellifera* (Vasquez & Olofsson 2009). The fermentative properties of LAB have been hypothesized to aid in the conversion of pollen to bee bread and the protection of bee bread from spoilage (Vasquez & Olofsson 2009). It has also been hypothesized that *Ap. mellifera* add LAB to collected nectar and that LAB ferment nectar, possibly aiding in the conversion of nectar to honey (Olofsson & Vasquez 2008). Lastly, it has recently been shown that bacteria associated with bumble bees can benefit their hosts. The microbiota of *Bombus terrestris*, which resembles the *Ap. mellifera* microbiota, protects the bee hosts from *Crithidia bombi*, a trypanosomatid gut parasite (Koch & Schmid-Hempel 2011b).

There exists, therefore, some evidence that *Ap. mellifera* and *Bombus* spp. associate with beneficial bacteria. However, these bacterial phylotypes appear to be generally lacking from other bees. Martinson *et al.* (2011) surveyed lineages across the bee phylogeny and found the bacterial phylotypes associated with *Ap. mellifera* to be almost completely lacking from bees outside the corbiculate apids. Instead, bees outside the corbiculate apids exhibited no predictable associations, besides the presence of a common environmental bacterium, *Burkholderia cepacia*, in most specimens. Why the corbiculate apids are consistently associated with a small community of putative mutualists, whereas other bees are not, is an open question. Martinson *et al.* (2011) suggested the hypothesis that bacteria associated with corbiculate apids are transferred vertically between generations and that eusociality facilitates this transmission, either via trophallaxis (within nest oral-oral sharing of food) or via the swarming (nest fission) method of colony foundation found in *Ap. mellifera*.

The fact that a core set of bacterial phylotypes are almost exclusively associated with corbiculate apids, while consistent with the notion of vertical transmission of bacteria from mother to offspring, does not reject the hypothesis that beneficial bacteria can be acquired environmentally. Bacterial phylotypes other than the core set, for example, have been repeatedly detected in association with *Ap. mellifera*, but not consistently across all surveys. Several studies have detected *Lactobacillus kunkeei* in association with *Ap. mellifera* (Mohr & Tebbe 2007; Olofsson & Vasquez 2008; Vasquez & Olofsson 2009; Vasquez *et al.* 2009), whereas several others have failed to detect *L. kunkeei* phylotypes (Jeyaprakash *et al.* 2003; Babendreier *et al.* 2007; Cox-Foster *et al.* 2007; Martinson *et al.* 2011). Although differences in methodology across surveys may explain this variation, environmental acquisition of *L. kunkeei* is another possible explanation. Additionally, LAB community composition appears to vary seasonally. *Lactobacillus kunkeei* was found to dominate the LAB microbiota of *Ap. mellifera* honey stomachs in both Sweden and the United States, but LAB community membership and abundance was found to vary with foraging activity (Vasquez *et al.* 2009). This response of the *Ap. mellifera* LAB community to foraging may also be explained by environmental acquisition of a subset of the LAB microbiota. It is therefore plausible that some LAB associated with *Ap. mellifera* may be environmentally transmitted, while others are maternally inherited (vertically transmitted).

Here, we address the hypothesis that some bee-associated, acidophilic bacteria are environmentally acquired. First, we use a 16S rDNA amplicon survey to study the microbiota of three species of halictid bees as well as the bacterial communities associated with the contents of their nests. We use this information to detect the presence of putative mutualists and pathogens and to explore bacterial community composition and diversity. Next, we present a phylogeny of the lactobacilli, combining our 16S rDNA sequence data with publicly available sequence information. We find that three of the four lactobacilli known from flowers have also been isolated from bees. Additionally, *L. kunkeei* has been isolated from flowers and *Apis* spp., and is closely related to lactobacilli isolated from sweat bees. Together, these data suggest that the acquisition of putatively beneficial bacteria from the environment may be more important than previously thought.

Study organisms and methods

The structure of halictid nests makes them particularly interesting for studies of bacterial communities. Halictid bees build brood cells that they line with the secretion of the Dufour's gland, which creates a waterproof lining (Duffield *et al.* 1981). Adult females then provision the brood cell with pollen and nectar, lay an egg on top of the provisions and seal the cell with a plug (Michener 1974). This brood cell typically remains sealed until the brood ecloses into an adult, although some halictids open brood cells for inspection (Michener 1974). The newly eclosed adult bee breaks through the plug, and the cell is either abandoned or cleaned and reused. Cells can also include microbes, mites, nematodes and springtails; they can therefore be thought of as miniature, ephemeral ecosystems (sensu Biani *et al.* 2009).

Collections

We sampled adults, brood, pollen provisions, frass, tunnel-wall scrapings and the nearby nesting substrate from the nests of three species of sweat bees. For all collections, we collected the samples directly into molecular-grade ethanol, with forceps that we flame-sterilized between handling each sample. We carefully opened nests and brood cells to minimize contamination.

First, we sampled the entire contents of a wild *Augochlora pura* nest found in a decaying log at Blandy Experimental Farm in Boyce, Virginia (N 39°03′20.63″ W 78°04′22.48″) on 14 August 2009. *Au. pura* is a solitary halictid in the tribe Augochlorini (Michener 1974). We dissected the nest in the field. For insect samples, we screened the entire insect; therefore, these communities represent both intestinal and surface microbes. We sampled the adult mother of the nest along with three brood and their brood cell contents. Lastly, we haphazardly sampled several portions of the tunnel walls from inside the nest, decayed wood that we haphazardly collected from the rotten log that housed the nest (within ~0.3 m of the nest), and opened a collecting vial to be exposed to air as a negative control.

We also sampled the entire contents of a laboratory nest of *Halictus ligatus* that was kept in an environmental chamber at the University of Virginia. Although *H. ligatus* is an obligate eusocial bee, workers collected from flowers will initiate nests in a solitary manner. The brood from these nests are usually entirely male, as most *H. ligatus* workers are not mated (Packer 1986). The mother of this nest was a worker collected at Blandy Experimental Farm from a floral head of *Carduus nuttans* on 4 August 2009, who initiated nesting in the laboratory on 19 August 2009. We built the nest with autoclaved soil sandwiched between two glass plates, based on the design by Michener (1974). We dissected the nest in a laminar flow hood on 13 September 2009. We collected three pupae that ranged in age from 12 to 15 days. We also collected a pollen provision that the female had not oviposited on, and the remains of a failed cell that had become infested with an unidentified fungus. We again collected tunnel-wall scrapings, soil from the surrounding substrate and a negative control. The laboratory nest, although not exposed to soil bacteria, was exposed to bacteria found on flowers (mostly ruderal composites) that the adults foraged on for nectar and pollen and to *Helianthus annus* pollen that we collected from garden grown plants and offered to the foraging bees in small dishes.

Lastly, we sampled the contents of two *Megalopta genalis* nests from Barro Colorado Island, Panama. *M. genalis* is a socially polymorphic, Neotropical sweat bee that nests in decaying branches and sticks in the forest understory and exhibits trophallaxis (Wcislo *et al.* 2004; Wcislo & Gonzalez 2006). We collected a solitary nest (N 09°09′35.4″ W 079°50′21.2″) on 24 January 2011 and a social nest (one queen and one worker, N 09°09′40.7″ W 079°50′26.7″) on 3 February 2011. To collect samples, we carried the nests back to the laboratory and dissected them inside a plastic container sanitized with 10% bleach solution followed by 100% ethanol, kept on its side and with a alcohol bulb burning in front of the tub in order to minimize the entrance of airborne microbes into the samples. We were unable to amplify DNA from three samples from the social nest. Our samples, however, still spanned developmental stages ranging from the pollen provision from an egg to the frass from a pupa to an adult bee. We again collected samples from the tunnel walls of the nest and wood and scrapings from different parts of the stick that did not directly contact the nest, and exposed a vial of ethanol to the air as a negative control. We also measured the pH of the pollen provisions with pH paper (Hydrion, Micro Essential Laboratory) and pollen mixed 1:1 with nanopure water.

Bacterial tag-encoded FLX 454 pyrosequencing (bTEFAP) and bioinformatics

Research and Testing Laboratories (Lubbock TX) conducted DNA extraction, PCR and sequencing using previously described protocols (Dowd *et al.* 2008; Sen *et al.* 2009; Ishak *et al.* 2011). We used the 28F (5′GAGT TTGATCNTGGCTCAG) and 519R (5′GTNTTACNGCG GCKGCTG) primer pair, which amplifies the V1, V2 and V3 regions of the 16S rRNA gene. After sequencing, we removed low-quality sequence ends, tags and primers. We then discarded any chimeric sequences as detected by the program B2C2 (Gontcharova *et al.* 2010) and denoised the remaining sequences. To assign reads to phylotypes, we compared the sequences via BLAST against a curated database of 16S rDNA sequences that was compiled from the National Center for Biotechnology and maintained by the Medical Biofilm Research Institute. We assigned reads to the finest taxonomic level possible based on the following similarity criteria: 96% and above for species, 94–96% for genus, 89–94% for family, 85–89% for order, 80–85% for class and 77–80% for phylum (Dowd *et al.* 2008; Sen *et al.* 2009; Ishak *et al.* 2011). As an additional validation of our taxonomic assignments, we conducted a separate BLASTn search on all of our 454 sequences against the SILVA database release 108 (Pruesse *et al.* 2007). To visualize the results, we processed the resulting BLAST table with MEGAN4 (Huson *et al.* 2011). Finally, to investigate the patterns of diversity, we calculated the probability of an interspecific encounter (PIE), an

estimator of the evenness of a community (Hurlbert 1971).

Firmicutes-specific reference database for lactobacilli analyses

The Medical Biofilm Research Institute's 16S rDNA database does not fully encompass bacterial diversity for certain groups, and we found that many of the most abundant bacteria were assigned to unknown taxonomic levels within the phylum Firmicutes (e.g. Bacilli, Lactobacillales, Lactobacillaceae). To gain a more detailed understanding of where these samples fell within the Firmicutes, we downloaded all of the 16S rDNA sequences that were labelled as belonging to the phylum Firmicutes, were not an unidentified species and were publicly available from NCBI as of 3 October 2011 (60027 sequences). We clustered sequences that were of 97% or greater sequence identity using CDhit (Huang *et al.* 2010), and removed sequences that were shorter than 300 bases or longer than 2000 bases, which resulted in a database of 6600 Firmicutes sequences. We extracted reads from our 454 data set that we had assigned to an unknown taxonomic position in the Firmicutes with our previous BLAST search, and used BLASTn to compare these reads to the Firmicutes only database.

Phylogenetic analyses

We conducted two separate phylogeny-based analyses: a Fast UniFrac analysis (Hamady *et al.* 2010) on all of the 454 sequence data and a maximum-likelihood analysis on amassed *Lactobacillus* sequences. For the Fast UniFrac analysis, we used a pipeline previously described by Ishak *et al.* (2011) to analyse our 454 sequence data. First, we used a custom Perl script to select sequences that were 300 bases or longer. As the number of reads from each sample varied from 508 to 11 027, we used another custom Perl script to randomly sample 1000 reads from each sample. Six of our thirty-four samples had less than 1000 reads, so all of the reads greater than 300 bases long were used for those samples (Q15-900 reads, Q16-894 reads, Q2A-694 reads, GeSoL4Bc0019-564 reads, GeSoTuBc0019-508 reads, GeEuSuBc0029-995 reads). Next, to cluster redundant sequences, we used CDhit (Huang *et al.* 2010), with the sequence identity cut-off set to 97%, as has been suggested to avoid inflated diversity estimates caused by sequencing error (Kunin *et al.* 2010). In addition, we used a custom Perl script to generate a sample ID mapping file that assigns the number of reads from each sample to a specific CDhit cluster. We used Mothur (Schloss *et al.* 2009) to align the sequences, with the SILVA database as a reference. As the SILVA database alignment results in many gaps (Ishak *et al.* 2011), we used the filter.seqs function in Mothur to remove these gaps. We then used a soft filter to remove insertions that appeared in <10% of the samples. The resulting alignment consisted of 2138 sequences with a length of 431 bases. To create an approximate maximum-likelihood tree for the Fast UniFrac analysis, we used the program FastTree (Price *et al.* 2010) under a general time reversible (GTR) model of sequence evolution. Lastly, to assess similarity in our sampled bacterial communities, we used the resulting maximum-likelihood phylogeny, a sample category mapping file and the sample ID mapping file as input for the Fast UniFrac analysis. In Fast UniFrac, we performed a weighted, normalized principal coordinate analyses based on phylogenetic distances between the sampled communities. To determine whether some of the samples were clustering because of the presence of abundant *Lactobacillus* spp., we executed a second analysis where we first removed *Lactobacillus* spp. reads, then ran the remaining sequences through the same Fast UniFrac pipeline as described above.

For the approximate maximum-likelihood phylogeny of *Lactobacillus*, we first extracted sequence reads from our data set that were identified as *Lactobacillus* by our BLAST searches. We again clustered redundant sequences within host species using CDhit (Huang *et al.* 2010), with the sequence identity cut-off set to 97%. An input of 19844 sequences resulted in the clustering of 85 phylotypes with read length >350 nucleotides.

To represent lactobacilli that associate with other species of bees, we obtained 16S rDNA sequences from GenBank, mostly from previous studies of *Apis mellifera* (Jeyaprakash *et al.* 2003; Babendreier *et al.* 2007; Mohr & Tebbe 2007; Olofsson & Vasquez 2008; Vasquez *et al.* 2009; Martinson *et al.* 2011), as well as from *Bombus* and other bees (Mohr & Tebbe 2007; Koch & Schmid-Hempel 2011a; Martinson *et al.* 2011; Tajabadi *et al.* 2011; Disayathanoowat *et al.* 2012). Next, we downloaded 226 *Lactobacillus* full-length 16S rDNA sequences from NCBI that were available as of 18 October 2011. After removing sequences that shared 97% or greater sequence identity and sequences from unidentified lactobacilli, the alignment consisted of 100 known *Lactobacillus* sequences, 71 sequences from other studies of bee-associated bacteria, 15 outgroup sequences and 85 of our 454 generated sequences. To align these sequences, we used MUSCLE (Edgar 2004) to create an initial alignment. Although conserved segments of 16S rDNA aligned well, several variable sections were difficult to align. We therefore used the program SATe to simultaneously estimate the alignment and phylogeny over 100 iterations, using the general time reversible with invari-

ant sites and gamma-distributed rate variation model of sequence evolution. SATe outperforms two-step alignment and phylogenetic estimation methods when working with genes with high variability and insertions and deletions, such as the 16S rRNA gene (Liu *et al.* 2009). We then compared the resulting alignment to a structural model of the 16S rRNA small subunit (Cannone *et al.* 2002), which allowed us to manually refine the conserved sections of the alignment in the program Mesquite (Maddison & Maddison 2010).

To determine the model of sequence evolution that best fit our alignment, we used the program Modeltest (Posada & Crandall 1998). We used the Akaike Information Criterion to select the most likely substitution model (GTR + I + G) for approximate maximum-likelihood analyses. We conducted 20 independent search replicates and 100 bootstrap pseudoreplicates in the program Garli (Zwickl 2006).

We additionally tested whether corbiculate apid and halictid bees are more closely associated with flower-associated lactobacilli than other lactobacilli. We first used the best-scoring maximum-likelihood tree to calculate the patristic distances from corbiculate apid-associated lactobacilli and halictid-associated lactobacilli to flower-associated lactobacilli and all other lactobacilli. After verifying that the distances were normally distributed, we conducted two t-tests to determine whether (i) the distances from corbiculate apid-associated lactobacilli to flower-associated lactobacilli differ from the distances from corbiculate apid-associated lactobacilli to all other lactobacilli and (ii) the distances from halictid-associated lactobacilli to flower-associated lactobacilli differ from the distances from halictid-associated lactobacilli to all other lactobacilli.

Lastly, we investigated divergence among paralogous copies of the 16S rRNA gene within lactobacilli genomes in order to determine whether this might affect our diversity analyses. We downloaded all of the annotated 16S rDNA copies from 10 lactobacilli that had complete genome sequences available. We aligned the 16S rDNA copies from each genome separately, in the program Mesquite (Maddison & Maddison 2010), and determined the number of polymorphic sites per genome.

Results

Bacterial survey based on 16S rDNA amplicon pyrosequencing

Individual bees of the species *Augochlora pura*, *Halictus ligatus* and *Megalopta genalis* and the contents of their nests hosted a rich diversity of microbes, with a total of 992 unique phylotypes identified across all samples from a total of 125 005 analysis quality reads (for the raw data, see Tables S1–S4, Supporting information). These bacteria included putative mutualists, pathogens and commensals. Rarefaction curves indicated that the majority of our sequencing was deep enough to characterize the bacterial communities, with the exception of several samples taken from the substrate surrounding the nests (Figs S1–S4, Supporting information).

Bacteria from various taxonomic levels within the phylum Firmicutes were abundant in all four bee nests, although to a lesser degree in the *Au. pura* nest (Fig. 1). We further scrutinized sequences that matched the phylum Firmicutes in our initial BLAST search with a second BLAST search to a sequence database representing the full diversity of the Firmicutes. This second BLAST search revealed that the sequences shared ≥89% of their 16S rDNA sequence with lactic acid bacteria (LAB) from the genus *Lactobacillus* (Table 1.). For example, reads that shared 94.8% and 92.4% sequence identity with *Lactobacillus kunkeei* were abundant in *Au. pura* and *M. genalis* samples, respectively. Our additional BLAST results against the SILVA database and subsequent MEGAN4 analysis verified these data, also showing *L. kunkeei* as the most abundant bacterium from our samples (Fig. S5, Supporting information). In general, lactobacilli often occurred at high abundance in samples from the brood cells but not from the adult insects, with the exception of the *H. ligatus* female (Fig. 1).

Many of the other common bacteria also come from acidophilic groups: members of the Acidobacteriaceae along with *Saccharibacter*, which is a member of the Acetobacteraceae (Fig. 1). Similar to the bee bread of *Apis mellifera* (Anderson *et al.* 2011), the pollen provisions from three *M. genalis* brood cells exhibited acidic pH (3.4, 3.4 and 3.7). In addition to acid-loving bacteria, we detected endosymbiotic bacteria in the genera *Wolbachia* and *Cardinium* (Zchori-Fein & Perlman 2004) from several insect samples and at low abundance in two frass samples (Fig. 1, Tables S1–S4, Supporting information).

Some of the acidophilic bacteria that we detected in association with sweat bees are closely related to bacteria that have also been detected at flowers. Bacteria with greater than 96% sequence identity to *Saccharibacter floricola* associated with the social *M. genalis* female (Table S3, Supporting information). *S. floricola* was described from pollen from Japanese flowers (Jojima *et al.* 2004) and is nested in the alpha-2 phylotype that forms part of the core *Ap. mellifera* microbiota (Martinson *et al.* 2011). Our SILVA BLAST searches also indicated that *S. floricola* was abundantly represented (Fig. S5, Supporting information). Although uncommon, several *M. genalis* reads had 97–98% sequence identity with a *Lactobacillus* that is known from flowers,

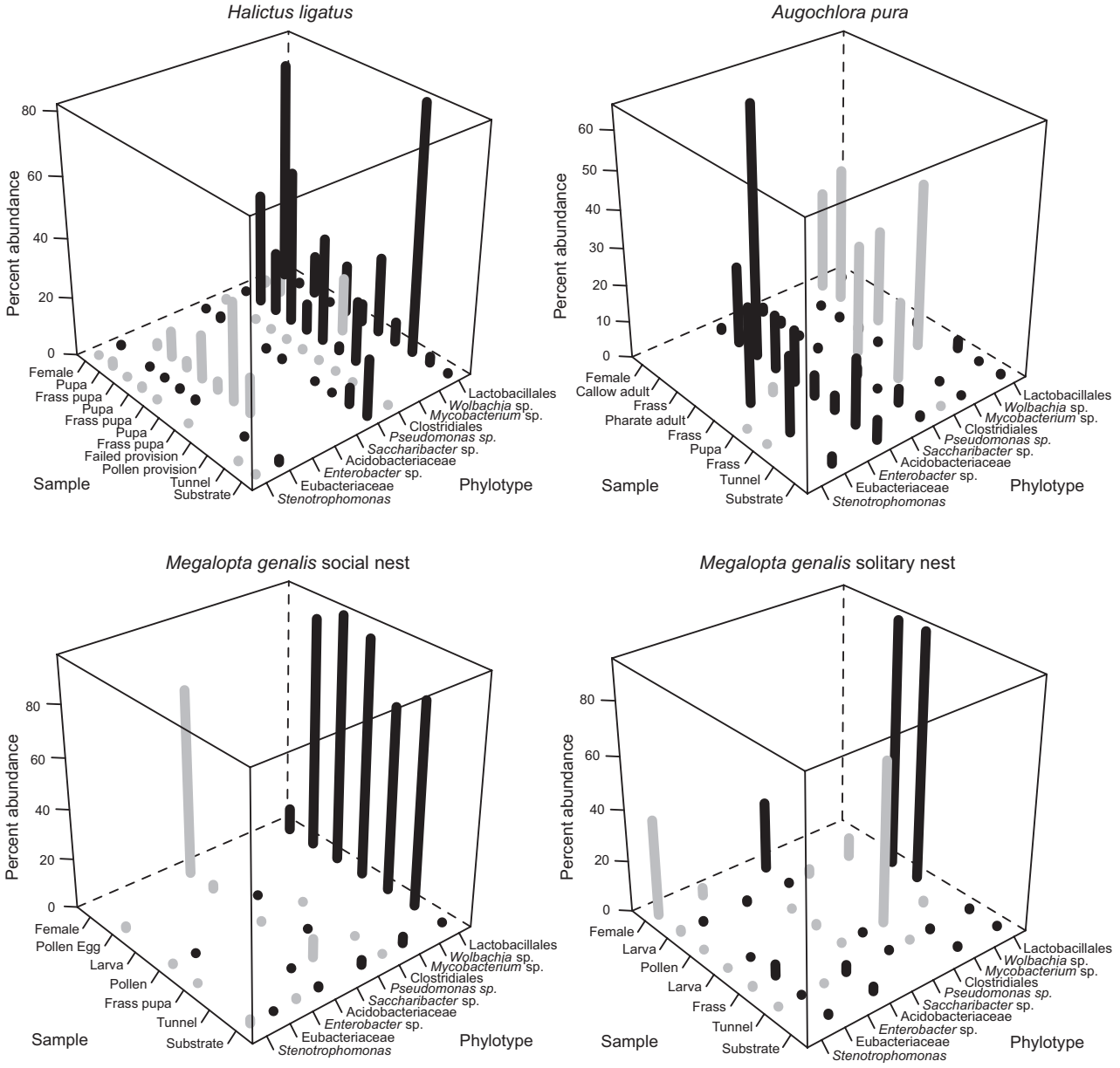

Fig. 1 Proportional abundance of the ten most commonly observed bacterial phylotypes from four sweat bee nests, as determined by BLAST searches against a database curated by the Medical Biofilm Research Institute. The first row, Firmicutes, includes the nested taxonomic levels Bacilli, Lactobacillales, Lactobacillaceae and *Lactobacillus*. Firmicute phylotypes all had top BLAST hits to *Lactobacillus*, but not all showed ≥97% sequence identity (Table 1). Rarefaction curves indicated that we sequenced deeply enough to accurately characterize most bacterial communities, with the exception of several samples taken from the substrate surrounding the nests (Figs S1–S4, Supporting information). For a heatmap version of this figure, see Fig. S7 (Supporting information).

Lactobacillus ozensis (Kawasaki *et al.* 2011). Finally, many of our reads had top BLAST hits to *L. kunkeei* (94.83% and 92.43% average sequence identity for the *Au. pura* and *M. genalis* associates, respectively), another species known from flowers (Table 1, Fig. S5, Supporting information, Endo *et al.* 2011).

We also detected a putative pathogen: an unidentified *Paenibacillus*, a genus that includes a pathogen of honey bees (Evans 2003). *Paenibacillus* was generally not abundant in our samples. In a failed *H. ligatus* brood cell, however, two unknown *Paenibacillus* phylotypes represented 11.7% and 5.7% of the bacterial sequences.

The microbiota of these miniature ecosystems exhibit interesting patterns of community composition and diversity. For example, the surrounding substrate showed the greatest richness and diversity from each nest, with the exception of the laboratory nest (Fig. 2). The soil that was used as the nesting substrate for the

Table 1 Top BLAST hit of sequences that were assigned to an unknown taxonomic position in the phylum Firmicutes against a reference database compiled from Firmicutes sequences from NCBI.

Host	Top BLAST hit	Average % identity	% abundance
Augochlora pura	*Lactobacillus kunkeei*	94.83	13.33
	Lactobacillus kefiri	89.72	0.22
	Lactobacillus fructivorans	89.67	0.12
	Lactobacillus lindneri	89.98	0.05
Halictus ligatus	*Lactobacillus fructivorans*	89.93	41.52
	Lactobacillus kefiri	89.43	25.59
	Lactobacillus lindneri	89.38	8.53
	Lactobacillus homohiochii	90.20	0.81
	Lactobacillus murinus	98.80	0.58
	Lactobacillus intestinalis	96.81	0.10
	Lactobacillus ozensis	90.21	0.01
	Lactobacillus tucceti	90.42	0.01
Megalopta genalis	*Lactobacillus kunkeei*	92.43	73.14
	Lactobacillus casei	98.74	3.37
	Lactobacillus rossiae	98.08	1.08
	Lactobacillus crispatus	98.45	0.10
	Lactobacillus ozensis	94.42	0.01
	Lactobacillus plantarum	97.62	0.01
	Lactobacillus sanfranciscensis	90.01	0.01

Average per cent identity represents the average sequence identity of those phylotypes that shared top hits to the same reference species. Per cent abundance represents the per cent abundance of a given phylotype within its respective host species, excluding the surrounding substrate samples. Only hits with a per cent abundance > 0.01% are shown.

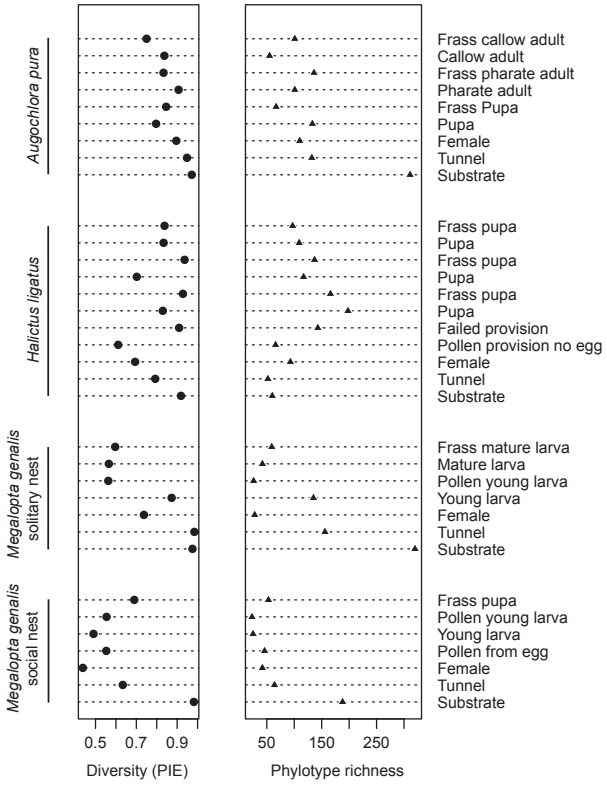

Fig. 2 Number of observed bacterial phylotypes and evenness (PIE) for each sample.

H. ligatus nest had been autoclaved, so it is not surprising that it exhibited some of the lowest phylotype richness. The *H. ligatus* soil sample, however, exhibited high evenness, indicating that no early colonizer was able to dominate the soil environment. The *M. genalis* nests exhibited lower evenness than the North American bee nests, with the female from the social *M. genalis* nest exhibiting the lowest evenness of all samples (Fig. 2).

The sweat bee–associated microbiota exhibit differing community composition, as determined by our Fast UniFrac analysis (Fig. 3). The *H. ligatus* and *Au. pura* microbiota clustered in different regions of principal coordinate 2, while several *M. genalis* samples were more widely dispersed. The *H. ligatus* nest was a laboratory nest; therefore, comparisons should be viewed cautiously. However, *Lactobacillus* dominated several *H. ligatus* and *M. genalis* samples that formed a cluster along coordinate 1 (Figs 2 and 3). When we ran the analysis with lactobacilli removed, these samples were widely dispersed (Fig. 4), showing that the dominance of lactobacilli in these samples drove the clustering in Fig. 3.

Phylogeny of the genus Lactobacillus *based on 16S rDNA*

Our phylogeny represents *Lactobacillus* from our 454-sequencing survey and publicly available *Lactobacillus*

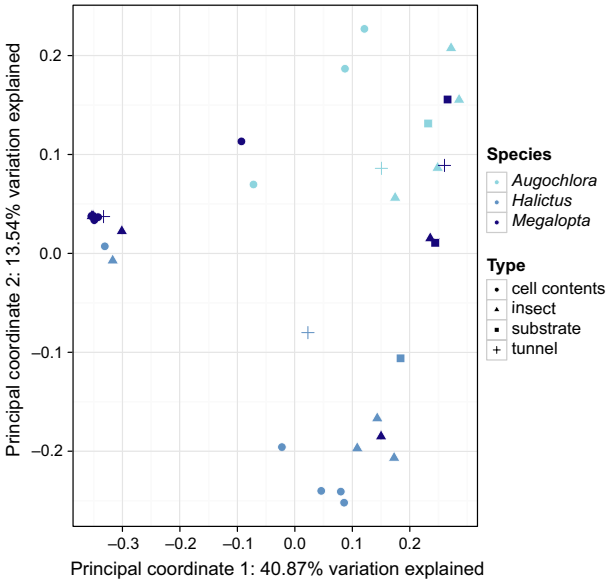

Fig. 3 First two principal coordinates from an abundance-weighted UniFrac analysis of the bacterial communities from four sweat bee nests.

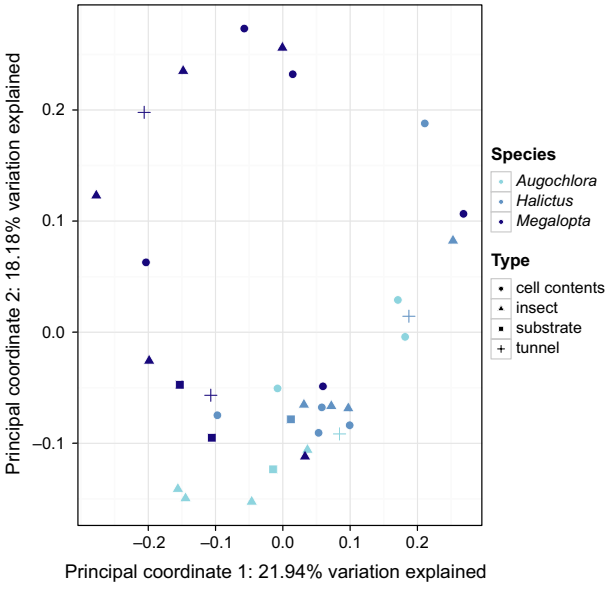

Fig. 4 First two principal coordinates from an abundance-weighted UniFrac analysis of the bacterial communities from four sweat bee nests, with phylotypes with top BLAST hits to *Lactobacillus* spp. removed.

sequences. We found that bee-associated lactobacilli are distributed across the *Lactobacillus* phylogeny. In general, membership of the clades in our phylogeny are in agreement with other published 16S rDNA phylogenies of *Lactobacillus* (Canchaya *et al.* 2006; Martinson *et al.* 2011). However, owing to the low bootstrap support for many of the clades, our phylogeny should not be used to interpret deeper branching patterns in the lactobacilli.

Several bee-associated *Lactobacillus* clades were entirely or nearly bee specific. For example, the F-3 clade (sensu Babendreier *et al.* 2007) was composed of sequences from corbiculate bee isolates only (Fig. 5). The F-4 clade was also mostly specific to *Apis* and *Bombus*; however, an isolate from a spacecraft assembly clean room (accession number GQ129909, La Duc *et al.* 2009) shared high sequence identity to F-4, as noted by Martinson *et al.* (2011). We also found the F-5 clade to be mostly *Bombus* and *Apis* specific, with the exception of isolates from *Colletes inaequalis*, a solitary bee (Fig. 5, Martinson *et al.* 2011), and *Anopheles stephensi* (accession number FJ608053, Rani *et al.* 2009) with a top BLAST hit (98% sequence identity) to sequences from the F-5 clade. F-5 lactobacilli therefore appear to associate with insects other than corbiculate bees.

Other bee-associated lactobacilli clustered with bacteria isolated from flowers. For example, *Lactobacillus kunkeei* and *L. ozensis*, both of which have been isolated from flowers (Kawasaki *et al.* 2011; Endo *et al.* 2011), fell into a clade that also included lactobacilli isolated from *Augochlora pura*, *Megalopta genalis*, *Caupolicana yarrowi*, *Diadasia opuntiae* and *Apis mellifera* (Fig. 5). The sister clade to the *L. kunkeei* clade was composed of lactobacilli isolated from *Halictus ligatus*. Finally, *L. floricola* clustered with bacteria isolated from *Ap. mellifera* and shared 99% sequence similarity with bacteria isolated from *Agapostemon virescens*.

The genetic distances within clades that include halictid-associated lactobacilli also support the floral transmission hypothesis. First, the genetic distance in the *L. kunkeei* clade (maximum/average patristic distance = 0.35/0.11) was in the range of the distances exhibited by the corbiculate apid-associated lactobacilli (maximum/average patristic distances: F4 clade = 0.54/0.25, F5 clade = 0.32/0.09, F3 clade = 0.14/0.08). Second, halictid-associated lactobacilli exhibited significantly smaller patristic distances to flower-associated lactobacilli than to other lactobacilli ($T = -16.8431$, $df = 370.08$, $P < 2.2e\text{-}16$), while corbiculate apids did not ($T = 1.734$, $df = 249.859$, $P = 0.08416$, Fig. S6, Supporting information). While these statistical tests support floral transmission of halictid-associated lactobacilli, comparisons based on phylogenetic tree topologies are difficult to interpret, and future work is still necessary.

Some bee-associated lactobacilli clustered with bacteria associated with humans and other mammals. Several *H. ligatus* isolates clustered with *L. murinus* and *L. animalis*, both of which were originally isolated from rodent intestinal tracts. Several other isolates from *M. genalis* clustered with *L. larvae*, an associate of *Ap. mellifera*, and *L. paracasei*, which is thought to be a human probiotic (Gardiner *et al.* 2000). The shared records of these lactobacilli from bees, rodents and

humans indicate that these bacteria likely occur in the environment.

Within-genome 16S rDNA variation

Finally, our analysis of within-genome variation in paralogous copies of the 16S rRNA gene from 10 lactobacilli genomes suggests that 16S rDNA serves as a reliable barcode for the genus. We found that the number of paralogous copies varied from 4 to 9 copies per genome, with a mode of five copies (Table S5, Supporting information). Although there was some polymorphism present in the paralogous copies, the amount of polymorphism did not exceed 1% within any of the genomes (Table S5, Supporting information).

Discussion

Augochlora pura, *Halictus ligatus* and *Megalopta genalis* associate with putative beneficial and pathogenic bacteria, and some of the putatively beneficial bacteria are possibly acquired by the bees from flowers. In addition to acquisition from flowers, our phylogeny agrees with studies that suggest that some lactobacilli appear to be specific to the corbiculate apids and may be vertically transmitted (Koch & Schmid-Hempel 2011a; Martinson *et al.* 2011). Bees may therefore acquire beneficial bacteria by at least two routes: one route from parent to offspring (corbiculate bees) and a second route via flowers and possibly other environmental sources.

Our analyses provide several lines of evidence suggesting that transmission of bacteria from flowers to wild bees may be more important than previously thought. First, our 16S rDNA amplicon pyrosequencing survey detected several bacterial phylotypes that either had at least 97% sequence identity to bacteria isolated from flowers (*Saccharibacter floricola* and *L. ozensis*) or had top BLAST hits at lower sequence identity to a bacterium known to occur on flowers (*L. kunkeei*). Additional support comes from other studies documenting bacteria that have been isolated from bees and that are also known to occur on flowers: *L. kunkeei* from *Apis mellifera* (Olofsson & Vasquez 2008); *S. floricola* from *Ap. mellifera*, *L. floricola* from *Agapostemon virescens* (Martinson *et al.* 2011); *L. floricola* from *Ap. mellifera* (NCBI accession number HM534770); and *L. ozensis* from *Ap. mellifera* (NCBI accession number HM534743). The shared association of these bacteria with both bees and flowers suggests that bees may be obtaining these bacteria from flowers, although experimental confirmation of this hypothesis is still necessary. Alternatively, bees may be the source of the lactobacilli and other acidophilic bacteria found at flowers, but this would still indicate that flowers are sites for potential horizontal transmission between bee lineages. Next, our sequence data show that *Lactobacillus* spp. and *S. floricola* can be found at high abundance in association with sweat bees, but that there is intra- and internest variance in the presence and abundance of these bacteria (Fig. 1). Therefore, if *Lactobacillus* and *S. floricola* are only vertically transmitted, transmission must be incomplete. Although symbionts with imperfect vertical transmission and no horizontal transmission can persist in a host population, for a symbiont with such imperfect vertical transmission to persist, it would need to have very strong positive effects on host fecundity and/or offspring survival (Fine 1975). While it is entirely plausible that *Lactobacillus* spp. and *S. floricola* have such strong effects, a more parsimonious explanation is that both positive effects on host fitness and environmental transmission play a role in the biology of bees.

Our data largely agree with a previous study that failed to detect the core gut microbiota of *Ap. mellifera* outside the corbiculate apids (Martinson *et al.* 2011). We did, however, detect one of the *Ap. mellifera* core gut microbes, *S. floricola*, in one *M. genalis* nest. Martinson *et al.* (2011) found that *Burkholderia cepacia*, a widespread bacterium, was commonly isolated from bees outside the corbiculate apids. While we detected *B. cepacia* in association with *H. ligatus*, it was present at very low abundance (<0.17%) and was absent from the *Au. pura* and *M. genalis* nests. Instead, many of our samples were dominated by lactobacilli that were related to bacteria that Martinson *et al.* (2011) isolated at low abundance from *Halictus patellatus*, *C. yarrowi*, *D. opuntiae* and *Hoplitis biscutellae* and that were closely related to *L. kunkeei* in our phylogenetic and BLAST analyses (Fig. 5, Table 1). The differences in abundance between our two studies may be explained by seasonal variation in the presence of *Lactobacillus* at flowers. Alternatively, the variation in abundance between the two studies may be explained by methodological differences; Martinson *et al.* (2011) surveyed abdomens while we surveyed entire bees and the contents of their nests. Our data agree with the findings of Martinson *et al.* (2011), in that *Lactobacillus* was often found in low abundance or missing from our adult insect samples, with the exception of the *H. ligatus* mother (Fig. 1). Through our surveys of the contents of brood cells, however, we were able to determine that lactobacilli can be abundant in the nests of bees outside the corbiculate apids.

Apis mellifera has been found to associate with *Lactobacillus kunkeei*, and floral transmission would explain why there has been variation in the presence and abundance of *Ap. mellifera*-associated *L. kunkeei* across surveys, although variation in PCR parameters or other methodological differences cannot be excluded. Floral

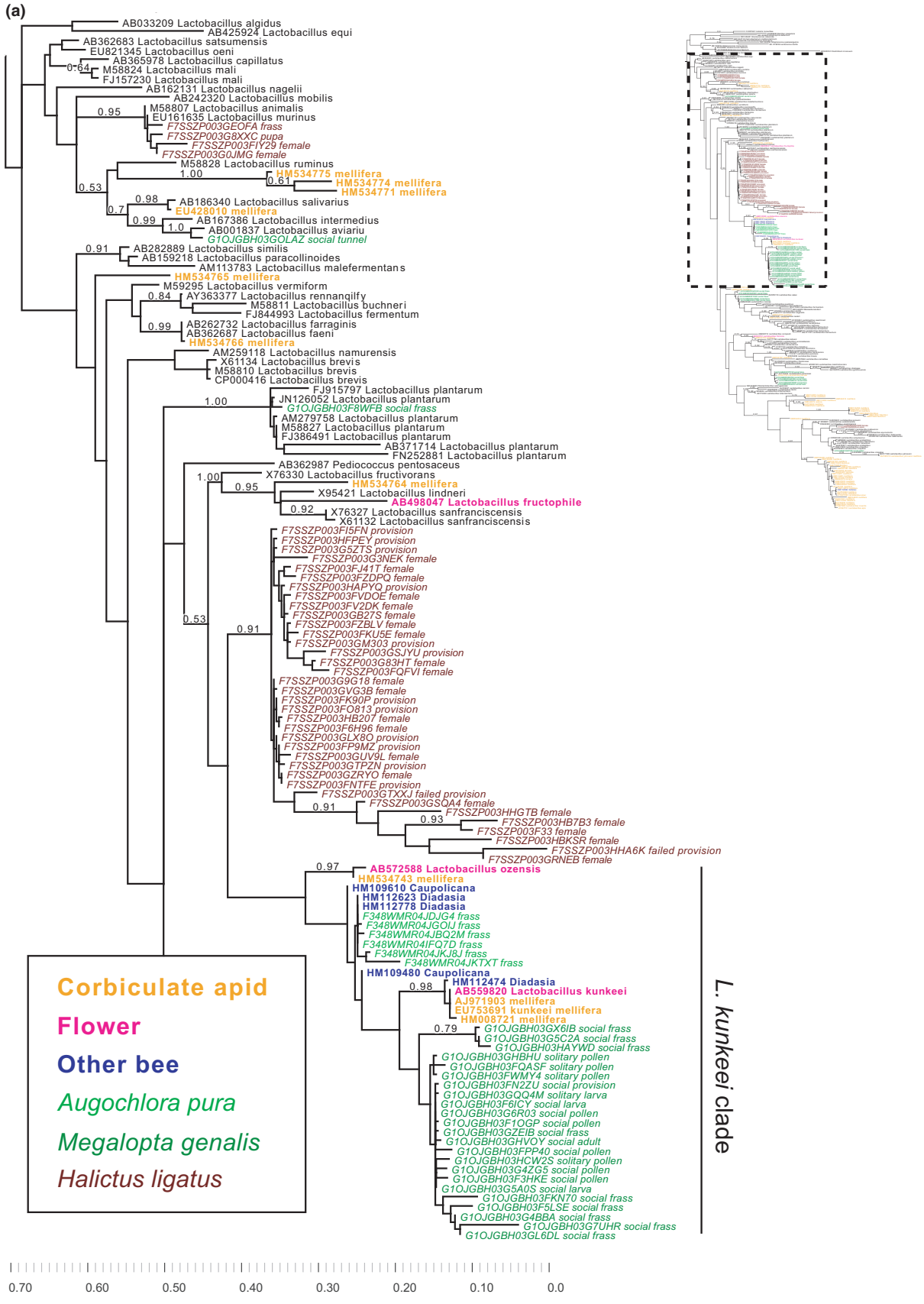

Fig. 5 Best-scoring *Lactobacillus* 16s rDNA tree of 20 maximum-likelihood search replicates, with bootstrap values from 100 pseudoreplicates placed on branches with greater than 50% support. Corbiculate apid-associated sequences are in boldface, orange font; flower-associated sequences are in boldface, pink font; other bee-associated sequences are in boldface, blue font; *Augochlora pura* sequences are in italicized, light green font; *Megalopta genalis* sequences are in italicized dark green font; and *Halictus ligatus* sequences are in italicized, brown font. Fig. 5a,b represent the respective outlined portion of the larger tree, as indicated by the smaller complete tree.

transmission, for example, could explain why *L. kunkeei* was absent in an *Ap. mellifera* hive surveyed in January, a time in winter when little forage is available (Martinson *et al.* 2011), while *L. kunkeei* was abundant in surveys of actively foraging colonies (Olofsson & Vasquez 2008). The floral transmission hypothesis was rejected by Olofsson & Vasquez (2008), as they sampled several species of flowers, honey and specimens of *Ap. mellifera* and detected *L. kunkeei* from the bees and honey only. The hypothesis deserves revisiting, however, as *L. kunkeei* has since been reported from azalea, narcissus, cosmos, morning glory and crape myrtle flowers (Endo *et al.* 2009, 2011).

Besides acidophilic bacteria, we found several other abundant bacteria associated with the sweat bees in our study. *Wolbachia* has been previously detected in sweat bees (Werren & Windsor 2000), was present in all three species and, besides very low levels detected in frass and pollen, was found only in the insect samples (Fig. 1). Another endosymbiotic bacterium that is known to distort the sex ratios of its host, *Cardinium* (Zchori-Fein & Perlman 2004), was found in high abundance in a couple of samples from the *M. genalis* solitary nest (Table S3, Supporting information). How these endosymbiotic bacteria interact with their bee hosts remains unknown. Bacteria that are commonly found in the soil, such as Clostridiales and *Pseudomonas* (Janssen 2006), were both found in two of the four nests, indicating that environmental bacteria can colonize bee nests.

If future studies show that sweat bee–associated lactobacilli or other acidophilic bacteria provide benefits to their hosts, environmental transmission has some interesting implications, because association with acidophilic bacteria is not guaranteed. Through secretion of acids, lactobacilli and other acidophilic bacteria may provide protection of the pollen provisions of bees such as *M. genalis*, *H. ligatus* and *Au. pura* from exploitation by other bacteria and fungi. If so, variation in the presence of these bacteria may explain some variation in wild bee health. Floral nectar pH varies widely, with extremes values from 3 to 10, although most nectars appear to be slightly acidic (Baker & Baker 1983). There may be variation in the presence of acidophilic bacteria in flowers, related to variation in nectar pH, or vice versa. Broad surveys of bacteria associated with flowers and experiments testing for the transmission of putative mutualists from flowers to bees would further elucidate these relationships. If flowers do vary in their suitability as reservoirs of acidophilic bacteria, planting those flowers that best provide acidophilic bacteria to wild bees may be an applied outcome of such surveys.

The fermentation of human foods by LAB does not affect the food's nutritional value, but rather spoilage and growth of pathogens is suppressed (Lindgren & Dobrogosz 1990). *L. sanfranciscensis*, for example, is thought to suppress fungi through the secretion of a variety of acids, thus prolonging the shelf life of sourdough breads (Corsetti *et al.* 1998). We found that the pH of the three pollen samples from *M. genalis* was quite low (3.4–3.7); however, whether the acidic provisions is a cause or a consequence of *Lactobacillus* presence remains to be tested. Additionally, some sweat bee–associated lactobacilli are related to *L. sanfranciscensis* and may inhibit mould growth on pollen provisions and frass inside of bee brood cells. However, further work is necessary to determine whether fermentative properties are conserved in bee-associated lactobacilli, as bacterial phenotype can be unpredictable from phylotype (Schloss & Westcott 2011).

It has been hypothesized that bacteria associated with *Apis* and *Bombus* represent a co-evolutionary relationship (Martinson *et al.* 2011); however, evidence of reciprocal selection is currently lacking. Our phylogeny largely agrees with the finding that some lactobacilli are specific to the corbiculate apids, a condition conducive to co-evolution. We found that the F-3 clade was completely *Apis*- and *Bombus*-specific, which is consistent with suggestions that these lactobacilli may be vertically transmitted (Koch & Schmid-Hempel 2011a; Martinson *et al.* 2011). Clades F-4 and F-5 appear to be mostly, but not entirely, associated with the corbiculate apids. The isolation of F-4 and F-5 sequences from a clean room, a solitary bee and a mosquito suggests that more comprehensive sampling may change future views of any of these relationships.

Our data suggest promising lines of future research. It has been hypothesized that social structure affects the acquisition of beneficial bacteria by bees (Koch & Schmid-Hempel 2011a; Martinson *et al.* 2011), and the first evidence supporting this hypothesis has recently been published. Exposure of laboratory-reared bumble bees to their nest mate's faeces allowed for the establishment of bacteria that protected the bees from the parasite *Crithidia bombi*, indicating that social contacts may be important in intracolony transmission (Koch & Schmid-Hempel 2011b). Although low replication means that our data should not be overinterpreted, *Lactobacillus* consistently occurred in samples from social nests and patchily occurred in samples from solitary nests (Fig. 1). *M. genalis* is known to exhibit trophallaxis (Wcislo & Gonzalez 2006), which may help to maintain the presence of lactobacilli in social nests (Martinson *et al.* 2011); however, trophallaxis has not been reported in *H. ligatus*. More work teasing out the role of environmental and social transmission of beneficial bacteria is needed. Additionally, understanding how microbes affect bee health is of timely importance. For example,

several fungi are known to spoil the provisions of wild bees (Batra *et al.* 1973), and experimentation, metagenomics and transcriptomics investigating interactions between fungi and bacteria with antifungal properties, such as *Lactobacillus* (Corsetti *et al.* 1998), will improve understanding of how microbiomes affect the host's phenotype.

Conclusion

Our data suggest that there are at least two different pathways for bees to form associations with putative mutualists: (i) Our phylogenetic analysis of the genus *Lactobacillus* is consistent with previous claims that some lactobacilli are specific to the corbiculate apids (Koch & Schmid-Hempel 2011a; Martinson *et al.* 2011). These previous studies suggested that host social structure might facilitate the vertical transmission of these putatively co-evolved lactobacilli. (ii) Using second-generation sequencing of bacterial 16S rDNA amplicons, we discovered lactobacilli and other acidophilic bacteria that are associated with *Augochlora pura*, *Halictus ligatus* and *Megalopta genalis*. Some of these acidophilic bacteria are also associated with flowers, while others are related to bacteria known to associate with flowers. These bacteria may therefore be transmitted at flowers. Although associating with adults, lactobacilli can also be found in high abundance in pollen provisions and frass inside the nest. As these lactobacilli are related to species that are known to suppress mould growth, it is plausible that they suppress mould growth inside the nest, contributing to the health of the bee. Bee-associated acidophilic bacteria may not need to be tightly co-evolved with their host to act as probiotics; facultative relationships may also arise, although they may not be as stable as co-evolved relationships.

Research aimed at understanding these relationships between microbes and wild bees should be a priority. If honey bees continue to decline, wild pollinators will become ever more important for agriculture. We must therefore begin to build an understanding of how pathogens and mutualists affect nonmanaged, wild pollinator populations so that we are better able to protect such resources.

Acknowledgements

We thank Sabrina Amador and three anonymous reviewers for comments on an earlier draft; Blandy Experimental Farm for access, logistical support and financial support; and the Smithsonian Tropical Research Institute for access to BCI and logistical support. This material is based on work supported by the National Science Foundation under Grant Nos. PRFB-1003133, DEB-0919519 and DDIG-0808454.

References

Anderson K, Sheehan T, Eckholm B, Mott B, DeGrandi-Hoffman G (2011) An emerging paradigm of colony health: microbial balance of the honey bee and hive (*Apis mellifera*). *Insectes Sociaux*, **58**, 431–444.

Babendreier D, Joller D, Romeis J, Bigler F, Widmer F (2007) Bacterial community structures in honeybee intestines and their response to two insecticidal proteins. *FEMS Microbiology Ecology*, **59**, 600–610.

Baker HG, Baker I (1983) A brief historical review of the chemistry of floral nectar. In: *The Biology of Nectaries* (eds Bentley B and Elias T), pp. 126–152. Colombia University Press, New York city, New York.

Batra L, Batra S, Bohart G (1973) The mycoflora of domesticated and wild bees (Apoidea). *Mycopathologia*, **49**, 13–44.

Biani NB, Mueller UG, Wcislo WT (2009) Cleaner mites: sanitary mutualism in the miniature ecosystem of neotropical bee nests. *The American Naturalist*, **173**, 841–847.

Bromenshenk J, Henderson C, Wick C, Stanford M, Zulich A (2010) Iridovirus and microsporidian linked to honey bee colony decline. *PLoS ONE*, **5**, e13181.

Cameron SA, Lozier JD, Strange JP *et al.* (2011) Patterns of widespread decline in North American bumble bees. *Proceedings of the National Academy of Sciences*, **108**, 662–667.

Canchaya C, Claesson MJ, Fitzgerald GF, Van Sinderen D, O'Toole PW (2006) Diversity of the genus *Lactobacillus* revealed by comparative genomics of five species. *Microbiology*, **152**, 3185–3196.

Cannone J, Subramanian S, Schnare M *et al.* (2002) The comparative RNA web (CRW) site: an online database of comparative sequence and structure information for ribosomal, intron, and other RNAs. *BMC Bioinformatics*, **3**, 2.

Corsetti A, Gobbetti M, Rossi J, Damiani P (1998) Antimould activity of sourdough lactic acid bacteria: identification of a mixture of organic acids produced by *Lactobacillus sanfrancisco* CB1. *Applied Microbiology and Biotechnology*, **50**, 253–256.

Cox-Foster DL, Conlan S, Holmes EC *et al.* (2007) A metagenomic survey of microbes in honey bee colony collapse disorder. *Science*, **318**, 283–287.

Disayathanoowat T, Young JPW, Helgason T, Chantawannakul P (2012) T-RFLP analysis of bacterial communities in the midguts of *Apis mellifera* and *Apis cerana* honey bees in Thailand. *FEMS Microbiology Ecology*, **79**, 273–281.

Dowd SE, Sun Y, Secor PR *et al.* (2008) Survey of bacterial diversity in chronic wounds using pyrosequencing, DGGE, and full ribosome shotgun sequencing. *BMC Microbiology*, **8**, 43.

Duffield R, Fernandes A, Lamb C, Wheeler J, Eickwort G (1981) Macrocyclic lactones and isopentenyl esters in the Dufour's gland secretion of halictine bees (Hymenoptera: Halictidae). *Journal of Chemical Ecology*, **7**, 319–331.

Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, **32**, 1792–1797.

Endo A, Futagawa-Endo Y, Dicks LMT (2009) Isolation and characterization of fructophilic lactic acid bacteria from fructose-rich niches. *Systematic and Applied Microbiology*, **32**, 593–600.

Endo A, Irisawa T, Futagawa-Endo Y *et al.* (2011) Characterization and emended description of *Lactobacillus kunkeei* as a fructophilic lactic acid bacterium. *International Journal of Systematic and Evolutionary Microbiology*. Published online ahead of print April 8, 2011.

Evans JD (2003) Diverse origins of tetracycline resistance in the honey bee bacterial pathogen *Paenibacillus larvae*. *Journal of Invertebrate Pathology*, **83**, 46–50.

Fine PEM (1975) Vectors and vertical transmission: an epidemiologic perspective. *Annals of the New York Academy of Sciences*, **266**, 173–194.

Forsgren E, Olofsson TC, Vasquez A, Fries I (2010) Novel lactic acid bacteria inhibiting *Paenibacillus larvae* in honey bee larvae. *Apidologie*, **41**, 99–108.

Gardiner G, O'Sullivan E, Kelly J *et al.* (2000) Comparative survival rates of human-derived probiotic *Lactobacillus paracasei* and *L. salivarius* strains during heat treatment and spray drying. *Applied and Environmental Microbiology*, **66**, 2605–2612.

Gontcharova V, Youn E, Wolcott RD *et al.* (2010) Black box chimera check (B2C2): a windows-based software for batch depletion of chimeras from bacterial 16S rRNA gene datasets. *The Open Microbiology Journal*, **4**, 47–52.

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

Huang Y, Niu B, Gao Y, Fu L, Li W (2010) CD-HIT suite: a web server for clustering and comparing biological sequences. *Bioinformatics*, **26**, 680–682.

Hurlbert SH (1971) The nonconcept of species diversity: a critique and alternative parameters. *Ecology*, **52**, 577–586.

Huson DH, Mitra S, Ruscheweyh HJ, Weber N, Schuster SC (2011) Integrative analysis of environmental sequences using MEGAN4. *Genome Research*, **21**, 1552–1560.

Ishak HD, Plowes R, Sen R *et al.* (2011) Bacterial diversity in *Solenopsis invicta* and *Solenopsis geminata* ant colonies characterized by 16S amplicon 454 pyrosequencing. *Microbial Ecology*, **61**, 821–831.

Janssen PH (2006) Identifying the dominant soil bacterial taxa in libraries of 16S rRNA and 16S rRNA genes. *Applied and Environmental Microbiology*, **72**, 1719–1728.

Jeyaprakash A, Hoy MA, Allsopp MH (2003) Bacterial diversity in worker adults of *Apis mellifera capensis* and *Apis mellifera scutellata* (Insecta: Hymenoptera) assessed using 16S rRNA sequences. *Journal of Invertebrate Pathology*, **84**, 96–103.

Johnson RM, Evans JD, Robinson GE, Berenbaum MR (2009) Changes in transcript abundance relating to colony collapse disorder in honey bees (*Apis mellifera*). *Proceedings of the National Academy of Sciences*, **106**, 14790–14795.

Jojima Y, Mihara Y, Suzuki S *et al.* (2004) *Saccharibacter floricola* gen. nov., sp. nov., a novel osmophilic acetic acid bacterium isolated from pollen. *International Journal of Systematic and Evolutionary Microbiology*, **54**, 2263–2267.

Kawasaki S, Kurosawa K, Miyazaki M *et al.* (2011) *Lactobacillus ozensis* sp. nov., a novel lactic acid bacteria isolated from Oze National Park in Japan. *International Journal of Systematic and Evolutionary Microbiology*, **61**, 2435–2438.

Koch H, Schmid-Hempel P (2011a) Bacterial communities in Central European bumblebees: low diversity and high specificity. *Microbial Ecology*, **62**, 121–133.

Koch H, Schmid-Hempel P (2011b) Socially transmitted gut microbiota protect bumble bees against an intestinal parasite. *Proceedings of the National Academy of Sciences*, **108**, 19288–19292.

Kunin V, Engelbrektson A, Ochman H, Hugenholtz P (2010) Wrinkles in the rare biosphere: pyrosequencing errors can lead to artificial inflation of diversity estimates. *Environmental Microbiology*, **12**, 118–123.

La Duc MT, Osman S, Vaishampayan P *et al.* (2009) Comprehensive census of bacteria in clean rooms by using DNA microarray and cloning methods. *Applied and Environmental Microbiology*, **75**, 6559–6567.

Lindgren SE, Dobrogosz WJ (1990) Antagonistic activities of lactic acid bacteria in food and feed fermentations. *FEMS Microbiology Letters*, **87**, 149–164.

Liu K, Raghavan S, Nelesen S, Linder CR, Warnow T (2009) Rapid and accurate large-scale coestimation of sequence alignments and phylogenetic trees. *Science*, **324**, 1561–1564.

Maddison WP, Maddison DR (2010) Mesquite: a modular system for evolutionary analysis. Version 2.74. Available from http://mesquiteproject.org. Accessed 2/21/2011.

Martinson VG, Danforth BN, Minckley RL *et al.* (2011) A simple and distinctive microbiota associated with honey bees and bumble bees. *Molecular Ecology*, **20**, 619–628.

Michener CD (1974) *The Social Behavior of the Bees*. The Belknap Press of the Harvard University Press, Cambridge, Massachusets.

Mohr KI, Tebbe CC (2006) Diversity and phylotype consistency of bacteria in the guts of three bee species (Apoidea) at an oilseed rape field. *Environmental Microbiology*, **8**, 258–272.

Mohr K, Tebbe C (2007) Field study results on the probability and risk of a horizontal gene transfer from transgenic herbicide-resistant oilseed rape pollen to gut bacteria of bees. *Applied Microbiology and Biotechnology*, **75**, 573–582.

Olofsson TC, Vasquez A (2008) Detection and identification of a novel lactic acid bacterial flora within the honey stomach of the honeybee *Apis mellifera*. *Current Microbiology*, **57**, 356–363.

Packer L (1986) The social organisation of *Halictus ligatus* (Hymenoptera; Halictidae) in southern Ontario. *Canadian Journal of Zoology/Revue Canadienne de Zoologie*, **64**, 2317–2324.

Posada D, Crandall KA (1998) Modeltest: testing the model of DNA substitution. *Bioinformatics*, **14**, 817–818.

Price MN, Dehal PS, Arkin AP (2010) FastTree 2 - approximately maximum-likelihood trees for large alignments. *PLoS ONE*, **5**, e9490.

Pruesse E, Quast C, Knittel K *et al.* (2007) SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Research*, **35**, 7188–7196.

Rani A, Sharma A, Rajagopal R, Adak T, Bhatnagar RK (2009) Bacterial diversity analysis of larvae and adult midgut microflora using culture-dependent and culture-independent methods in lab-reared and field-collected *Anopheles stephensi*-an Asian malarial vector. *BMC Microbiology*, **9**, 96.

Schloss PD, Westcott SL (2011) Assessing and improving methods used in operational taxonomic unit-based approaches for 16S rRNA gene sequence analysis. *Applied and Environmental Microbiology*, **77**, 3219–3226.

Schloss PD, Westcott SL, Ryabin T *et al.* (2009) Introducing mothur: open-source, platform-independent, community-supported

software for describing and comparing microbial communities. *Applied and Environmental Microbiology*, **75**, 7537–7541.

Sen R, Ishak HD, Estrada D *et al.* (2009) Generalized antifungal activity and 454-screening of *Pseudonocardia* and *Amycolatopsis* bacteria in nests of fungus-growing ants. *Proceedings of the National Academy of Sciences*, **106**, 17805–17810.

Tajabadi N, Mardan M, Abdul Manap MY *et al.* (2011) Detection and identification of *Lactobacillus* bacteria found in the honey stomach of the giant honeybee *Apis dorsata*. *Apidologie*, **42**, 642–649.

Vasquez A, Olofsson TC (2009) The lactic acid bacteria involved in the production of bee pollen and bee bread. *Journal of Apicultural Research*, **48**, 189–195.

Vasquez A, Olofsson TC, Sammataro D, Hartfelder K (2009) A scientific note on the lactic acid bacterial flora in honeybees in the USA: a comparison with bees from Sweden. *Apidologie*, **40**, 26–28.

Wcislo WT, Gonzalez VH (2006) Social and ecological contexts of trophallaxis in facultatively social sweat bees, *Megalopta genalis* and *M. ecuadoria* (Hymenoptera, Halictidae). *Insectes Sociaux*, **53**, 220–225.

Wcislo WT, Arneson L, Roesch K *et al.* (2004) The evolution of nocturnal behaviour in sweat bees, *Megalopta genalis* and *M. ecuadoria* (Hymenoptera: Halictidae): an escape from competitors and enemies? *Biological Journal of the Linnean Society*, **83**, 377–387.

Werren JH, Windsor DM (2000) *Wolbachia* infection frequencies in insects: evidence of a global equilibrium? *Proceedings of the Royal Society of London. Series B: Biological Sciences*, **267**, 1277–1285.

Zchori-Fein E, Perlman SJ (2004) Distribution of the bacterial symbiont *Cardinium* in arthropods. *Molecular Ecology*, **13**, 2009–2016.

Zwickl DJ (2006) *Genetic Algorithm Approaches for the Phylogenetic Analysis of Large Biological Sequence Datasets under the Maximum Likelihood Criterion*. The University of Texas, Austin, Texas.

Q.S.M. is interested in how interactions with microbes, other organisms and the environment affect wild bees. W.T.W. studies evolution and behavior in social insects, mainly at the cusp of sociality. D.R.T. studies how evolution is influenced by population structure and multi-level selection, mostly in plant reproductive systems. H.D.I. is interested in how coevolution and symbioses shape the ecology of insects. S.E.D. is interested in microbial ecology in diverse environments. U.G.M. studies the ecology and evolution of cooperative and antagonistic interactions.

Data accessibility

16S rDNA amplicon reads: available on NCBI's sequence read archive (SRA), accession number SRP009083. Phylogenetic data: http://purl.org/phylo/treebase/phylows/study/TB2:S12120, also available as online supplemental material.

Supporting information

Additional supporting information may be found in the online version of this article.

Fig S1 Rarefaction curves for *Augochlora pura* samples.

Fig S2 Rarefaction curves for *Halictus ligatus* samples.

Fig S3 Rarefaction curves for *Megalopta genalis* social nest samples.

Fig S4 Rarefaction curves for *Megalopta genalis* solitary nest samples.

Fig S5 Results of BLAST searches of 16S rDNA amplicons against the SILVA database, as visualized in MEGAN4.

Fig S6 Average patristic distance from bee-associated lactobacilli to flower-associated lactobacilli and from bee-associated lactobacilli to all other lactobacilli, excluding flower-associated lactobacilli. Error bars are 95% confidence intervals.

Fig S7 Natural log of the proportional abundance of the ten most commonly observed bacterial phylotypes from four sweat bee nests.

Table S1 Halictus ligatus. Percent abundance of phylotypes per sample according to BLAST matches at >96% for species, 94% > 96% for genus, 89% > 94% for family, 85% > 89% for order, 80% > 85% for class and 77% > 80% for phylum.

Table S2 *Augochlora pura*. Percent abundance of phylotypes per sample according to BLAST matches at >96% for species, 94% > 96% for genus, 89% > 94% for family, 85% > 89% for order, 80% > 85% for class and 77% > 80% for phylum.

Table S3 Megalopta genalis social nest. Percent abundance of phylotypes per sample according to BLAST matches at >96% for species, 94% > 96% for genus, 89% > 94% for family, 85% > 89% for order, 80% > 85% for class and 77% > 80% for phylum.

Table S4 Megalopta genalis solitary nest. Percent abundance of phylotypes per sample according to BLAST matches at >96% for species, 94% > 96% for genus, 89% > 94% for family, 85% > 89% for order, 80% > 85% for class and 77% > 80% for phylum.

Table S5 Within genome variation of 16S copies.

Data S1 Alignment of 16S rDNA from lactobacilli. Sequences were both obtained from NCBI GenBank and generated for this study.

Data S2 Alignment of 16S rDNA from sweat bee-associated bacteria, used in FastTree and Fast UniFrac analyses.

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

MOLECULAR ECOLOGY

Molecular Ecology (2012) **21**, 2552–2553 doi: 10.1111/j.1365-294X.2012.05584.x

Corrigendum

QUINN S. McFREDERICK, WILLIAM T. WCISLO, DOUGLAS R. TAYLOR, HEATHER D. ISHAK, SCOT E. DOWD and ULRICH G. MUELLER

After publication, we realized that an error had occurred in the alignment of the 16S rDNA sequences that were produced for McFrederick *et al.* (2012). Specifically, after the sequences had been aligned, we overlooked a screening step that would have removed sequences that were poorly aligned and had been truncated by a filtering step. This alignment (Supporting information Data S2 in McFrederick *et al.* 2012) was used in the analysis that resulted in Fig. 3 (McFrederick *et al.* 2012). To determine if this oversight affected downstream analyses, we screened out truncated sequences, and additionally trimmed the 3 prime end of the alignment. We then reran the downstream analyses, and found that while there were minor differences between the original (Fig. 1, corrigendum) and corrected figures (Fig. 2, corrigendum), the differences were small and did not affect our interpretation of the results. In addition to the alignment that was used in the analysis for Fig. 3, we also generated an alignment with *Lactobacillus* sequences removed, and further scrutiny of this alignment confirmed that we had correctly performed the post-alignment screening steps for this file. Downstream analyses of the alignment lacking the *Lactobacillus* sequences were therefore not affected (Fig. 4 in McFrederick *et al.* 2012). We apologize for this oversight, and refer readers to the corrected FASTA file, which is available as Data S1 (Supporting information) in the online version of this corrigendum.

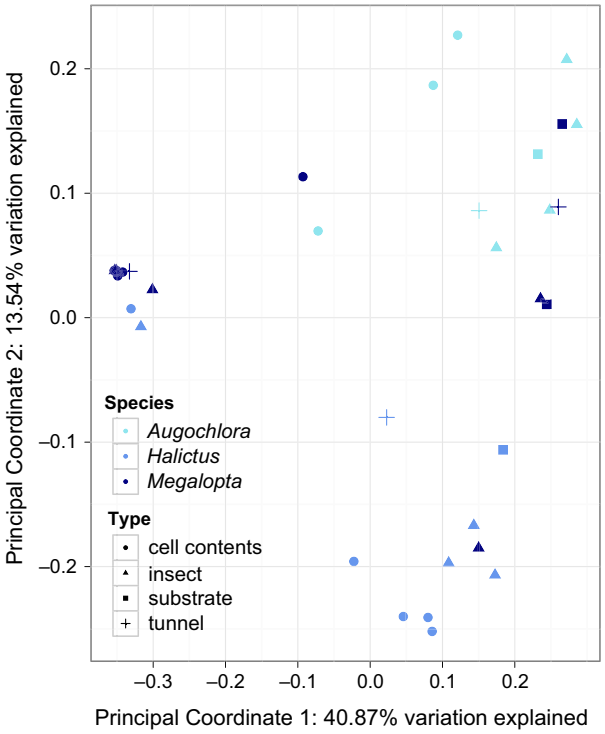

Fig. 1 Original graph, based on alignment with truncated sequences.

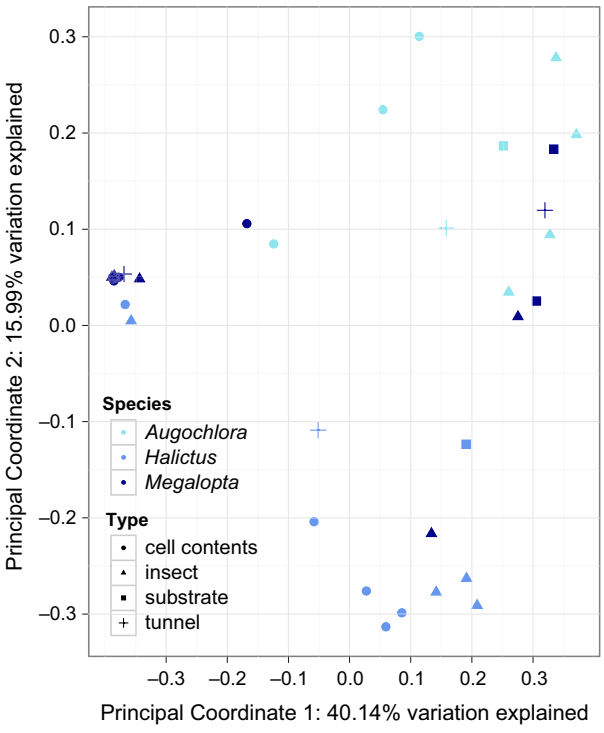

Fig. 2 Corrected graph, based on alignment with truncated sequences removed.

Correspondence: Quinn S. McFrederick, Fax: +1 (512) 471-3878; E-mail: quinnmcfrederick@gmail.com

Reference

McFrederick QS, Wcislo WT, Taylor DR, Ishak HD, Dowd SE, Mueller UG (2012) From environment or kin: whence do bees obtain acidophilic bacteria? *Molecular Ecology*, **21**, 1754–1768.

Supporting information

Additional supporting information may be found in the online version of this article.

Data S1 Corrected FASTA file.

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