Received: 14 January 2022 | Revised: 14 September 2022 | Accepted: 28 September 2022

DOI: 10.1111/mec.16736

ORIGINAL ARTICLE

MOLECULAR ECOLOGY WILEY

Incipiently social carpenter bees (*Xylocopa*) host distinctive gut bacterial communities and display geographical structure as revealed by full-length PacBio 16S rRNA sequencing

Madeline Y. Handy[1] | Dino L. Sbardellati[2] | Michael Yu[3] | Nicholas W. Saleh[4] | Madeleine M. Ostwald[5] | Rachel L. Vannette[1]

[1]Department of Entomology and Nematology, University of California Davis, Davis, California, USA

[2]Microbiology Graduate Group, University of California Davis, Davis, California, USA

[3]Department of Ecology and Evolutionary Biology, University of California, Los Angeles, California, USA

[4]Entomology and Nematology Department, Fort Lauderdale Research and Education Center, University of Florida, Davie, Florida, USA

[5]School of Life Sciences, Arizona State University, Tempe, Arizona, USA

Correspondence
Rachel L. Vannette, Department of Entomology and Nematology, University of California Davis, 367 Briggs Hall, Davis, California 95616, USA.
Email: rlvannette@ucdavis.edu

Funding information
National Science Foundation, Grant/Award Number: 1929516; University of California Natural Reserve System

Abstract

The gut microbiota of bees affects nutrition, immunity and host fitness, yet the roles of diet, sociality and geographical variation in determining microbiome structure, including variant-level diversity and relatedness, remain poorly understood. Here, we use full-length 16S rRNA amplicon sequencing to compare the crop and gut microbiomes of two incipiently social carpenter bee species, *Xylocopa sonorina* and *Xylocopa tabaniformis*, from multiple geographical sites within each species' range. We found that *Xylocopa* species share a set of core taxa consisting of *Bombilactobacillus*, *Bombiscardovia* and *Lactobacillus*, found in >95% of all individual bees sampled, and *Gilliamella* and *Apibacter* were also detected in the gut of both species with high frequency. The crop bacterial community of *X. sonorina* comprised nearly entirely *Apilactobacillus* with occasionally abundant nectar bacteria. Despite sharing core taxa, *Xylocopa* species' microbiomes were distinguished by multiple bacterial lineages, including species-specific variants of core taxa. The use of long-read amplicons revealed otherwise cryptic species and population-level differentiation in core microbiome members, which was masked when a shorter fragment of the 16S rRNA (V4) was considered. Of the core taxa, *Bombilactobacillus* and *Bombiscardovia* exhibited differentiation in amplicon sequence variants among bee populations, but this was lacking in *Lactobacillus*, suggesting that some bacterial genera in the gut may be structured by different processes. We conclude that these *Xylocopa* species host a distinctive microbiome, similar to that of previously characterized social corbiculate apids, which suggests that further investigation to understand the evolution of the bee microbiome and its drivers is warranted.

KEYWORDS
Apidae, bee microbiome, gut microbiome, PacBio sequencing, sociality

Madeline Y. Handy and Dino L. Sbardellati are co-first authors.

1 | INTRODUCTION

Gut microbial communities can be important mediators of host health and fitness (McFall-Ngai et al., 2013). Many bee species (Hymenoptera: Apoidea: Anthophila) host distinct and functionally important bacterial communities in the gastrointestinal (GI) tract (Engel et al., 2016; Lee et al., 2015; Martinson et al., 2011; Moran, 2015). As a low-diversity and tractable experimental system, the bee gut represents an excellent model to examine metabolic specialization, function and coexistence within host-associated microbial communities (Engel et al., 2016). However, bee species vary in microbiome composition, including the presence of specialized taxa and the relative abundance of environmental bacteria (Kwong et al., 2017; McFrederick et al., 2017). The factors that predict this variation among species and microbial functions remain poorly understood (Engel et al., 2016), but sociality has been proposed as an important driver of gut microbiome evolution for bees, as with other macroorganisms (Kwong et al., 2017; Moeller et al., 2016; Moran et al., 2019).

Social corbiculate bees in the subfamily Apinae ("pollen basket" bees), including honey bees and bumble bees, are characterized by distinctive gut microbial communities that are relatively consistent among individuals within a species (Kwong et al., 2017). In honey bees and bumble bees, gut bacterial communities are consistent among individuals and transmitted by social interactions (Billiet et al., 2017; Koch & Schmid-Hempel, 2011; Powell et al., 2014). By contrast, in nonsocial bee species, including those closely related to social corbiculates (McFrederick & Rehan, 2019), individuals host more variable and less distinctive microbiomes, probably driven by environmental rather than social acquisition of microbes (Cohen et al., 2020; McFrederick et al., 2012; McFrederick & Rehan, 2019). However, key tests of the sociality hypothesis using bee species in the genera *Megalopta* and *Halictus* (both of which contain solitary and social species) found limited influence of sociality on bacterial composition (McFrederick et al., 2014; Rubin et al., 2018). These results raise the possibility that other traits instead of or in addition to sociality may be more important in shaping microbiome composition and specialization among bees.

Carpenter bee species in the genus *Xylocopa* (Apidae: Xylocopini) offer a unique system to study the relative role of sociality in structuring microbiome composition. *Xylocopa* are large-bodied bees and close phylogenetic relatives of social corbiculate apids (Bossert et al., 2019). *Xylocopa* are locally common and economically important pollinators in some systems (Giannini et al., 2015; Keasar, 2010), and nest in timber or dead stalks of plants (Barrows, 1980). Of particular note, several species of carpenter bees have been characterized as facultatively or incipiently social (Gerling et al., 1989; Michener, 1990). In characterized species, the oldest female in a nest maintains reproductive dominance and feeds younger nestmates via trophallaxis (Lucia et al., 2015), but cooperative brood care is rarely documented. One species in which sociality has been well-studied is *Xylocopa sonorina*, which lives in small, fluid, dominance-based societies with reproductive division of labour, where the proportion of individuals nesting socially is temporally dynamic (Ostwald et al., 2020). Moreover, in all *Xylocopa* species in which sociality has been examined, both social and solitary nests are present within the same population (Gerling et al., 1989). Investigation of the microbiome of carpenter bees—close relatives of corbiculates with contrasting social structure—may offer insights into the role of sociality in the evolution of the bee microbiome. Recent sequencing efforts in *Xylocopa tenuiscapa* document gut communities dominated by Lactobacillales, Enterobacteriaceae and Bifidobacteriacae (Subta et al., 2020), while three of the four *Xylocopa* species from the central and eastern USA hosted *Bifidobacterium*, *Bombiscardovia*, *Bombilactobacillus*, *Apilactobacillus* and *Lactobacillus*, in striking similarity to social corbiculate microbiomes (Holley et al., 2022). These results suggest that the bacterial communities of some *Xylocopa* may resemble those of social corbiculates, but it is unclear if other *Xylocopa* species with well-studied social structure exhibit similar microbiome composition. Further, geographical or population-level variation in microbial communities within a species may inform processes that shape microbiome composition among and within *Xylocopa* species.

Here, we examine the composition of bacterial full-length 16S rRNA genes in *X. sonorina* Smith [previously *X. varipuncta* Patton] and *Xylocopa tabaniformis orpifex* Smith (Bezark, 2013), two carpenter bee species common in western North America. These *Xylocopa* species often co-occur locally, collect nectar and pollen from the same plant hosts, and show a similar seasonal phenology in activity and reproduction. We utilize PacBio Sequencing and a sample inference method with single-nucleotide resolution (Callahan et al., 2019) to examine phylogenetic relationships among bacterial variants and previously characterized microbial taxa. We examined microbiome composition in two tissues, the crop (foregut) and gut (combined midgut and hindgut), which are disparate in function and separated by a proventricular valve. We hypothesized that the crop would be variable in bacterial composition among individuals, due to frequent intake of food including pollen, low microbial biomass and a predominance of environmentally sourced microbes (Anderson et al., 2013). We hypothesized that if *Xylocopa* is similar to social apids, the gut would host a core microbiome distinctive from the crop that was consistent among individuals (Moran, 2015), or if *Xylocopa* microbiomes resemble solitary bees sampled to date, the gut would host a variable microbiome with high similarity to the crop (Voulgari-Kokota et al., 2019).

To address these hypotheses, we sampled bees from three geographical locations and first compared how bacterial alpha and beta diversity differ among tissues, species and geographical locations. We also examined if sex or foraging status was associated with bacterial composition. Next, we defined the core bacterial taxa for these species and examined phylogenetic patterns among variants within these core clades. Lastly, to examine if shorter regions of the 16S rRNA gene could also detect these patterns, we repeated analyses using the full-length data that had been trimmed to the V4 region only. Our results suggest that carpenter bees host distinctive gut bacterial communities including bacterial

clades previously detected in corbiculates. In addition, *Xylocopa* species host phylogenetically and geographically distinct lineages within core clades, which are revealed by the full-length 16S rRNA but not by the V4 region alone.

2 | MATERIALS AND METHODS

2.1 | Sample collection

Between 2019 and 2020, 33 *Xylocopa sonorina* and 22 *X. tabaniformis* adults were collected. Bees were captured in one of three ways: netted while foraging, caught using traps over the nest entrance (Figure S1) or through excavation of logs to sample entire nests. The type of capture was recorded, with bees caught foraging in flight denoted as "foraging" and those captured within or exiting a log denoted as "nest-caught." Samples were obtained from Davis, California (21 *X. sonorina*; 14 *X. tabaniformis*), Anza-Borrego Desert State Park in southern California (one *X. sonorina*, eight *X. tabaniformis*) and Tempe, Arizona (11 *X. sonorina*). Samples from Davis and Anza-Borrego were captured during late summer and early autumn while samples from Tempe were collected in early summer (see data file for collection date). Bee species can be distinguished morphologically, and identifications were verified using voucher specimens at the Bohart Museum of Entomology. Captured carpenter bees were photographed, then killed by placing them in a −20°C freezer where they were stored until dissection.

2.2 | Sample processing and DNA extraction

Before dissection, carpenter bee samples were rinsed in 70% ethanol for 15s. They were then air dried and placed in a sterile Petri plate for dissection. Dissecting tools were flame sterilized before dissection and before each organ removal. The crop and the rest of the gut (combined midgut and hindgut) were separated and stored at −20°C until DNA extraction.

Microbial DNA was extracted from individual crop and gut samples separately and kit reagents only as a blank extraction control, using the Qiagen DNeasy PowerSoil Kit with slight modifications (Rubin et al., 2014). Modifications include adding four magnetic beads per PowerBead Tube after tissue samples had been added and beating tubes in a BeadBlaster 24 Homogenizer for three cycles of speed 7 for 20s per cycle. Then 60 µl of Solution C1 and 2 µl Proteinase K solution (600 mAU ml^{-1}; Qiagen Tissue and Blood) were added to each tube and tubes were incubated overnight at 56°C. The following day tubes were beaten once more using the same cycle settings and the rest of the protocol followed the manufacturer's protocol beginning at step 6. Extracted DNA was sent for library preparation and sequencing at Dalhousie University Integrated Microbiome Resource Facility. Briefly, the full-length 16S rRNA region was amplified in duplicate using 27F AGRGTTYGATYMTGGCTCAG (Paliy et al., 2009) and 1492R RGYTACCTTGTTACGACTT (Lane, 1991) using full-length fusion primers (PacBio adapters + barcodes + specific regions). PCR (polymerase chain reaction) products were visualized using a Hamilton Nimbus Select robot using Coastal Genomics Analytical Gels. PCR products were pooled within a sample, cleaned and normalized using a Charm Biotech Just-a-Plate 96 Well Normalization kit and quantified fluorometrically before sequencing. PacBio Sequel 2 chemistry was used in sequencing, performed by the Dalhousie University Integrated Microbiome Resource facility (Halifax, Canada).

2.3 | Bioinformatics

Preliminary processing and filtering of raw full-length 16S rRNA reads into amplicon sequence variants (ASVs) was performed in R version 4.1.0 (R Core Team, 2021) using DADA2 (version 1.20.0; Callahan et al., 2019). Primer sequences were removed, and reads were filtered by size and quality to yield sequences ranging from 1000 to 1600 bp with no ambiguous bases, two maximum expected errors and a minimum quality score of 3. Filtered reads were then dereplicated, and sequencing errors were inferred using the PacBioErrfun function and removed. Chimeras were inferred with a minFoldParentOverAbundance value of 3.5 and removed using sequence consensus as a method. Finally, taxonomy was assigned using the BEExact database (Daisley & Reid, 2021) and SILVA version 138.1 (Quast et al., 2012); the resulting taxonomy was often identical but BEExact allowed more specific annotations for a few groups, so we present assignments from BEExact below.

The ASV and taxonomy tables generated from the DADA2 pipeline outlined above were merged with metadata using PHYLOSEQ (McMurdie & Holmes, 2013). ASVs classified as chloroplast or mitochondria were removed and any ASVs found in the control sample were removed, resulting in the removal of a single ASV annotated as *Entomomonas* that was only found in the negative control. The final data set included only samples with more than 500 total sequences (see sampling curves in Figure S2).

2.4 | Statistical analysis

2.4.1 | Alpha and beta diversity

All statistical analyses were performed using R version 4.1.0. We examined rarefaction curves to assess if sampling depth was adequate and all samples plateaued (Figure S2). Alpha diversity was quantified using Chao1 as an estimate of species richness, and Shannon diversity which also accounts for evenness of ASVs in a sample. To determine if host species, tissue type, their interactions, or sex and foraging type significantly affect diversity, a linear mixed effect model was constructed using *lmer* (Bates et al., 2015), with separate models for Chao1 and Shannon diversity. In both models, individual bee was included as a random effect and *p*-values were calculated using the *lmerTest* package (Kuznetsova et al., 2017).

To examine drivers of species composition among samples, we first examined if predictor variables, including species, tissue, geographical location or sex contributed significantly to differences in species composition, using permutational ANOVA (PERMANOVA) using *adonis2* and assessed variation in dispersion among groups using *betadisper* (Oksanen et al., 2012). We present results using Bray–Curtis dissimilarities in the main text, but results are nearly identical to those based on weighted Unifrac dissimilarities (both presented in Table S1). Differences between GI tract tissues were confirmed using PERMANOVA on a subset of the data containing only bees from which both the crop and gut sequenced successfully.

Because not all species of bee were sampled at each site, we conducted additional analyses to validate detected species and geographical effects in our data set using stratified PERMANOVA. First, we subset the data set to only samples collected in Davis, where both bee species were sampled, and examined effects of bee species. Second, we examined the effect of sampling location separately for each bee species. In both cases, we used PERMANOVA implemented in *adonis2*.

Additionally, to examine if bees caught foraging or nesting differed in bacterial diversity or composition, we repeated all alpha and beta diversity analyses on only *X. sonorina* samples, the only bee species for which we had nesting and foraging individuals. Models included sampling location, tissue, behaviour and sex.

2.4.2 | Assessing bacteria that distinguish bee species and are core to *Xylocopa*

To identify bacterial taxa that differ between host species and tissue type, we compared relative abundance of ASVs between bee species and tissues using DESEQ2 (Love et al., 2013), with a false discovery rate (FDR) <0.05.

To identify core gut bacterial communities, individual ASVs were first collapsed at the genus–species level using *tax_glom()*. The function *eulerr::core_members()* was then used to identify genus–species bins present at ≥0.01% relative abundance in at least 95% of gut samples and heatmaps were constructed to visualize the number of shared core taxa for both *Xylocopa* species together and separately.

2.4.3 | Phylogenetic tree construction

We constructed phylogenies to compare our sequences to previously characterized bacteria from bees or other related habitats for Lactobacillaceae and Bifidobacteria. Previously published 16S rRNA sequences were downloaded from NCBI's 16S rRNA RefSeq database. Our list consisted of previous studies of corbiculate apid and non-*Apis* bee-associated bacteria with published full-length 16S rRNA sequences or genomes, as well as related outgroups (Lugli et al., 2017; Zheng et al., 2020).

To build trees, sequences were aligned using *AlignSeqs* (Wright, 2015). Pairwise distances of aligned sequences were calculated and used in the construction of first a neighbour-joining tree then fitting a GTR+G+I maximum-likelihood tree using the neighbour-joining tree as a starting point (Schliep, 2011) and visualized using GGTREE (Yu et al., 2017).

2.4.4 | Assessing differences between full-length and V4 data sets

To generate a short-read data set from our full-length sequences, we repeated the DADA2 pipeline but used V4 universal bacterial primers (V4_F = GTGCCAGCMGCCGCGGTAA V4_R = GGACTACHVGGGTWTCTAAT) in place of PacBio sequencing primers, and then repeated all downstream bioinformatics steps. To compare inference between data sets, we produced new phylogenetic trees and repeated beta diversity analyses using the V4 data set. The effects of different predictors on the beta diversity of our V4 data set were assessed via *adonis* PERMANOVA tests, as with full-length sequences.

2.4.5 | Comparative visualization

To contextualize our sequencing results, we added data from the current study to a previously published summary table (Shell & Rehan, 2022). Briefly, we calculated mean relative abundance of bacterial genera present in *X. sonorina* and *X. tabaniformis* in more than 50% of samples with an average relative abundance of >0.5%. To compare among bee taxa, the top bacterial genera by abundance were selected from 71 additional species spanning 13 bee tribes, adding bacterial composition and count data from additional data sets to the level of genus (Figueroa et al., 2021; Holley et al., 2022; Shell & Rehan, 2022; Tola et al., 2021). Bacterial genera were classified as "Corbiculate core," "Corbiculate associates and Lactobacillus" or "Environmental or insect-associated" following previous work (Kwong et al., 2017).

3 | RESULTS

3.1 | Sequencing results

From the 110 bee samples submitted, one negative control and one extraction liquid sample for PacBio long amplicon sequencing, 532,208 raw reads and 326,740 high-quality sequences were retained after processing. From these sequences, 811 unique ASVs were inferred, 80 of which were classified as chloroplast, mitochondria or present in the negative control and were removed. The samples containing fewer than 500 total sequences (17, all crop and extraction or negative controls) were removed. Overall, the pipeline

resulted in a final data set containing 731 ASVs and 93 samples (55 gut and 38 crop) with a median of 2472 sequences per sample. Sampling curves (Figure S2) were saturating, indicating sufficient sampling depth. The most abundant bacterial genera within *Xylocopa* crops included *Apilactobacillus*, *Acinetobacter* and *Saccharibacter*, while the gut was dominated by bacterial genera often documented in social corbiculates including *Apibacter*, *Bombilactobacillus*, *Bifidobacterium*, *Bombiscardovia* and *Gilliamella*, as well as *Entomomonas*, *Fructobacillus* and members of the Lactobacillaceae (Figure 1).

3.2 | Drivers of alpha and beta diversity in *Xylocopa* microbiomes

Xylocopa gut samples hosted ~219% more ASVs and had greater evenness compared with crop samples (Figure 2; Chao1 lmer $F_{1,54} = 151$, $p < .001$; Shannon lmer $F_{1,45} = 169$, $p < .001$) across bee species. *Xylocopa sonorina* hosted marginally greater species richness than *X. tabaniformis* ($F_{1,56} = 4.18$, $p = .046$) and significantly greater diversity ($F_{1,51} = 16.43$, $p < .001$), while the interaction of tissue and species was not significant (Chao1 $F_{1,54} = 0.16$, $p = .69$; Shannon $F_{1,45} = 0.029$, $p = 0.87$). Within *X. sonorina*, female bees had significantly greater species richness and diversity within the gut than male bees (Chao1 $p = 0.001$, Shannon $p = 0.04$; Figure S3), but no difference was detected in *X. sonorina* between sexes in the crop (Chao1 $p > .05$; Shannon $p > .05$), nor between sexes in either tissue for *X. tabaniformis* ($p > .2$).

Bacterial species composition varied with sampling location, tissue type, species, their interactions as well as the sex of bees (Figure 2; full model Bray–Curtis PERMANOVA $p < .05$ for all; Table S1, Figure S4). Paired analyses conducted on bees where both the crop and gut were sequenced confirmed that crop and gut communities differed in both bee species (Table S2; *X. sonorina* Bray $F = 5.63$, $R^2 = .098$, $p < .001$; *X. tabaniformis* Bray $F = 3.87$, $R^2 = .16$, $p < .001$). The gut and crop communities differed in dispersion. Using Bray–Curtis dissimilarities, the crop samples were more variable than the gut for both species (betadisper *X. sonorina*, $p < .05$; *X. tabaniformis* $p < .01$; Figure 2c; Figure S4).

In our validation analysis, species differed in bacterial composition when samples from a single location where both species were sampled were considered (Davis), for both the crop (Table S3, Bray $R^2 = .10$, $F = 2.56$, $p = .01$) and the gut tissues (Bray $R^2 = .21$, $F = 8.53$, $p < .001$).

3.3 | Geographical variation within species

Collection location and its interaction with tissue type were highly significant in the full model (Table S1) but to account for unbalanced sampling and understand drivers of these interactions, we also compared the effect of location within each species and tissue type separately. For *X. sonorina*, samples collected in Davis and Tempe differed in species composition for both the crop and gut (Table S4; Bray–Curtis crop $R^2 = .16$, $F = 4.47$, $p < .001$; gut $R^2 = .094$, $F = 3.10$ $p < .001$). For *X. tabaniformis*, samples from Davis and Anza-Borrego differed in species composition, but only significantly in the gut, but to a lesser extent in the crop (Bray–Curtis dissimilarity gut $R^2 = .17$, $F = 4.18$ $p < .001$, crop $R^2 = .15$, $F = 1.58$ $p = .072$).

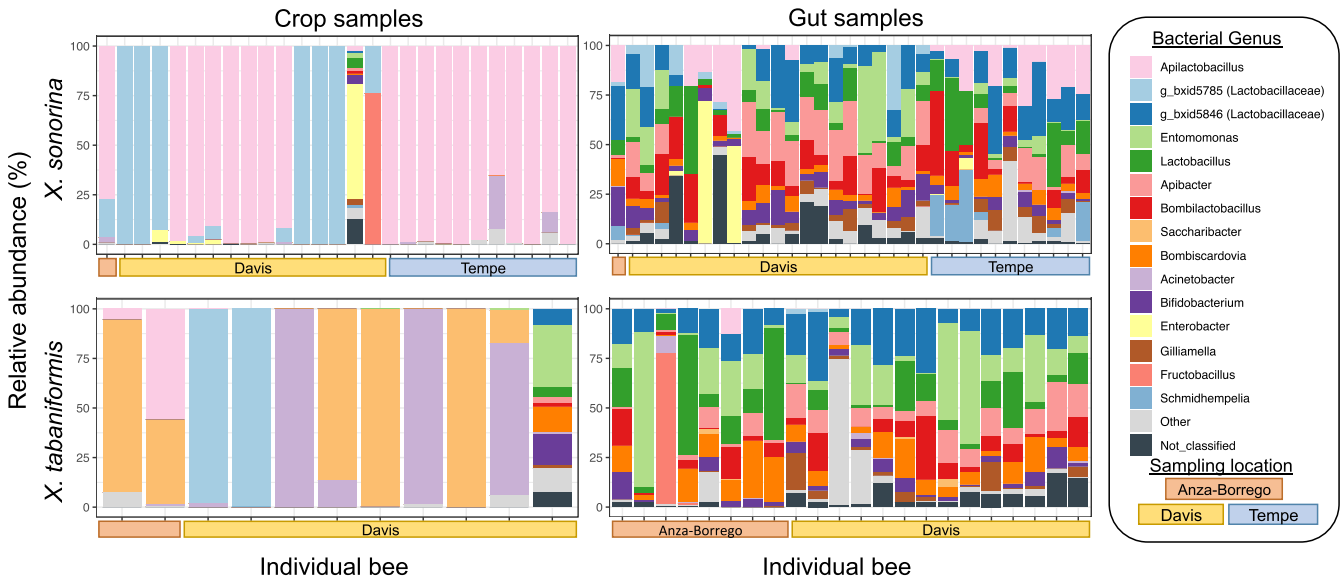

FIGURE 1 Comparison of the 15 most abundant bacterial genera across all samples in the crop and gut in *Xylocopa sonorina* and *X. tabaniformis* described using full-length 16S rRNA PacBio sequencing. Vertical bars represent individual bee samples, for either the bee crop or gut (midgut and hindgut). Bars under individual bee samples indicate the geographical location where the bee was collected. Individual sequence variants (ASVs) were collapsed at the genus level. Category "other" includes ASVs classified but not in the most abundant 15 genera. Category "Not classified" includes ASVs not classified to the genus level. [Colour figure can be viewed at wileyonlinelibrary.com]

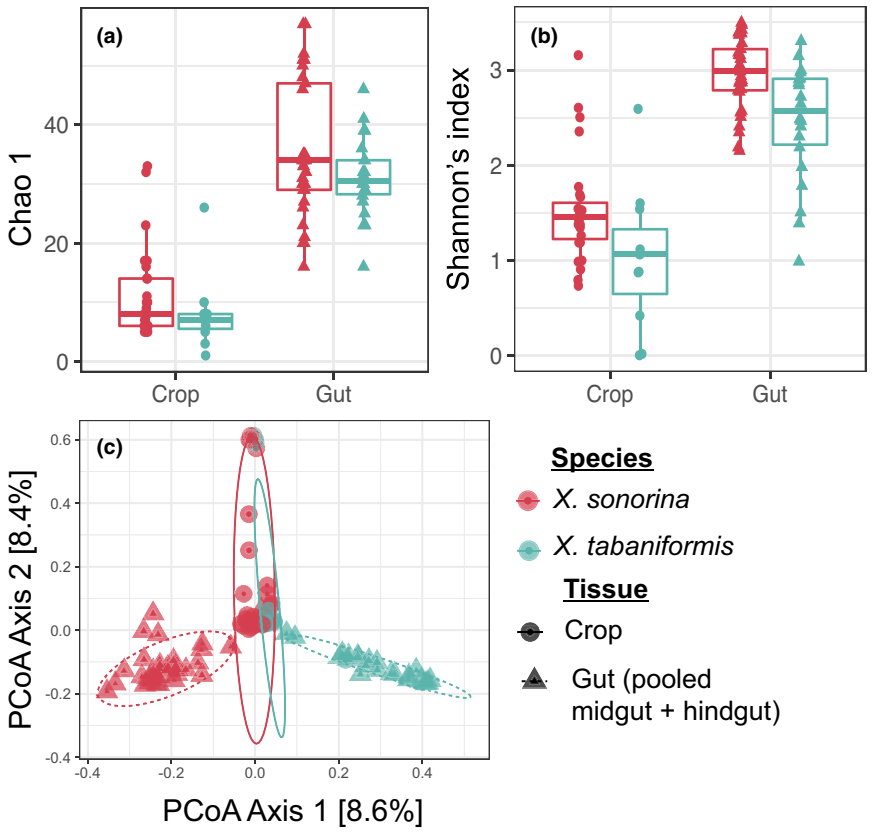

FIGURE 2 Comparison of alpha and beta diversity between bee species, *Xylocopa sonorina* and *Xylocopa tabaniformis*, and between crop and gut tissues (see text for *p*-values). (a) Bacterial species Chao 1 richness and (b) Shannon's index diversity. (c) Species and tissues differ in beta diversity estimated using Bray–Curtis dissimilarity, with the gut communities tightly clustering and differentiated between bee species. [Colour figure can be viewed at wileyonlinelibrary.com]

3.4 | Effects of foraging behaviour on bacterial composition

Within *X. sonorina* (the only species in which we sampled bees from nests and while foraging), bees caught foraging had greater richness compared to nest-caught bees ($F_{1,49} = 4.23$, $p = .045$), and with a trend towards higher diversity (Figure S3; $F_{1,26} = 2.80$, $p = .11$). Bacterial composition differed between foraging bees and nest-caught bees (PERMANOVA $p = .009$) and between male and female bees ($p = .04$) as well as with location and tissue (both $p < .001$). No male bees were captured while foraging and a significant interaction between tissue and foraging type was observed ($p = .021$). Foraging bees exhibited a lower distance to the centroid compared to nest-caught bees (betadisper behaviour $p = .022$) and male bees had a lower distance to the centroid compared to females (betadisper sex $p = .0046$). Notably, the crop of nest-inhabiting bees was comprised nearly entirely *Apilactobacillus*, while the crop of foragers varied among individuals, with some foraging bees containing a high abundance of known nectar inhabitants including *Acinetobacter nectaris* and *Neokomagataea thailandica*.

3.5 | Bacterial ASVs that distinguish species and tissue types

In the crop, 15 ASVs differed significantly between *X. sonorina* and *X. tabaniformis*, 13 of which were classified as *Apilactobacillus* (DESEQ2 FDR < 0.05, Figure S5). In the gut, 58 ASVs differed between *Xylocopa* species, 15 of which were classified as Bifidobacteriaceae (mostly *Bombiscardovia*), 27 as Lactobacillaceae (mostly *Bombilactobacillus*) and six as Orbaceae (*Gilliamella*) (DESEQ2 FDR < 0.05, Figure S5).

In addition ASVs also differentiated the crop and gut in *X. sonorina* and *X. tabaniformis* (Figure S6). The crop was enriched in *Apilactobacillus* (*X. sonorina* only) and another undefined genus within the Lactobacillaceae (both species). In contrast, the gut of both species was enriched in many ASVs including from the genera *Entomomonas*, *Bombiscardovia*, *Bifidobacterium* and *Apibacter*, with additional genera differentially abundant depending on the bee species considered, with more taxa enriched in the gut within *X. sonorina*.

3.6 | Core *Xylocopa* gut bacteria

Of the genus–species bins present in the gut of *Xylocopa* species (Figure 3), *Bombiscardovia coagulans*, *Bombilactobacillus bombi* and *Lactobacillus bxid5692* (BEExact ID number) were present in at least 95% of gut samples of both species as well as in *X. sonorina*. In addition to these taxa, *X. tabaniformis* also contained *Entomomonas*, bxid5846_bxid5850 (Lactobacilliaceae) and *Bifidobacterium* in 95% of samples.

3.7 | Phylogenetic analysis: core taxon ASVs display a high degree of host specificity

Phylogenetic and variant-level analysis suggests that the majority of the ASVs comprising core gut bacteria are host-specific (Figures 3

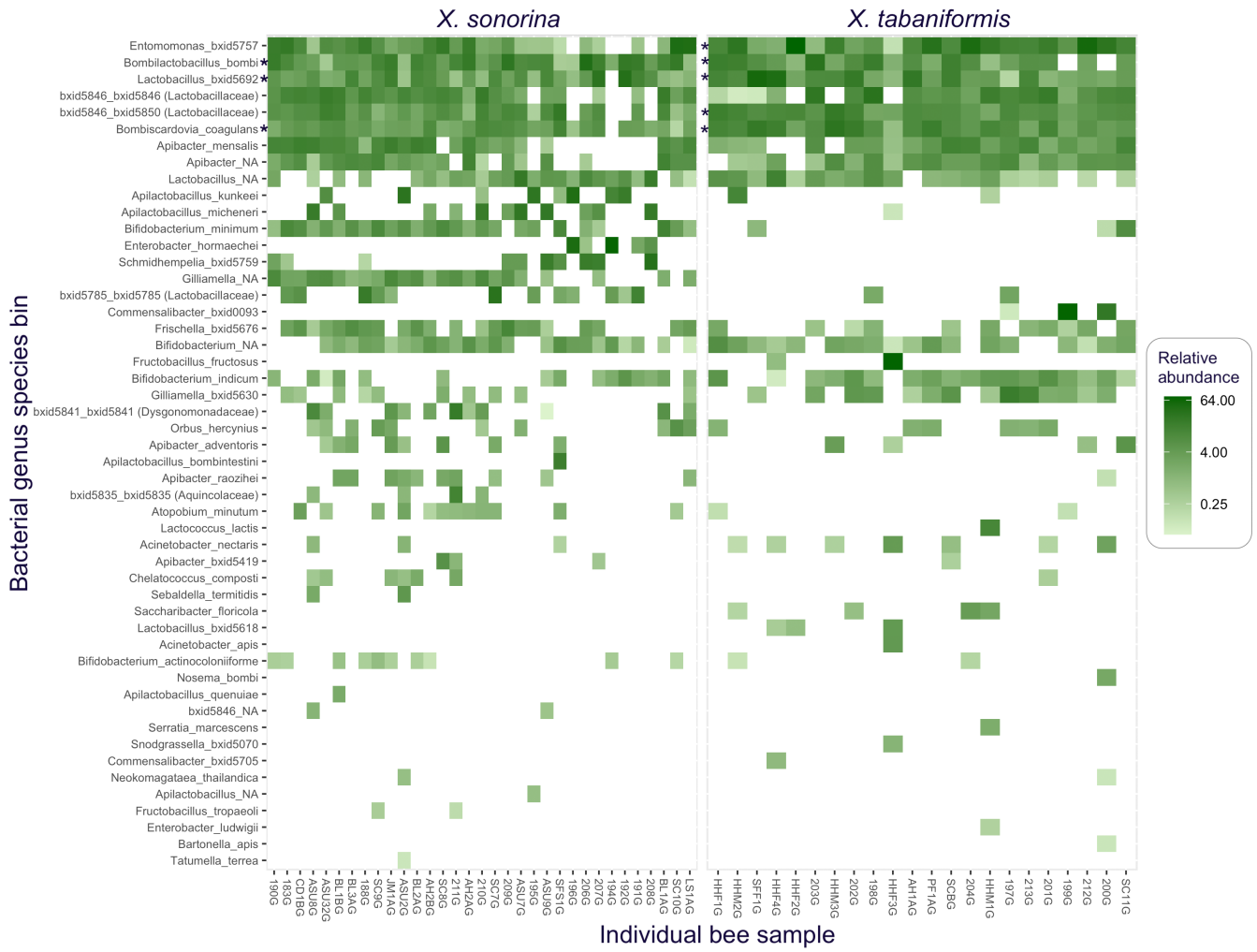

FIGURE 3 Heatmap of the relative abundance of bacterial genus–species bins from *Xylocopa sonorina* and *X. tabaniformis* gut samples. Bacterial taxa are ordered by descending mean relative abundance. Color refers to the relative abundance (out of 100) in an individual sample. Genus–species bins that were designated as core taxa (see main text for details) are indicated with stars. [Colour figure can be viewed at wileyonlinelibrary.com]

and 4; Figures S7 and S8). In contrast only a few ASVs span host species and locations, including the most abundant *Lactobacillus* sequence (ASV9). Moreover, while some ASVs are found at multiple geographical sampling locations, some are highly location-specific, particularly in *Bombilactobacillus* and *Bombiscordovia* (Figures S7 and S8).

Within the Lactobacillaceae, we detected ASVs closely related to corbiculate symbionts and those more closely related to taxa found across many social and solitary bee species in the Apilactobacillaceae (Figure 4). Most ASVs were limited to a single species (Figure 4) and a single population (Figure S7), although species-specificity varied among clades. Within the Bifidobacteriaceae, ASVs annotated as *Bombiscardovia coagulans* were closely related to isolates from *Bombus* and *Xylocopa violacea* (Figure 5). Sequence variants were exclusively restricted to a single species and were sometimes location-specific (Figure S8).

3.8 | Comparison between long and short amplicon reads

Amplicon length influenced bacterial diversity and ecological inference. When only V4 regions were considered, 174 unique ASVs were detected (637 fewer than the full-length 16S rRNA data set at this

FIGURE 4 Phylogenetic placement of Lactobacillaceae ASVs detected in *Xylocopa sonorina* or *X. tabaniformis* crop or gut samples, with members of the family previously isolated from social corbiculate bees (either *Apis* or *Bombus*), noncorbiculate bees, flowers, and other hosts and environments, and outgroup following Zheng et al. (2020). Phylogenetic relationships were first estimated using neighbour joining, then optimized using maximum likelihood, where nodes with >75% bootstrap support from 1000 resampled trees are indicated with a black point. Branch length is proportional to the number of sequence differences. Only ASVs present at >5% relative abundance in at least a single sample are included. [Colour figure can be viewed at wileyonlinelibrary.com]

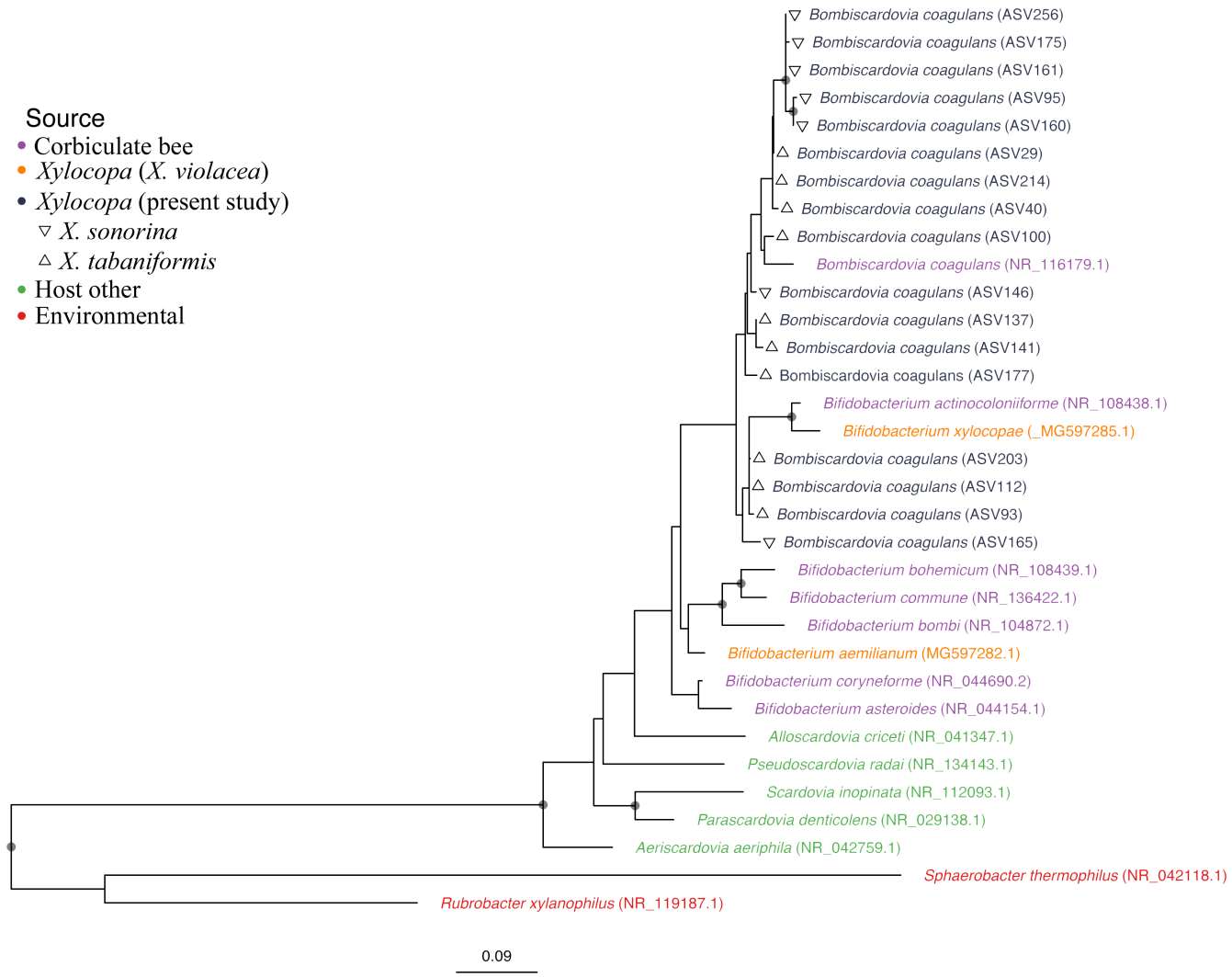

FIGURE 5 Phylogenetic placement of gut *Xylocopa*-associated Bifidobacteriaceae ASVs, with outgroups selected from Lugli et al. (2017) and other related taxa including species isolated from social corbiculate bees (*Bombus* or *Apis*) or other carpenter bees (*Xylocopa violacea*). The neighbour-joining phylogenetic tree was optimized with internal maximum likelihood, where nodes with >75% bootstrap support from 1000 resampled trees are indicated with a black dot. Branch length is proportional to the number of sequence differences. Only ASVs present at >2.5% relative abundance in at least a single sample are included. [Colour figure can be viewed at wileyonlinelibrary.com]

step), 42 of which were classified as chloroplast, as mitochondria or were detected in the negative controls and removed. Finally, 18 samples containing fewer than 500 sequences were removed, resulting in a final V4 data set containing 92 samples with a median of 2959.5 sequences per sample. Reduced sequence length in the V4 region resulted in the clustering of sequence variants from the full-length data set: ASVs that in the full-length data set distinguished host species or geographical location were considered a single ASV in the V4 data set (Figures S9 and S10). However, despite a lower number of sequences, beta diversity inference was quite similar when either full-length or V4 regions was considered, with interactions among sex and sampling location less pronounced in the V4 data set (Tables S1–S4).

In addition, V4-region phylogenies contain fewer ASVs, masking species and population-level differentiation in variants revealed by the full-length 16S rRNA region (Figures S9 and S10).

4 | DISCUSSION

The microbiome of the incipiently social *Xylocopa* species examined here is similar to previously characterized social corbiculates in at least two ways. First, *Xylocopa* individuals host a distinctive set of bacterial taxa found consistently in bees that span geographical sampling locations, sexes and individuals displaying different behaviours (nest-caught vs. foraging), unlike most previously characterized solitary bees (Voulgari-Kokota et al., 2019). Second, many of the bacterial lineages detected at high abundance in the gut have been previously described in social corbiculates (Figure 6). Below, we discuss the specific taxa found and the implications for the effect of sociality on the evolution of the bee microbiome.

Many bacterial taxa in the gut of most *Xylocopa* are considered core bacterial symbionts of corbiculate bees (Figures 4–6;

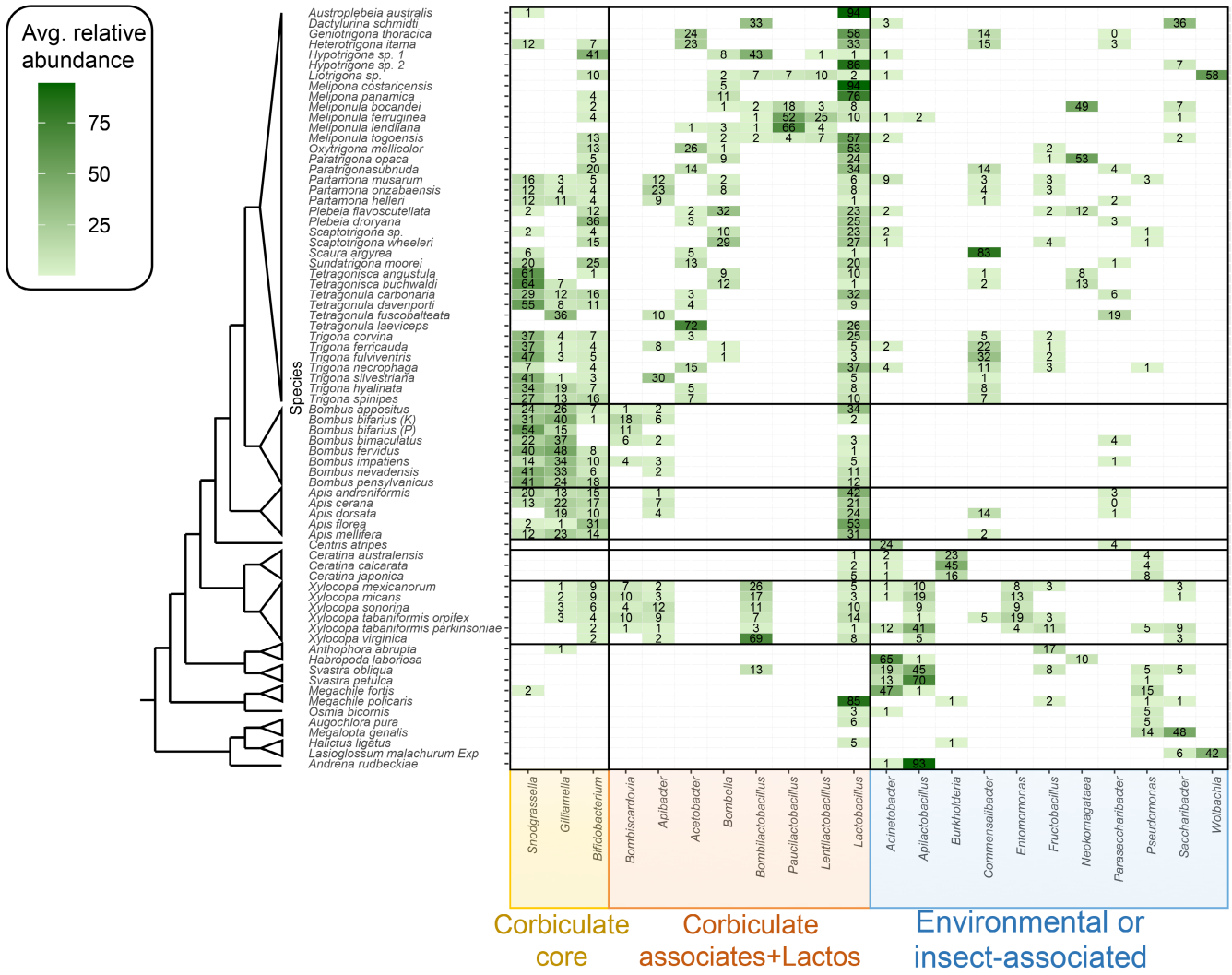

FIGURE 6 Average relative abundances of bacterial genera across diverse bee species complied from Kwong et al. (2017), Shell and Rehan (2022) and the current study. Gut bacterial communities of bee species, including bacterial genera consistently present within corbiculate apids (here, "Corbiculate Core"), those often present within corbiculate apids ("Corbiculate associates + Lactobacillus," see note below), and environmental or insect-associated genera. Bacterial genera were assigned to categories following Kwong et al. (2017) and Shell and Rehan (2022). The 22 most abundant bacterial genera across studies are shown. Bacterial ASVs or operational taxonomic units (OTUs) were aggregated at the genus level based on taxonomy in each study. Average relative abundance values are rounded to the nearest whole number, with a minimum average relative abundance >0.5% displayed. We note that the genus *Lactobacillus* includes taxa that were separated into new genera in newer studies but not prior to taxonomy changes (*Apilactobacillus*, *Bombilactobacillus*, etc.), and *Pseudomonas* probably includes taxa in the newly created genus *Entomomonas* in some but not all studies. Data from Tola et al. (2021), Figueroa et al. (2021) and Holley et al. (2022) were added to previously summarized data in Shell and Rehan (2022). The number of bee individuals sampled per species was typically at least five but often higher (see referenced studies for details). Some previous studies were omitted due to insufficient microbial taxonomic resolution. Bee phylogeny (on the *y*-axis) illustrates phylogenetic relationships among bee tribes so branch lengths are not intended to be interpreted quantitatively. Bee species are listed within the tri Meliponini, Bombini, Apini, Centridini, Ceratinini, Xylocopini, Anthophorini, Eucerini, Megachilini, Osmiini, Augochlorini, Halictini and Andrenini (from top to bottom) following Hedtke et al. (2013) and Gibbs et al. (2012). [Colour figure can be viewed at wileyonlinelibrary.com]

Subta et al., 2020, Holley et al., 2022). *Bombilactobacillus* (previously *Lactobacillus* Firm4), *Bombiscardovia* (Bifidobacteriaceae) and *Lactobacillus bxid5692* (related to *Lactobacillus* Firm5) were detected in nearly every individual, similar to previously characterized *Apis* and *Bombus* species (Killer et al., 2010; Kwong et al., 2017; Parmentier et al., 2018), suggesting these genera have a larger host range than previously recognized (Alberoni et al., 2019; Hammer et al., 2021; Kwong et al., 2017). In *Bombus terrestris*, *Bombilactobacillus* and *Bifidobacterium* sp. are socially transmitted (Billiet et al., 2017) and colonize the hindgut, and in *Apis mellifera* are implicated in saccharide breakdown and fermentation (Lee et al., 2015). Future studies are required to examine if social transmission and bacterial function are similar in *Xylocopa*.

Despite similarities to the *A. mellifera* and *Bombus* microbiome (Figure 6), the *Xylocopa* microbiomes characterized here are distinct from those of *Apis* and *Bombus* in a few ways. First, most *Xylocopa*

consistently host *Apibacter* (Bacteriodetes) and *Entomomonas* (Proteobacteria), in contrast to their sporadic presence in *Apis* (Kwong et al., 2018) and *Bombus* (Figure 6). These bacteria contain limited metabolic capability (Kwong et al., 2018; Wang et al., 2020). Notably, *Xylocopa* also notably lacks *Snodgrassella* (we detected a single ASV in low abundance) and hosts very low relative abundance of *Gilliamella*, similarly to recently characterized social stingless *Melipona* species (Cerqueira et al., 2021) and some members of the Meliponini (Figure 6). In *X. sonorina*, the crop was highly dominated by *Apilactobacillus* (Figure 1; Figures S5a and S6a), a common bacterial taxon found in solitary bee gut and provisions and in the crop of bees (McFrederick et al., 2018). *Apilactobacillus* have been hypothesized to inhibit pathogen growth or prevent spoilage of stored pollen (Kapheim et al., 2021; McFrederick et al., 2018; Vásquez & Olofsson, 2009).

Since *Xylocopa* are incipiently social rather than classically eusocial (Gerling et al., 1981), their distinctive microbiome raises questions about the mechanisms required for effective social transmission of the microbiome. Dominant *Xylocopa* females feed newly emerged nestmates via trophallaxis, and allow consumption of the stored provision (Gerling et al., 1989; Ostwald et al., 2021; Vickruck & Richards, 2021). In this way, *Xylocopa* exhibit behavioural similarities to *Apis*, which engage in trophallaxis, and *Bombus*, which do not engage in trophallaxis yet feed from shared food resources and engage in coprophagy (Näpflin & Schmid-Hempel, 2016). In addition, *Xylocopa* individuals migrate among nests, including those of nonkin (Ostwald et al., 2021, Vickruck & Richards, 2021). We hypothesize that the feeding and reproductive biology of *Xylocopa*, as well as the relatively long lifespan (1–2 years), and large body size could help explain the maintenance of specialized microbial taxa. Differences between *Xylocopa* species in microbiome composition and richness also suggest areas for future study. Although we lack details on *X. tabaniformis* sociality, this species is probably incipiently social like *X. sonorina* (Breed, 1976) yet differs in breeding systems: males patrol flowers for mating, while X. *sonorina* males host nonresource-based territories (Marshall & Alcock, 1981). This could explain the differences in sex-specific microbiomes between species. Additional information on social structure or nesting biology may inform microbiome differences observed between *X. tabaniformis* and *X. sonorina*, and *X. virginica* and the other characterized *Xylocopa* (Holley et al., 2022).

It is notable that the bee genus *Ceratina*, a sister clade to *Xylocopa*, is mostly subsocial, interacting with the developing brood and occasionally overwintering with adult offspring (Rehan & Richards, 2010), yet hosts a fundamentally distinct microbiome lacking corbiculate core or associated taxa (Figure 6). In previous work, the microbiome of halictid bees that nest socially, including those that exhibit trophallaxis and eusociality (Kapheim et al., 2016), did not differ substantially from those that nest nonsocially (McFrederick et al., 2014; Rubin et al., 2018), and resembled solitary bee microbiomes (Voulgari-Kokota et al., 2019), which are primarily environmentally acquired (Kapheim et al., 2021; McFrederick et al., 2012). Combined, these studies suggest that not only social behaviours but perhaps also additional other biological or ecological differences may be required for the maintenance of a distinctive bee microbiome.

Geographical populations *of Xylocopa* were differentiated by *Bombilactobacillus* and *Bombiscardovia* ASVs (Figures S7 and S8), and the bacterial genus *Schmidhempelia* was frequently present in southern but not northern populations of *X. sonorina* (Figure 1). Sampling location explained 9%–16% of the variation in bacterial composition, more than previous studies examining geographical signatures in honey bees (Ge et al., 2021), stingless bees (Liu et al., 2021) and even some solitary *Osmia* (Rothman et al., 2020). Although long-read sequences probably enable us to detect such patterns (Table S1), distinctive *Xylocopa* sociality and patterns of microbial transmission may also contribute to geographical structuring. Notable population-level differentiation of *Bombilactobacillus* and *Bombiscardovia* but not *Lactobacillus* may suggest a different frequency or mode of transmission of bacterial genera among bees.

As expected, we found that restricting our analysis to the V4 region resulted in fewer taxa detected, a loss of phylogenetic resolution (Figures S9 and S10), and the loss of genetic information that could distinguish bacterial taxa between host species and among geographical locations (Table S4). Due to reduced sequence length, ASVs that previously distinguished species and locations were collapsed into fewer ASVs that largely masked such variation (Figures S9 and S10), yet species and geographical location could still be distinguished using V4 rRNA sequences only (Table S1). This comparison suggests that short-read amplicon sequencing may often be able to detect drivers of microbial community composition, but fail to show the extent of variant-level differentiation that exists among populations and species. However, we caution that our comparative approach does not account for realistic primer bias or sequencing bias (Quail et al., 2012; Tedersoo et al., 2018) and, as a result, may overestimate the similarity of these regions and their ecological inference in our simulated V4 data. Moreover, we caution that within-genome variation among 16S rRNA copies can inflate diversity estimates especially in some bacterial clades (Větrovský & Baldrian, 2013) and longer reads generated by PacBio sequencing may capture more of this variation than short-read amplicon sequencing. Nevertheless, our data support the conclusion that long reads enable enhanced ecological insights into the variant-level composition and evolution of the microbiome, suggesting that despite its greater cost, this approach may be warranted to examine population-level variation in bacterial clades (Tedersoo et al., 2021) and generate hypotheses about drivers of such variation.

Overall, our results provide evidence that the microbiomes of species with simple social groups can have characteristics typically associated with the more complex eusociality of the corbiculate bees. Further work will be necessary to determine the role of sociality, such as trophallaxis, as well as other ecological factors that determine symbiont gain and loss to uncover the functional consequences of a specialized microbiome.

AUTHOR CONTRIBUTIONS

MH and RLV conceived the study, MH, MO and NS collected and processed samples, and MH performed laboratory work and wrote the first draft of the paper. MH, DS, RLV and MY performed bioinformatic sequence processing and statistical analyses and DS, MH and RLV made figures with contributions from MO. DS performed phylogenetic analyses. MH and RLV wrote the manuscript, and all authors contributed to revisions.

ACKNOWLEDGEMENTS

We thank the UC Davis Arboretum, UC Anza Borrego Desert Research Center, Arizona State University and Davis homeowners for allowing us access to collect *Xylocopa* samples, the Bohart Museum of Entomology for access to voucher specimens, and members of the Vannette laboratory for comments on previous versions of the manuscript. We acknowledge funding to MH from UC Davis Provost's Undergraduate Research Fellowship and the UC Natural Reserve System through the Field Science Fellowship, and to RLV from National Science Foundation DEB #1929516.

CONFLICT OF INTEREST

All authors declare no conflict of interests.

DATA AVAILABILITY STATEMENT

All sequence data are accessible at NCBI SRA PRJNA863317, and all metadata and code are available via Dryad and Zenodo https://doi.org/10.25338/B8C05S.

ORCID

Rachel L. Vannette https://orcid.org/0000-0002-0447-3468

REFERENCES

Alberoni, D., Gaggìa, F., Baffoni, L., Modesto, M. M., Biavati, B., & Di Gioia, D. (2019). Bifidobacterium xylocopae sp. nov. and Bifidobacterium aemilianum sp. nov., from the carpenter bee (*Xylocopa violacea*) digestive tract. *Systematic and Applied Microbiology*, *42*, 205–216.

Anderson, K. E., Sheehan, T. H., Mott, B. M., Maes, P., Snyder, L., Schwan, M. R., Walton, A., Jones, B. M., & Corby-Harris, V. (2013). Microbial ecology of the hive and pollination landscape: Bacterial associates from floral nectar, the alimentary tract and stored food of honey bees (*Apis mellifera*). *PLoS One*, *8*, e83125.

Barrows, E. M. (1980). Results of a survey of damage caused by the carpenter bee *Xylocopa virginica* (Hymenoptera: Anthophoridae). *Proceedings of the Entomological Society of Washington*, *82*, 44.

Bates, D., Maechler, M., & Bolker, B. (2015). Fitting Linear Mixed-Effects ModelsUsing lme4, *Journal of Statistical Software*, *67*, 1–48.

Bezark, L. G. (2013). Field guide to the common bees of California: Including bees of the Western United States. *The Pan-Pacific Entomologist*, *89*, 266.

Billiet, A., Meeus, I., Van Nieuwerburgh, F., Deforce, D., Wäckers, F., & Smagghe, G. (2017). Colony contact contributes to the diversity of gut bacteria in bumblebees (*Bombus terrestris*). *Insect Sci.*, *24*, 270–277.

Bossert, S., Murray, E. A., Almeida, E. A. B., Brady, S. G., Blaimer, B. B., & Danforth, B. N. (2019). Combining transcriptomes and ultraconserved elements to illuminate the phylogeny of Apidae. *Molecular Phylogenetics and Evolution*, *130*, 121–131.

Breed, M. D. (1976). The evolution of social behavior in primitively social bees: A multivariate analysis. *Evolution*, *30*, 234–240.

Callahan, B. J., Wong, J., Heiner, C., Oh, S., Theriot, C. M., Gulati, A. S., McGill, S. K., & Dougherty, M. K. (2019). High-throughput amplicon sequencing of the full-length 16S rRNA gene with single-nucleotide resolution. *Nucleic Acids Research*, *47*, e103.

Cerqueira, A. E. S., Hammer, T. J., Moran, N. A., Santana, W. C., Kasuya, M. C. M., & da Silva, C. C. (2021). Extinction of anciently associated gut bacterial symbionts in a clade of stingless bees. *The ISME Journal*, *15*, 2813–2816.

Cohen, H., McFrederick, Q. S., & Philpott, S. M. (2020). Environment shapes the microbiome of the blue orchard bee, *Osmia lignaria*. *Microbial Ecology*, *80*, 897–907.

Daisley, B. A., & Reid, G. (2021). BEExact: A metataxonomic database tool for high-resolution inference of bee-associated microbial communities. *Msystems*, *6*, e00082-21.

Engel, P., Kwong, W. K., McFrederick, Q., Anderson, K. E., Barribeau, S. M., Chandler, J. A., Cornman, R. S., Dainat, J., De Miranda, J. R., & Doublet, V. (2016). The bee microbiome: Impact on bee health and model for evolution and ecology of host-microbe interactions. *MBio*, *7*, e02164-15.

Figueroa, L. L., Maccaro, J. J., Krichilsky, E., Yanega, D., & McFrederick, Q. S. (2021). Why did the bee eat the chicken? Symbiont gain, loss, and retention in the vulture bee microbiome. *MBio*, *12*, e0231721.

Ge, Y., Jing, Z., Diao, Q., He, J.-Z., & Liu, Y.-J. (2021). Host species and geography differentiate honeybee gut bacterial communities by changing the relative contribution of community assembly processes. *MBio*, *12*, e0075121.

Gerling, D., Hurd, P. D., Jr., & Hefetz, A. (1981). In-nest behaviour of the carpenter bee, *Xylocopa pubescens* Spinola (Hymenoptera: Anthophoridae). *Journal of the Kansas Entomological Society*, *54*, 209–218.

Gerling, D., Velthuis, H., & Hefetz, A. (1989). Bionomics of the large carpenter bees of the genus *Xylocopa*. *Annual Review of Entomology*, *34*, 163–190.

Giannini, T., Boff, S., Cordeiro, G., Cartolano, E., Veiga, A., Imperatriz-Fonseca, V., & Saraiva, A. (2015). Crop pollinators in Brazil: A review of reported interactions. *Apidologie*, *46*, 209–223.

Gibbs, J., Brady, S. G., Kanda, K., & Danforth, B. N. (2012). Phylogeny of halictine bees supports a shared origin of eusociality for Halictus and Lasioglossum (Apoidea: Anthophila: Halictidae). *Molecular phylogenetics and evolution*, *65*(3), 926–939.

Hammer, T. J., Le, E., Martin, A. N., & Moran, N. A. (2021). The gut microbiota of bumblebees. *Insectes Sociaux*, *68*, 287–301.

Hedtke, S. M., Patiny, S., & Danforth, B. N. (2013). The bee tree of life: a supermatrix approach to apoid phylogeny and biogeography. *BMC Evol Biol*, *13*, 138. https://doi.org/10.1186/1471-2148-13-138

Holley, J.-A. C., Jackson, M. N., Pham, A. T., Hatcher, S. C., & Moran, N. A. (2022). Carpenter Bees (Xylocopa) harbor a distinctive gut microbiome related to that of honey bees and bumble bees. *Applied and Environmental Microbiology*, *88*, e0020322.

Kapheim, K. M., Chan, T.-Y., Smith, A., Wcislo, W. T., & Nonacs, P. (2016). Ontogeny of division of labor in a facultatively eusocial sweat bee *Megalopta genalis*. *Insectes Sociaux*, *63*, 185–191.

Kapheim, K. M., Johnson, M. M., & Jolley, M. (2021). Composition and acquisition of the microbiome in solitary, ground-nesting alkali bees. *Scientific Reports*, *11*, 2993.

Keasar, T. (2010). Large carpenter bees as agricultural pollinators. *Psyche*, *2010*, 7. https://doi.org/10.1155/2010/927463

Killer, J., Kopečný, J., Mrázek, J., Havlík, J., Koppová, I., Benada, O., Rada, V., & Kofroňová, O. (2010). *Bombiscardovia coagulans* gen. nov., sp. nov., a new member of the family Bifidobacteriaceae isolated from the digestive tract of bumblebees. *Systematic and Applied Microbiology*, *33*, 359–366.

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

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*, 1–26.

Kwong, W. K., Medina, L. A., Koch, H., Sing, K.-W., Soh, E. J. Y., Ascher, J. S., Jaffé, R., & Moran, N. A. (2017). Dynamic microbiome evolution in social bees. *Science Advances*, *3*, e1600513.

Kwong, W. K., Steele, M. I., & Moran, N. A. (2018). Genome sequences of *Apibacter* spp., gut symbionts of Asian honey bees. *Genome Biology and Evolution*, *10*, 1174–1179.

Lane, D. J. (1991). 16S/23S rRNA sequencing. In E. Stackebrandt & M. Goodfellow (Eds.), *Nucleic acid techniques in bacterial systematics* (pp. 115–175). John Wiley and Sons.

Lee, F. J., Rusch, D. B., Stewart, F. J., Mattila, H. R., & Newton, I. L. (2015). Saccharide breakdown and fermentation by the honey bee gut microbiome. *Environmental Microbiology*, *17*, 796–815.

Liu, H., Hall, M. A., Brettell, L. E., Halcroft, M., Wang, J., Nacko, S., Spooner-Hart, R., Cook, J. M., Riegler, M., & Singh, B. (2021). Gut microbial diversity in stingless bees is linked to host wing size and is influenced by geography. *bioRxiv:2021.2007.2004.451070*.

Love, M., Anders, S., & Huber, W. (2013). *Differential analysis of count data, the DESeq2 package*.

Lucia, M., Gonzalez, V. H., & Abrahamovich, A. H. (2015). Systematics and biology of *Xylocopa* subgenus Schonnherria (Hymenoptera, Apidae) in Argentina. *ZooKeys*, 129–167.

Lugli, G. A., Milani, C., Turroni, F., Duranti, S., Mancabelli, L., Mangifesta, M., Ferrario, C., Modesto, M., Mattarelli, P., Jiří, K., van Sinderen, D., & Ventura, M. (2017). Comparative genomic and phylogenomic analyses of the Bifidobacteriaceae family. *BMC Genomics*, *18*, 568.

Marshall, L. D., & Alcock, J. (1981). The evolution of the mating system of the carpenter bee *Xylocopa varipuncta* (Hymenoptera: Anthophoridae). *Journal of Zoology*, *193*, 315–324.

Martinson, V. G., Danforth, B. N., Minckley, R. L., Rueppell, O., Tingek, S., & Moran, N. A. (2011). A simple and distinctive microbiota associated with honey bees and bumble bees. *Molecular Ecology*, *20*, 619–628.

McFall-Ngai, M., Hadfield, M. G., Bosch, T. C., Carey, H. V., Domazet-Lošo, T., Douglas, A. E., Dubilier, N., Eberl, G., Fukami, T., & Gilbert, S. F. (2013). Animals in a bacterial world, a new imperative for the life sciences. *Proceedings of the National Academy of Sciences*, *110*, 3229–3236.

McFrederick, Q. S., & Rehan, S. M. (2019). Wild bee pollen usage and microbial communities co-vary across landscapes. *Microbial Ecology*, *77*, 513–522.

McFrederick, Q. S., Thomas, J. M., Neff, J. L., Vuong, H. Q., Russell, K. A., Hale, A. R., & Mueller, U. G. (2017). Flowers and wild megachilid bees share microbes. *Microbial Ecology*, *73*, 188–200.

McFrederick, Q. S., Vuong, H. Q., & Rothman, J. A. (2018). *Lactobacillus micheneri* sp. nov., *Lactobacillus timberlakei* sp. nov. and *Lactobacillus quenuiae* sp. nov., lactic acid bacteria isolated from wild bees and flowers. *International Journal of Systematic and Evolutionary Microbiology*, *68*, 1879–1884.

McFrederick, Q. S., Wcislo, W. T., Hout, M. C., & Mueller, U. G. (2014). Host species and developmental stage, but not host social structure, affects bacterial community structure in socially polymorphic bees. *FEMS Microbiology Ecology*, *88*, 398–406.

McFrederick, Q. S., Wcislo, W. T., Taylor, D. R., Ishak, H. D., Dowd, S. E., & Mueller, U. G. (2012). Environment or kin: Whence do bees obtain acidophilic bacteria? *Molecular Ecology*, *21*, 1754–1768.

McMurdie, P. J., & Holmes, S. (2013). Phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One*, *8*, e61217.

Michener, C. (1990). Castes in xylocopine bees. In W. Engel (Ed.), *Social insects: An evolutionary approach to castes and reproduction* (pp. 123–146). Springer Verlag.

Moeller, A. H., Foerster, S., Wilson, M. L., Pusey, A. E., Hahn, B. H., & Ochman, H. (2016). Social behavior shapes the chimpanzee pan-microbiome. *Science Advances*, *2*, e1500997.

Moran, N. A. (2015). Genomics of the honey bee microbiome. *Current Opinion in Insect Science*, *10*, 22–28.

Moran, N. A., Ochman, H., & Hammer, T. J. (2019). Evolutionary and ecological consequences of gut microbial communities. *Annual Review of Ecology, Evolution, and Systematics*, *50*, 451–475.

Näpflin, K., & Schmid-Hempel, P. (2016). Immune response and gut microbial community structure in bumblebees after microbiota transplants. *Proceedings of the Royal Society B: Biological Sciences*, *283*, 20160312.

Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., O'Hara, R. B., Simpson, G. L., Solymos, P., Stevens, M. H., & Wagner, H. (2012). *vegan: Community Ecology Package*.

Ostwald, M., Dahan, R. A., Shaffer, Z., & Fewell, J. H. (2021). Fluid nest membership drives variable relatedness in groups of a facultatively social bee. *Frontiers in Ecology and Evolution*, *9*, 893.

Ostwald, M., Lyman, B., Shaffer, Z., & Fewell, J. (2020). Temporal and spatial dynamics of carpenter bee sociality revealed by CT imaging. *Insectes Sociaux*, *67*, 203–212.

Paliy, O., Kenche, H., Abernathy, F., & Michail, S. (2009). High-throughput quantitative analysis of the human intestinal microbiota with a phylogenetic microarray. *Applied and Environmental Microbiology*, *75*, 3572–3579.

Parmentier, A., Meeus, I., Van Nieuwerburgh, F., Deforce, D., Vandamme, P., & Smagghe, G. (2018). A different gut microbial community between larvae and adults of a wild bumblebee nest (*Bombus pascuorum*). *Insect Sci.*, *25*, 66–74.

Powell, J. E., Martinson, V. G., Urban-Mead, K., & Moran, N. A. (2014). Routes of acquisition of the gut microbiota of the honey bee *Apis mellifera*. *Applied and Environmental Microbiology*, *80*, 7378–7387.

Quail, M. A., Smith, M., Coupland, P., Otto, T. D., Harris, S. R., Connor, T. R., Bertoni, A., Swerdlow, H. P., & Gu, Y. (2012). A tale of three next generation sequencing platforms: Comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. *BMC Genomics*, *13*, 1–13.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., & Glöckner, F. O. (2012). The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Research*, *41*, D590–D596.

R Core Team. (2013). R: A language and environment for statistical computing.

Rehan, S. M., & Richards, M. H. (2010). Nesting biology and subsociality in *Ceratina calcarata* (Hymenoptera: Apidae). *The Canadian Entomologist*, *142*, 65–74.

Rothman, J. A., Cox-Foster, D. L., Andrikopoulos, C., & McFrederick, Q. S. (2020). Diet breadth affects bacterial identity but not diversity in the pollen provisions of closely related polylectic and oligolectic bees. *Insects*, *11*, 645.

Rubin, B. E., Sanders, J. G., Hampton-Marcell, J., Owens, S. M., Gilbert, J. A., & Moreau, C. S. (2014). DNA extraction protocols cause differences in 16S rRNA amplicon sequencing efficiency but not in community profile composition or structure. *MicrobiologyOpen*, *3*, 910–921.

Rubin, B. E. R., Sanders, J. G., Turner, K. M., Pierce, N. E., & Kocher, S. D. (2018). Social behaviour in bees influences the abundance of *Sodalis* (Enterobacteriaceae) symbionts. *Royal Society Open Science*, *5*, 180369.

Schliep, K. P. (2011). phangorn: Phylogenetic analysis in R. *Bioinformatics*, *27*, 592–593.

Shell, W. A., & Rehan, S. M. (2022). Comparative metagenomics reveals expanded insights into intra-and interspecific variation among wild bee microbiomes. *Communications Biology*, *5*, 1–12.

Subta, P., Yodsuwan, P., Yongsawas, R., In-on, A., Warrit, N., Panha, S., Khongphinitbunjong, K., Chantawannakul, P., Attasopa, K., & Disayathanoowat, T. (2020). Bacterial communities in three parts of intestinal tracts of carpenter bees (*Xylocopa tenuiscapa*). *Insects*, *11*, 497.

Tedersoo, L., Albertsen, M., Anslan, S., & Callahan, B. (2021). Perspectives and benefits of high-throughput long-read sequencing in microbial ecology. *Applied and Environmental Microbiology*, *87*, e00626–e00621.

Tedersoo, L., Tooming-Klunderud, A., & Anslan, S. (2018). PacBio metabarcoding of Fungi and other eukaryotes: Errors, biases and perspectives. *New Phytologist*, *217*, 1370–1385.

Tola, Y. H., Waweru, J. W., Ndungu, N. N., Nkoba, K., Slippers, B., & Paredes, J. C. (2021). Loss and gain of gut bacterial phylotype symbionts in Afrotropical stingless bee species (Apidae: Meliponinae). *Microorganisms*, *9*, 2420.

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

Větrovský, T., & Baldrian, P. (2013). The variability of the 16S rRNA gene in bacterial genomes and its consequences for bacterial community analyses. *PLoS One*, *8*, e57923.

Vickruck, J. L., & Richards, M. H. (2021). Competition drives group formation and reduces within nest relatedness in a facultatively social carpenter bee. *Frontiers in Ecology and Evolution*, *9*, 840.

Voulgari-Kokota, A., McFrederick, Q. S., Steffan-Dewenter, I., & Keller, A. (2019). Drivers, diversity, and functions of the solitary-bee microbiota. *Trends in Microbiology*, *27*, 1034–1044.

Wang, J., Su, Q., Zhang, X., Li, C., Luo, S., Zhou, X., & Zheng, H. (2020). *Entomomonas moraniae* gen. nov., sp. nov., a member of the family Pseudomonadaceae isolated from Asian honey bee gut, possesses a highly reduced genome. *International Journal of Systematic and Evolutionary Microbiology*, *70*, 165–171.

Wright, E. S. (2015). DECIPHER: Harnessing local sequence context to improve protein multiple sequence alignment. *BMC Bioinformatics*, *16*, 1–14.

Yu, G., Smith, D. K., Zhu, H., Guan, Y., & Lam, T. T. Y. (2017). ggtree: An R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods in Ecology and Evolution*, *8*, 28–36.

Zheng, J., Wittouck, S., Salvetti, E., Franz, C. M., Harris, H., Mattarelli, P., O'Toole, P. W., Pot, B., Vandamme, P., & Walter, J. (2020). A taxonomic note on the genus Lactobacillus: Description of 23 novel genera, emended description of the genus Lactobacillus Beijerinck 1901, and union of Lactobacillaceae and Leuconostocaceae. *International Journal of Systematic and Evolutionary Microbiology*, *70*, 2782–2858.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Handy, M. Y., Sbardellati, D. L., Yu, M., Saleh, N. W., Ostwald, M. M., & Vannette, R. L. (2023). Incipiently social carpenter bees (*Xylocopa*) host distinctive gut bacterial communities and display geographical structure as revealed by full-length PacBio 16S rRNA sequencing. *Molecular Ecology*, *32*, 1530–1543. https://doi.org/10.1111/mec.16736