EVOLUTIONARY BIOLOGY

Dynamic microbiome evolution in social bees

Waldan K. Kwong,^{1,2}* Luis A. Medina,² Hauke Koch,^{2†} Kong-Wah Sing,^{3‡} Eunice Jia Yu Soh,⁴ John S. Ascher,⁴ Rodolfo Jaffé,^{5,6} Nancy A. Moran²*

The highly social (eusocial) corbiculate bees, comprising the honey bees, bumble bees, and stingless bees, are ubiquitous insect pollinators that fulfill critical roles in ecosystem services and human agriculture. Here, we conduct wide sampling across the phylogeny of these corbiculate bees and reveal a dynamic evolutionary history behind their microbiota, marked by multiple gains and losses of gut associates, the presence of generalist as well as host-specific strains, and patterns of diversification driven, in part, by host ecology (for example, colony size). Across four continents, we found that different host species have distinct gut communities, largely independent of geography or sympatry. Nonetheless, their microbiota has a shared heritage: The emergence of the eusocial corbiculate bees from solitary ancestors appears to coincide with the acquisition of five core gut bacterial lineages, supporting the hypothesis that host sociality facilitates the development and maintenance of specialized microbiomes.

INTRODUCTION

Social behavior is one of the great evolutionary innovations of life, permitting its adopters to realize novel organizational complexities and providing competitive advantages and fitness benefits to groups working together; these interactions have been fundamental to the origins of multicellularity (1), developmental specialization (2), and culture and technology (3). Sociality affects all aspects of an organism's biology, and there is accumulating evidence that this includes host-associated gut microbial communities as well. The reliability of transmission brought about by social contact potentially allows for stable, long-term microbial assemblages to establish and for host-adapted symbiont lineages to arise (4-10). Nonetheless, how gut microbiomes evolve—through changes in community composition, diversity, and functional capabilities—is not well understood, and the forces guiding these evolutionary trajectories remain unclear.

The eusocial Western honey bee (*Apis mellifera*) harbors a distinctive gut microbiota, composed of <10 bacterial phylotypes that are transmitted through contact between nestmates (*11–13*). These bacteria are extremely specialized, usually found in bee guts or the hive environment but not elsewhere (*14, 15*). However, *A. mellifera* is but one of more than 775 species of social corbiculate bees—a clade dating to about 80 million years and containing other ecologically important pollinators such as the Eastern honey bee (*Apis cerana*), bumble bees (tribe Bombini, all in genus *Bombus*), and stingless bees (tribe Meliponini) (*16*). Previous studies found that diverse bumble bee species harbor gut associates related to those of *A. mellifera* (*14, 17–21*), suggesting that an evolutionarily ancient microbiota is conserved among the eusocial corbiculates.

Here, we compare the gut microbiomes of honey bees (tribe Apini, all in genus *Apis*), bumble bees, and stingless bees—the three major

+Present address: Royal Botanic Gardens, Kew, Richmond, Surrey TW9 3AB, U.K.

*Present address: State Key Laboratory for Biology of Plant Diseases and Insect Pests, Institute of Plant Protection, Chinese Academy of Agricultural Sciences, Beijing 100193, People's Republic of China. 2017 © The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. Distributed under a Creative Commons Attribution NonCommercial License 4.0 (CC BY-NC).

groups of eusocial corbiculates—using samples collected from multiple locations on four continents (Fig. 1). This approach enables us to (i) define the normal gut microbiota of corbiculate bees, (ii) determine whether gut community structure is influenced by phylogenetic relatedness and geographic co-occurrence, (iii) infer shifts in the microbiota over host evolutionary history, and (iv) evaluate how host ecology shapes microbiome composition and diversity. We also use in vivo experiments with cultured bacterial strains to evaluate potential barriers to cross-host gut symbiont exchange. Our results show that, as with humans and several other social animals, the eusocial corbiculate bees have highly characteristic gut communities whose origins and maintenance may be facilitated by their social nature.

RESULTS

We examined the gut microbiota of 472 individual adult bees representing 25 species of corbiculates and two outgroup bee species (data file S1). The V4 region of the bacterial 16S rRNA gene was amplified and sequenced on the Illumina MiSeq platform, generating, on average, 26,672 reads per sample (range, 2313 to 80,286).

Corbiculate bees harbor a small, recurring set of bacterial phylotypes

Gut communities were profiled at a depth of 1900 reads per sample. Despite the antiquity of this clade and the span of geographic regions, habitats, and nesting behaviors represented, the eusocial corbiculate bees have markedly low gut community diversity at 97% operational taxonomic unit clustering (OTU₉₇), with only 199 OTUs in total. Individual specimens have from 1 to 22 OTUs₉₇ [comparatively, humans have 500 to 1000 OTUs₉₇ (22) and termites have $\sim 10^2$ to $> 10^3$, depending on species (7)]. The same phylotypes [phylogenetically related OTUs, according to Martinson et al. (14)] are found across diverse corbiculate species: For instance, OTUs corresponding to the genus Snodgrassella were detected in all Apis and Bombus species surveyed, as well as in 9 of 13 stingless bee species. Similarly, Gilliamella, Bifidobacterium, Lactobacillus Firm-4, and Lactobacillus Firm-5 are prevalent across the three major corbiculate clades, suggesting that these taxa comprise the core of the corbiculate gut microbiome (Fig. 2 and data file S2). Some phylotypes appear lineagespecific: Bartonella apis (23) and Frischella (24) in honey bees, Bombiscardovia (25) and Schmidhempelia (26) in bumble bees, and an Acetobacter-like OTU in the stingless bees. These bacteria are largely absent in our outgroups, the solitary Centris atripes (an oil-collecting bee in a tribe that is

¹Department of Ecology and Evolutionary Biology, Yale University, New Haven, CT 06511, USA. ²Department of Integrative Biology, University of Texas, Austin, Austin, TX 78712, USA. ³Institute of Biological Sciences, Faculty of Science, University of Malaya, 50603 Kuala Lumpur, Malaysia. ⁴Department of Biological Sciences, National University of Singapore, Singapore 117543, Singapore. ⁵Vale Institute of Technology, Sustainable Development, 66055-090 Belém PA, Brazil. ⁶Department of Ecology, Universidade de São Paulo, Rua do Matão 321, 05508-090 São Paulo SP, Brazil.

^{*}Corresponding author. Email: waldan.kwong@yale.edu (W.K.K.); nancy.moran@austin. utexas.edu (N.A.M.)

SCIENCE ADVANCES | RESEARCH ARTICLE



Fig. 1. Distribution of the eusocial corbiculate bees and sample collection sites. The honey bees (*Apis* spp.) are largely restricted to South and East Asia, with the exception of *A. mellifera* and an introduced population of *A. cerana* in Australia. Stingless bees (pictured; *H. itama*) are found in tropical and subtropical regions, including Africa. Bumble bees (pictured; *B. impatiens*) in this study were collected from several sites in the United States. Biogeographical data are from previous studies (*34, 60, 120, 121*).

sister to the corbiculates) and the miner bee *Anthophora abrupta*, and were also not found in other noncorbiculate bees and wasps (14).

Gut communities differ among host species

Although low diversity and the presence of recurring phylotypes appear to be common characteristics uniting corbiculate gut communities, the microbiome of each host species remains distinctive when compared against that of other species. Principal coordinates analysis (PCoA) and nonmetric multidimensional scaling (NMDS) analysis of dissimilarities (Fig. 3A and figs. S2A and S3) and unweighted pair-group method with arithmetic mean clustering (UPGMA) (fig. S4) show clear separation between the three eusocial corbiculate tribes, with the greatest compositional variation within the Meliponini (PERMDISP, P < 0.001). Gut communities assorted according to host, rather than geography. For example, community makeup was not significantly different in A. cerana sampled across five countries (Fig. 3B). Sympatric bee species maintain distinct gut communities (Fig. 3C and fig. S2B), despite having ample opportunities to swap microbes at shared floral resources, during hive robbing, or in construction of interspecific colonies (27-29). Linear modeling corroborates the greater weight of host identity over sampling location in predicting gut community composition (fig. S5; likelihood ratio test, $P < 1 \times 10^{-15}$). These results do not rule out the existence of geographic factors but suggest that host identity is a much more important determinant of microbiome composition.

In East Asia, *A. cerana* has historically been the most economically important bee species, but introductions of *A. mellifera*, which produces greater honey yields, have recently become common (30–32). These introductions, combined with the close genetic relationship and lifestyle

of the two species, have facilitated a disastrous transmission of pathogenic microbes in both directions, with various mites, viruses, and microsporidian parasites devastating naïve hosts (*33*). In 2007, *A. cerana* was inadvertently introduced to Australia at Cairns, Queensland, sparking concern for the existing apicultural industry based around *A. mellifera* (*34*). Whether the normal gut microbiotas of the two species have also become homogenized was unknown. We analyzed samples of co-occurring *A. mellifera* and *A. cerana* in four countries in 2014, as well as of historical *A. cerana* samples from Cairns, and found the gut communities of *A. mellifera* and *A. cerana* to be clearly distinguishable (Fig. 3C and fig. S2). This held true across different locations, implying the existence of persistent barriers to gut symbiont exchange between these closely related species.

Gut community evolution is dynamic

In terms of both phylotype prevalence and relative abundance, there is considerable variation in gut communities among bee species (Fig. 2). To understand why different bees harbor different microbiomes, we examined how gut community composition varies with degree of host relatedness. A strong correlation between pairwise community dissimilarities and host divergence was found, with closely related species having more similar microbiomes (Fig. 4A; Mantel test, r = 0.782, P < 0.0001). This correlation persisted when controlling for geography (partial Mantel test, r = 0.753, P < 0.0001), whereas the effect of geography when controlling for host relatedness was minimal (partial Mantel test, r = 0.046, P = 0.0079).

These results suggest that host phylogeny is key to explaining gut community composition and that social corbiculate microbiomes



Bacterial taxa

Fig. 2. Gut microbial community profiles of sampled bee species. Bacterial taxa were categorized on the basis of the OTU_{97} taxonomic assignments (see data file S2). Relative abundances, as averaged over all samples for each host species, are given as percentages. Prevalence heat map indicates the proportion of samples per host species carrying a bacterial taxon in >0.5% abundance. Absolute abundances of 16S rRNA gene copies represent means of quantitative polymerase chain reaction (qPCR) measures (see data file S1 and fig. S1). n = number of individuals per bee species sampled. *Bombus* data are from this study and the work of Powell *et al.* (93). *, species identified on the basis of the closest mitochondrial DNA match but unverified by morphology.

reflect descent with modification rather than arbitrary assembly within each host. That no single bacterial phylotype is present across all bee species (Fig. 2) is indicative of bacterial lineage gain/loss being a major mechanism by which the bee microbiome evolves. Assuming that gut communities are highly heritable, shifts in the gut microbiome during the evolution of their hosts can be estimated by mapping the community

compositions of extant bees onto a corbiculate phylogeny (Fig. 4B and fig. S6). This parsimony-based inference shows that the five core phylotypes were likely present at the base of the corbiculates. Every sampled corbiculate species, save *Austroplebeia australis*, harbors at least one of the core bacteria at >1% abundance, and each core phylotype is found within members of *Apis, Bombus*, and the Meliponini (Fig. 2).



Fig. 3. Gut community similarities, compared using PCoA. (A) All corbiculate bee samples, compared at a depth of 1900 reads per sample [see fig. S2A for plot excluding Powell *et al.*'s set (93)]. (B) *A. cerana* samples from different geographic locations, compared at a depth of 4500 reads. (C) Comparison of gut communities in sympatric bees, at a depth of 2900 (Austin, USA) or 4500 reads. See figs. S2 and S3 for additional PCoA and NMDS plots. All plots are derived from binary Sørensen-Dice dissimilarities of samples at OTU₉₇ clustering. Analysis of similarities (ANOSIM) and permutational multivariate analysis of variance (PERMANOVA) were used for statistical testing of group similarities. The proportion of variation explained for each axis is given in parentheses. Tick marks, 0.2.



Fig. 4. Gut microbiome evolution. (A) Microbiome dissimilarity as a function of host divergence, showing that closely related bee species have more similar gut communities. (B) Phylogenetic relationships of sampled bee species, with parsimony-inferred shifts in the microbiota over corbiculate evolution. Only gains and losses of the corbiculate core and corbiculate-specific lineages are plotted. Taxa with <1% abundance and <20% prevalence are considered losses. See Materials and Methods for details on tree inference and microbiome reconstruction.

In addition to the core members, other bacteria are often found in the corbiculate gut community. Some of these may be environmental in origin, acquired during foraging or from hive materials. However, others appear to be distinctively associated with particular hosts. One of the main differentiators between the *A. mellifera* and *A. cerana* microbiomes is the consistent presence of *Apibacter* in *A. cerana* (in 84.3% of individuals, compared to 4.8% in *A. mellifera*). *Apibacter* is also characteristic of the *Apis dorsata* microbiome (85.2% prevalence), albeit with a different strain from that found in *A. cerana* (data file S3) (35).

Gains of microbiota lineages are also reflected in broader host clades. For instance, we detected *B. apis* in all sampled *Apis* species, but not in *Bombus* or the Meliponini, suggesting that *B. apis* is more host-restricted than *Apibacter*, which is found in bees across all three clades. The gains of new microbial lineages have been tempered by apparent losses. Of the 25 sampled corbiculate species, 13 lacked detectable levels of at least one core phylotype. The extremely low abundance of *Snodgrassella* in *A. dorsata* (0.3%) may be an example of a loss in progress, whereby a previously abundant gut symbiont is gradually eliminated. It is also possible that OTU prevalence fluctuates with host age or season (*36*), for which further sampling is needed to ascertain.

Strain-level diversity is largely host-specific

Clustering highly conserved, slow-evolving genes such as 16S rRNA at 97% identity obscures strain-level variation. Strain-level variants can be a source of functional diversity (*37*, *38*) and can represent specialists adapted to particular hosts or environments. For example, even though *Snodgrassella* strains occur in both *Apis* and *Bombus*, and are >98% similar in their 16S rRNA sequences, strains in one host are divergent from those in another and contain different genomic repertoires (*39*).

Clustering at 99.5% identity (OTU_{99.5}) resolves a number of distinct strains within each phylotype, mainly assorting with host species (Fig. 5 and data file S3). The presence of host-specific strains supports the hypothesis of a coevolved, vertically transmitted gut microbiota (*18*) and weighs against the idea that the microbiota passes freely between bee species.

The diversity of *Lactobacillus* Firm-5, one of the most ubiquitous and abundant members of the corbiculate microbiota, was examined at greater resolution by phylogenetic analysis of the protein-coding genes *rpoA* and *tuf*, cloned from individual bees (Fig. 5 and figs. S7 to S9). We found that the diversity represented by previously characterized Firm-5 from *A. mellifera* (40–43) encompasses but a small portion of this bacterial clade (Fig. 5) and that different host species tend to carry unique Firm-5 lineages. Closely related hosts generally harbor closely related strains, suggesting that some degree of host-microbe codiversification helps drive strain-level evolution (fig. S10). Firm-5 strains also resolve into clusters that reflect higher host-level taxonomies (for example, strains from bumble bees, honey bees, and the Old World and New World stingless bees group separately), albeit with lower support at deeper nodes.

In *A. mellifera* and *A. cerana*, cloned Firm-5 sequences from a single sample fall across a range of phylogenetic positions, indicating cohabitation of multiple Firm-5 strains within individual bees (fig. S9). This was true even in *A. cerana* from Cairns, Australia, a highly inbred population founded by a single colony (*34*). There was also no clustering discernible along intraspecies morphoclusters, which potentially represent different host genetic backgrounds (for example, *A. cerana* in Cairns versus that in Seoul) (*44*). Hence, the entire diversity

of a host-specific phylotype might be largely represented within a single bee colony regardless of geography, at least for *A. mellifera* and *A. cerana.*

Despite the high degree of observed host specificity, the phylogenetic association between Firm-5 and host bees is not perfectly congruent (Fig. 5 and fig. S10), ruling out strict host-microbe codivergence. Gut microbiome evolution is complex and dynamic: Entire phylotypes can be gained and lost from host lineages, and it is likely that host switching of strains occurs at an appreciable frequency.

Generalist strains can colonize multiple hosts

Although host specificity appears to be prevalent, generalist strains can also be found in the corbiculate microbiota. A *tuf* gene phylogeny of *Snodgrassella* reveals clades in which closely related strains associate with multiple host species (Fig. 6A). For example, similar strains are shared between the Asian honey bees *A. cerana, Apis andreniformis,* and *Apis florea.* Other recurring, related strains are found across some bumble bee clades (*18*). These apparent generalists are useful for probing the barriers to gut microbe exchange and the phenomenon of host specialization.

A. mellifera does not naturally harbor Snodgrassella strains from A. cerana, A. andreniformis, or A. florea (Fig. 5). To determine whether this segregation among hosts reflects the inability of strains to colonize other host species, we performed experiments using monoinoculations with cultured strains in germ-free workers of A. mellifera. The results show that cross-host microbe transfer is possible between Apis spp. (Fig. 6B). In contrast, there was strict host fidelity in transfers between Apis and Bombus (Fig. 6B), as previously observed (39). Hence, the barrier to host switching is not due to direct physiological incompatibility, at least between closely related hosts.

Despite the innate colonization potential of "foreign" strains, their absence in field-collected *A. mellifera* could also be explained if "native" strains are consistently superior competitors. We examined this hypothesis with co-inoculation trials (Fig. 6C). Surprisingly, all three foreign strains we tested were able to simultaneously colonize *A. mellifera* together with a native strain, albeit with different efficacies. Typically, only one strain became dominant in any single gut, but the native *A. mellifera*-derived strain was not always the most competitive. Varying the inoculation ratios by 10-fold produced shifts in final ratios of only ~20% (Fig. 6C), suggesting that competitiveness is a robust trait not easily perturbed.

The apparent credibility of generalist lifestyles complicates our understanding of how bee species maintain distinctive gut communities. Other barriers may be important in enforcing host fidelity. Possibly, for instance, foreign strains may have greater difficulty invading the more complex, heterogeneous gut communities found under natural conditions. The absence of generalist strains colonizing distantly related hosts could also be due to differences in host physiologies that are too great to be bridged. Additionally, host behaviors or biogeographic distribution likely helps circumscribe the microbiota exchange opportunities between bee species.

Host ecology drives microbiota diversity

Causal factors behind large-scale trends in community diversity are often difficult to pinpoint. Our results suggest that the corbiculate gut provides a superb habitat for a distinctive set of bacterial colonists, yet there is substantial variation in gut community composition between bee species and between individuals within species. For instance, we find that gut communities in bumble bees are, on average, less diverse than those



Bacterial taxa, 99.5% OTU clustering (16S rRNA)

Fig. 5. Strain-level diversity in the bee gut microbiota is largely host-specific. Relative abundances of all reads per bee species (rows) corresponding to particular OTUs_{99,5} (columns) are shown for selected bacterial taxa. See data file S3 for OTU_{99,5} diversity summaries of other taxa. For *Lactobacillus* Firm-5, the conserved, single-copy *tuf* gene was cloned and sequenced from representative samples to visualize Firm-5 diversification and relationship with host taxa. Dashed lines and coloring indicate host origins. Clades containing previously cultured members are denoted by *"*"* [*Lactobacillus apis* (40)] and *"t"* (41, 42). Bayesian posterior probabilities and maximum likelihood bootstrap supports, in that order, are shown at select nodes as percentage values. Displayed tree is based on Bayesian inference; bar represents nucleotide substitutions per site. See figs. S7 and S8 for uncondensed trees. Tip taxa in host phylogeny are in the same order as in Fig. 4. Cophylogenetic analysis of hosts and Firm-5 is shown in fig. S10.

in honey bees (Fig. 7A, nested ANOVA OTUs₉₇, P = 0.0420; fig. S11, Shannon's H OTU_{99,5}, P = 0.0045). Bumble bee gut communities are also more erratic than those of honey bees (Fig. 7B), with individuals harboring communities ranging from having relatively even abundances to being dominated by a few phylotypes. For the stingless bees, there was significant variation between species in both diversity and evenness.

The species-area relationship (45) postulates that larger habitats harbor greater diversity. We examined total community size (absolute bacterial abundance) as a function of host size (length of average worker bee) and uncover a positive log-linear relationship in the corbiculates, whereby larger bees host more bacteria (Fig. 7C). Bacterial diversity itself does not correlate with either community size or bee size when



Fig. 6. Diversity of *Snodgrassella alvi* and its impact on host switching. (A) Bayesian phylogeny of *S. alvi* clones and cultured isolates is shown, with the strains used for colonization assays indicated by colored symbols. Bayesian posterior probabilities and maximum likelihood bootstrap supports, in that order, are shown at select nodes as percentages. Bar represents nucleotide substitutions per site. (B) Monoinoculations demonstrate that *S. alvi* strains from other *Apis* spp. are able to colonize *A. mellifera* [bars, mean; detection limit, 5 × 10⁵ colony-forming units (CFUs)]. (C) Co-inoculations suggest variability in competitiveness and that *A. mellifera*-derived strains are not always dominant (error bars, SEM).

comparing across all corbiculates (Fig. 7D and fig. S12; *F* test, *P* > 0.05). However, weak positive correlations are found when the corbiculate clades are considered separately (Fig. 7D and fig. S12; OLS) and when corrected for phylogeny (Fig. 7D and fig. S12; PGLS).

The peculiarities of host ecology may help explain these diversity patterns. Unlike honey bees and stingless bees, which grow by colony fission (46, 47), bumble bee colonies are founded by single queens (48), imposing potential bottlenecks on microbiota diversity. For yet unknown reasons, the *Bombus* gut microbiota also appears more prone to perturbation and displacement by environmental bacteria (19, 20, 49), and despite their larger body size, bumble bees often form colonies that are orders of magnitude smaller than those of *Apis* or the Meliponini (10¹ to 10² versus 10³ to 10⁴ individuals). Hence, the effective "habitat size" for *Bombus*-associated microbes may be considerably smaller than that for *Apis* or the Meliponini, leading to lower overall gut community diversity (Fig. 7A and fig. S11).

To test this hypothesis, consider that an analog to habitat size in such interconnected social systems may not simply be an individual bee but also the colony or the local host population among which microbes are shared among conspecific individuals. If this is true, bee species with larger colony sizes would be expected to have greater gut microbial diversity. In agreement, we find a strong correlation with gut community diversity when host bacterial load is adjusted by colony size (Fig. 8, A and B). Linear models suggest that both gut community size (or, collinearly, bee size) and colony size are significant predictors of gut microbiome diversity in the eusocial corbiculates (Fig. 8C and fig. S13).

DISCUSSION

Recent years have seen a proliferation of studies on gut communities in efforts to understand the microbiota's role in the health and development of their animal hosts. However, scant attention has been paid to how the microbiota itself arises and evolves (50-53). The evolutionary context of a microbiome may reveal crucial insights not evident when focusing on a single host, but the complexity of most gut communities has thus far hindered this type of investigation. Here, we show that the corbiculate bees are a tractable system for studying changes in gut microbiota composition across a clade's evolutionary history.

We find that the emergence of the eusocial corbiculate bees likely coincided with their acquisition of five core gut bacterial lineages, Snodgrassella, Gilliamella, Lactobacillus Firm-4, Lactobacillus Firm-5, and Bifidobacterium (Fig. 4B). This is supported by their widespread incidence in all three corbiculate groups examined (Fig. 2), as well as by deeply branching divergences marking host-specific clades in some of these bacteria, notwithstanding occasional host switching (Figs. 5 and 6A and figs. S8 and S10). These bacteria are largely absent in our outgroup bees, although the presence of trace levels of Snodgrassella and Gilliamella (Fig. 2) could represent occasional horizontal acquisition from corbiculate bees or environmental pools of related strains [for example, Gilliamella-related Orbaceae strains are found in other insects (54, 55)]. Gains of other bacterial lineages, as well as losses of existing ones, appear to happen with frequency. For instance, we identify a recurring Acetobacter-like OTU associated with many, but not all, Meliponini bees. This OTU also matches sequences



Fig. 7. Relationship between host identity, bacterial load, and microbiota diversity. (**A**) Number of OTUs at 97% clustering for each bee species. The *Bombus* microbiota is, on average, less diverse than that of *Apis*, whereas there is substantial variation in the Meliponini (nested ANOVA). (**B**) Gut community evenness as measured by Shannon's *E* on OTUs₉₇ for each bee species. The *Bombus* microbiota is more erratic than that of *Apis*, with more variation in evenness (Levene's test). n.s., not significant. (**C**) Mean bacterial abundance versus average worker size for each bee species. Larger bees harbor more bacteria (*F* test, excluding noncorbiculates). Dashed lines, 95% confidence interval. (**D**) Gut microbial diversity, as measured by Shannon's *H* on OTUs_{99.5}, versus mean bacterial abundance for each bee species. There is a weak positive correlation between diversity and community size when each corbiculate tribe is considered separately [ordinary least squares (OLS) *F* test, colored lines] but not together (OLS *F* test, black line). Correcting for phylogeny improves fit [phylogenetic generalized least squares (PGLS)]. Host taxa in (A) and (B) are in the same order as Fig. 2 and fig. S1; boxes and whiskers represent 25th to 75th and 5th to 95th percentiles, respectively, of individual samples within each bee species; bar indicates median. See fig. S11 and S12 for additional plots.

recovered from African stingless bees (18), suggesting that this uncharacterized bacterium is affiliated with a breadth of stingless bee diversity but has been lost in some host lineages.

When explaining microbiota composition in extant corbiculate species, a variety of potential mechanisms should be considered. Although conspecific transmission of gut symbionts has been experimentally demonstrated in honey bees and bumble bees (12, 18), strict host-microbe codiversification is unlikely the norm. The gain of a gut symbiont in a host clade may have occurred more recently than in the last common ancestor, with the observed distribution resulting from promiscuous cross-host transfers rather than codivergence. Greater phylogenetic scrutiny at the strain level will help decipher the signatures of host switching, coevolution, and independent introductions, and their roles in shaping the dynamic corbiculate microbiome.

Regardless of the mechanistic details of these processes, the outcomes are divergent microbial communities most strongly distinguishable by host of origin (Fig. 3). Similar observations have been made in mammals and other social insects, where vertical inheritance is thought to play a large role in gut community assembly (8, 50, 52, 56, 57). It is also probable that, through both neutral and adaptive mechanisms, the microbiota differs between allopatric bee populations (11). However, given our sample sizes and level of phylogenetic resolution, we did not detect a strong effect of geographic location on microbiota composition or strain diversity (Fig. 3 and figs. S2, S3, S5, and S9B). An analysis of *Snodgrassella* and *Gilliamella* 16S rRNA gene sequences across *Bombus* and *Apis* hosts by Koch *et al.* (18) also found poor correlation with geography, with a small significant effect present only in *Gilliamella*.

Within host species, there are instances of variable prevalence of the core microbiota members, suggesting that particular hosts may be more prone to compositional shifts influenced by geography, colony of origin, food resources, or physiological status. This intraspecific variation seems to be more pronounced in bumble bees and in some species of stingless bees. For instance, *Gilliamella* was prevalent in *Bombus bifarius* from Utah, but not that from Colorado (Fig. 2),



Fig. 8. Large bee colony size is associated with higher microbiome diversity. (A) Plot of microbiome diversity (number of OTUs at OTU_{97} clustering) against gut community size (16S rRNA gene copies) × bee colony size. (B) Plot of microbiome diversity (Shannon's *H* at $OTU_{99.5}$ clustering) against gut community size (16S rRNA gene copies) × bee colony size. Large dots, species averages; small dots, individual samples. Best-fit lines for different regression analyses are shown. (C) Linear modeling of the effects of gut community size (16S rRNA gene copies), host phylogenetic affiliation (tribe), and bee colony size on microbiome diversity (Shannon's *H* at $OTU_{99.5}$ clustering). Models that include an interaction of gut community size and bee colony size have the greatest relative likelihood. Goodness of fit between models was compared with likelihood ratio tests. See fig. S13 for linear model details.

and there was differential presence of *Snodgrassella* and *Parasacchar-ibacter* in *Tetragonula carbonaria* between two collection sites (data file S2). Similarly, a previous survey of Australian stingless bees found that microbiomes differed between colonies (58).

Compared to *Apis* and *Bombus*, the stingless bee microbiome appears to span a much greater range of possible states, being highly variable in composition, richness, and evenness (Figs. 2, 3A, and 7, and figs. S1 and S2). Perhaps not coincidentally, the Meliponini are also the most diverse of the corbiculate groups, comprising ~500 species in ~50 genera (59, 60) compared to ~9 in *Apis* and ~250 in *Bombus* (48). The variability found in the Meliponini microbiota may be an evolutionary hallmark of their hosts' correspondingly diverse life histories, morphologies, and behaviors (60). A fourth corbiculate tribe, the Euglossini (orchid bees), contains species with differing degrees of social organization, from solitary to communal (61), but their microbiomes remain uncharacterized.

Host sociality facilitates the development of heritable microbial communities by providing a semiclosed system with reliable transmission routes between host "islands" of similar environments. There is evidence for this in some clades of primates, rodents, and ants (8-10, 62), but asocial hosts, such as Drosophila (55, 63), generally lack characteristic gut microbiotas unless particular adaptations are adopted (for example, maternal transmission and environmental filtering) (64, 65). Longterm, transgenerational habitat stability becomes a suitable backdrop against which microbial specialists can evolve to best exploit emergent niches. For example, the transition to sociality in the wood roaches/ termites enabled the development of a nutritionally specialized gut microbiota and a shift from an omnivorous to a wood-based diet (7, 66, 67). The corbiculate bee microbiota also appears to be metabolically optimized toward its host's diet, with symbiont lineages such as Gilliamella, Lactobacillus, and Bifidobacterium able to enzymatically break down and ferment the sugars found in pollen, honey, and nectar (37, 39, 43, 68, 69). Other gut bacteria have specialized to particular physical locations, including *Snodgrassella*, which occupies the gut wall of the ileum (70); Frischella, which colonizes a small section of the pylorus (71); and Para-

Kwong et al., Sci. Adv. 2017; 3:e1600513 29 March 2017

saccharibacter, which is found in relative abundance in larval food and worker hypopharyngeal glands (72).

In addition to opening up new niches for novel bacterial lineages, a stable gut habitat may promote strain-level diversification within existing community members by fostering large bacterial population sizes through which diversity can be maintained (73). The degree to which this diversity is functional or neutral is unknown (37, 74), but strain-level variants are common features of gut communities (38, 75). In animals and plants, taxon richness is often a function of habitat size, as modeled by the species-area relationship $S = c \times A^z$ or $S = \log(c \times A^z)$, where S is a measure of diversity (for example, species or OTU richness), c is a constant particular to the system in question, A is the habitat size, and z is the correspondence between S and A (slope in log space) (45).

This relationship has been observed in some free-living microbial systems (76-78) but has been largely ignored for gut microbiomes, partly due to the lack of a direct analog to the habitat size metrics (for example, area) used for macroorganisms. Gut volume has been considered (79), as has body weight, as in a previous study that found no relation between microbiota diversity and weight of host individuals within a single bumble bee species (80). Here, we compare across corbiculate bee species and find a weak correlation with microbiota diversity by using absolute community size or bee size as proxies for habitat size (Fig. 7D and fig. S12). However, these proxies are incomplete: In social organisms, individuals are constantly exchanging microbes such that, if we wish to compare between host species, a meaningful habitat scale is likely larger than the gut of any individual. A more accurate measure would take into account host characteristics (for example, colony size, migration, degree of social interactions, and bottlenecking effects) so that the time-averaged suitable habitat for a microbial community to evolve within is approximated-a "nominal habitat size."

We find support for this idea by showing that the inclusion of bee colony size as an explanatory factor greatly improves the correlation with microbiome diversity (Fig. 8 and fig. S13). Hence, we propose that a species-area relationship for host-associated microbiomes may be expressed as $S = c \times H^{z}$, where nominal habitat size (H) is proportional not only to the metrics of individual hosts (for example, body size) but also to the particular characteristics of the host species. If such a relationship holds, then a better understanding of host biology, including life history and population structure, will be a key component toward explaining ecological patterns in gut microbial communities. Larger H may increase microbial population sizes, decreasing extinction and maintaining neutral variation, or give rise to allopatric diversification due to incomplete mixing and emergence of novel niches (81). However, we note that habitat size itself may only be one of many forces that influence the observed trends and may not be the factor that is most directly responsible. Further work will be needed to uncover the underlying mechanisms that drive microbial species-area relations and to help unite a long-standing macroecological concept with microbiome biology.

Although not examined in this study, the social nature of the corbiculates also presents opportunities to investigate the effect of caste (82, 83) and differing social organizations (84) on microbiome composition. Sociality alone is not enough for the establishment of host-specific microbiotas, as exemplified by the lack of consistent microbiomes in some clades of ants and bees (8, 84, 85). Measuring the contributions of additional factors, such as behavior and diet, will be necessary to better understand the evolution of persistent host-microbe associations.

Other fundamental yet understudied aspects of microbial ecology and evolution (51, 53, 86) await examination from the perspective of gut communities. For example, it is unclear what barriers prevent the spread of generalist strains into bees that have more specialized microbiomes (Fig. 6). Interaction between the microbiota and foreign invaders, such as parasites, and disease states is another area that remains poorly understood (87, 88). A major advantage of the corbiculate bee microbiome is its amenability to experimental hypothesis testing using cultivated strains and microbiota transplants. Together with a relatively low complexity compared to mammalian systems (22, 89), this makes the study of microbiome evolution in bees a feasible endeavor, with the potential to provide novel insights into the symbiotic gut communities of social animals.

CONCLUSION

We collected bees at multiple sites across four continents. A small, recurring set of bacteria was found across the eusocial corbiculates, including within the previously sparsely sampled stingless bees. Microbiome composition strongly aligned with host identity, more so than with geographic origin, and closely related hosts had more similar microbiomes. We show that bee microbiome evolution is dynamic, with multiple gains and losses of bacterial lineages; nonetheless, a core corbiculate community exists, the acquisition of which coincides with the origin of eusociality. Within bacterial lineages, strain-level diversity is common and mostly host-specific. How this host specificity is maintained is unclear, but we are able to test hypotheses, such as host incompatibility and interbacterial competition, using in vivo microbiota transplant experiments.

There are large-scale patterns in bee gut community diversity. For instance, we discovered that bumble bee microbiomes are less diverse and more erratic in composition than those of honey bees. Microbiome diversity was weakly correlated with gut community size and bee size and more strongly associated with bee colony size. These findings indicate that host ecology drives large-scale trends in microbiome diversity and that a form of species-area relationship may be applicable to gut communities.

Overall, this study offers evidence that vertical inheritance, through social contact, is a major force that shapes the corbiculate bee microbiome over the clade's evolutionary history. Our results support the emerging hypothesis that host sociality facilitates the development and maintenance of specialized microbiomes by providing favorable ecological conditions and reliable transmission mechanisms.

MATERIALS AND METHODS

Specimen collection and identification

Adult worker bees were obtained from sites across seven countries (Australia, Malaysia, Singapore, Hong Kong, South Korea, United States, and Brazil), mostly between May 2013 and October 2014 (data file S1). Specimens were kept in 100% ethanol or frozen at -80° C for DNA preservation. For a subset of samples, the guts were removed, homogenized, and then frozen in 19% glycerol, allowing for preservation of live bacteria from which strains were later isolated.

Of the approximately nine recognized *Apis* species, the five most common species were collected. Both the Neotropical and Indo-Malay/Australasia clades of the diverse stingless bees (90) were sampled, including species used in commercial meliponiculture such as *Heterotrigona itama*, *Geniotrigona thoracica*, and *T. carbonaria* (91, 92). Bumble bees were collected from sites in the United States (Colorado, New Jersey, Texas, and Utah) and included new samples and samples from the study of Powell *et al.* (93). As outgroups for comparative analyses, we obtained specimens of *C. atripes* [tribe Centridini, a close relative of the corbiculates (94)] and *A. abrupta* [tribe Anthophorini, which is a more distant outgroup but still belongs to the "Apine line" of subfamily Apinae (95)]. Where available, sympatric bees were collected at the same time and location.

Bee species were identified by morphology. For most stingless bee samples, the mitochondrial 16S rRNA gene was also sequenced, following Rasmussen and Cameron (90). For *C. atripes* and *A. abrupta*, the 28S rRNA gene was sequenced, following Martins *et al.* (94). Sequences were classified on the basis of the closest BLAST hits in the GenBank nonredundant (nr) database (data file S1). To verify that samples classified as the same species were related, phylogenies based on their sequences were built and inspected in Geneious R9 (Biomatters Ltd.).

DNA extraction

DNA was extracted with a protocol adapted from Powell et al. (12). Entire guts were removed from bee abdomens and crushed in 100 µl of CTAB buffer [0.1 M tris-HCl (pH 8.0), 1.4 M NaCl, 20 mM EDTA, and cetrimonium bromide (20 mg/ml)], resuspended in a total of 728 µl of CTAB buffer and 20 µl of proteinase K [20 mM tris-HCl, 1 mM CaCl₂, 50% glycerol, and proteinase K (20 mg/ml)], and transferred to a capped vial containing 500 µl of 0.1-mm Zirconia beads (BioSpec Products Inc.). 2-Mercaptoethanol (2 µl) and 2 µl of RNase A cocktail (Invitrogen Corp.) were added, and samples were bead-beated for 3 × 2 min. Samples were digested overnight at 50°C before the addition of 750 µl of phenol/chloroform/isoamyl (25:24:1, pH 8.0). Samples were mixed and centrifuged for 15 min at 4°C and 20,000g. The aqueous layer was removed, and the DNA was precipitated at -20° C for 0.5 hours with 700 μ l of isopropanol and 70 μ l of sodium acetate (pH 5.3). Precipitated samples were spun at 4°C at 20,000g for 30 min, and the supernatant was decanted. DNA pellets were washed with -20°C

ethanol and spun for an additional 5 min at 4°C. The ethanol wash was removed by pipetting, and the DNA pellets were dried at 50°C and then resuspended in 50 μ l of water. Final DNA samples were stored at –20°C.

16S rRNA gene amplification and sequencing

Extracted DNA samples were quantified fluorometrically (Qubit, Thermo Fisher Scientific Inc.) and then diluted to 10 ng/µl for use as PCR template. For all Bombus samples and samples prefixed with "WKsample" (data file S1), Illumina barcoded primers, designed according to Caporaso et al. (96), were used to amplify a 291-base pair (bp) fragment encompassing the V4 region of bacterial 16S rRNA. Triplicate 25-µl reactions were run for 25 to 30 cycles of amplification at an annealing temperature of 65°C, using 1 µl of DNA template and the NEBNext DNA Library Prep Master Mix according to the manufacturer's protocols. The PCR products were cleaned using AxyPrep Mag PCR Cleanup (Axygen Scientific Inc.) with 0.8× volume of beads to remove primer dimers. Samples were pooled to equimolar concentration and sequenced on the Illumina MiSeq 2×250 bp platform with primers Read1F (5'-TATGGTAATTGTGTGCCAGCMGCCGCGGTAA), Read2R (5'-AGTCAGTCAGCCGGACTACHVGGGTWTCTAAT), and Index (5'-ATTAGAWACCCBDGTAGTCCGGCTGACTGACT).

For samples prefixed with "A" (data file S1), the Illumina Nextera kit was used for library preparation of V4 region amplicons generated with the adapter-primer pairs Hyb515F_rRNA (5'-TCGTCGGCAGCGT-CAGATGTGTATAAGAGACAGGTGYCAGCMGCCGCGGTA) and Hyb806R_rRNA (5'-GTCTCGTGGGCTCGGAGATGTGTA-TAAGAGACAGGGACTACHVGGGTWTCTAAT), using the above amplification protocol; these were then sequenced on the Illumina MiSeq 2 × 250 bp platform. The two primer sets are designed to amplify 251 and 252 bp (excluding primer sequence), corresponding to the V4 region of *Escherichia coli* 16S rRNA, which completely overlap save for 1 bp at the 5' end. All sequencing was performed at the University of Texas Genomic Sequencing and Analysis Facility.

Reads were demultiplexed on the basis of the barcode sequences, allowing for one mismatch. Control samples using purified DNA from *E. coli* K12 or a synthetic bee gut community, composed of *G. apicola* wkB1, *S. alvi* wkB2, and *Lactobacillus* Firm-5 wkB8, were included in each sequencing run. Between 2313 and 80,286 reads were acquired for each sample.

16S rRNA-based gut community analysis

Demultiplexed reads were processed on the QIIME 1.9.1 pipeline (97). The following scripts were run sequentially: join_paired_ends.py with the SeqPrep option to join forward and reverse reads; split_libraries_fastq.py for quality filtering set at phred $q \ge 30$, maximum N = 0, and read length fraction minimum of 0.8; and read_length_filter.py to retain only reads between 230 and 270 bp (expected read length of ~250 bp after primer trimming).

For 97% OTU clustering (OTU₉₇), pick_de_novo_otus.py was used with default settings (uclust algorithm) and taxonomy assignment based on the Silva119 reference database (98). Sequences were then checked for chimeras with ChimeraSlayer in QIIME 1.9.1. Each sample was further filtered to remove OTUs that comprise $\leq 0.5\%$ of reads (due to potential misassignent of barcodes between samples during multiplexed sequencing). Each sequence comprising this final set of representative OTU₉₇ sequences was then manually checked by BLAST against the GenBank nr database to refine taxonomic assignment (data file S2). For 99.5% OTU clustering (OTU_{99.5}), de novo OTU picking was done with pick_otus.py and pick_rep_set.py using default settings (uclust algorithm). Singleton OTUs were filtered out, as were samples with read depths of less than 5000. Remaining samples were then rarefied to 5000 reads. $OTU_{99.5}$ taxonomy was assigned with parallel_ assign_taxonomy_blast.py using the manually refined taxonomy from OTU_{97} as reference. Representative sequences from OTU_{97} 16S rRNA clustering were deposited in GenBank (data file S4), and representative sequences from OTU_{97} and $OTU_{99.5}$ clustering are presented in data file S5.

To determine relative abundance and prevalence of bacterial taxa in Fig. 2, nonbacterial OTUs were removed from the OTU_{97} table, and remaining OTUs were removed from the OTU97 table, and remaining OTUs were subsumed into broader categories based on phylogenetic relatedness (see data file S2). In Fig. 3 and fig. S2, PCoA ordinations on binary Sørensen-Dice dissimilarities of samples at the OTU_{97} clustering were determined in QIIME 1.9.1. To test for group similarities, ANOSIM using 999 permutations was performed in QIIME 1.9.1. The ANOSIM test statistic *R* indicates group similarity, where 0 = indistinguishable and 1 = dissimilar. NMDS ordinations on the Sørensen-Dice dissimilarity matrices were generated with the metaMDS function in the "vegan" package for R v.3.2.2 (99, 100).

Diversity metrics (number of unique OTUs, Shannon's H, and Shannon's E) were calculated in QIIME 1.9.1 (Fig. 7 and figs. S11 and S12). At OTU_{99.5} clustering, Shannon's H is superior to OTU counts as a measure of diversity because clustering at this level results in OTU count inflation from sequencing errors; Shannon's H places less weight on low-abundance (potentially erroneous) OTUs. To test for diversity differences between groups (Fig. 7A and fig. S11), nested ANOVA was conducted in R v.3.2.2, with species as subgroups. To test for differences in group variation (Fig. 7B), Levene's test in the "car" package for R v.3.2.2 was used (*101*). Worker size (length in millimeters) and bee colony size represent mean of size ranges reported in the literature (table S1).

Absolute bacterial quantification

A subset of samples (data file S1) was selected for quantification of 16S rRNA gene copy number, as a proxy for the number of bacteria present per bee gut. qPCR using universal bacterial 16S rRNA gene primers was performed as in the work of Cariveau *et al.* (19). Standards were cloned into pGEM-T Easy vector (Promega Corp.), linearized with Apa I, and fluorometrically quantified before generation of standard curves.

Ancestral microbiome reconstruction

To estimate the changes in microbiome composition over corbiculate bee evolution (Fig. 4B), a presence/absence matrix of *Snodgrassella*, *Gilliamella*, *Lactobacillus* Firm-4, *Lactobacillus* Firm-5, *Bifidobacterium*, *B. apis*, *Frischella*, *Bombiscardovia*, *Schmidhempelia*, and the *Acetobacter*like OTU was mapped against the host phylogeny (cladogram). Taxa with <1% abundance and <20% prevalence (Fig. 2) were considered absent. Gains and losses across host lineages were determined by asymmetrical Wagner parsimony in Count v.10.04 (*102*) with gain and loss penalties of 1.5 and 1, respectively.

The true phylogeny of the corbiculate bees and related taxa is not well resolved. Although the species relationships within *Apis* and *Bombus* and the reciprocal monophyly of Bombini and Meliponini are generally accepted, the relationships within the Meliponini, as well as the placement of the Euglossini with respect to the other three corbiculate tribes, are contentious (103). The host taxa relationships presented here (Figs. 4 and 5) were obtained by joining the results of multigene phylogenies within *Apis* (104), *Bombus* (105, 106), and the Meliponini (90), together in a supertree based on their most highly supported relations with each other and with the outgroups (94, 103, 107). For bee species that were not included in these previous studies, placement was based on that of their respective genera. Because these data sources are not directly comparable, phylogenetic distances were omitted from these analyses.

Microbiome dissimilarity and host relatedness

The impact of host phylogeny on microbial community dissimilarity was investigated using a Mantel test of correlation between matrices of pairwise Sørensen-Dice dissimilarities and host divergences in the "vegan" package in R v.3.2.2 with 10⁴ permutations (99). Because phylogenetic distances or divergence times were not available for our sampled bee species, a nodal distance approach based on the work of Webb (108) was used, whereby the number of nodes separating bee species on our tree (Fig. 4B) was counted as a proxy for host divergence. Although this method only gives a relative measure of divergence and is subject to sampling bias effects (108), it enables analysis with a tree topology that is consistent with that of published reports.

We performed partial Mantel tests to determine the correlation between microbiome dissimilarity and host phylogeny while controlling for geography and the correlation between microbiome dissimilarity and geography while controlling for host phylogeny. A geographic distance matrix generated from sample collection locations was used; 10^4 permutations were run for each analysis.

Linear models testing

To further examine the contributions of geography and host identity on microbiome composition (fig. S5), linear mixed-effects models were generated with the lmer function in the "lme4" package for R v.3.2.2 (109). The NMDS coordinates of each sample (fig. S3A) were used as proxies for microbiome composition. "Locations" were designated as collection sites within approximately 15-km-radius regions. To test for the contributions of host identity, host body size, bee colony size, and microbiota community size on microbiome diversity (fig. S13), linear models were generated with the lm function in R v.3.2.2. Bee body sizes and colony sizes were obtained from literature sources (table S1). Samples where diversity values (number of OTUs₉₇, Shannon's H) or bee colony size information were not available were excluded from the analysis. Optimal model selection using Akaike's Information Criterion was performed with the model.sel function in the "MuMIn" package for R v.3.2.2 (110). The goodness of fit between alternative models was tested with the ANOVA likelihood ratio test in R v.3.2.2.

To account for the effects of phylogenetic nonindependence in comparisons of species-level covariates (Figs. 7, C and D, and 8 and fig. S12), PGLS analyses were conducted in addition to OLS tests. The bee phylogeny was obtained as described above, with all internal and terminal branches set at equal lengths. The packages "ape" (111), "geiger" (112), "nlme" (113), and "phytools" (114) for R v.3.2.2 were used to construct and test PGLS models with the gls function under a Brownian motion model of evolution ($\lambda = 1$).

Phylogenetic analysis of *Lactobacillus* Firm-5 and *Snodgrassella alvi*

To obtain greater phylogenetic resolution of the *Lactobacillus* Firm-5 group, two conserved single-copy genes were cloned and sequenced.

The gene encoding ribosomal polymerase α subunit, *rpoA*, was amplified with primers wk121 (5'-GAATTTGAAAAACCAAATATTAC-CG) and wk122 (5'-CGTACACGCATCATATCTGC), whereas the gene encoding translation elongation factor Tu, *tuf*, was amplified with wk123 (5'-GAACAAAGCCACACGTAAAC) and wk124 (5'-GAC-CAGCACCAACAGTACGA). These produced products of 815 and 1103 bp, respectively. Primers were designed to be specific for Firm-5 sequences, on the basis of existing genomes for this clade (*42*, *43*), and for other *Lactobacillus* species (*115*).

For *S. alvi*, *tuf* was amplified with primers wk135 (5'-TGGCTAAA-GAAAAATTCGAGCGG) and wk136 (5'-AAGCGATAACTTTAG-CAACCACAC), producing a 1141-bp amplicon. There are two *tuf* copies in *S. alvi* genomes.

PCRs were carried out using Phusion polymerase (New England Biolabs Inc.) and the following cycling conditions: initial denaturing at 98°C for 30 s; 30 cycles at 98°C for 10 s, at 60°C for 20 s, and at 72°C for 30 s; and a final extension at 72°C for 2 min. PCR products were cleaned using AxyPrep Mag PCR Cleanup (Axygen Scientific Inc.), A-tailed with Taq polymerase (New England Biolabs Inc.), and cloned into pGEM-T Easy vector (Promega Corp.). The resulting plasmids were electroporated into *E. coli* DH5 α , and colonies carrying inserts were identified by blue-white screening. Cloned inserts were sequenced by the Sanger method using vector-specific primers T7 and SP6.

Forward and reverse sequences were assembled, and ambiguities were manually corrected in Geneious R9 (Biomatters Ltd.). A total of 184 and 55 MUSCLE-aligned tuf sequences were used for Firm-5 and S. alvi phylogenetic analysis, respectively (figs. S7 and S8). Appropriate models of evolution were determined using jModelTest2 (116). Phylogenies based on Bayesian inference were produced with MrBayes 3.2 (117) using GTR+gamma+inv for each codon position, a run-time of 2×10^7 generations with tree sampling every 1000 generations, and a burn-in of 0.25. Convergence of Bayesian Markov chain Monte Carlo runs was assessed with AWTY (118). For maximum likelihood analyses, the GTR+gamma+inv model was used for each codon position in RAxML v8 (119) with 1000 bootstrap replicates. For Lactobacillus Firm-5 rpoA (fig. S9A), a total of 46 MUSCLE-aligned sequences were used to build phylogenies as above, except a 2×10^6 generation run-time with tree sampling every 500 generations was used in the Bayesian analysis.

Isolation of gut symbiont strains

Frozen glycerol stocks of homogenized guts were plated out on heart infusion agar or Columbia agar supplemented with 5% sheep's blood (Hardy Diagnostics) and incubated at 35°C and 5% CO₂ for 3 to 7 days. Colonies were screened by colony PCR to identify *S. alvi*, using primers described in previous studies (*14*, *18*). Follow-up sequencing of 16S rRNA genes using the universal primers 27F and 1492R was used to confirm the identity of isolates (*54*).

Colonization and competition assays

S. alvi strains isolated from *A. mellifera* (wkB2, wkB332), *A. cerana* (wkB298), *A. florea* (wkB273), *A. andreniformis* (wkB237), and *Bombus bimaculatus* (wkB12) were used to colonize laboratory-reared *A. mellifera*. Approximately 10² cells were fed to germ-free, newly emerged workers; after 5 days, CFUs were counted, as in Kwong *et al.* (*39*).

For competition assays, germ-free, newly emerged A. *mellifera* workers were co-inoculated with two strains of S. *alvi*, as described by Kwong *et al.* (39). Bees were fed $\sim 10^6$ cells in 1:1 ratios of strains

derived from *A. mellifera* (wkB2 or wkB332) and those from other *Apis* species. To determine the effect of the initial dosage on strain competitiveness, ratios of 1:10 and 10:1 were also used in a subsequent assay between wkB2 and wkB298/wkB273. Strains were differentiated by transferring colonies onto blood heart infusion agar plates with tetracycline (25 μ g/ml). Strains wkB2 and wkB332 are tetracycline-resistant, whereas the other strains are susceptible. Count data are presented in table S2.

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at http://advances.sciencemag.org/cgi/ content/full/3/3/e1600513/DC1

fig. S1. Absolute bacterial abundance per bee gut, as measured by 16S rRNA gene copies. fig. S2. Gut community similarities, compared using PCoA of binary Sørensen-Dice

dissimilarities at OTU₉₇ clustering.

fig. S3. NMDS plots of gut community Sørensen-Dice dissimilarities at OTU₉₇ clustering.

fig. S4. UPGMA clustering of communities based on Sørensen-Dice and unweighted UniFrac distances.

fig. S5. Linear mixed models testing of the effects of sampling location and host identity on microbiome composition.

fig. S6. Ancestral microbiome reconstructions using parsimony for the major corbiculate bacterial phylotypes.

fig. S7. Bayesian phylogeny of Lactobacillus Firm-5 based on the tuf gene.

fig. S8. Maximum likelihood phylogeny of *Lactobacillus* Firm-5 based on the *tuf* gene. fig. S9. Phylogeny of *Lactobacillus* Firm-5 based on *rpoA*, and geographic associations of strains.

fig. S10. Tanglegram of host bee and *Lactobacillus* Firm-5 phylogenetic topologies. fig. S11. Gut microbial diversity, as measured by Shannon's *H* on OTUs clustered at 99.5% similarity.

fig. S12. Relationships between bee body size, gut community size (165 copies), and microbiome diversity.

fig. S13. Linear models testing of the effects of bee body size, gut community size (165 copies), host phylogenetic affiliation (tribe), and bee colony size on microbiome diversity.

table S1. Bee size and colony size estimation and literature sources.

table S2. Snodgrassella co-inoculation trials count data.

data file S1. List of samples and associated metadata used in this study.

data file S2. \mbox{OTU}_{97} tables and curated taxonomic assignments.

data file S3. $OTU_{99.5}$ tables summarized by bacterial taxa, from samples rarefied to 5000 reads. data file S4. List of sequences deposited in GenBank.

data file S5. Representative sequences from OTU_{97} and $OTU_{99.5}$ clustering. References (122–148)

REFERENCES AND NOTES

- J. E. Strassmann, D. C. Queller, The social organism: Congresses, parties, and committees. Evolution 64, 605–616 (2010).
- G. F. Oster, E. O. Wilson, *Caste and Ecology in the Social Insects* (Princeton Univ. Press, 1978).
- 3. K. N. Laland, W. Hoppitt, Do animals have culture? Evol. Anthropol. 12, 150-159 (2003).
- H. Ochman, M. Worobey, C.-H. Kuo, J. B. Ndjango, M. Peeters, B. H. Hahn, P. Hugenholtz, Evolutionary relationships of wild hominids recapitulated by gut microbial communities. *PLOS Biol.* 8, e1000546 (2010).
- P. L. Oh, A. K. Benson, D. A. Peterson, P. B. Patil, E. N. Moriyama, S. Roos, J. Walter, Diversification of the gut symbiont *Lactobacillus reuteri* as a result of host-driven evolution. *ISME J.* 4, 377–387 (2010).
- P. Stallforth, D. A. Brock, A. M. Cantley, X. Tian, D. C. Queller, J. E. Strassmann, J. Clardy, A bacterial symbiont is converted from an inedible producer of beneficial molecules into food by a single mutation in the *gacA* gene. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 14528–14533 (2013).
- C. Dietrich, T. Köhler, A. Brune, The cockroach origin of the termite gut microbiota: Patterns in bacterial community structure reflect major evolutionary events. *Appl. Environ. Microbiol.* **80**, 2261–2269 (2014).
- J. G. Sanders, S. Powell, D. J. C. Kronauer, H. L. Vasconcelos, M. E. Frederickson, N. E. Pierce, Stability and phylogenetic correlation in gut microbiota: Lessons from ants and apes. *Mol. Ecol.* 23, 1268–1283 (2014).
- J. Tung, L. B. Barreiro, M. B. Burns, J.-C. Grenier, J. Lynch, L. E. Grieneisen, J. Altmann, S. C. Alberts, R. Blekhman, E. A. Archie, Social networks predict gut microbiome composition in wild baboons. *eLife* 4, e05224 (2015).

- A. H. Moeller, S. Foerster, M. L. Wilson, A. E. Pusey, B. H. Hahn, H. Ochman, Social behavior shapes the chimpanzee pan-microbiome. *Sci. Adv.* 2, e1500997 (2016).
- N. A. Moran, A. K. Hansen, J. E. Powell, Z. L. Sabree, Distinctive gut microbiota of honey bees assessed using deep sampling from individual worker bees. *PLOS ONE* 7, e36393 (2012).
- J. E. Powell, V. G. Martinson, K. Urban-Mead, N. A. Moran, Routes of acquisition of the gut microbiota of the honey bee *Apis mellifera*. *Appl. Environ. Microbiol.* **80**, 7378–7387 (2014).
- W. K. Kwong, N. A. Moran, Gut microbial communities of social bees. *Nat. Rev. Microbiol.* 14, 374–384 (2016).
- V. G. Martinson, B. N. Danforth, R. L. Minckley, O. Rueppell, S. Tingek, N. A. Moran, A simple and distinctive microbiota associated with honey bees and bumble bees. *Mol. Ecol.* 20, 619–628 (2011).
- K. E. Anderson, T. H. Sheehan, B. M. Mott, P. Maes, L. Snyder, M. R. Schwan, A. Walton, B. M. Jones, V. Corby-Harris, 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 (2013).
- 16. D. Grimaldi, M. S. Engel, Evolution of the Insects (Cambridge Univ. Press, 2005).
- H. Koch, P. Schmid-Hempel, Bacterial communities in central European bumblebees: Low diversity and high specificity. *Microb. Ecol.* 62, 121–133 (2011).
- H. Koch, D. P. Abrol, J. Li, P. Schmid-Hempel, Diversity and evolutionary patterns of bacterial gut associates of corbiculate bees. *Mol. Ecol.* 22, 2028–2044 (2013).
- D. P. Cariveau, J. E. Powell, H. Koch, R. Winfree, N. A. Moran, Variation in gut microbial communities and its association with pathogen infection in wild bumble bees (*Bombus*). *ISME J.* 8, 2369–2379 (2014).
- J. Li, J. E. Powell, J. Guo, J. D. Evans, J. Wu, P. Williams, Q. Lin, N. A. Moran, Z. Zhang, Two gut community enterotypes recur in diverse bumblebee species. *Curr. Biol.* 25, R652–R653 (2015).
- H. C. Lim, C.-C. Chu, M. J. Seufferheld, S. A. Cameron, Deep sequencing and ecological characterization of gut microbial communities of diverse bumble bee species. *PLOS ONE* 10, e0118566 (2015).
- P. B. Eckburg, E. M. Bik, C. N. Bernstein, E. Purdom, L. Dethlefsen, M. Sargent, S. R. Gill, K. E. Nelson, D. A. Relman, Diversity of the human intestinal microbial flora. *Science* 308, 1635–1638 (2005).
- L. Kešnerová, R. Moritz, P. Engel, Bartonella apis sp. nov., a honey bee gut symbiont of the class Alphaproteobacteria. Int. J. Syst. Evol. Microbiol. 66, 414–421 (2016).
- P. Engel, W. K. Kwong, N. A. Moran, *Frischella perrara* gen. nov., sp. nov., a gammaproteobacterium isolated from the gut of the honey bee, *Apis mellifera*. *Int. J. Syst. Evol. Microbiol.* **63**, 3646–3651 (2013).
- J. Killer, J. Kopečný, J. Mrázek, J. Havlík, I. Koppová, O. Benada, V. Rada, O. Kofroňová, Bombiscardovia coagulans gen. nov., sp. nov., a new member of the family Bifidobacteriaceae isolated from the digestive tract of bumblebees. Syst. Appl. Microbiol. 33, 359–366 (2010).
- V. G. Martinson, T. Magoc, H. Koch, S. L. Salzberg, N. A. Moran, Genomic features of a bumble bee symbiont reflect its host environment. *Appl. Environ. Microbiol.* 80, 3793–3803 (2014).
- C. D. Michener, The Social Behavior of the Bees: A Comparative Study (Harvard Univ. Press, 1973).
- J. C. Nieh, The evolution of honey bee communication: Learning from Asian species. Formosan Entomol. 31, 101–115 (2011).
- J. P. Cunningham, J. P. Hereward, T. A. Heard, P. J. De Barro, S. A. West, Bees at war: Interspecific battles and nest usurpation in stingless bees. *Am. Nat.* 184, 777–786 (2014).
- P. Akratanakul, *Beekeeping in Asia* (FAO Agricultural Services Bulletin 68/4, Food and Agriculture Organization of the United Nations, 1990); www.fao.org/docrep/x0083E/ X0083E00.htm.
- W. S. Li, Quantitative declination of Chinese bee colonies and its causes. *Chinese J. Ecol.* 10, 50–53 (1991).
- Z. H. Xie, Managed honeybee colonies and honey production in China grew during the last five decades. J. Apic. Sci. 55, 77–85 (2011).
- D. P. Abrol, in *Pollination Biology: Biodiversity Conservation and Agricultural Production* (Springer, 2012), pp. 651–652.
- A. H. Koetz, Ecology, behaviour and control of *Apis cerana* with a focus on relevance to the Australian incursion. *Insects* 4, 558–592 (2013).
- W. K. Kwong, N. A. Moran, *Apibacter adventoris* gen. nov., sp. nov., a member of the phylum *Bacteroidetes* isolated from honey bees. *Int. J. Syst. Evol. Microbiol.* 66, 1323–1329 (2016).
- 36. V. Corby-Harris, P. Maes, K. E. Anderson, The bacterial communities associated with honey bee (*Apis mellifera*) foragers. *PLOS ONE* **9**, e95056 (2014).
- P. Engel, V. G. Martinson, N. A. Moran, Functional diversity within the simple gut microbiota of the honey bee. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 11002–11007 (2012).

- S. Greenblum, R. Carr, E. Borenstein, Extensive strain-level copy-number variation across human gut microbiome species. *Cell* 160, 583–594 (2015).
- W. K. Kwong, P. Engel, H. Koch, N. A. Moran, Genomics and host specialization of honey bee and bumble bee gut symbionts. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 11509–11514 (2014).
- J. Killer, S. Dubná, I. Sedláček, P. Švec, Lactobacillus apis sp. nov., from the stomach of honeybees (Apis mellifera), having an in vitro inhibitory effect on the causative agents of American and European foulbrood. Int. J. Syst. Evol. Microbiol. 64, 152–157 (2014).
- 41. T. C. Olofsson, M. Alsterfjord, B. Nilson, E. Butler, A. Vásquez, Lactobacillus apinorum sp. nov., Lactobacillus mellifer sp. nov., Lactobacillus mellis sp. nov., Lactobacillus melliventris sp. nov., Lactobacillus kimbladii sp. nov., Lactobacillus helsingborgensis sp. nov. and Lactobacillus kullabergensis sp. nov., isolated from the honey stomach of the honeybee Apis mellifera. Int. J. Syst. Evol. Microbiol. 64, 3109–3119 (2014).
- W. K. Kwong, A. L. Mancenido, N. A. Moran, Genome sequences of *Lactobacillus* spp. strains wkB8 and wkB10, members of the 'Firm-5' clade, from honey bee guts. *Genome Announc.* 2, e01176-14 (2014).
- K. M. Ellegaard, D. Tamarit, E. Javelind, T. C. Olofsson, S. G. Andersson, A. Vásquez, Extensive intra-phylotype diversity in lactobacilli and bifidobacteria from the honeybee gut. *BMC Genomics* 16, 284 (2015).
- S. E. Radloff, C. Hepburn, H. R. Hepburn, S. Fuchs, S. Hadisoesilo, K. Tan, M. S. Engel, V. Kuznetsov, Population structure and classification of *Apis cerana*. *Apidologie* 41, 589–601 (2010).
- E. F. Connor, E. D. McCoy, The statistics and biology of the species-area relationship. Am. Nat. 113, 791–833 (1979).
- B. P. Oldroyd, S. Wongsiri, Asian Honey Bees: Biology, Conservation and Human Interactions (Harvard Univ. Press, 2006).
- 47. A. Wille, Biology of the stingless bees. Ann. Rev. Entomol. 28, 41-64 (1983).
- 48. C. D. Michener, The Bees of the World (Johns Hopkins Univ. Press, ed. 2, 2007).
- L. Parmentier, I. Meeus, H. Mosallanejad, D. C. de Graaf, G. Smagghe, Plasticity in the gut microbial community and uptake of *Enterobacteriaceae* (Gammaproteobacteria) in *Bombus terrestris* bumblebees' nests when reared indoors and moved to an outdoor environment. *Apidologie* 47, 237–250 (2016).
- R. E. Ley, C. A. Lozupone, M. Hamady, R. Knight, J. I. Gordon, Worlds within worlds: Evolution of the vertebrate gut microbiota. *Nat. Rev. Microbiol.* 6, 776–788 (2008).
- E. K. Costello, K. Stagaman, L. Dethlefsen, B. J. M. Bohannan, D. A. Relman, The application of ecological theory toward an understanding of the human microbiome. *Science* 336, 1255–1262 (2012).
- A. H. Moeller, Y. Li, E. Mpoudi Ngole, S. Ahuka-Mundeke, E. V. Lonsdorf, A. E. Pusey, M. Peeters, B. H. Hahn, H. Ochman, Rapid changes in the gut microbiome during human evolution. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 16431–16435 (2014).
- Q. Zeng, J. Sukumaran, S. Wu, A. Rodrigo, Neutral models of microbiome evolution. *PLOS Comput. Biol.* 11, e1004365 (2015).
- 54. W. K. Kwong, N. A. Moran, Cultivation and characterization of the gut symbionts of honey bees and bumble bees: Description of *Snodgrassella alvi* gen. nov., sp. nov., a member of the family *Neisseriaceae* of the *Betaproteobacteria*, and *Gilliamella apicola* gen. nov., sp. nov., a member of *Orbaceae* fam. nov., *Orbales* ord. nov., a sister taxon to the order *'Enterobacteriales'* of the *Gammaproteobacteria*. Int. J. Syst. Evol. Microbiol. **63**, 2008–2018 (2013).
- J. A. Chandler, J. M. Lang, S. Bhatnagar, J. A. Eisen, A. Kopp, Bacterial communities of diverse *Drosophila* species: Ecological context of a host-microbe model system. *PLOS Genetics.* 7, e1002272 (2011).
- Y. Hongoh, P. Deevong, T. Inoue, S. Moriya, S. Trakulnaleamsai, M. Ohkuma,
 C. Vongkaluang, N. Noparatnaraporn, T. Kudo, Intra- and interspecific comparisons of bacterial diversity and community structure support coevolution of gut microbiota and termite host. *Appl. Environ. Microbiol.* **71**, 6590–6599 (2005).
- J. G. Sanders, A. C. Beichman, J. Roman, J. J. Scott, D. Emerson, J. J. McCarthy, P. R. Girguis, Baleen whales host a unique gut microbiome with similarities to both carnivores and herbivores. *Nat. Commun.* 6, 8285 (2015).
- S. D. Leonhardt, M. Kaltenpoth, Microbial communities of three sympatric Australian stingless bee species. *PLOS ONE* 9, e105718 (2014).
- J. S. Ascher, J. Pickering, Discover Life bee species guide and world checklist (Hymenoptera: Apoidea); www.discoverlife.org/mp/20q?guide=Apoidea_species.
- P. Vit, S. R. M. Pedro, D. W. Roubik, Eds., *Pot-Honey: A Legacy of Stingless Bees* (Springer, 2013).
 S. Cardinal, B. N. Danforth, The antiquity and evolutionary history of social behavior
- in bees. *PLOS ONE* 6, e21086 (2011).
 62. C. Ubeda, L. Lipuma, A. Gobourne, A. Viale, I. Leiner, M. Equinda, R. Khanin, E. G. Pamer, Familial transmission rather than defective innate immunity shapes the distinct intestinal microbiota of TLR-deficient mice. *J. Exp. Med.* 209, 1445–1456 (2012).
- A. C.-N. Wong, J. M. Chaston, A. E. Douglas, The inconstant gut microbiota of *Drosophila* species revealed by 16S rRNA gene analysis. *ISME J.* 7, 1922–1932 (2013).
- R. V. M. Rio, M. Maltz, B. McCormick, A. Reiss, J. Graf, Symbiont succession during embryonic development of the European medicinal leech, *Hirudo verbana. Appl. Environ. Microbiol.* **75**, 6890–6895 (2009).

- P. Engel, N. A. Moran, The gut microbiota of insects—Diversity in structure and function. *FEMS Microbiol. Rev.* 37, 699–735 (2013).
- S. Otani, A. Mikaelyan, T. Nobre, L. H. Hansen, N. A. Koné, S. J. Sørensen, D. K. Aanen, J. J. Boomsma, A. Brune, M. Poulsen, Identifying the core microbial community in the gut of fungus-growing termites. *Mol. Ecol.* 23, 4631–4644 (2014).
- A. E. Pérez-Cobas, E. Maiques, A. Angelova, P. Carrasco, A. Moya, A. Latorre, Diet shapes the gut microbiota of the omnivorous cockroach *Blattella germanica*. *FEMS Microbiol. Ecol.* **91**, fiv022 (2015).
- F. J. Lee, D. B. Rusch, F. J. Stewart, H. R. Mattila, I. L. Newton, Saccharide breakdown and fermentation by the honey bee gut microbiome. *Environ. Microbiol.* 17, 796–815 (2015).
- H. Zheng, A. Nishida, W. K. Kwong, H. Koch, P. Engel, M. I. Steele, N. A. Moran, Metabolism of toxic sugars by strains of the bee gut symbiont *Gilliamella apicola. mBio* 7, e01326-16 (2016).
- V. G. Martinson, J. Moy, N. A. Moran, Establishment of characteristic gut bacteria during development of the honey bee worker. *Appl. Environ. Microbiol.* 78, 2830–2840 (2012).
- P. Engel, K. D. Bartlett, N. A. Moran, The bacterium *Frischella perrara* causes scab formation in the gut of its honeybee host. *mBio.* 6, e00193-15 (2015).
- V. C. Corby-Harris, L. A. Snyder, M. R. Schwan, P. Maes, Q. S. McFrederick, K. E. Anderson, Origin and effect of *Acetobacteraceae* Alpha 2.2 in honey bee larvae and description of *Parasaccharibacter apium*, gen. nov., sp. nov. *Appl. Environ. Microbiol.* **80**, 7460–7472 (2014).
- B. Charlesworth, Effective population size and patterns of molecular evolution and variation. *Nat. Rev. Genet.* 10, 195–205 (2009).
- 74. P. Engel, R. Stepanauskas, N. A. Moran, Hidden diversity in honey bee gut symbionts detected by single-cell genomics. *PLOS Genet.* **10**, e1004596 (2014).
- S. Schloissnig, M. Arumugam, S. Sunagawa, M. Mitreva, J. Tap, A. Zhu, A. Waller, D. R. Mende, J. R. Kultima, J. Martin, K. Kota, S. R. Sunyaev, G. M. Weinstock, P. Bork, Genomic variation landscape of the human gut microbiome. *Nature* **493**, 45–50 (2013).
- M. C. Horner-Devine, M. Lage, J. B. Hughes, B. J. M. Bohannan, A taxa-area relationship for bacteria. *Nature* **432**, 750–753 (2004).
- T. Bell D. Ager, J.-I. Song, J. A. Newman, I. P. Thompson, A. K. Lilley, C. J. van der Gast, Larger islands house more bacterial taxa. *Science* **308**, 1884 (2005).
- L. Zinger, A. Boetius, A. Ramette, Bacterial taxa–area and distance–decay relationships in marine environments. *Mol. Ecol.* 23, 954–964 (2014).
- Q. Yan, C. J. van der Gast, Y. Yu, Bacterial community assembly and turnover within the intestines of developing zebrafish. *PLOS ONE* 7, e30603 (2012).
- H. Koch, G. Cisarovsky, P. Schmid-Hempel, Ecological effects on gut bacterial communities in wild bumblebee colonies. J. Anim. Ecol. 81, 1202–1210 (2012).
- J. Hortal, K. A. Triantis, S. Meiri, E. Thébault, S. Sfenthourakis, Island species richness increases with habitat diversity. *Am. Nat.* **174**, E205–E217 (2009).
- D. R. Tarpy, H. R. Mattila, I. L. G. Newton, Development of the honey bee gut microbiome throughout the queen-rearing process. *Appl. Environ. Microbiol.* 81, 3182–3191 (2015).
- K. M. Kapheim, V. D. Rao, C. J. Yeoman, B. A. Wilson, B. A. White, N. Goldenfeld, G. E. Robinson, Caste-specific differences in hindgut microbial communities of honey bees (*Apis mellifera*). *PLOS ONE* **10**, e0123911 (2015).
- Q. S. McFrederick, W. T. Wcislo, M. C. Hout, U. G. Mueller, Host species and developmental stage, but not host social structure, affects bacterial community structure in socially polymorphic bees. *FEMS Microbiol. Ecol.* **88**, 398–406 (2014).
- Q. S. McFrederick, J. J. Cannone, R. R. Gutell, K. Kellner, R. M. Plowes, U. G. Mueller, Specificity between lactobacilli and hymenopteran hosts is the exception rather than the rule. *Appl. Environ. Microbiol.* **79**, 1803–1812 (2013).
- D. R. Nemergut, S. K. Schmidt, T. Fukami, S. P. O'Neill, T. M. Bilinski, L. F. Stanish, J. E. Knelman, J. L. Darcy, R. C. Lynch, P. Wickey, S. Ferrenberg, Patterns and processes of microbial community assembly. *Microbiol. Mol. Biol. Rev.* 77, 342–356 (2013).
- H. Koch, P. Schmid-Hempel, Gut microbiota instead of host genotype drive the specificity in the interaction of a natural host-parasite system. *Ecol. Lett.* **15**, 1095–1103 (2012).
- C. A. Lozupone, M. Li, T. B. Campbell, S. C. Flores, D. Linderman, M. J. Gebert, R. Knight, A. P. Fontenot, B. E. Palmer, Alterations in the gut microbiota associated with HIV-1 infection. *Cell Host Microbe* 14, 329–339 (2013).
- A. F. Andersson, M. Lindberg, H. Jakobsson, F. Bäckhed, P. Nyrén, L. Engstrand, Comparative analysis of human gut microbiota by barcoded pyrosequencing. *PLOS ONE* 3, e2836 (2008).
- C. Rasmussen, S. A. Cameron, Global stingless bee phylogeny supports ancient divergence, vicariance, and long distance dispersal. *Biol. J. Linn. Soc.* **99**, 206–232 (2010).
- N. Kelly, M. S. N. Farisya, T. K. Kumara, P. Marcela, Species diversity and external nest characteristics of stingless bees in meliponiculture. *Trop. Agr. Sci.* 37, 293–298 (2014).
- T. A. Heard, A. Dollin, Stingless beekeeping in Australia, snapshot of an infant industry. Bee World 82, 116–125 (2000).
- E. Powell, N. Ratnayeke, N. A. Moran, Strain diversity and host specificity in a specialized gut symbiont of honeybees and bumblebees. *Mol. Ecol.* 25, 4461–4471 (2016).

- A. C. Martins, G. A. R. Melo, S. S. Renner, The corbiculate bees arose from New World oil-collecting bees: Implications for the origin of pollen baskets. *Mol. Phylogenet. Evol.* 80, 88–94 (2014).
- A. Roig-Alsina, C. D. Michener, Studies of the phylogeny and classification of long-tongued bees (Hymenoptera: Apoidea). *Univ. Kansas Sci. Bull.* 55, 124–162 (1993).
- J. G. Caporaso, C. L. Lauber, W. A. Walters, D. Berg-Lyons, J. Huntley, N. Fierer, S. M. Owens, J. Betley, L. Fraser, M. Bauer, N. Gormley, J. A. Gilbert, G. Smith, R. Knight, Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J.* 6, 1621–1624 (2012).
- J. G. Caporaso, J. Kuczynski, J. Stombaugh, K. Bittinger, F. D. Bushman, E. K. Costello, N. Fierer, A. G. Peña, J. K. Goodrich, J. I. Gordon, G. A. Huttley, S. T. Kelley, D. Knights, J. E. Koenig, R. E. Ley, C. A. Lozupone, D. McDonald, B. D. Muegge, M. Pirrung, J. Reeder, J. R. Sevinsky, P. J. Turnbaugh, W. A. Walters, J. Widmann, T. Yatsunenko, J. Zaneveld, R. Knight, QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* 7, 335–336 (2010).
- C. Quast, E. Pruesse, P. Yilmaz, J. Gerken, T. Schweer, P. Yarza, J. Peplies, F. O. Glöckner, The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Res.* 41, D590–D596 (2013).
- J. Oksanen, R. Kindt, P. Legendre, B. O'Hara, M. H. H. Stevens, M. J. Oksanen; MASS Suggests, The vegan package. *Community Ecology Package* 10, 631–637 (2007).
- R Core Team, R: A Language and Environment for Statistical Computing (R Foundation for Statistical Computing, 2013).
- 101. J. Fox, S. Weisberg, D. Adler, D. Bates, G. Baud-Bovy, S. Ellison, D. Firth, M. Friendly, G. Gorjanc, S. Graves, R. Heiberger, R. Laboissiere, G. Monette, D. Murdoch, H. Nilsson, D. Ogle, B. Ripley, W. Venables, D. Winsemius, A. Zeileis, car, S. Weisberg, A. Zeleis: Companion to applied regression, R Package version 1.2-16 (2009); http://cran.r-project. org/web/packages/car/index.html.
- M. Csüös, Count: Evolutionary analysis of phylogenetic profiles with parsimony and likelihood. *Bioinformatics* 26, 1910–1912 (2010).
- 103. J. Romiguier, S. A. Cameron, S. H. Woodard, B. J. Fischman, L. Keller, C. J. Praz, Phylogenomics controlling for base compositional bias reveals a single origin of eusociality in corbiculate bees. *Mol. Biol. Evol.* **33**, 670–678 (2016).
- N. Lo, R. S. Gloag, D. L. Anderson, B. P. Oldroyd, A molecular phylogeny of the genus *Apis* suggests that the Giant Honey Bee of the Philippines, *A. breviligula* Maa, and the Plains Honey Bee of southern India, *A. indica* Fabricius, are valid species. *Syst. Entomol.* 35, 226–233 (2009).
- S. A. Cameron, H. M. Hines, P. H. Williams, A comprehensive phylogeny of the bumble bees (*Bombus*). *Biol. J. Linn. Soc.* 91, 161–188 (2007).
- H. M. Hines, Historical biogeography, divergence times, and diversification patterns of bumble bees (Hymenoptera: Apidae: *Bombus*). Syst. Biol. 57, 58–75 (2008).
- A. C. Martins, G. A. R. Melo, The New World oil-collecting bees *Centris* and *Epicharis* (Hymenoptera, Apidae): Molecular phylogeny and biogeographic history. *Zool. Scr.* 45, 22–33 (2016).
- C. O. Webb, Exploring the phylogenetic structure of ecological communities: An example for rain forest trees. Am. Nat. 156, 145–155 (2000).
- D. Bates, M. Maechler, B. Bolker, S. Walker, Ime4: Linear mixed-effects models using Eigen and S4. R package version 1.1-7 (2014); http://CRAN.R-project.org/package=Ime4.
- K. Bartoń. MuMln: Multi-model Inference. R package version 1.15.6 (2016); http://CRAN. R-project.org/package=MuMln.
- E. Paradis, J. Claude, K. Strimmer, APE: Analyses of Phylogenetics and Evolution in R language. *Bioinformatics* 20, 289–290 (2004).
- L. J. Harmon, J. T. Weir, C. D. Brock, R. E. Glor, W. Challenger, GEIGER: Investigating evolutionary radiations. *Bioinformatics* 24, 129–131 (2008).
- J. Pinheiro, D. Bates, S. DebRoy, D. Sarkar, R Core Team, nlme: Linear and Nonlinear Mixed Effects Models. R package version 3.1-124 (2016); http://CRAN.R-project.org/ package=nlme.
- 114. L. J. Revell, phytools: An R package for phylogenetic comparative biology (and other things). *Methods Ecol. Evol.* **3**, 217–223 (2012).
- 115. M. Ventura, C. Canchaya, V. Meylan, T. R. Klaenhammer, R. Zink, Analysis, characterization, and loci of the *tuf* genes in *Lactobacillus* and *Bifidobacterium* species and their direct application for species identification. *Appl. Environ. Microbiol.* **69**, 6908–6922 (2003).
- 116. D. Darriba, G. L. Taboada, R. Doallo, D. Posada, jModelTest 2: More models, new heuristics and parallel computing. *Nat. Methods.* **9**, 772 (2012).
- 117. F. Ronquist, M. Teslenko, P. van der Mark, D. L. Ayres, A. Darling, S. Höhna, B. Larget, L. Liu, M. A. Suchard, J. P. Huelsenbeck, MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* **61**, 539–542 (2012).
- J. A. A. Nylander, J. C. Wilgenbusch, D. L. Warren, D. L. Swofford, AWTY (are we there yet?): A system for graphical exploration of MCMC convergence in Bayesian phylogenetics. *Bioinformatics* 24, 581–583 (2008).

- A. Stamatakis, RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30, 1312–1313 (2014).
- C. D. Michener, Classification of the Apidae (Hymenoptera). Appendix: *Trigona genalis* Friese, a hitherto unplaced New Guinea species. *Univ. Kansas Sci. Bull.* 54, 75–163 (1990).
- P. H. Williams, An annotated checklist of bumble bees with an analysis of patterns of description (Hymenoptera: Apidae, Bombini). *Bull. Nat. Hist. Mus. Lond. (Entomol.)* 67, 79–152 (1998).
- 122. S. Boongird, Aspects of culturing, reproductive behavior, and colony formation in the stingless bee *Tetragonula fuscobalteata* (Hymenoptera: Apidae: Meliponini). J. Kansas Entornol. Soc. 84, 190–196 (2011).
- 123. T. X. Chinh, M. J. Sommeijer, W. J. Boot, C. D. Michener, Nest and colony characteristics of three stingless bee species in Vietnam with the first description of the nest of *Lisotrigona carpenteri* (Hymenoptera: Apidae: Meliponini). *J. Kansas Entomol. Soc.* **78**, 363–372 (2005).
- 124. F. A. L. Contrera, J. C. Nieh, V. L. Imperatriz-Fonseca, Temporal and climatological influences on flight activity in the stingless bee *Trigona hyalinata* (Apidae, Meliponini). *Revista Tecnologia e Ambiente* **10**, 35–43 (2004).
- A. E. Dollin, L. J. Dollin, S. F. Sakagami, Australian stingless bees of the genus *Trigona* (Hymenoptera: Apidae). *Invert. Taxon.* **11**, 861–896 (1997).
- 126. M. A. M. El-Aw, K. A. Draz, K. S. A. Eid, H. F. I. Abou-Shaara, Measuring the morphological characters of honey bee (*Apis mellifera* L.) using a simple semi-automatic technique. J. Am. Sci. 8, 558–564 (2012).
- 127. V. H. Gonzalez, T. L. Griswold, Two new species of *Paratrigona* and the male of *Paratrigona ornaticeps* (Hymenoptera, Apidae). *ZooKeys* **120**, 9–25 (2011).
- M. Halcroft, R. Spooner-Hart, P. Neumann, Behavioral defense strategies of the stingless bee, *Austroplebeia australis*, against the small hive beetle, *Aethina tumida. Insectes Soc.* 58, 245–253 (2011).
- G. A. Hobbs, Ecology of species of *Bombus* Latr. (Hymenoptera: Apidae) in Southern Alberta. V. subgenus *Subterraneobombus* Vogt. *Canad. Entomol.* 98, 288–294 (1966).
- T. Inoue, S. Salmah, I. Abbas, E. Yusuf, Foraging behavior of individual workers and foraging dynamics of colonies of three Sumatran stingless bees. *Res. Popul. Ecol.* 27, 373–392 (1985).
- R. Jaffé, F. C. Pioker-Hara, C. F. Dos Santos, L. R. Santiago, D. A. Alves, A. de M. P. Kleinert, T. M. Francoy, M. C. Arias, V. L. Imperatriz-Fonseca, Monogamy in large bee societies: A stingless paradox. *Naturwissenschaften* **101**, 261–264 (2014).
- C. D. Michener, Notes on the habits of some Panamanian stingless bees (Hymenoptera, Apidæ). J. New York Entomol. Soc. 54, 179–197 (1946).
- C. D. Michener, S. Boongird, A new species of *Trigona* from Peninsular Thailand (Hymenoptera: Apidae: Meliponini). *J. Kansas Entomol. Soc.* 77, 143–146 (2004).
- 134. T. B. Mitchell, Bees of the Eastern United States, II. Tech. Bull. 152, 1–557 (1962).
- S. O'Donnell, M. Reichardt, R. Foster, Individual and colony factors in bumble bee division of labor (*Bombus bifarius nearcticus* Handl; Hymenoptera, Apidae). *Insectes Soc.* 47, 164–170 (2000).
- S. R. M. Pedro, J. M. F. Camargo, Meliponini Neotropicais: O gênero Partamona Schwarz, 1939 (Hymenoptera, Apidae). Rev. Bras. Entomol. 47, 1–117 (2003).
- D. W. Roubik, Foraging behavior of competing Africanized honeybees and stingless bees. *Ecology* 61, 836–845 (1980).
- S. F. Sakagami, Stingless bees (excl. *Tetragonula*) from the continental Southeast Asia in the collection of Bernice P. Bishop museum, Honolulu (Hymenoptera, Apidae). *J. Fac. Sci. Hokkaido Univ. Ser. VI Zool.* **20**, 49–76 (1975).
- 139. S. F. Sakagami, T. Inoue, S. Yamane, S. Salmah, Nests of the myrmecophilous stingless bee, *Trigona moorei*: How do bees initiate their nests within an arboreal ant nest? *Biotropica* **21**, 265–274 (1989).
- H. F. Schwarz, Results of the Oxford University Sarawak (Borneo) expedition: Bornean stingless bees of the genus *Trigona. Bull. Am. Mus. Nat. Hist.* 73, 281–328 (1937).
- 141. H. F. Schwarz, Stingless bees (Meliponidae) of the western hemisphere. *Lestrimelitta* and the following subgenera of *Trigona*: *Trigona*, *Paratrigona*, *Schwarziana*, *Parapartamona*, *Cephalotrigona*, *Oxytrigona*, *Scaura*, and *Mourella*. *Bull. Am. Mus. Nat. Hist.* **90**, 1–546 (1948).
- T. E. Shelly, S. L. Buchmann, E. M. Villalobos, M. K. O'Rourke, Colony ergonomics for a desert-dwelling bumblebee species (Hymenoptera: Apidae). *Ecol. Entomol.* 16, 361–370 (1991).
- 143. F. Smith, Catalogue of the hymenopterous insects collected at Sarawak, Borneo; Mount Ophir, Malacca; and at Singapore, by A. R. Wallace. J. Linn. Soc. London Zool. 2, 42–88 (1857).
- E. Tóth, D. C. Queller, A. Dollin, J. E. Strassmann, Conflict over male parentage in stingless bees. *Insectes Soc.* 51, 1–11 (2004).
- P. H. Williams, R. W. Thorp, L. L. Richardson, S. R. Colla, Bumble Bees of North America: An Identification Guide (Princeton Univ. Press, 2014).
- 146. N. Wilson-Rich, The Bee: A Natural History (Princeton Univ. Press, 2014).
- 147. M. L. Winston, The Biology of the Honey Bee (Harvard Univ. Press, 1991).

 S. Witter, B. Blochtein, Espécies de Abelhas Sem Ferrão de Ocorrôncia no Rio Grande do Sul (Centro Ecológico, 2009).

Acknowledgments: We thank D. Cariveau (Rutgers University), J. P. Strange (United States Department of Agriculture, Logan, Utah), L. Lach (James Cook University), R. Gloag (University of Sydney), T. Heard (Commonwealth Scientific and Industrial Research Organisation, Australia), J. J. Wilson (University of Malaya), J. X. Q. Lee (Singapore), the Po Sang Yuen Bee Farm (Hong Kong), and A. Kwong (Hong Kong) for providing samples, reagents, and/or advice for this project. In addition, we thank E. Frederick for technical assistance and J. E. Powell and N. Ratnayeke for *Bombus* data. **Funding:** This work was supported by Yale University, the Yale University Department of Ecology and Evolutionary Biology Chair's Fund, the Sigma Xi Grants-in-Aid of Research, and the Canadian Natural Sciences and Engineering Research Council Postgraduate Scholarship (to W.K.K.); the Swiss National Science Foundation Postdoctoral Fellowship (140157 and 147881 to H.K.); Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP 2012/13200-5 and 2013/23661-2 to R.J.); and the U.S. National Science Foundation Dimensions of Biodiversity (awards 1046153 and 1415604 to N.A.M.). **Author contributions:** W.K.K. conceived the study, designed sampling and experiments, conducted sampling, performed experiments, analyzed data, and wrote the paper. LA.M. isolated bacterial strains and performed experiments

and qPCR. H.K., K-W.S., EJ.Y.S., J.S.A, and R.J. designed and conducted sampling. N.A.M. advised on study design, provided samples and reagents, and wrote the paper. All authors participated in manuscript revisions. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** Sequences from bee identification, phylogenetic analysis of *rpoA* and *tuf* genes, and OTU₉₇ 16S rRNA clustering are deposited under GenBank accession numbers KU571725 to KU572279 (data file S4). Raw 16S rRNA gene amplicon reads are deposited in the National Center for Biotechnology Information Sequence Read Archive under accession SRP071118. All other data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Additional data related to this paper may be requested from the authors.

Submitted 10 March 2016 Accepted 10 February 2017 Published 29 March 2017 10.1126/sciadv.1600513

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

Science Advances Dynamic microbiome evolution in social bees

NAAAS

Waldan K. Kwong, Luis A. Medina, Hauke Koch, Kong-Wah Sing, Eunice Jia Yu Soh, John S. Ascher, Rodolfo Jaffé and Nancy A. Moran (March 29, 2017) *Sci Adv* 2017, 3:. doi: 10.1126/sciady.1600513

This article is publisher under a Creative Commons license. The specific license under which this article is published is noted on the first page.

For articles published under CC BY licenses, you may freely distribute, adapt, or reuse the article, including for commercial purposes, provided you give proper attribution.

For articles published under CC BY-NC licenses, you may distribute, adapt, or reuse the article for non-commerical purposes. Commercial use requires prior permission from the American Association for the Advancement of Science (AAAS). You may request permission by clicking here.

The following resources related to this article are available online at http://advances.sciencemag.org. (This information is current as of March 29, 2017):

Updated information and services, including high-resolution figures, can be found in the online version of this article at: http://advances.sciencemag.org/content/3/3/e1600513.full

Supporting Online Material can be found at: http://advances.sciencemag.org/content/suppl/2017/03/27/3.3.e1600513.DC1

This article **cites 130 articles**, 32 of which you can access for free at: http://advances.sciencemag.org/content/3/3/e1600513#BIBL

Science Advances (ISSN 2375-2548) publishes new articles weekly. The journal is published by the American Association for the Advancement of Science (AAAS), 1200 New York Avenue NW, Washington, DC 20005. Copyright is held by the Authors unless stated otherwise. AAAS is the exclusive licensee. The title Science Advances is a registered trademark of AAAS