MOLECULAR ECOLOGY

Molecular Ecology (2014) **23**, 1268–1283 doi: 10.1111/mec.12611

SPECIAL ISSUE: NATURE'S MICROBIOME

Stability and phylogenetic correlation in gut microbiota: lessons from ants and apes

JON G. SANDERS,* SCOTT POWELL,† DANIEL J. C. KRONAUER,*‡
HERALDO L. VASCONCELOS,§ MEGAN E. FREDERICKSON¶ and NAOMI E. PIERCE*
*Department of Organismic and Evolutionary Biology, Harvard University, 26 Oxford Street, Cambridge, MA 02138, USA, †Department of Biological Sciences, George Washington University, Lisner Hall Room 340, 2023 G Street, NW, Washington, DC 20052, USA, ‡Laboratory of Insect Social Evolution, The Rockefeller University, 1230 York Avenue, New York, NY 10065, USA, §Instituto de Biologia, Universidade Federal de Uberlândia (UFU), Caixa Postal 593, 38400-902 Uberlândia, MG, Brazil, ¶Department of Ecology and Evolutionary Biology, University of Toronto, 25 Harbord Street, Toronto, ON M5S3G5, Canada

Abstract

Correlation between gut microbiota and host phylogeny could reflect codiversification over shared evolutionary history or a selective environment that is more similar in related hosts. These alternatives imply substantial differences in the relationship between host and symbiont, but can they be distinguished based on patterns in the community data themselves? We explored patterns of phylogenetic correlation in the distribution of gut bacteria among species of turtle ants (genus *Cephalotes*), which host a dense gut microbial community. We used 16S rRNA pyrosequencing from 25 *Cephalotes* species to show that their gut community is remarkably stable, from the colony to the genus level. Despite this overall similarity, the existing differences among species' microbiota significantly correlated with host phylogeny. We introduced a novel analytical technique to test whether these phylogenetic correlations are derived from recent bacterial evolution, as would be expected in the case of codiversification, or from broader shifts more likely to reflect environmental filters imposed by factors such as diet or habitat. We also tested this technique on a published data set of ape microbiota, confirming earlier results while revealing previously undescribed patterns of phylogenetic correlation. Our results indicated a high degree of partner fidelity in the *Cephalotes* microbiota, suggesting that vertical transmission of the entire community could play an important role in the evolution and maintenance of the association. As additional comparative microbiota data become available, the techniques presented here can be used to explore trends in the evolution of host-associated microbial communities.

Keywords: co-evolution, insects, microbiota, mutualism, primates, pyrosequencing

Received 11 March 2013; revision received 22 November 2013; accepted 27 November 2013

Introduction

Gut microbes have had an enormous impact on animal evolution (McFall-Ngai *et al.* 2013). In addition to their well-documented nutritional role in herbivorous mammals, gut bacteria have recently been implicated in processes ranging from brain development in mice (Diaz Heijtz *et al.* 2011) to sexual selection in *Drosophila* (Sharon *et al.* 2010). But our understanding of the mechanisms governing the ecology and evolution of these communities is still hampered by a paucity of comparative data.

In mammals, hosts to perhaps the best-studied animal microbiota, comparative analysis has revealed remarkably consistent correlations between host diet and microbiota composition (Ley *et al.* 2008; Muegge *et al.* 2011), with host lineages sharing convergent gut physiology and diet also tending to share similar gut microbes. Microbiota composition has also been shown

Correspondence: Jon G. Sanders, Fax: +1 (617) 495 5667;
E-mail: jsanders@oeb.harvard.edu

to correlate with host phylogeny in the great apes (Ochman *et al.* 2010), perhaps mediated by maternal (i.e. vertical) transmission of microbes, as has been observed in humans for *Helicobacter pylori* (Falush *et al.* 2003). However, subsequent studies of ape microbiota have found the picture to be substantially more complicated than simply one dominated by vertical transmission: maternal lineage explained only a small proportion of variance among chimpanzee microbiota (Degnan *et al.* 2012), while individual chimpanzees switched gut community 'enterotypes' over time (Moeller *et al.* 2012). While humans have recently been shown to maintain individual microbial taxa across several years (Faith *et al.* 2013) and to share a somewhat greater proportion of microbial taxa with relatives than with unrelated individuals (Turnbaugh *et al.* 2009; Faith *et al.* 2013), how these dynamics scale across millions of years remains uncertain.

Evidence for consistent trends in the evolution of insect gut microbiota is mixed. Compared to vertebrates, insects tend to have less diverse and potentially more labile gut microbial associations (Dillon & Dillon 2004; Engel & Moran 2013). Recent studies have examined gut microbiota of flies (Chandler *et al.* 2011), mosquitos (Osei-Poku *et al.* 2012) and leaf beetles (Kelley & Dobler 2010) using next-generation sequencing in a comparative framework; but in each case, microbiota composition has shown little if any correlation with host phylogeny. Where targeted sequencing approaches have revealed extracellular gut symbionts that closely track the evolutionary history of their hosts, it has been in highly specialized cases: stink bugs that sequester symbiont monocultures in gut crypts (Kikuchi *et al.* 2009) and termites whose codiversifying bacteria are themselves physically associated with protists (Noda *et al.* 2007). Detailed comparative surveys of the entire gut microbiota in these organisms are still lacking, limiting our ability to understand how these patterns translate from microbial lineage to microbiota.

For all of these comparative analyses of microbiota, assessing the importance of phylogenetic correlation at the community level is hampered by the potential input of multiple causative factors. Even while some microbial lineages (such as *Helicobacter* in humans) are reliably passed vertically from mother to offspring, the bulk of the community may be acquired horizontally from environmental sources. Random horizontal transmission would simply obscure underlying patterns of phylogenetic correlation. But if the ecological success of particular horizontally acquired gut microbes is dependent on selective conditions that are themselves strongly conserved across the host phylogeny (e.g. diet, habitat or immunity), phylogenetic correlation would be observed even in the absence of vertical transmission.

In this study, we present one of the first comparative studies of ant gut microbiota using next-generation sequencing, along with a new approach designed to untangle the factors underlying correlation between host phylogeny and microbiota composition. Among insects, ants make a particularly appealing system for studying gut microbiota ecology and evolution. Like mammals, ants utilize a broad range of diets (Davidson *et al.* 2003), permitting comparisons among convergently evolved hosts. Like many other insects, some ant lineages harbour vertically transmitted endosymbionts, which may eventually permit tests comparing the roles of intracellular and extracellular symbioses. We target the arboreal ant genus *Cephalotes*, known as 'turtle ants' for their shell-like exoskeletal armour, which host a dense gut microbiota in their midgut and morphologically elaborated hindgut (Fig. S1, Supporting information; Roche & Wheeler 1997; Bution & Caetano 2008; Anderson *et al.* 2012).

Although no unequivocal demonstration has been published to date, circumstantial evidence suggests that *Cephalotes* relies upon its specialized gut microbiota to complement a nutritionally imbalanced diet (Jaffe *et al.* 2001). While the details of their feeding biology are still not well understood, *Cephalotes* are, like many other tree-nesting ants, thought to subsist largely on plant-derived nutrients (de Andrade & Baroni Urbani 1999; Davidson *et al.* 2003), including pollen, sap from leaf wounds, insect honeydew and extrafloral nectar (Byk & Del-Claro 2010; Gordon 2012). This is complemented by feeding on vertebrate waste products and especially bird droppings (Weber 1957; Adams 1990; Jaffe *et al.* 2001; Powell 2008). Captive colonies fed artificial diets maintain their microbial communities over long timescales (Russell *et al.* 2009; Hu *et al.* in press), and many of the bacterial lineages appear to be specific to the genus (Anderson *et al.* 2012), giving this association two of the hallmarks of a co-evolved mutualistic relationship.

Utilizing 454 sequencing of bacterial 16S genes from 25 *Cephalotes* and several outgroup species, we show that cephalotine ants (*Cephalotes* plus their sister genus *Procryptocerus*) host a relatively simple and remarkably stable microbiota. In contrast both to the noncephalotine ants in this study, we find that *Cephalotes* nestmates harbour very similar communities. Finally, we use a novel application of sensitivity analysis of microbiota clustering to investigate the role of host phylogeny in explaining the distribution of microbial diversity within the genus *Cephalotes*. This novel approach is further tested and validated using a reanalysis of the phylogenetically correlated ape microbiota data set from Ochman *et al.* (2010) for comparison.

Methods

Sample collection and preservation

We collected samples to permit comparisons across the *Cephalotes* host phylogeny and also to explore the influence of geography and colony structure on microbiota composition. To that end, we concentrated novel collections at two field sites with high *Cephalotes* species diversity: the Panga Ecological Station (Estação Ecológica do Panga) in Minas Gerais, Brazil (19°10′ S, 48°23′ W); and the Los Amigos research station (Centro de Investigación y Capacitación de Rio Los Amigos) in Madre de Dios, Peru (12°33′ S, 70°8′ W). These two sites represent very different ecosystems, contributing both to the potential phylogenetic diversity of host species and to ecological diversity with respect to exogenous microbes. The Brazil site is a dry Neotropical savannah biome or 'cerrado', with 3- to 5-m-tall trees providing 30–50% canopy cover. The Peru site is dominated by primary wet tropical forest, approaching 100% canopy cover at a canopy height of ~30 m. In total, we were able to include 25 *Cephalotes* colonies, representing 17 unique species, from these two locations (for detailed collections information, see Table S1). Four species (*Cephalotes minutus*, *Cephalotes atratus*, *Cephalotes maculatus* and *Cephalotes clypeatus*) are represented in the data set by a colony collected in each location. To permit additional intraspecific, intercolony comparisons, two abundant species from the Brazil site (*Cephalotes pusillus* and *Cephalotes persimilis*) are represented by three colonies each. Three individuals from *C. pusillus* colony #12 were included on two separate sequencing rounds, serving as example technical replicates. Most colonies were discovered by baiting trees with nitrogen-rich baits (as in Powell 2008), with some additional colonies discovered by twig snapping.

This primary sample set was augmented with additional samples to provide broader coverage of the *Cephalotes* phylogeny as well as outgroups for genus-level comparisons (Table S1). We collected one additional *Cephalotes* colony (*C. rohweri*) from desert scrubland in Arizona and included individual ants from seven additional *Cephalotes* species from ethanol-preserved museum collections. For outgroups, we included two ethanol-preserved individuals from separate museum collections of *Procryptocerus* (the sister genus to *Cephalotes*), as well as one colony from each of the arboreal ants *Crematogaster*, *Azteca* and *Pseudomyrmex* collected from the Brazil field site concurrently with *Cephalotes* colonies.

We processed samples to maximize the representation of gut microbes while trying to limit the influence of exogenous contamination from either environmental or host-derived, nongut microbes. Thus, for the colonies collected specifically for this study (excepting museum samples and workers from the *C. rohweri* colony, which were kept for several weeks in the laboratory before being dissected and immediately extracted without intermediate preservation), guts were dissected in the field within 2 days of collection. Adult workers were killed in 100% ethanol, transferred within 30 min to a 1:10 solution of bleach in distilled water (final concentration ~0.5% sodium hypochlorite) for 30–60 s and rinsed in filter-sterilized phosphate-buffered saline (PBS). Dissections took place in sterile PBS. Dissected guts were transferred immediately to filter-sterilized RNALater nucleic acid preservative (Ambion, Inc). To maximize consistency among samples, only the midgut and ileum (excluding the fragile crop and rectum) were extracted for sequencing. For freshly collected specimens, three workers were sequenced separately per colony.

As ethanol dehydration renders internal structures extremely fragile, museum specimens were not dissected. Instead, whole individual ants were transferred from ethanol to bleach solution as above, rinsed in PBS and just the gasters retained for analysis. Due to limited specimen availability, only a single worker was sequenced per colony for these species. To help evaluate the potential biases imposed by alternative preservation methods, three additional workers from one of the freshly collected *C. pusillus* colonies were preserved in ethanol and processed in the same manner as the museum specimens.

DNA extraction and sequencing

We employed a relatively intensive extraction protocol, based on a method developed for sampling termite gut microbes (Matson *et al.* 2007), to minimize the potential for biases against difficult-to-lyse microbes (e.g. Firmicutes, see Willner *et al.* 2012) and to decrease the potential influence of inhibitors on downstream enzymatic reactions. Briefly, tubes containing preserved guts in RNALater were diluted ~1:1 with sterile water (to decrease solution density and dissolve any precipitated salts) and spun for 10 min at 14 000 rcf. Supernatant was removed and replaced with 700 μL of TLS-C sample lysis buffer (MPBio, Inc). Tubes were then vortexed at maximum speed for 1 min to resuspend pelleted material. Contents were transferred to sterile lysis tubes containing a bead mixture (Lysis matrix A, MPBio) and 500 μL phenol:chloroform:iso-amyl alcohol, pH 8, and mechanically lysed for 40 s at maximum speed on an MPBio FastPrep-20. After lysis, tubes were centrifuged

at 8000 rcf for 1 min, and the aqueous phase removed and washed with 500 μL chloroform. The remaining aqueous phase was column-purified using Qiagen DNeasy Blood and Tissue extraction columns, starting with addition of equal volumes of buffer AL and 100% molecular-grade ethanol to the aqueous phase and application to the column. The remainder of the cleanup was performed according to manufacturer's instructions. Finally, purified extracts were concentrated by isopropanol precipitation, resuspended in 32 μL TE and quantified using a Qubit fluorometer (Invitrogen).

To characterize the microbial community, an approximately 500-bp fragment spanning the V1-V3 regions of the bacterial 16S gene was amplified with universal primers 27F and 515R (Kumar *et al.* 2011) and sequenced using 454 Titanium chemistry at a commercial facility (Research and Testing Laboratories, Lubbock, TX, USA) according to previously published protocols (Dowd *et al.* 2008). Briefly, amplifications were performed in 25 μL reactions using 1 μL of each 5 μM primer and 1 μL of template. Reactions were performed under the following thermal profile: 95 °C for 5 min, then 35 cycles of 94 °C for 30 s, 54 °C for 40 s, 72 °C for 1 min, followed by one cycle of 72 °C for 10 min.

16S rRNA sequence filtering and clustering

All microbiota community sequences were processed to limit the effect of amplification and sequencing artefacts using AMPLICONNOISE v1.25 under parameters recommended for 454 Titanium chemistry (Quince *et al.* 2009; see Appendix S1, Supporting information for a brief discussion of this process). Denoised sequences were then analysed using QIIME v1.4.0 (Caporaso *et al.* 2010). Some samples initially yielded fewer sequences than we specified, and the same amplicon pools were rerun by the sequencing facility; in these cases, we excluded sequences from the poorer initial run from further analysis. Specific analysis scripts are detailed in Supporting information, but with the exception of modifications to sequence clustering detailed below, most steps were performed using default parameters. We assigned taxonomy to sequences with the RDP classifier (Wang *et al.* 2007) using the curated GreenGenes 16S rRNA database dated 4 February 2011, available from the QIIME website (Werner *et al.* 2012).

We clustered sequences into operational taxonomic units (OTUs) using two clustering algorithms and a total of four different similarity thresholds (93%, 95%, 97% and 99%). Chimeric sequences were removed using both de novo and reference-based chimera removal (using the Gold 16S database described in Haas *et al.* 2011) with UCHIME (Edgar *et al.* 2011). Clustering sequences into OTUs reduces the impact of sequencing error, speeds computation and permits analyses unbiased by assumptions about bacterial taxonomy. However, different OTU clustering algorithms may give very different—and sometimes overstated—estimates of bacterial diversity in a sample (Huse *et al.* 2010). Furthermore, while different percentage similarity thresholds should theoretically correspond to different average evolutionary divergence times, various clustering algorithms can yield surprisingly different results at a given similarity threshold (Sun *et al.* 2012). Thus, we performed separate analyses at 93%, 95%, 97% and 99% similarity thresholds using the CD-HIT (Li & Godzik 2006) and UCLUST (Edgar 2010) algorithms implemented in QIIME. Both algorithms were run under default parameters.

Comparative analyses

We calculated basic descriptive and comparative statistics for microbiota data under all clustering parameter combinations using QIIME 1.4.0. To permit summarized comparisons across the broadest possible range of samples, alpha-diversity estimates (including observed species richness, Shannon diversity, the Chao1 nonparametric richness estimator and whole-tree phylogenetic diversity) were calculated using sample by OTU abundance observation tables (OTU tables) rarified to 1000 observations per sample and excluding samples with less than 1000 high-quality sequences. Beta-diversity metrics (including abundance-weighted and unweighted UniFrac distances and Sørensen, Jaccard and Bray–Curtis dissimilarities) were calculated using OTU tables rarified to 1000 observations, but retaining samples with fewer sequences (Aguirre de Cárcer *et al.* 2011). Alpha- and beta-diversity calculations, including collectors' curves, PCoA calculations and OTU network tables, were generated in QIIME. OTU networks were visualized in CYTOSCAPE v2.8.1 (Shannon *et al.* 2003). We also calculated the 'core' OTUs (present in ≥50% of *Cephalotes* samples) for 93%, 95% and 97% thresholds using QIIME v1.6.0.

We used a number of methods to explore the impacts of colony structure, host phylogeny and geography on microbiota composition. Between-sample geographic distances were calculated from sample locality information using the AMNH geographic distance calculator tool (Ersts 2013). Host genetic distances were calculated as patristic distances in PyCogent (Knight *et al.* 2007) using the time-calibrated *Cephalotes* phylogeny of Price *et al.* (2013) and modified by hand to include the outgroup genera *Pseudomyrmex*, *Azteca* and *Crematogaster* with approximate branching times as indicated in the phylogeny of the ants by Moreau *et al.* (2006).

Beta-diversity dissimilarity matrices, along with host geographic and genetic distances, were imported into R (R Development Core Team 2012) for further analysis. To test for differences in average between-sample distances among sample categories [e.g. across levels of host colony, species, clade (*sensu* de Andrade & Baroni Urbani 1999) and genus], we used Monte Carlo permutations of category labels to generate null distributions of between-sample distances appropriate to each comparison; direct comparisons of pairwise distances violate assumptions of independence for most other statistical tests. Specific permutation designs are described in the Supporting information. To test for significant associations between bacterial community dissimilarities and host genetic and geographic distances, we used partial Mantel tests, as implemented in the vegan package in R (Oksanen *et al.* 2012).

Beta-diversity clustering sensitivity analysis

Clustering of microbial communities by similarity has been shown to recapitulate host phylogeny in a few cases, such as between species of great apes and their microbiota (Ochman *et al.* 2010), lending support to the idea that gut communities co-evolve with their hosts. However, sample clustering is strongly influenced by OTU picking parameters and choice of diversity metric (Hamady & Knight 2009). Changes in the pattern of support under these different parameter conditions may be useful in interpreting biological significance. For microbial communities that correlate with host phylogeny, we would expect measures of beta-diversity to be differentially affected by changes to OTU clustering, depending on how these correlations arose (Fig. 1). For example, if neutral codiversification was the sole force shaping the gut community of a 15-Myr-old host genus and mutations in 16S rRNA accumulated uniformly among gut microbes at a rate of 0.1% per million years, homologous gut microbes—meaning, specifically, those derived from a shared ancestral microbe—across the genus should be at least 98.5% identical (or at most 1.5% different) at the 16S locus. In this case, clustering microbial OTUs at 97% similarity or below would cause all hosts in the genus to appear to host identical gut communities, and microbiota from closely related hosts would not be expected to appear more similar. Clustering OTUs at 99% would start to reveal phylogenetically correlated microbiota structure, with host clades separated by <10 Myr grouping together. Thus, the constraint on genetic distances between microbes imposed by the age of the hosts' most recent common ancestor would be reflected in a threshold OTU clustering width, below which beta-diversity metrics would no longer be reflective of host phylogeny. By contrast, when correlation of microbiota structure with host phylogeny is mediated by environmental factors (e.g. differences in diet that in turn correlate with host phylogeny), we would expect beta-diversity metrics to reflect host phylogeny with much less sensitivity to OTU clustering width.

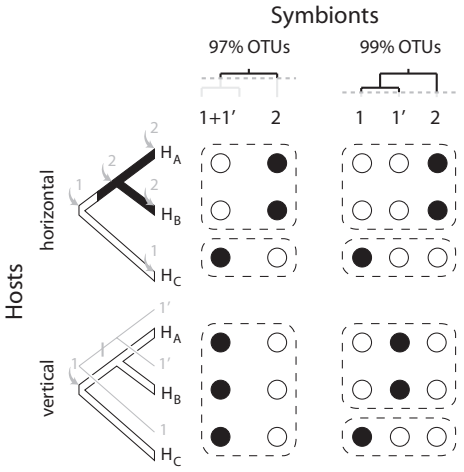

Fig. 1 Conceptual illustration of the beta-diversity sensitivity analysis. Host ($H_{A,B,C}$) phylogenies are depicted on the left, symbiont (1, 1′, 2) phylogenies above. In each scenario, presence/absence of symbionts in each host is depicted by filled/open circles, respectively. Similar host communities for each scenario are grouped by a dotted border. Top row: hosts select symbionts horizontally from the environment with each generation. A change in host diet (depicted by black fill on host tree) leads to the replacement of symbiont 1 by the symbiont 2 in H_A and H_B. As symbionts 1 and 2 diverged long ago, they fall into different OTUs at both 97% and 99% clustering widths, so grouping of microbiota H_A and H_B separately from H_C is insensitive to clustering width. Bottom row: hosts acquire symbionts vertically from parent generation. A mutation in symbiont 1′ (depicted by a vertical hash) causes sequence divergence sufficient to cluster separately from symbiont 1 under 99% OTU clustering, but not under 97%. Consequently, grouping of microbiota H_A and H_B separately from H_C is sensitive to clustering width.

We would also expect different measures of beta-diversity to perform differently in these cases. 'Star phylogeny' measures of beta-diversity, such as Sorensen and Jaccard dissimilarities, weight each OTU equally, regardless of how closely related two different OTUs might be (Lozupone & Knight 2008). These methods effectively increase the sensitivity of the metric to recent bacterial evolution, as divergence just exceeding the OTU clustering threshold will have the same effect as much older splits. By contrast, the UniFrac metric is designed to minimize the effects of such recent bacterial evolution by weighting the longer, internal branches of the bacterial phylogeny. Thus, we would expect communities differing primarily due to recent diversification to separate

more clearly using Jaccard dissimilarities than UniFrac distances.

Borrowing a technique from systematics (Sanders 2010), we visualized the sensitivity of beta-diversity-based sample clustering to various parameter combinations using a series of grids overlaid on the host phylogeny. We generated UPGMA-clustered dendrograms of 100 jackknifed OTU tables for all five beta-diversity metrics and four OTU clustering thresholds and compared these to the host phylogeny using the tree_compare.py script in QIIME. We performed this analysis for our entire ant data set, using OTU tables summarized by colony. For context, we also performed this analysis on the great apes microbiota data set of Ochman *et al.* (2010) mentioned above. Because the ape data were sequenced at greater depth, we repeated the analysis at both a level of rarefaction close to that used for the *Cephalotes* data (1000 sequences/sample) and at much higher coverage (15 000 sequences/sample) to approximate previously published analyses of these data. Partial Mantel correlations were performed as for the ants, using patristic distances from the time-calibrated whole-genome phylogeny of Prado-Martinez *et al.* (2013) and geographic distances estimated from Fig. 1 in Ochman *et al.* (2010) using Google Earth.

Results

Sequencing results and taxonomic composition

The 454 sequencing reveals that *Cephalotes* host a relatively simple microbiota that is remarkably conserved. After denoising, clustering and chimera-checking, we generated a total of 241 519 sequences from 102 specimens in a nested design, permitting comparisons within colonies, among conspecific colonies, among *Cephalotes* species, among geographic areas and between *Cephalotes* and four other genera (see Table S1, Supporting information for per-sample sequence counts).

Cephalotes gut microbiota from across the genus were dominated by Verrucomicrobia and Proteobacteria (Fig. 2). This is the first genus-wide survey with broad phylogenetic sampling for *Cephalotes* and is consistent with previous results from small numbers of species (*C. atratus*, *C. rohweri* and *C. varians*; see Russell *et al.* 2009; Anderson *et al.* 2012; Kautz *et al.* 2013). These communities appear to be relatively simple, averaging just 20 unique 97% OTUs per 1000 sequences (Fig. S2, Table S3, Supporting information). Many of these OTUs were widely distributed across the genus, occurring in more than 50% of samples. All of these 'core OTUs' were close matches to sequences from clades that have previously been described as *Cephalotes* specific (Fig. S3, Supporting information; Anderson *et al.* 2012). Verrucomicrobia sequences dominated most *Cephalotes* gut samples, although the combined effects of tissue choice and preservative had strong effects on relative abundance (Fig. 2; see Supporting information for additional discussion). The two *Procryptocerus* samples were broadly similar to *Cephalotes* in both taxonomic composition and measures of species richness and alpha-diversity (Fig. 2 and Fig. S2, Supporting information). By contrast, gut microbiota from noncephalotines showed very little taxonomic overlap with those from *Cephalotes* (Fig. 2), and at least for *Azteca* and *Pseudomyrmex*, were substantially more diverse (Fig. S2, Table S3, Supporting information). Additional results, including notable trends in relative abundance for particular microbial taxa, can be found in the Supporting information.

Beta-diversity: effects of colony structure

Cephalotes nestmates had gut communities considerably more similar to one another than to the gut communities of conspecifics from other colonies or to communities from other species. At the 97% OTU clustering threshold, the average Jaccard dissimilarity among *Cephalotes* nestmates was 0.66 (SD = 0.16, $n = 90$), indicating that approximately one-third of OTUs were shared between individuals from the same colony (Fig. 3). This probably overestimates the divergence between samples, as the three technical replicates in our sample showed a comparable level of dissimilarity (mean of 0.74, SD = 0.15, $n = 3$). Several factors are likely to have contributed to the apparently high technical variance component, perhaps dominated by the relatively high PCR cycle number performed by the sequencing facility (see Supporting information for additional discussion). Given that caveat, nestmates appear to share quite similar communities. By comparison, just 21% of OTUs (Jaccard dissimilarity = 0.79) were shared between conspecific individuals from separate colonies (SD = 0.078, $n = 108$), and only marginally more than the 14% shared between heterospecific *Cephalotes* (SD = 0.072, $n = 270$; $P = 0.012$). UniFrac distances between samples showed a similar pattern (Fig. S2, Supporting information).

In contrast to the consistency observed among *Cephalotes* nestmates, just 8% of OTUs were shared between nestmates of the outgroup genera *Pseudomyrmex*, *Crematogaster* and *Azteca* (Jaccard dissimilarity = 0.92, SD = 0.030, $n = 9$). Dissimilarities between heterospecifics among these three genera were higher than within-colony dissimilarities (mean = 0.98, SD = 0.030, $n = 93$), although not significantly so ($P = 0.095$).

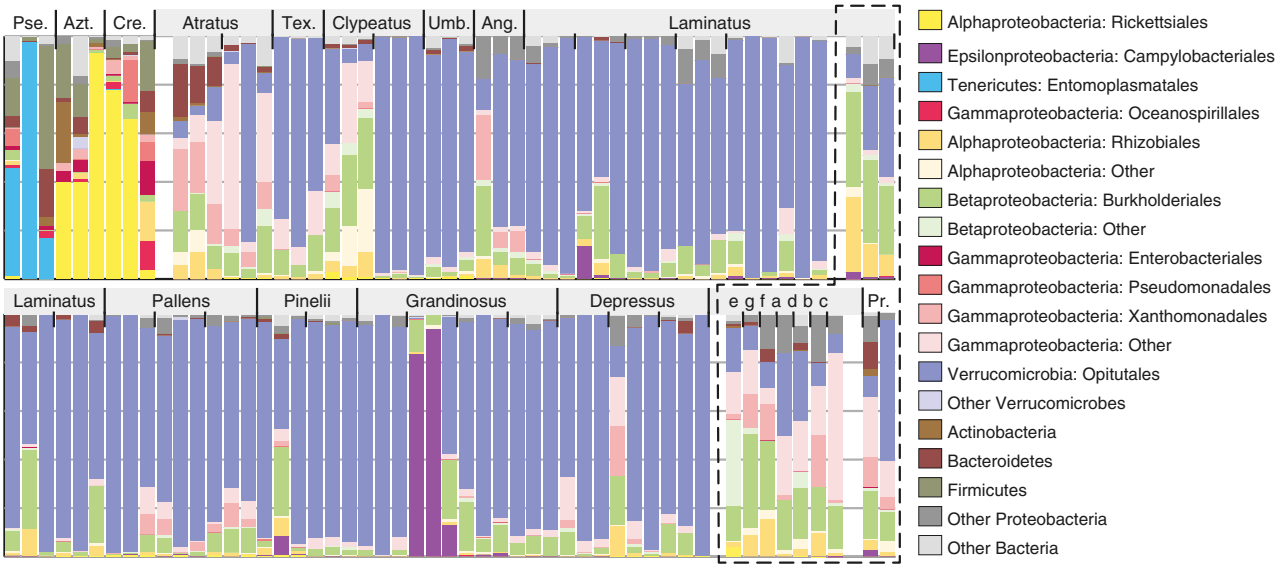

Fig. 2 Distribution of class-level taxonomic diversity across sampled ant gut microbiota. Bars are placed roughly according to host phylogeny, with small ticks between colonies and large ticks separating host genera/clades. Museum specimens are labelled with colony letter (see Table 1 for reference). Blank spaces are left purely to assist in visual separation of groups. Museum samples derived from ethanol-preserved gasters are circled by a dashed line to emphasize the apparent effect on relative abundance of Verrucomicrobia sequences. Three freshly collected individuals were also preserved in ethanol and sequenced from whole gasters; these are placed on the top right, next to dissected gut-derived samples from the same nest.

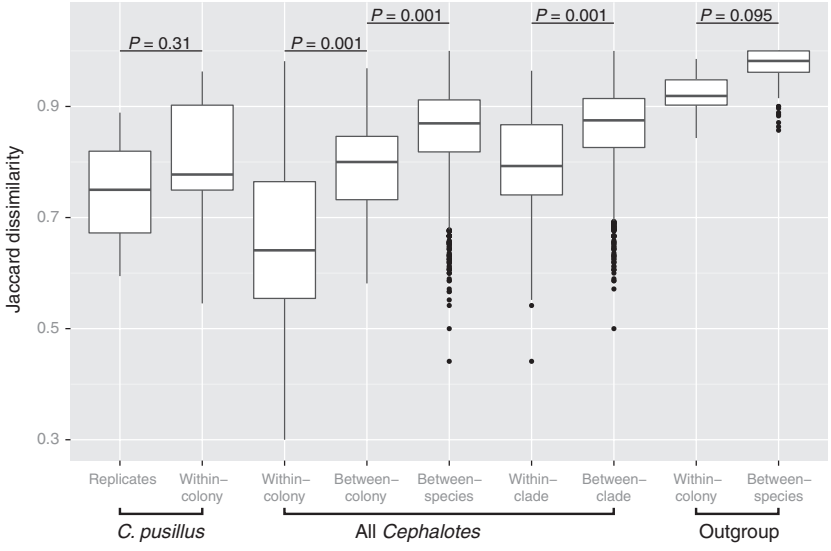

Fig. 3 Dissimilarity boxplots, comparing community Jaccard dissimilarities within and among groups of ants for 97% OTUs. Significance values for between-group comparisons were calculated using Monte Carlo permutation tests (see Supporting information for details on permutation structure).

Beta-diversity: effects of host phylogeny and biogeography

Network visualization of shared OTUs (Fig. 4) suggested that host phylogeny plays an important role in structuring the *Cephalotes* gut community. In the network analysis of 99% OTUs, samples were positioned using a spring-embedded edge-weighted algorithm, which places individual samples closer together solely based on the number of shared OTUs. By colouring sample nodes according to their position on the host phylogeny, clear divisions were visible among *Cephalotes* clades. Divisions between the cephalotine ants and the outgroups, and between *Cephalotes* and *Procryptocerus*, were even more apparent at the wider 97% OTU clustering width (Fig. S5, Supporting information).

Principal coordinates analysis of beta-diversity dissimilarities (Figs S6–S9, Supporting information) largely recapitulated the network analysis, with *Cephalotes*

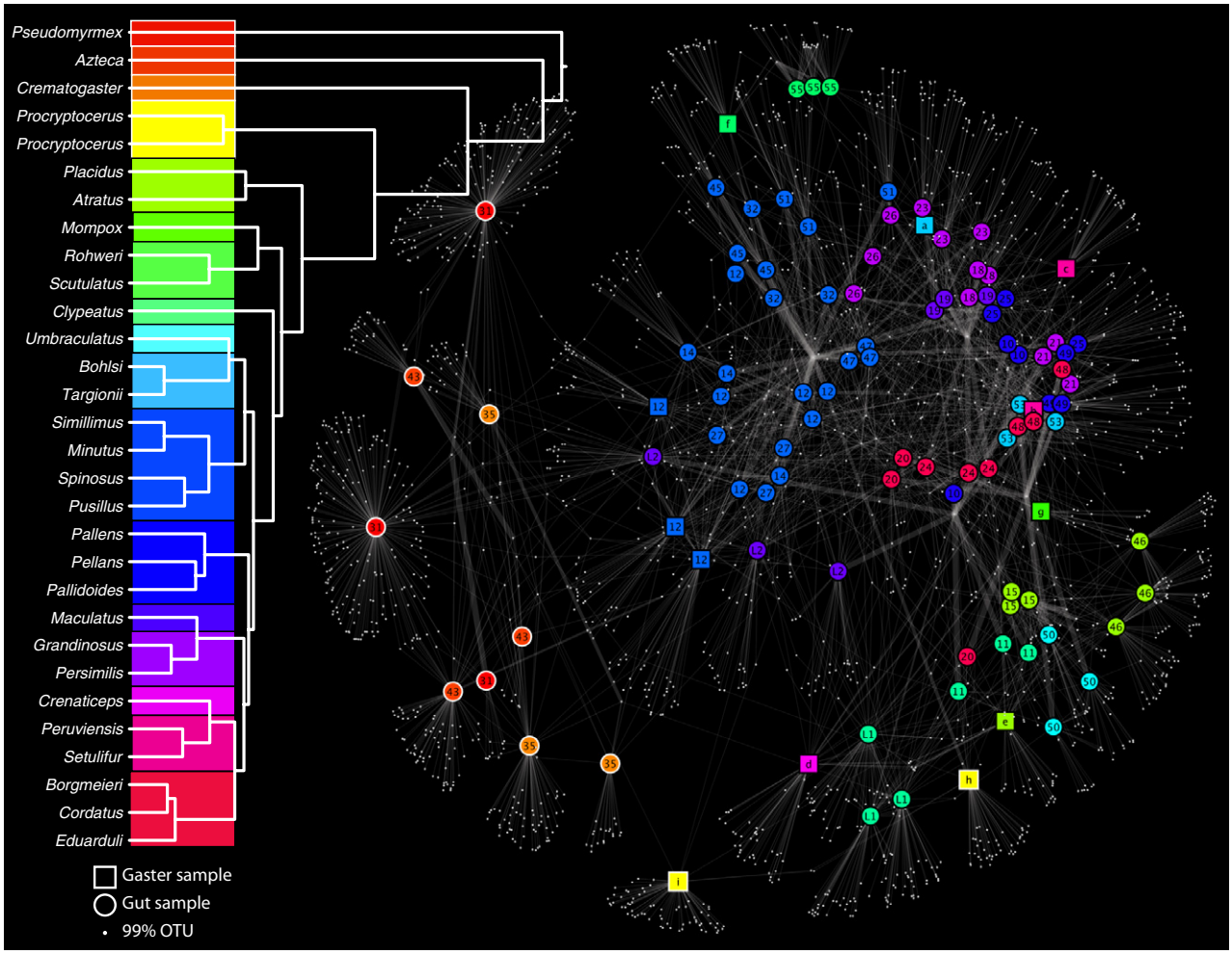

Fig. 4 OTU network showing relationships among ant gut microbiota. Large, coloured nodes represent individual ant samples, while small grey nodes represent individual 99% OTUs. Edges connect OTUs with each host sample in which they occur. Nodes are placed according to a weighted, spring-embedded algorithm, causing host nodes that share more OTUs to appear close together. Host nodes are labelled by colony (see Table 1) and coloured by clades (*sensu* de Andrade & Baroni Urbani 1999) in rainbow order according to phylogeny (inset), with non-*Cephalotes* host nodes outlined in white for emphasis.

samples from the same clade tending to group together. This effect was particularly apparent with narrower OTU picking thresholds and beta-diversity metrics, like Jaccard dissimilarity, dependent solely on the number of shared OTUs (Fig. S6, Supporting information). Unweighted UniFrac, which takes into account the phylogenetic similarity of shared OTUs, clearly separated cephalotine samples from outgroups, but was less likely to group samples from related host species, especially among the more recently diverged clades. In general, the separation between the earlier branching groups (especially *C. atratus*) and the remainder of the *Cephalotes* phylogeny was apparent across a broader range of clustering widths and diversity measures. Neither geography (Fig. S7, Supporting information), preservation method (Fig. S8, Supporting information) nor sequencing quadrant (Fig. S9, Supporting information) appeared to have strong effects relative to host phylogeny with these metrics. Most variance in abundance-weighted metrics, such as Bray–Curtis and weighted UniFrac, appeared to be driven primarily by differences in the relative abundance of Verrucomicrobia, which in turn was strongly affected by tissue and/or preservative (see Supporting information).

Partial Mantel tests indicated that the majority of the variance in community beta-diversity could be explained by host genetic distance, both for the data set as a whole and for the subset of *Cephalotes* samples (Fig. 5C). For the whole data set, after accounting for geographic distance, correlation between host genetic distance and Jaccard community dissimilarity increased at wider OTU clustering thresholds and ultimately accounted for as much as 79% of variance using 93% OTUs (Fig. 5C; correlation between UniFrac community

dissimilarity and host genetic distance was insensitive to OTU clustering width; Fig. S10, Supporting information). For comparisons among *Cephalotes*, however, narrower OTU clustering thresholds explained a greater proportion of the total variance, suggesting that much of the among-*Cephalotes* phylogenetically correlated variance is a consequence of recent bacterial evolution. Genetic distance and community dissimilarity were highly significantly correlated at every OTU clustering threshold for both the whole and *Cephalotes* data sets ($P < 0.001$). Geographic distance also explained a small proportion of the community dissimilarity after correcting for genetic distance, with generally higher correlation at lower OTU clustering thresholds (Fig. 5C). Results for the ape microbiota data were substantially different (Fig. 5D), showing a much lower overall impact of phylogeny and only marginal significance at all OTU clustering widths.

Beta-diversity clustering sensitivity

To summarize the expectations detailed above, we predicted that communities whose patterns of similarity arose primarily through recent bacterial evolution would be grouped more often at narrower OTU clustering widths than wide ones (Fig. 1) and more often using 'star phylogeny' measures of beta-diversity, such as Jaccard or Sorensen dissimilarities, than using UniFrac.

Our sensitivity analysis approach recovered just such a pattern for the internal nodes of the great ape data set from Ochman *et al.* (2010), recapitulating the results of that study (Fig. 5B). As shown previously, the nodes grouping chimpanzees, chimpanzees + bonobos and chimpanzees + bonobos + humans were all recovered primarily using 99% OTU clustering and star phylogeny diversity measures and only at the deeper level of rarefaction. Our analysis also reveals patterns that were not apparent in the earlier, parsimony-based analysis. Strikingly, most of the nodes grouping samples at the tips of the tree—that is, from the same species or subspecies—were recovered robustly under most parameter combinations. This suggests the presence of some common selective filter, such as diet, generating cohesion in these communities (Fig. 1).

Our analysis of *Cephalotes* microbiota showed a similar pattern to that observed for the internal nodes of the ape phylogeny (Fig. 5A), suggesting that recent evolution of gut bacteria may mirror host evolution. Using Jaccard dissimilarities, 15 of 28 internal nodes on the *Cephalotes* phylogeny were recovered in at least one jackknife replicate. Of these, 9 nodes were recovered more often under 99% OTU clustering than under any other threshold. In contrast to the ape data, nodes grouping conspecifics were generally not broadly supported, implying that species-specific selective filters do not play a large role in differentiating *Cephalotes* microbiota from each other. Two notable exceptions to this were the node grouping the two *C. atratus* colonies and that grouping *C. pallidoides* and *C. pellans* (Fig. 5A), both of which were supported more often under wider OTU thresholds—suggesting that similarity among these microbiota may reflect broader shifts in the communities. The nodes separating *Cephalotes* from *Procryptocerus*, and separating the cephalotines from the outgroup genera, were supported at higher frequency and across more measures of beta-diversity, reflective of the greater differences between these communities.

To ensure that these results were not simply due to chance, we also compared 99% Jaccard support among clustered *Cephalotes* microbiota to 100 host trees with randomly permuted tips. No nodes were supported under these parameters in 56% of these randomized data sets, with a maximum of three supported in two of the permutations (Fig. S10, Supporting information).

Discussion

Correlation between phylogeny and the gut microbiota

Our results show that the gut microbiota of *Cephalotes* ants are very stable—perhaps exceptionally so among ants, as even individuals from different *Cephalotes* species typically shared microbiota more similar to one another than did nestmates of the other ant genera we tested (Fig. 3). Our findings are the first to characterize gut microbiota across a broad fraction of this genus, one of the most diverse ant lineages in the Neotropics (Fernández & Sandoya 2004).

Remarkably, despite this overall similarity among *Cephalotes* gut communities, we observed a substantial and significant effect of host phylogeny. This effect was readily apparent in the clustering of closely related species in OTU network diagrams (Fig. 4 and Fig. S5, Supporting information) and PCoA ordinations (Fig. S6, Supporting information). Under some parameter sets, host phylogenetic distance accounted for the majority of total variation in beta-diversity among *Cephalotes* microbiota (Fig. 5C). By comparison, geographic distance accounted for far less of this variation, although our samples were not collected for biogeographic comparisons and are thus limited in this regard.

That closely related host species harbour similar microbiota suggests, but does not necessarily demonstrate, some degree of codiversification between hosts and microbes. In mammals, for example, gut microbes are highly correlated with diet (Ley *et al.* 2008). Similar microbes tend to inhabit the guts of unrelated hosts

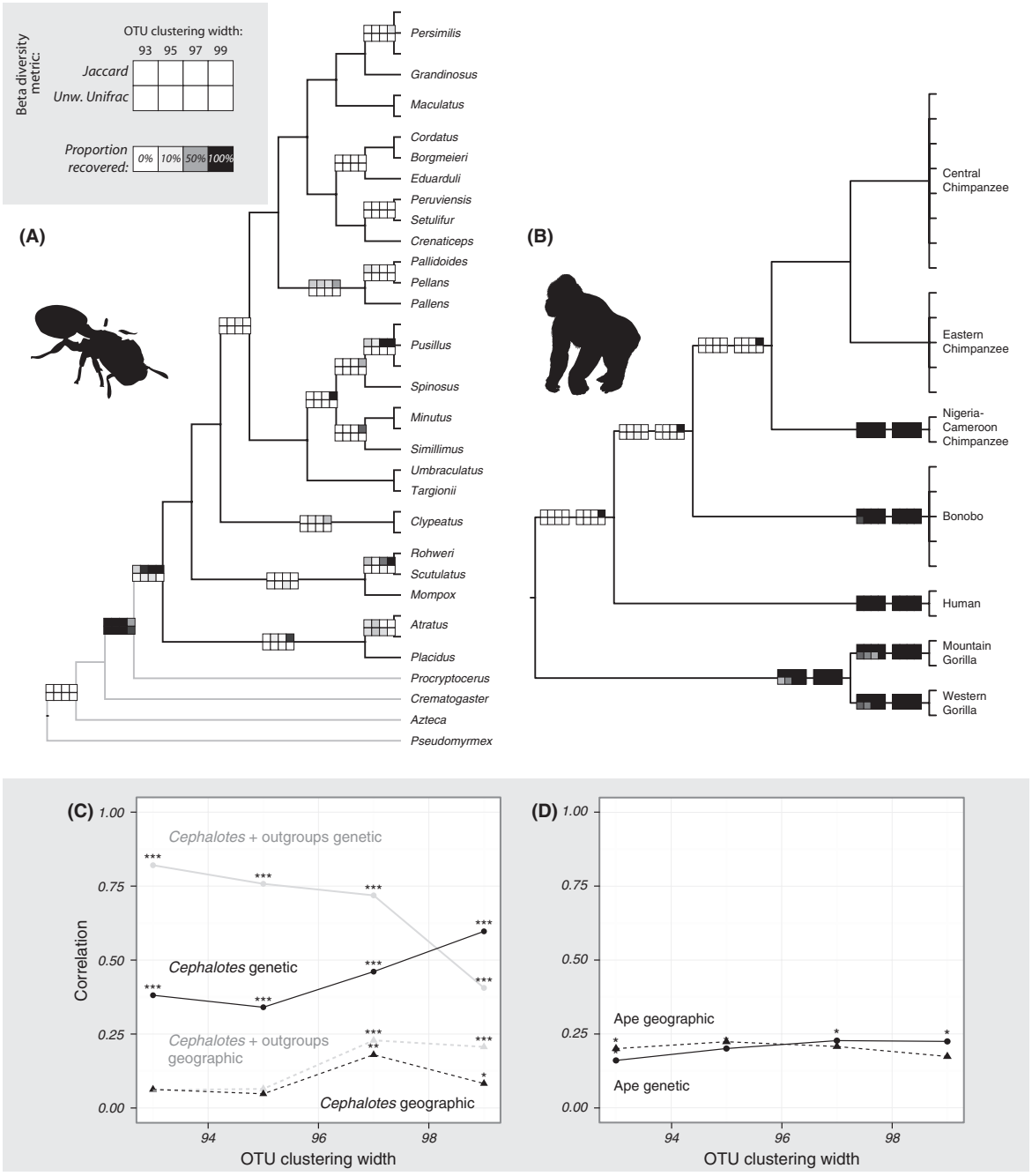

Fig. 5 Contrasting patterns of phylogenetically correlated microbiota in ants and apes. Beta-diversity clustering sensitivity analysis of ant (A) and ape (B) samples using the Jaccard and UniFrac dissimilarity indices (other indices omitted for clarity; for these results, see Fig. S11, Supporting information). Each grid represents the support for a particular sample grouping across a jackknifed data set, with each square representing a continuous gradient of support for that grouping at a given combination of parameters; black = 100%, white = 0%. Parameter combinations are given in the inset key at left. For ape data, grids are displayed for two levels of sampling depth, with results for 1000 sequence rarefactions at left and 15,000 sequence rarefactions at right. Only nodes recovered at least once have grids displayed. For ant data, terminals represent individual colonies; *Cephalotes persimilis*, *Cephalotes maculatus*, *Cephalotes pusillus*, *Cephalotes minutus*, *Cephalotes clypeatus* and *Cephalotes atratus* are all represented by multiple colonies. Results from partial Mantel tests for ants (C) and apes (D). Correlation coefficients between Jaccard dissimilarity matrices and genetic and geographic distances, respectively, are plotted at four different OTU clustering thresholds. *P*-values from partial Mantel tests indicated by asterisks (***$P = 0.001$; **$0.001 < P < 0.01$; *$0.01 < P < 0.05$). Ant microbiota (C) show a strong influence of phylogeny in explaining variation. Results plotted for *Cephalotes*-only and for *Cephalotes* plus outgroups (grey lines). Ape microbiota (D) show less correlation with phylogeny across all OTU clustering widths.

with convergently evolved diets, indicating that host switching occurs relatively frequently compared to the rate of microbial diversification; thus, very few microbial lineages appear to be restricted to monophyletic groups of mammals (Muegge *et al.* 2011; Delsuc *et al*. in press). In cases where changes in diet or other environmental filters are phylogenetically correlated rather than convergent, clear patterns of correlation between microbiota and host phylogeny could thus be due more to these filters than to a history of codiversification with the host.

We developed our beta-diversity clustering sensitivity analysis as a way to help distinguish between these alternative mechanisms—phylogenetically correlated environmental filtering and shared evolutionary history—in explaining correlation between microbial community composition and host phylogeny. The intuition for this approach is based on the assumption that, in either scenario, more distantly related hosts will have more distantly related microbes; but in the case where historical codiversification is the sole or primary factor leading to similarity among microbiota, the age of the last common ancestor of the hosts will constrain the genetic distance between the symbionts. In other words, recent host speciation should be reflected by recent symbiont speciation. By contrast, in the case where host diet selects for different microbes, the most recent common ancestor of a pair of microbes in the two hosts may far pre-date the last common ancestor of the hosts. As a hypothetical example, a particular lineage of bacteria may have diversified into herbivore-gut and carnivore-gut specialist lineages along with the evolution of herbivory in terrestrial vertebrates in the Carboniferous (Sues & Reisz 1998), far pre-dating the evolution of dietary specialization among placental mammals in the Cretaceous (Bininda-Emonds *et al.* 2007).

To test this approach, we reanalysed the ape microbiota data set from Ochman *et al.* (2010), a frequently cited example of codiversification in microbiota. Confirming its utility, our method recapitulated the earlier findings: most internal nodes of the ape phylogeny were recovered under 99% OTU clustering with Jaccard dissimilarity (Fig. 5B). By testing additional parameter combinations, though, our method goes further: that these internal nodes were not recovered under wider OTU clustering thresholds implies that the information grouping different ape species and subspecies is primarily a product of recent bacterial evolution. This was not previously known, as the earlier analysis only reported results from 99.5% OTUs (Ochman *et al.* 2010). That we were only able to recover support for these internal nodes at greater sequencing depth also suggests that the microbial taxa supporting these groupings make up a relatively small proportion of the community.

Perhaps more surprisingly, the additional parameter combinations in our analysis revealed that the grouping of gut communities from conspecific hosts was highly insensitive to parameter choice—the information grouping communities within subspecies (and, for gorillas, between subspecies as well) is retained even when obscuring many millions of years of bacterial divergence in very broad OTUs and at relatively shallow depth of sampling. This pattern of broad support suggests an important role for horizontal acquisition of microbiota through species-specific filters, such as diet or immune selectivity. The partial Mantel correlations (Fig. 5D), which show that phylogenetic distance explains only a small and marginally significant proportion of ape microbiota beta-diversity, echo this pattern.

Taken together, our results would be consistent with a model in which apes acquire species-specific microbiota largely horizontally, while retaining a small proportion of vertically transmitted microbes over longer timescales. Our findings may help to reconcile the apparent patterns of codiversification found by Ochman *et al*. (2010) with subsequent studies that found larger roles for factors such as social group affiliation and geography (Degnan *et al.* 2012; Moeller *et al.* 2012; Moeller *et al.* 2013). Notably, we were able to detect both patterns using only the original data set.

As observed for the internal nodes of the ape phylogeny, our sensitivity analysis suggests that phylogenetic correlation in the *Cephalotes* microbiota is driven in large part by recent bacterial evolution. The results of the partial Mantel tests (Fig. 5C) reinforce this finding. Within *Cephalotes*, microbiota at 99% OTU clustering showed much greater correlation with host phylogeny than at wider clustering thresholds. Furthermore, the much greater overall proportion of beta-diversity variance explained by the *Cephalotes* phylogeny relative to that explained by the ape phylogeny suggests that such recent evolution (potentially, the result of codiversification) plays an overall greater role in structuring the *Cephalotes* microbiota than it does in the apes'. This interpretation is further supported by the fact that we did not see a shift to broad, parameter-insensitive support—possibly indicative of some sort of environmental filtering—at the tips of the *Cephalotes* tree. And while only about half of *Cephalotes* nodes were recovered in the analysis, we suspect that the high level of similarity among microbiota across the genus limits the overall level of support in the beta-diversity clustering analysis, as variability within species overlaps substantially with variability between species (Fig. 3). Repeating this analysis with more sensitive techniques, such as the low-error amplicon sequencing approach recently developed by Faith *et al.* (2013), would help to determine whether the lack of support for the remaining

internal nodes is due to biological variation or is an artefact of the relatively high level of technical error we observed in our data set (see Supporting information). Future techniques exploring patterns of codiversification in particular microbial lineages will provide further context to these community-level trends.

The apparent impact of recent bacterial evolution does not mean that niche-driven ecology or environmental filtering does not also play a role in structuring the *Cephalotes* microbiota. Rather, our analysis suggests that these factors are unlikely to be driving the bulk of the observed phylogenetic correlation. As noted above, some differentiation among clades is still apparent even at lower clustering widths (Fig. S6, Supporting information), indicating that there have probably been some more substantial shifts in the composition of the *Cephalotes* microbiota. These shifts may be apparent in the branches leading to *Cephalotes atratus* and to the *Cephalotes pallens* clade, for which community similarity patterns reflected host relatedness primarily at wider OTU clustering thresholds. Future studies, perhaps incorporating shotgun metagenomic data, will help to clarify the functional significance of these changes.

The primary strength of the sensitivity analysis approach we present here lies in its ability to effectively visualize correlation with microbiota composition across particular nodes of the host phylogeny. In this respect, it is especially complementary to more frequently applied techniques, like Mantel tests, which assess the overall strength of phylogenetic correlation in the data set, but not its distribution. This is especially apparent in the ape data set, where phylogenetic correlation at internal nodes was much more sensitive to parameter choice than at terminal nodes.

The usefulness of this approach extends beyond exploring phylogenetic correlation. The sensitivity analysis we have performed here is really a formalization of what is implicitly done whenever one compares the results of different beta-diversity measures. Different diversity measures emphasize different properties of the underlying data, and performing such comparisons has been recommended as a general practice for understanding ecological patterns (Anderson *et al.* 2011). While we consider support for phylogenetic grouping of samples, in principle, the same technique could be applied to any sample grouping hierarchy, allowing quick examination of support for a given hypothesis.

Social transmission and the stabilization of mutualism

Theory suggests that mutualisms should be vulnerable to cheating (Sachs *et al.* 2004). Despite this, evidence for breakdown of mutualistic lifestyles is comparatively rare (Sachs & Simms 2006; Sachs *et al.* 2011). Several mechanisms have been proposed to explain the evolutionary stability of mutualism (Sachs *et al.* 2004; Archetti *et al.* 2011), of which two—partner choice and partner fidelity—are particularly relevant to co-evolution in gut microbiota. Partner choice mechanisms limit cheating by detecting and favouring interactions with cooperators; partner fidelity links fitness outcomes of partners, favouring shared investment over cheating.

Gut microbes present a challenging problem for mutualism: microbial mutualists localized to the gut lumen must be maintained in the face of a constant influx of food-associated microbes, potentially limiting partner fidelity, and are physically distant from epithelial-associated immune factors often associated with partner choice mechanisms (Nyholm & Graf 2012). In insects, partner fidelity via vertical transmission is typically associated with obligate intracellular symbionts (Moran *et al.* 2008). Many of the invertebrates that are known to rely on extracellular microbes for defined benefits—such as light production in squid (Nyholm & McFall-Ngai 2004) and nutrition in stinkbugs (Kikuchi *et al.* 2007; Matsuura *et al.* 2012)—have highly specific associations with just one or a few microbial lineages and enhance the efficacy of partner choice mechanisms by physically sequestering these microbes for part or all of their life cycle. While vertebrates do maintain a very complex lumenal gut community, it has been suggested that the vertebrate adaptive immune system may have evolved in part as a partner choice mechanism for dealing with this complexity (McFall-Ngai 2007). Such an immune-mediated mechanism could be responsible for the broad support for the grouping of conspecific apes in our sensitivity analysis (Fig. 5B) by imposing a host-specific selective filter on the acquisition of gut microbes from the environment.

In insects, social transmission could function to maintain relatively complex gut mutualisms by augmenting partner fidelity, effectively playing a similar role for an entire microbiota to that of ovarial transmission for individual insect endosymbionts. The high similarities we observed among *Cephalotes* nestmate microbiota indicate efficient homogenization of the colony's gut microbiota, and young *Cephalotes* queens presumably inherit the gut microbiota of their mother colony with similarly high fidelity, passing these on to their own offspring in turn. While experiments are ongoing to characterize the mutualistic, commensalistic or parasitic nature of the various players in the *Cephalotes* microbiota, our data show that partner fidelity may be sufficiently strong to result in phylogenetic correlation of a substantial fraction of the microbiota with the host across tens of millions of years—and across one of the most significant Neotropical ant adaptive radiations.

Notably, the noncephalotine ants in our study hosted microbiota that were quite divergent among nestmates (Fig. 3), suggesting that eusociality alone is not sufficient to generate this degree of homogeneity. Others have proposed anal trophallaxis (i.e. adult feeding from anal secretions) as a mechanism for transmission of microbes in *Cephalotes* (Wheeler 1984; Russell *et al.* 2009; Anderson *et al.* 2012). This behaviour has been observed in *Cephalotes* and *Procryptocerus* (Wilson 1976; Wheeler 1984) and is reportedly rare among ants generally (Hölldobler & Wilson 1990), consistent with the much greater variance observed in the three unrelated outgroup genera in our study. Anal trophallaxis has also been shown to be critical for transmission of beneficial microbes in bumblebee colonies (Koch & Schmid-Hempel 2011) and termites (Kitade 2004; Köhler *et al.* 2012), both of which host microbiota that appear to be more stable across their respective phylogenies than are those of their more solitary relatives (Martinson *et al.* 2011; Colman *et al.* 2012).

Conclusions

Our results, the first to explore *Cephalotes* gut microbiota in the broader context of host evolution, demonstrate remarkable lineage-wide stability. Many of the members of this community appear to have been present since the diversification of the host genus in the Eocene and perhaps since before it split with *Procryptocerus* in the Cretaceous (Price *et al.* 2013). Using a novel application of sensitivity analysis, we have shown that correlation between these microbiota and their hosts' phylogeny appears to be driven largely by relatively recent bacterial evolution, suggesting it may be the result of codiversification. Notably, our reanalysis of great ape microbiota showed a substantially different pattern: while we still see patterns consistent with codiversification, much of the phylogenetic correlation might be better explained by phylogenetically correlated selective forces such as diet or immunity.

We have presented an approach here that enables us to look at whole-community dynamics, while permitting some insight into the potential underlying drivers. Future techniques capable of identifying patterns of codiversification between hosts and individual members of complex communities will help us to better understand the composition of these broad patterns of similarity and perhaps provide additional insight into the processes of transmission and co-evolution underlying the patterns.

Acknowledgements

We thank Antonio Coral, Lina Arcila Hernandez, Gabriel Miller, Kleber del Claro and Flavio Camarota for assistance with field work; Shauna Price for providing early access to the *Cephalotes* phylogeny; Andrew Richardson, Yoel Stuart and the class of OEB 210 for helpful comments on an early draft of this manuscript; and editor Jacob Russell and four anonymous reviewers for constructive criticism. Brazilian material was obtained under Material Transfer Agreement #08/2009 between the Universidade Federal de Uberlandia and Harvard University and Peruvian material under permits 394-2009-AG-DGFFS-DGEFFS and 79-2008-INRENA-IFFS-DCB to MEF from the Peruvian Ministry of Agriculture. This work was supported by Putnam Expedition Grants from the Museum of Comparative Zoology to JGS, a grant from the Milton Fund and the Harvard Society of Fellows to DJCK, a Doctoral Dissertation Improvement grant from the National Science Foundation DDIG 1110515 to NEP and JGS, NSF Graduate Student Research Fellowship DGE 1144152 to JGS, and NSF Awards DEB 0842144 and IOS 0841756 to SP.

References

Adams ES (1990) Interaction between the ants *Zacryptocerus maculatus* and Azteca trigona: interspecific parasitization of information. *Biotropica*, **22**, 200–206.

Aguirre de Cárcer D, Denman SE, McSweeney C, Morrison M (2011) Evaluation of subsampling-based normalization strategies for tagged high-throughput sequencing data sets from gut microbiomes. *Applied and Environmental Microbiology*, **77**, 8795–8798.

Anderson MJ, Crist TO, Chase JM *et al.* (2011) Navigating the multiple meanings of β diversity: a roadmap for the practicing ecologist. *Ecology Letters*, **14**, 19–28.

Anderson KE, Russell JA, Moreau CS *et al.* (2012) Highly similar microbial communities are shared among related and trophically similar ant species. *Molecular Ecology*, **21**, 2282–2296.

de Andrade ML, Baroni Urbani C (1999) *Diversity and Adaptation in the Ant Genus Cephalotes, Past and Present*. Stuttgarter Beiträge zur Naturkunde Series B (Geologie und Palaëontologie), Stuttgart.

Archetti M, Scheuring I, Hoffman M *et al.* (2011) Economic game theory for mutualism and cooperation. *Ecology Letters*, **14**, 1300–1312.

Bininda-Emonds ORP, Cardillo M, Jones KE *et al.* (2007) The delayed rise of present-day mammals. *Nature Materials*, **446**, 507–512.

Bution ML, Caetano FH (2008) Ileum of the *Cephalotes* ants: a specialized structure to harbor symbionts microorganisms. *Micron*, **39**, 897–909.

Byk J, Del-Claro K (2010) Nectar- and pollen-gathering *Cephalotes* ants provide no protection against herbivory: a new manipulative experiment to test ant protective capabilities. *acta ethologica*, **13**, 33–38.

Caporaso JG, Kuczynski J, Stombaugh J *et al.* (2010) QIIME allows analysis of high-throughput community sequencing data. *Nature Methods*, **7**, 335–336.

Chandler JA, Morgan Lang J, Bhatnagar S, Eisen JA, Kopp A (2011) Bacterial communities of diverse *Drosophila* species: ecological context of a host–microbe model system. *PLoS Genetics*, **7**, e1002272.

Colman DR, Toolson EC, Takacs-Vesbach CD (2012) Do diet and taxonomy influence insect gut bacterial communities? *Molecular Ecology*, **21**, 5124–5137.

Davidson DW, Cook SC, Snelling RR, Chua TH (2003) Explaining the abundance of ants in lowland tropical rainforest canopies. *Science*, **300**, 969–972.

Degnan PH, Pusey AE, Lonsdorf EV *et al.* (2012) Factors associated with the diversification of the gut microbial communities within chimpanzees from Gombe National Park. *Proceedings of the National Academy of Sciences*, **109**, 13034–13039.

Delsuc F, Metcalf JL, Wegener Parfrey L *et al.* (in press) Convergence of gut microbiomes in myrmecophagous mammals. *Molecular Ecology*.

Diaz Heijtz R, Wang S, Anuar F *et al.* (2011) Normal gut microbiota modulates brain development and behavior. *Proceedings of the National Academy of Sciences*, **108**, 3047–3052.

Dillon R, Dillon V (2004) The gut bacteria of insects: nonpathogenic interactions. *Annual Review of Entomology*, **49**, 71–92.

Dowd SE, Callaway TR, Wolcott RD *et al.* (2008) Evaluation of the bacterial diversity in the feces of cattle using 16S rDNA bacterial tag-encoded FLX amplicon pyrosequencing (bTEFAP). *BMC Microbiology*, **8**, 125.

Edgar RC (2010) Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*, **26**, 2460–2461.

Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R (2011) UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics*, **27**, 2194–2200.

Engel P, Moran NA (2013) The gut microbiota of insects—diversity in structure and function. *FEMS Microbiology Reviews*, **37**, 699–735.

Ersts PJ (2013) Geographic Distance Matrix Generator. American Museum of Natural History, Center for Biodiversity and Conservation. Available from http://biodiversityinformatics.amnh.org/open_source/gdmg.

Faith JJ, Guruge JL, Charbonneau M *et al.* (2013) The long-term stability of the human gut microbiota. *Science*, **341**, 1237439.

Falush D, Wirth T, Linz B *et al.* (2003) Traces of human migrations in *Helicobacter pylori* populations. *Science*, **299**, 1582–1585.

Fernández F, Sandoya S (2004) Syninimic list of Neotropical ants (Hymenoptera: Formicidae). *Biota Colombiana*, **5**, 3–105.

Gordon DM (2012) The dynamics of foraging trails in the tropical arboreal ant *Cephalotes goniodontus*. *PLoS One*, **7**, 1–7.

Haas BJ, Gevers D, Earl AM *et al.* (2011) Chimeric 16S rRNA sequence formation and detection in Sanger and 454-pyrosequenced PCR amplicons. *Genome Research*, **21**, 494–504.

Hamady M, Knight R (2009) Microbial community profiling for human microbiome projects: tools, techniques, and challenges. *Genome Research*, **19**, 1141–1152.

Hölldobler B, Wilson EO (1990) *The Ants*. Harvard University Press, Cambridge.

Hu Y, Lukasik P, Moreau CS, Russell JA (in press) Correlates of gut community composition across an ant species (*Cephalotes varians*) elucidate causes and consequences of symbiotic variability. *Molecular Ecology*.

Huse SM, Welch DM, Morrison HG, Sogin ML (2010) Ironing out the wrinkles in the rare biosphere through improved OTU clustering. *Environmental Microbiology*, **12**, 1889–1898.

Jaffe K, Caetano F, Sanchez P *et al.* (2001) Sensitivity of ant (*Cephalotes*) colonies and individuals to antibiotics implies feeding symbiosis with gut microorganisms. *Canadian Journal of Microbiology*, **79**, 1120–1124.

Kautz S, Rubin BER, Russell JA, Moreau CS (2013) Surveying the microbiome of ants: comparing 454 pyrosequencing with traditional methods to uncover bacterial diversity. *Applied and Environmental Microbiology*, **79**, 525–534.

Kelley ST, Dobler S (2010) Comparative analysis of microbial diversity in *Longitarsus* flea beetles (Coleoptera: Chrysomelidae). *Genetica*, **139**, 541–550.

Kikuchi Y, Hosokawa T, Fukatsu T (2007) Insect-microbe mutualism without vertical transmission: a stinkbug acquires a beneficial gut symbiont from the environment every generation. *Applied and Environmental Microbiology*, **73**, 4308–4316.

Kikuchi Y, Hosokawa T, Nikoh N *et al.* (2009) Host-symbiont co-speciation and reductive genome evolution in gut symbiotic bacteria of acanthosomatid stinkbugs. *BMC Biology*, **7**, 2.

Kitade O (2004) Comparison of symbiotic flagellate faunae between termites and a wood-feeding cockroach of the genus *Cryptocercus*. *Microbes and Environments*, **19**, 215–220.

Knight R, Maxwell P, Birmingham A *et al.* (2007) PyCogent: a toolkit for making sense from sequence. *Genome Biology*, **8**, R171.

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

Köhler T, Dietrich C, Scheffrahn RH, Brune A (2012) High-resolution analysis of gut environment and bacterial microbiota reveals functional compartmentation of the gut in wood-feeding higher termites (*Nasutitermes* spp.). *Applied and Environmental Microbiology*, **78**, 4691–4701.

Kumar PS, Brooker MR, Dowd SE, Camerlengo T (2011) Target region selection is a critical determinant of community fingerprints generated by 16S pyrosequencing. *PLoS One*, **6**, e20956.

Ley RE, Hamady M, Lozupone C *et al.* (2008) Evolution of mammals and their gut microbes. *Science*, **320**, 1647–1651.

Li W, Godzik A (2006) CD-Hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics*, **22**, 1658–1659.

Lozupone CA, Knight R (2008) Species divergence and the measurement of microbial diversity. *FEMS Microbiology Reviews*, **32**, 557–578.

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

Matson EG, Ottesen EA, Leadbetter JR (2007) Extracting dna from the gut microbes of the termite (*Zootermopsis angusticollis*) and visualizing gut microbes. *Journal of Visualized Experiments*, **4**.

Matsuura Y, Kikuchi Y, Hosokawa T *et al.* (2012) Evolution of symbiotic organs and endosymbionts in lygaeid stinkbugs. *The ISME Journal*, **6**, 397–409.

McFall-Ngai M (2007) Adaptive immunity: care for the community. *Nature Materials*, **445**, 153.

McFall-Ngai M, Hadfield MG, Bosch TCG *et al.* (2013) Animals in a bacterial world, a new imperative for the life sciences. *Proceedings of the National Academy of Sciences*, **110**, 3229–3236.

Moeller AH, Degnan PH, Pusey AE *et al.* (2012) Chimpanzees and humans harbour compositionally similar gut enterotypes. *Nature Communications*, **3**, 1179.

Moeller AH, Peeters M, Ndjango JB *et al.* (2013) Sympatric chimpanzees and gorillas harbor convergent gut microbial communities. *Genome Research*, **23**, 1715–1720.

Moran NA, McCutcheon JP, Nakabachi A (2008) Genomics and evolution of heritable bacterial symbionts. *Annual Review of Genetics*, **42**, 165–190.

Moreau CS, Bell CD, Vila R, Archibald SB, Pierce NE (2006) Phylogeny of the ants: diversification in the age of angiosperms. *Science*, **312**, 101–104.

Muegge BD, Kuczynski J, Knights D *et al.* (2011) Diet drives convergence in gut microbiome functions across mammalian phylogeny and within humans. *Science*, **332**, 970–974.

Noda S, Kitade O, Inoue T *et al.* (2007) Cospeciation in the triplex symbiosis of termite gut protists (*Pseudotrichonympha* spp.), their hosts, and their bacterial endosymbionts. *Molecular Ecology*, **16**, 1257–1266.

Nyholm SV, Graf J (2012) Knowing your friends: invertebrate innate immunity fosters beneficial bacterial symbioses. *Nature Reviews Microbiology*, **10**, 815–827.

Nyholm SV, McFall-Ngai MJ (2004) The winnowing: establishing the squid-*Vibrio* symbiosis. *Nature Reviews Microbiology*, **2**, 632–642.

Ochman H, Worobey M, Kuo C-H *et al.* (2010) Evolutionary relationships of wild hominids recapitulated by gut microbial communities. *PLoS Biology*, **8**, e1000546.

Oksanen J, Blanchet FG, Kindt R *et al.* (2012) VEGAN: Community Ecology Package. Available from http://CRAN.R-project.org/package=vegan.

Osei-Poku J, Mbogo CM, Palmer WJ, Jiggins FM (2012) Deep sequencing reveals extensive variation in the gut microbiota of wild mosquitoes from Kenya. *Molecular Ecology*, **21**, 5138–5150.

Powell S (2008) Ecological specialization and the evolution of a specialized caste in Cephalotes ants. *Functional Ecology*, **22**, 902–911.

Prado-Martinez J, Sudmant PH, Kidd JM *et al.* (2013) Great ape genetic diversity and population history. *Nature Materials*, **499**, 471–475.

Price SL, Powell S, Kronauer DJC *et al.* (2013) Renewed diversification is associated with new ecological opportunity in the Neotropical turtle ants. *Journal of Evolutionary Biology*. doi: 10.1111/jeb.12300 (in press).

Quince C, Lanzén A, Curtis TP *et al.* (2009) Accurate determination of microbial diversity from 454 pyrosequencing data. *Nature Methods*, **6**, 639–641.

R Development Core Team (2012) R: A Language and Environment for Statistical Computing. Available from http://www.R-project.org/.

Roche R, Wheeler D (1997) Morphological specializations of the digestive tract of *Zacryptocerus rohweri* (Hymenoptera: Formicidae). *Journal of Morphology*, **234**, 253–262.

Russell JA, Moreau CS, Goldman-Huertas B *et al.* (2009) Bacterial gut symbionts are tightly linked with the evolution of herbivory in ants. *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 21236–21241.

Sachs JL, Simms EL (2006) Pathways to mutualism breakdown. *Trends in Ecology & Evolution*, **21**, 585–592.

Sachs J, Mueller U, Wilcox T, Bull J (2004) The evolution of cooperation. *Quarterly Review of Biology*, **79**, 135–160.

Sachs JL, Skophammer RG, Regus JU (2011) Evolutionary transitions in bacterial symbiosis. *Proceedings of the National Academy of Sciences*, **108**(Suppl 2), 10800–10807.

Sanders JG (2010) Program note: Cladescan, a program for automated phylogenetic sensitivity analysis. *Cladistics*, **26**, 114–116.

Shannon P, Markiel A, Ozier O *et al.* (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Research*, **13**, 2498–2504.

Sharon G, Segal D, Ringo JM *et al.* (2010) Commensal bacteria play a role in mating preference of *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences*, **107**, 20051–20056.

Sues HD, Reisz RR (1998) Origins and early evolution of herbivory in tetrapods. *Trends in Ecology & Evolution*, **13**, 141–145.

Sun Y, Cai Y, Huse SM *et al.* (2012) A large-scale benchmark study of existing algorithms for taxonomy-independent microbial community analysis. *Briefings in Bioinformatics*, **13**, 107–121.

Turnbaugh PJ, Hamady M, Yatsunenko T *et al.* (2009) A core gut microbiome in obese and lean twins. *Nature Materials*, **457**, 480–484.

Wang Q, Garrity GM, Tiedje JM, Cole JR (2007) Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Applied and Environmental Microbiology*, **73**, 5261–5267.

Weber NA (1957) The nest of an anomalous colony of the arboreal ant *Cephalotes atratus*. *Psyche: A Journal of Entomology*, **64**, 60–69.

Werner JJ, Koren O, Hugenholtz P *et al.* (2012) Impact of training sets on classification of high-throughput bacterial 16s rRNA gene surveys. *The ISME Journal*, **6**, 94–103.

Wheeler DE (1984) Behavior of the ant, *Procryptocerus scabriusculus* (Hymenoptera: Formicidae), with comparisons to other Cephalotines. *Psyche: A Journal of Entomology*, **91**, 171–192.

Willner D, Daly J, Whiley D *et al.* (2012) Comparison of DNA extraction methods for microbial community profiling with an application to pediatric bronchoalveolar lavage samples. *PLoS One*, **7**, e34605.

Wilson EO (1976) A social ethogram of the neotropical arboreal ant *Zacryptocerus varians* (Fr. Smith). *Animal Behaviour*, **24**, 354–363.

J.G.S., N.E.P., S.P. and D.J.C.K. designed research; J.G.S., S.P., M.E.F. and H.L.V. contributed to field work; J.G.S. designed and performed data analysis; J.G.S. conceived and implemented novel methodology; J.G.S., D.J.C.K. and N.E.P. contributed reagents; S.P., D.J.C.K. and M.E.F. contributed additional data and specimens; J.G.S., N.E.P., D.J.C.K., S.P., M.E.F. and H.L.V. wrote the manuscript.

Data accessibility

Raw reads from this study are available on the NCBI Short Read Archive under the Project Accession no. PRJNA227767. Additionally, computer scripts and example data are deposited in the Dryad data repository under DOI:10.5061/dryad.023s6, and processed sequence reads and metadata are available in MG-RAST under project number 6249.

Supporting information

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

Appendix S1 Supplemental methods.

Table S1 Sample information.

Table S2 Number of sequences for each sample in this study.

Table S3 Alpha-diversity values (mean and standard error) under UCLUST and CD-HIT OTU clustering methods.

Table S4 Breakdown of changes made to our data set during the denoising process in AMPLICONNOISE, after Gaspar and Thomas (2013).

Fig. S1 (a) *Cephalotes pallidoides* worker and soldier (top). The elaborated, disc-shaped head of the soldier caste functions as a 'door' to protect the nest entrance. (b) *Cephalotes* gastrointestinal tract. Clockwise from right: midgut (mg), ileum (il), rectum (r). The straw-coloured organ at centre is not part of the GI tract. The midgut and ileum harbour dense populations of microbes; the bulbous, enlarged ileum is unusual in ants (de Andrade & Baroni Urbani 1999). (c) Fluorescence micrograph of bacteria from *Cephalotes* midgut, illustrating morphological diversity.

Fig. S2 Average alpha-diversity rarefaction curves for different categories of samples.

Fig. S3 Maximum likelihood tree of 'core' *Cephalotes* OTUs (as defined by presence in ≥50% of *Cephalotes* samples) and representative '*Cephalotes*-specific' OTUs [one representative sequence per 97% OTU as published in Anderson *et al.* (2012)].

Fig. S4 Dissimilarity boxplots, comparing community UniFrac distances within and among groups of ants.

Fig. S5 OTU network showing relationships among ant gut microbiota.

Fig. S6 PCoA plots showing relationships among samples, for four different beta-diversity measures (Jaccard, Bray–Curtis, unweighted UniFrac and weighted UniFrac) and four OTU clustering widths (93%, 95%, 97% and 99% identity).

Fig. S7 As in Fig. S6, but with points and polygons coloured according to location of origin.

Fig. S8 As in Fig. S6, but with points and polygons coloured according to preservative.

Fig. S9 As in Fig. S6, but with points and polygons coloured according to sequencing quadrant.

Fig. S10 Results from partial Mantel tests show the influence of phylogeny in explaining variation among ant microbiomes.

Fig. S11 Histograms showing the results of beta-diversity sensitivity analyses using host trees with randomly permuted tips for comparison.

Fig. S12 Version of Fig. 5 with additional beta-diversity indices noted in legend.