E V O L U T I O N A R Y  B I O L O G Y

Dynamic microbiome evolution in social bees

Waldan K. Kwong,1,2a Luis A. Medina,2 Hauke Koch,21 Kong-Wah Sing,34 Eunice Jia Yu Soh,4 John S. Ascher,4 Rodolfo Jaffé,5,6 Nancy A. Moran2a

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 sympathy. 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 sociology 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 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 OTUs in a single OTU97 (comparatively, humans have 500 to 1000 OTUs; 22) and termite OTUs have ~102 to ~104, 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 Lactococcus 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 lineage-specific: Bartonella apis (23) and Frischella (24) in honey bees, Bombiscardovia (25) and Schmidtrompella (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...
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
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 Austrolebeia 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 declades. The gains of new microbial lineages have been tempered by a source of functional diversity (36). 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 (OTU99.5) resolves a number of distinct strains within each phyotype, 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 coevolution 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 phyotype 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 efficiencies. 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...
in honey bees (Fig. 7A, nested ANOVA OTUs97, P = 0.0420; fig. S11, Shannon’s H OTU99.5, 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...
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 formed 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 (10^2 to 10^5 versus 10^6 to 10^8 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, collinearity, 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 occurrence in all three corbiculate groups examined (Fig. 2), as well as by deep branching diversions 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. 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 as select nodes as percentages. Bar represents nucleotide substitutions per site. **(B)** Mono-inoculations demonstrate that S. alvi strains from other Apis spp. are able to colonize A. mellifera [bars, mean; detection limit, 5 × 10^3 colony-forming units (CFUs)]. **(C)** Co-inoculations suggest variability in competitiveness and that A. mellifera-derived strains are not always dominant (error bars, SEM).
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 by Koch (16). 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),
and there was differential presence of Snodgrassella and Parasaccharibacter 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 (39, 60). 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 (orcid 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). Long-term, generational 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 from 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 for habitat size (Fig. 7D and fig. S12). However, these proxies are incomplete: In social organisms, individuals are constantly exchanging

Fig. 8. Large bee colony size is associated with higher microbiome diversity. (A) Plot of microbiome diversity (number of OTUs at OTU97 clustering) against gut community size (16S rRNA gene copies) × bee colony size. (B) Plot of microbiome diversity (Shannon’s H at OTU99.5 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 OTU99.5 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.

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 plants and animals, taxon richness is often a function of habitat size, as modeled by the species-area relationship $S = c A$ or $S = \log(c + A^z)$, 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 bottlenecks) 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. Sociability alone is not enough for the establishment of host-specific microbiota, 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 SI). 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. atrieps* [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. atrieps* 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 µl of isopropanol and 70 µl 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 μl 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′-TATGGTAATTTGTTGCGACMGCCGCGGTAA), Read2R (5′-AGTCAGTCAGCGGCACTACHVGGGTTCTAAT), and Index (5′-ATTAGAWACCCBDGTAGTCCGGCTAGCTACTGACT).

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′-TCTGCGCAGCGT-CAGATGTGTATAAGAGACAGGTGYCAGCMGCCGCGGTAA) and Hyb806R_rRNA (5′-TCTGCGCAGCGT-CAGATGTGTATAAGAGACAGGTGYCAGCMGCCGCGGTA). For samples prefixed with “WK” (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′-TCTGCGCAGCGT-CAGATGTGTATAAGAGACAGGTGYCAGCMGCCGCGGTAA) and Hyb806R_rRNA (5′-TCTGCGCAGCGT-CAGATGTGTATAAGAGACAGGTGYCAGCMGCCGCGGTA). 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′-TCTGCGCAGCGT-CAGATGTGTATAAGAGACAGGTGYCAGCMGCCGCGGTAA) and Hyb806R_rRNA (5′-TCTGCGCAGCGT-CAGATGTGTATAAGAGACAGGTGYCAGCMGCCGCGGTA). 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′-TCTGCGCAGCGT-CAGATGTGTATAAGAGACAGGTGYCAGCMGCCGCGGTAA) and Hyb806R_rRNA (5′-TCTGCGCAGCGT-CAGATGTGTATAAGAGACAGGTGYCAGCMGCCGCGGTA).

For 99.5% OTU clustering (OTU99.5), de novo OTU picking was done with pick_otus.py and pick_rep_set.py using default settings (uchlt algorithm). Singleton OTUs were filtered out, as were samples with read depths of less than 5000. Remaining samples were then rarefied to 5000 reads. OTU99.5 taxonomy was assigned with parallel_assign_taxonomy_blast.py using the manually refined taxonomy from OTU97 as reference. Representative sequences from OTU97 16S rRNA clustering were deposited in GenBank (data file S4), and representative sequences from OTU97 and OTU99.5 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 OTU97 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 OTU97 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 OTU99.5 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 Smadgrassella, Gilliamella, Lactobacillus Firm-4, Lactobacillus Firm-5, Bifidobacterium, B. apis, Frischella, Bombiscardovia, Schmidhemelia, 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 10^4 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 OTUs, 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 “MuMln” 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 (κ = 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’-GAATTTGAAAACCAAAATTACCG) and wk122 (5’-CGTACGCATCATCATCGC), whereas the gene encoding translation elongation factor Tu, *tuf*, was amplified with wk123 (5’-GAACAAACGCCACGTAAAC) and wk124 (5’-GACAGCCACACAGTACGA). 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’-TGCGTAAA-GAAAAATTIGCCGCG) and wk136 (5’-AAGCGTAAACTTTGCAACCACAC), producing a 1114-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* DH5α, 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 scored 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 bifaculatus* (wkB12) were used to colonize laboratory-reared *A. mellifera*. Approximately 10⁵ 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 ~10⁵ 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 μg/mL). Strains wkB2 and wkB332 are tetracycline-resistant, whereas the other strains are susceptible. Count data are presented in table S2.

**SUPPLEMENTARY MATERIALS**

Supplemental material for this article is available at https://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 OTU1, clustering.

fig. S3. NMDS plots of gut community Sørensen-Dice dissimilarities at OTU10, 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 strains.

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 *Firm-5* phylogenetic topologies.

fig. S10. Tanglegram of host bee and *Lactobacillus Firm-5* phylogenetic topologies.

fig. S11. Gut microbial diversity, as measured by Shannon's *H* on OTU clusters 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. OTU10 tables and curated taxonomic assignments.

data file S3. OTU10 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 OTU10 and OTU20, clustering.

References (122–148)
Kwong et al., Sci. Adv. 2017;3:e160513. 29 March 2017

14 of 16
N. Lo, R. S. Gloag, D. L. Anderson, B. P. Oldroyd, A molecular phylogeny of the genus Centris (Hymenoptera: Apidae). *Univ. Kansas Sci. Bull.* **55**, 124–162 (1993).

J. Fox, S. Weisberg, D. Adler, D. Bates, G. Baud-Bovy, S. Ellison, D. Firth, M. Friendly, D. Darriba, G. L. Taboada, R. Doallo, D. Posada, jModelTest 2: More models, new options, and their direct application for species identification. *Bioinformatics* **29**, 289–290 (2004).

J. L. 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.

K. M. Fabricius, are valid species. *Zool. Scr.* **45**, 335–336 (2016).

A. C. Martins, G. A. R. Melo, The New World oil-collecting bees Implications for the origin of pollen baskets. *Mol. Phylogenet. Evol.* **80**, 88–94 (2014).

A. Roig-Ausina, 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, J. Kuczynski, J. Stombaugh, K. Bittinger, F. D. Bushman, E. K. Costello, N. Fierer, A. G. Peña, J. K. Goodrich, J. I. Gordon, A. H. Huttenhower, S. T. Kelley, D. Knight, J. E. Koenig, C. L. Lauber, W. A. Walters, J. Widmann, T. Yatsunenko, J. Zaneveld, M. A. Suchard, J. P. Huelsenbeck, MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* **61**, 312–1313 (2012).

J. A. A. Nylander, J. C. Wilgenbusch, D. L. Warren, D. L. Swoford, AWTY (are we there yet?): A system for graphical exploration of MCMC convergence in Bayesian phylogenetics. *Bioinformatics* **24**, S81–S83 (2008).

119. A. Stamatakis, RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1314 (2014).

120. 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).

121. 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 fuscohostensis (Hymenoptera: Apidae: Meliponini). *J. Kansas Entomol. Soc.* **84**, 196–196 (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 Tecnologe E Ambiente* **10**, 35–43 (2004).

125. A. E. Dolell, L. J. Dolell, 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. 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 omtacito (Hymenoptera, Apidae). *Zool. Keys* **120**, 9–25 (2011).

128. M. Halcroft, R. Spooner-Hart, P. Neumann, Behavioral defense strategies of the stingless bee, Austrolepeus australis, against the small hive beetle, Aethina tumida. *Insects Soc.* **58**, 245–253 (2011).

129. G. A. Hobbs, Ecology of species of Bombus Latr. (Hymenoptera: Apidae) in Southern Alberta. V. subgenus Subterraneobombus Vogt. *Canad. Entomol.* **98**, 288–294 (1966).

130. 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).

131. R. Jaffe, F. C. Pisker-Hara, C. F. Dos Santos, D. A. Alves, A. de M. P. Kleinert, T. M. Francoy, M. C. Aries, V. L. Imperatriz-Fonseca, Monogamy in large bee societies: A stingless paradox. *Naturscience* **101**, 261–264 (2014).

132. C. D. Michener, Notes on the habits of some Panamanian stingless bees (Hymenoptera, Apidae). *J. New York Entomol. Soc.* **54**, 179–197 (1946).

133. 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).

135. S. O'Donnell, M. Reichardt, R. Foster, Individual and colony factors in bumble bee division of labor (Bombus bifarius nearticus Handl; Hymenoptera, Apidae). *Insects Soc.* **47**, 164–170 (2000).

136. S. R. M. Pedro, J. M. F. Camargo, Meliponini Neotropicais: O gênero Trigona (Hymenoptera: Apidae). *Zool. Scr.* **19**, 896–898 (1997).

137. S. F. Sakagami, T. Inoue, S. Yamane, S. Salmah, Nests of the myrmecophilous stingless bee, Trigona moarei: How do bees initiate their nests within an arboreal ant nest? *Biotropica* **21**, 265–274 (1989).

138. 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).

139. H. F. Schwarz, Stingless bees (Meliponidae) of the western hemisphere. *Lestrimelitta* and the following subgenera of Trigona: Trigona, Paratrigona, Schwarziana, Paraparataguna, Cephalotrigona, Oxytrigona, Scuara, and Mouarella. *Bull. Am. Mus. Nat. Hist.* **90**, 1–546 (1948).

140. 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).

141. 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).

142. E. Töth, D. C. Queller, A. Dolell, J. E. Strassmann, Conflict over male parental care in stingless bees. *Insect. Soc.* **51**, 1–11 (2004).

143. P. H. Williams, R. W. Thor, L. L. Richardson, S. R. Colla, Bumble Bees of North America: An Identification Guide (Princeton Univ. Press, 2014).

144. N. Wilson-Rich, The Bee: A Natural History (Princeton Univ. Press, 2014).

145. M. L. Winston, The Biology of the Honey Bee (Harvard Univ. Press, 1991).

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

15 of 16
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. L.A.M. isolated bacterial strains and performed experiments and qPCR. H.K., K.-W.S., E.J.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).