Localization of bacterial communities within gut compartments across Cephalotes turtle ants

Peter J. Flynn1*, Catherine L. D’Amelio², Jon G. Sanders³, Jacob A. Russell², Corrie S. Moreau³,⁴

1University of Chicago, Committee on Evolutionary Biology, Chicago, IL, 60605 USA
2*Corresponding author
3Drexel University, Department of Biodiversity, Earth and Environmental Science, Philadelphia, PA, 19104 USA
4Cornell University, Department Ecology and Evolutionary Biology, Ithaca, NY, 14850 USA
5Cornell University, Department of Entomology, Ithaca, NY, 14850 USA

Correspondence: Peter J. Flynn, Committee on Evolutionary Biology, University of Chicago, 5801 S Ellis Ave, Chicago, IL 60637 E-mail: pflynn@uchicago.edu

Keyword(s): gut microbes, turtle ants, localization, host-microbe, symbionts
ABSTRACT

Microbial communities within the animal digestive tract often provide important functions for their hosts. The composition of eukaryotes’ gut bacteria can be shaped by host diet, vertical bacterial transmission, and physiological variation within the digestive tract. In several ant taxa, recent findings have demonstrated that nitrogen provisioning by symbiotic bacteria makes up for deficiencies in herbivorous diets. Using 16S rRNA amplicon sequencing and qPCR, this study examined bacterial communities at a fine scale across one such animal group, the turtle ant genus *Cephalotes*. We analyzed the composition and colonization density across four portions of the digestive tract to understand how bacterial diversity is structured across gut compartments, potentially allowing for specific metabolic functions of benefit to the host. In addition, we aimed to understand if caste differentiation or host relatedness influences the gut bacterial communities of *Cephalotes* ants. Microbial communities were found to vary strongly across *Cephalotes* gut compartments in ways that transcend both caste and host phylogeny. Despite this, caste and host phylogeny still have detectable effects. We demonstrated microbial community divergence across gut compartments, possibly due to the varying function of each gut compartment for digestion.

IMPORTANCE

Gut compartments play an important role in structuring the microbial community within individual ants. The gut chambers of the turtle ant digestive tract differ remarkably in symbiont abundance and diversity. Furthermore, caste type explains some variation in the microbiome composition. Finally, the evolutionary history of the *Cephalotes* species structures the microbiome in our study, which elucidates a trend in which related ants maintain related
microbiomes, conceivably owing to co-speciation. Amazingly, gut compartment-specific
signatures of microbial diversity, relative abundance, composition, and abundance have been
conserved over *Cephalotes* evolutionary history, signifying that this symbiosis has been largely
stable for over 50 million years.
INTRODUCTION

Bacterial communities in animal guts play significant roles in animal evolution (1, 2).

The diversity of eukaryotic gut bacteria is driven by host diet, distinct physiological niches across the gut habitat favoring differing microbes, and the capacity of organisms to transfer bacteria to their progeny (3–5). Bacterial gut symbionts can promote effective digestion, immunologic defense, and metabolic regulation (6).

Ants are an extremely successful clade of animals both in terms of their sheer biomass as well as their diversity, and their associated microbes play an important functional role for species of varying dietary types. For example, microbial symbionts have been hypothesized to largely enable the ant’s ecological dominance throughout nutrient-limited rainforest canopies (7, 8). In addition, broad molecular surveys have illustrated that symbiotic bacteria are correlated with the evolution of herbivory across ant species (9), while both experimentation and genomics have illustrated that herbivorous ants benefit from specialized, nitrogen recycling gut bacteria (10–13).

There are striking differences between microbial communities within predatory versus herbivorous ants (14). Army ants exhibit a carnivorous diet and harbor a small number of seemingly specialized bacterial species that are common across army ant lineages (15).

Herbivorous ants from Camponotini obligately use Blochmannia gut symbionts to nutritionally upgrade their diet with critical amino acids and possibly aid in recycling nitrogen (11). The benefits of the Blochmannia gut symbionts could be a reason for the evolutionary success of the Camponotini tribe (16).

In recent years, the bacterial communities of ants from the Cephalotini tribe have been studied extensively (17–21). This tribe consists of two sister genera of ants: Cephalotes and
Both genera are canopy nesters and foragers which consume extrafloral nectar, fungi, pollen, and occasionally mammal urine and bird droppings (22, 23). The diet of the adult workers is almost exclusively liquid food (24). There are 119 extant species within the turtle ant genus *Cephalotes* and 44 species of *Procryptocerus*. Across *Cephalotes*, these ants exhibit a remarkably conserved bacterial community composition. But bacterial communities differ slightly across colonies and, seemingly, species, primarily in the changing relative abundance and presence/absence of OTUs from core symbiotic taxa (21). Turtle ants also demonstrate a conserved N-recycling role of their core gut symbionts. This was empirically shown in *Cephalotes varians*, and through metagenomic analyses of core symbiont gene content across the genus within this same study (12). Together, these findings uncovered a nutritional mutualism between turtle ants and their gut microbes.

*Cephalotes* microbes are highly conserved within their hosts, similarly to other eusocial arthropod groups (25). Various factors have been hypothesized to explain this conservation, including oral-anal trophallaxis, in which newly eclosed adult ants consume anal secretions of older adult siblings. This trophallactic behavior appears to inoculate symbiont-free, young adults with adult-enriched, specialized core symbionts (21, 26, 27). It is relatively uncommon in most ant clades, but is prevalent in *Cephalotes* and *Procryptocerus* (28, 29).

After trophallactic colonization in early adulthood, specialized, adult-enriched symbionts appear to colonize multiple gut chambers, proliferating to establish large populations within adult Cephalotine ants (18, 26, 30–33). But the precise composition of communities across gut regions remains obscure outside of two *Cephalotes* species (18, 26). Detailing the precise trends of localization across a broader range of Cephalotine taxa will aid in our attempts to understand
the evolution of ant-symbiont associations, and the ways by which symbionts function to improve host ant fitness.

The basic structure of the adult digestive tract is conserved across all ants, consisting of three main subdivisions: the foregut, midgut, and hindgut. The ant foregut and hindgut are further subdivided into functionally discrete sub-compartments: the esophagus and crop within the foregut and the ileum and rectum within the hindgut (34). As part of foraging, ant workers collect liquid foods, which are stored temporarily in their crop. The crop is often thought of as a “social stomach” since the undigested food can then be regurgitated and shared among nestmates (28, 35). Due to the function of the crop in temporary food storage, it is large and often dilated within *Cephalotes*.

One reason that *Cephalotes* exhibit unusually high partner fidelity with their gut bacteria over evolutionary time may be due to their highly modified proventriculus, the valve separating the crop and midgut (36). The Cephalotine proventriculus consists of a mushroom-like nodule covered in a layer of hard cuticular fibers, rather than a simple tube-like structure found in most other ant clades. It functions as a “passive dam” with a rigid morphology and small opening and provides a filtering mechanism that prevents most bacteria from entering the midgut (26, 31). *Cephalotes rohweri* filters out particles smaller than most bacteria, however, the filtering layer does not fully form until several days into adulthood (26). Therefore, it is argued that sterile callow adults gain their core microbes through oral-anal trophallaxis during the short window between eclosion and proventriculus maturation. Subsequently, *Cephalotes* ants may use their proventricular filter to maintain the integrity of their symbiotic bacterial communities, warding off the invasion of potential competitors and pathogens (26).
Immediately posterior to the proventriculus lies the midgut, which functions as the primary site of nutrient digestion and absorption in insects (3). Within Cephalotes, the midgut is covered by a peritrophic matrix which protects it against injury and concentrates digestive enzymes (31, 33). In addition, the peritrophic matrix compartmentalizes the midgut, which may provide the bacteria protection against enzymatic lysis (33).

The ileum and the rectum comprise the remaining portions of the ant gut. These chambers play a major role in both water and nutrient absorption. Within Cephalotes, the ileum is the longest portion of the digestive tract, containing a large number of folds. The increased surface area allows for more efficient reabsorption while providing more room for bacterial attachment (31, 32). The rectum is the most distal sub-compartment of the hindgut, used to stores feces, while enabling further water reabsorption before waste exits the body (3).

Given the broad differences in physiology and host function across the digestive system, and a lack of precise knowledge on how symbionts with increasingly defined functions are tied to gut regions, we aimed to comprehensively understand bacterial communities across a phylogenetically broad span of Cephalotes species. Through use of 16S rRNA amplicon sequencing and qPCR, we investigated symbiont community composition in the crop, midgut, ileum, and rectum. Through these approaches, and through replicated sampling across eleven Cephalotes species and one species of Procryptocerus, we inferred broad-scale patterns of Cephalotine microbiome composition and variability across gut chambers. We, finally, studied the impacts of the host phylogeny on the microbiome, and the potential for microbiomes to vary across castes.

MATERIALS & METHODS
Specimens were collected from the Southeastern United States and both Central and South America in 2015-2016. Twenty-eight ant colonies were collected from eleven species of *Cephalotes* and one species of *Procryptocerus*. Voucher specimens for all species are deposited in the biological collections of the Cornell University Insect Collection or Field Museum of Natural History. Alate (winged) queens, dealate queens, soldiers, and minor worker castes were all collected when possible.

These specimens were either stored in 95-100% ethanol or RNA later in the field and kept at -20°C until extraction, or collected in the field, then lab-reared and subsequently live dissected. Adult ant workers, queens, and soldiers were rinsed in 70% ethanol and then sterile deionized water before dissection (17). Under a microscope, ant digestive compartments were dissected using sterile forceps. Between each individual dissection, forceps were washed with a 6% bleach solution and then with sterile deionized water. The individually dissected gut compartments (crop, ileum, midgut, and rectum) or entire gasters were placed into separate sterile 1.5-mL tubes with 180 uL enzymatic lysis buffer (20 mM Tris·Cl, pH 8.0, 2 mM sodium EDTA, 1.2% Triton X-100, Immediately before use, added lysozyme to 20 mg/ml), then crushed using sterile pestles. After grinding, samples were incubated at 37°C for 30 minutes and subsequently extracted using Qiagen DNEasy Blood & Tissue Kit using the protocol for tissue and the Pretreatment for Gram Positive Bacteria. After DNA extraction, 492 samples were prepped for sequencing. In addition, four negative controls were included for 16S rRNA amplification and library sequencing, for a total of 496 samples.

Amplicon sequencing of the bacterial community was performed using the V4 region of the 16S rRNA using the primers 515F (5’-GTGCCAGCMGCCGCGGTAA – 3’) and 806R (5’-GGACTACHVGGGTWTCTAAT – 3’), described in Caporaso *et al.* (2012), following the
Earth Microbiome Project (EMP) protocol. Paired-end 151 bp reads were then sequenced on the Illumina MiSeq platform (Illumina, Inc., San Diego, CA) at Argonne National Laboratory (Lemont, Illinois, USA).

Real-time quantitative polymerase chain reaction (qPCR) was performed to estimate the abundance of bacteria in each sample. The universal bacterial 16S rRNA primers 515F and 806R were used (37). All qPCR reactions were conducted in triplicate on a CFX Connect Real-Time System (Bio-Rad, Hercules, CA) using SsoAdvanced 2X SYBR green supermix (Bio-Rad) and 2uL of sample DNA extraction. Standard curves were generated from serial dilutions of linearized plasmid-containing inserts of Escherichia coli 16S rRNA (38, 39). We required that all qPCR standard curves have efficiency values between 90% to 110% and R² values above 0.9.

Each sample was also quantified via the Qubit Fluorometer using the Qubit dsDNA High Sensitivity assay kit (Thermo Fisher, Waltham, MA). The cycle number was log-transformed in order to yield a linear starting copy number. These log-transformed mean SQ values (starting quantity estimate for the sample) from the qPCR data were standardized by total DNA concentration as determined by Qubit fluorometry (40). We used ggplot2 (41) to plot all qPCR results in R and tested the correlation of qPCR estimates of bacterial abundance using linear mixed models (LMMs) with the lme4 package in R (42). For these LMMs, gut compartment and caste were treated as fixed effects and colony nested within host species were treated as random effects. Akaike Information Criteria was used to select the best model. Pairwise comparisons to assess significance of differences within specific gut compartments were made with the Emmeans package in R using the Tukey’s HSD method (42).

To determine bacterial composition in each gut compartment from across species in the Cephalotini, 16S rRNA amplicon sequencing data were analyzed with the open-source QIIME2
Sequences were quality filtered using the DADA2 algorithm (Callahan et al., 2016). This algorithm joins paired-end reads, then uses a quality-aware correcting model for amplicon data, which denoises, removes chimeric sequences and residual PhiX reads, dereplicates DNA reads and subsequently calls amplicon sequence variants (ASVs). ASVs are used as a proxy for bacterial ‘species’ and are similar to OTUs (operational taxonomic units) but at a finer scale resolution (100% similarity). Paired-end sequence reads were trimmed (both reads were trimmed at 12 and truncated at 150). An alignment was made using MAFFT (46) and an ASV phylogeny was inferred from these sequences using RAxMLv8.1.16 (47). These samples were normalized with a sampling depth of 13,335 reads. Sequences were assigned to taxonomic groups using the SILVA_132_QIIME database to train our specific classifier (48, 49).

Contaminants from negative controls were filtered using the Decontam package (50). To create a second dataset, which only contained Cephalotes species and excluded samples from the sister genus (Procryptocerus), all the Procryptocerus samples were excluded and the dataset was rarefied at a sampling depth of 13,814. This rarefied dataset with Procryptocerus samples excluded is the one used for subsequent analyses unless noted.

Alpha diversity metrics were computed to measure the richness of the communities within samples: Shannon Index and Faith’s Phylogenetic Diversity Index, Pielou’s evenness, and ASV richness (51–54). These metrics were all computed in QIIME2 using the Kruskal-Wallis Pairwise Comparison Test. Relative abundance plots were created using Phyloseq (55) to assess the bacterial community composition among all samples. Variables such as gut compartment and caste type were tested under these alpha diversity metrics. Caste type samples were tested as both pooled samples based on caste as well as being subdivided by gut compartment then tested by caste type within that specific compartment.
Differences in beta diversity between samples were measured with a permutational multivariate analysis of variance (PERMANOVA) on weighted and unweighted Unifrac distance matrices as well as a Bray-Curtis dissimilarity matrix. Principal coordinates analysis ordination was calculated based on the Bray-Curtis dissimilarity matrix using Adonis (56) from the Vegan package as well as in QIIME2 (57), which performed a PERMANOVA based on Bray-Curtis and weighted UniFrac distance matrices (999 permutations). The subsequent PCoA plots were then visualized using ggplot2 (55). This was used to test for differences in beta diversity among ant species, castes, colony, and sample type – i.e. particular gut compartments or whole gasters/abdomens, containing the entire gut. The R package ecdist (58) was used to test for host phylogenetic signal correlating with similarities in distribution and abundance by comparing the Cephalotes ant phylogeny and the Jaccard distance matrix from the ASV table, using the Mantel test (59). To compare to the host phylogeny we used the ant phylogeny from Price et al. 2014 (60), with the drop.tip function in ape (61) to prune the phylogeny to only the twelve ant species examined in this study (Top left phylogeny in Fig. 6). We used the subset_taxa function within Phyloseq to subset the rarefied datasets to only contain specific bacterial orders and subsequently test them for host phylogenetic signal using the Mantel test.

Using the Past3 program, similarity percentage analysis (SIMPER) was implemented to determine specific ASV contributions to the structure of the bacterial communities (62, 63). Subsequently, boxplots were created for what SIMPER identified as the top ten most important ASVs contributing to the bacterial community composition within Cephalotes gut compartments. All boxplots were made using the plugin ggplot2 within R and statistical tests between groups were performed using the Wilcoxon rank-sum test with FDR correction (p < 0.05) (41). The code for this study is available at github.com/peterjflynn/cephalotes_gut_localization.
RESULTS

A total of 492 samples were sequenced with four control samples (Dataset S1). The 16S rRNA amplicon sequencing raw reads are available from NCBI via BioProject record SAMN16421870. There were a total of 10,380,596 reads from the raw dataset. Within the rarefied dataset, there were no crop samples for *C. cordatus* samples, due to very low abundance of ASVs in those samples. For the rarefied dataset with the *Procryptocerus* samples included, at a sampling depth of 13,335, there were 5,334,039 reads after rarefaction with a total of 400 samples after 92 samples were removed (along with the four control samples). For the rarefied dataset with the *Procryptocerus* samples excluded, at a sampling depth of 13,814 there were a total of 5,055,971 reads left after rarefaction with a total of 81 samples removed due to rarefaction (along with the four controls and 45 *Procryptocerus* samples) for a total dataset of 366 samples. We excluded *Procryptocerus* samples from the main analyses, to focus on the similarity and differences across digestive tissues within the *Cephalotes* samples. This dataset was used when an outgroup was needed for the analyses.

**Alpha Diversity Metrics**

Across all samples, alpha diversity metrics including Shannon diversity and Pielou’s Evenness were significantly higher in the ileum, rectum, and total gaster samples than in the midgut and crop samples (Kruskal–Wallis ANOVA, df=4; H=227.93 and P < 0.001 for Shannon diversity, H=224.10 and P < 0.001 for Pielou’s Evenness; corrected P-values < 0.001 for all Kruskal–Wallis pairwise tests; Fig. 1A,C). Crop, ileum, rectum, and gaster samples were all significantly higher in alpha diversity than the midgut samples using the Faith’s phylogenetic
diversity metric as well as ASV richness (H=94.74 and P < 0.001 for Faith’s PD and H=153.72 and P < 0.001 for ASV richness; Fig. 1B,D). For the alpha diversity analyses of gut compartment all caste types were combined.

No alpha diversity metric tested found soldier vs worker or soldier vs queen samples significantly different (Fig. S1A,B,C,D). Across all samples, ASV richness was significantly different based on caste type (H=7.2, p=0.021; Fig. S1D). ASV richness was significantly higher in worker samples than in the queen samples (H=6.7 and p=0.009; Fig. S1D). Shannon diversity was different (though not significantly) when grouped by caste type (H=5.5 and p=0.064; Fig. S1A). Specifically, Shannon diversity was significantly higher when only looking at workers compared to queens (H=5.5, p=0.019; Fig. S1A). Pielou’s Evenness and Faith’s PD were not significantly different over all the samples (H=4.38, p=0.112 for Faith’s PD and H=4.14, p=0.126 for Pielou’s Evenness; Fig. S1B,C). However, for Pielou’s Evenness the workers were significantly more even than the queens (H=4.05, p=0.044; Fig. S1C). When caste type was subdivided first by gut compartment, the only significant results were that Shannon (H=2.64, p=0.0247) and Evenness (H=2.74, p=0.0184) were significantly higher in workers than in queens within the midgut and ASV richness (H=2.64, p=0.0252) was significantly higher in workers than in queens within the ileum. All statistics for alpha diversity metrics are found in Dataset S2.

**Bacterial Abundance and Quantification**

The top six most abundant bacterial orders overall within this 16S rRNA amplicon dataset were Burkholderiales, Enterobacteriales, Opitutales, Rhizobiales, Rickettsiales, and Xanthomonadales (Fig. 2A). The relative abundance plot in Fig. 2A illustrates that the samples from the midgut, ileum, and rectum, and gaster exhibited consistent bacterial orders across all
samples within their respective gut compartments. On the other hand, the crop samples were not consistent and on average the largest contributors to the relative abundance of the microbial communities were Rhizobiales, which make up 18.6% and Opitutales which make up 16.3% of the overall bacterial community, respectively (Fig. 3B). Most crop samples of *C. auricomus* were dominated by Rickettsiales (60.49%) with 100% being composed of the genus *Wolbachia*. 

**Cephalotes similimus** crop microbial communities were predominantly made up of Rhizobiales (74.05%). Crop samples from the remaining *Cephalotes* and *Procryptocerus* species exhibited no distinct pattern (Fig. 3B), harboring a range of specialized symbionts and rare, apparent transient non-specialists. Though the majority of the crop samples had low raw read number, the crop samples with higher raw read number, and presumably higher density, were comprised of bacteria from the same orders as specialized core *Cephalotes* symbionts (Opitutales, Xanthomonadales, Burkholderiales, and Rhizobiales) (Fig. 2B).

The midgut samples were dominated by Opitutales (bacterial genus *Cephaloticoccus*; Fig. 3C). Bacteria from this taxon made up 84% on average of the relative abundance of the microbial community within the midgut (Fig. 3C). The microbial community composition of the ileum and the rectum samples were similar (Fig. 2A), with the most abundant orders being Burkholderiales (26.3% ileum vs. 25.7% rectum), Xanthomonadales (20.89% ileum vs. 20.15% rectum), Opitutales (12.59% ileum vs. 12.15% rectum), and Rhizobiales (12.43% ileum vs. 11.6% rectum; Figs 3D-E). A co-occurrence Venn diagram illustrating the degree of overlap of bacterial ASVs demonstrated that bacterial ASVs differed among the four gut compartments (Fig. S2). In addition, it further indicates that the ileum and the rectum have the largest number of ASVs in common of any two gut compartments, the crop has a proportionally lower number
of ASVs in common with other gut compartments, and that there is a relatively large group of
core ASVs common to all gut compartments.

Based on the qPCR results, the crop samples exhibited an average of 4,928 copies of
bacterial 16S rRNA genes per sample (Fig. 4A). The midgut samples exhibited an average of
280,386 copies of bacterial 16S rRNA genes per sample (Fig. 4A). The ileum samples exhibited
an average of 157,463 copies of bacterial 16S rRNA genes per sample (Fig. 4A). The rectum
samples exhibited an average of 23,778 copies of bacterial 16S rRNA genes per sample (Fig.
4A). Finally, the gaster samples exhibited an average of 482,949 copies of bacterial 16S rRNA
genes per sample (Fig. 4A). These results mirrored read numbers from our raw 16S rRNA data
(Fig. 2B).

All Linear Mixed Models (LMMs) incorporating caste as a fixed effect had higher AIC
values, which indicates that caste does not correlate strongly with qPCR values. The best fit
LMM (lowest AIC value) incorporated gut compartment as the sole fixed effect (Table S1). Gut
compartment correlated strongly with qPCR values (P<0.0001; Table S1). The mean bacterial
qPCR abundances within the midgut, ileum, and the gaster had significantly more bacterial
abundance than the crop and the rectum (Fig. 4A). The crop and the rectum were the only gut
compartments that were not statistically different from each other with regard to bacterial
abundance (p = 0.8272, Fig. 4A; Table S2). When these values are plotted by both species and
gut compartment, it looks as though there is variation by species, but that crop and rectum have
significantly lower bacterial abundance for every species sampled (Fig. 4B).

**Beta Diversity Metrics**
We performed PERMANOVA on weighted UniFrac (wUniFrac) and Bray-Curtis distances calculated from the rarefied dataset without *Proryptocerus* to test for dissimilarities in microbial community composition among samples based on ASVs and the variables: gut compartment, species, caste, and colony (Table S3). Using a Principal coordinates analysis (PCoA) for visualization, the wUniFrac distances demonstrated that the ASV samples significantly clustered by gut compartment type (Fig. 5A; Pseudo-F for gut compartment = 157.143, P-value >0.001; Table S3). Though species, colony, and caste type all were significant, the gut compartment variable had by far the largest Pseudo-F statistic and helps explain the clustering of the samples by gut compartment (Table S3). Whole gaster microbial community samples overlapped with every other gut compartment. In addition, the ileum and the rectum samples were not distinct from each other, whereas every other gut compartment was significantly distinct from one another based on pairwise PERMANOVAs (Table 1). When we removed the whole gaster samples from the dataset, the wUniFrac PCoA again clumped samples by gut compartment, with the ileum and rectum overlapping and the midgut and crop samples both being distinct from every other gut compartment based on their microbial community composition over all axes (Fig. 5B, Pseudo-F for gut compartment = 298.537, P-value >0.001).

For the PCoA using the Bray-Curtis dissimilarity matrix, host species had the largest F statistic though gut compartment, caste type, and colony variables were also significant (Fig. S3; Table S4). Visually, all of the samples clustered together except for the *C. texanus* samples, which formed a distinct cluster (Fig. S3; Pseudo-F for Species = 77.777, P-value >0.001). To further examine this pattern, we analyzed the PCoA using the Bray-Curtis dissimilarity matrix for the gaster (Fig. 6A), crop (Fig. 6B), midgut (Fig. 6C), ileum (Fig. 6D), and rectum (Fig. 6E).

In the midgut, ileum, rectum, and gaster plots, we found that the *C. texanus* samples cluster
together away from every other *Cephalotes* species, which form a non-distinct cluster in the
gaster and midgut PCoA plots (Fig. 6A,C), and a more *Cephalotes* species-specific cluster in the
ileum and rectum PCoA plots (Fig. 6D,E). For the crop PCoA plot, the only samples that cluster
significantly by species are *C. simillius* samples due to the dominance of Rhizobiales within
these crop samples (Fig. 6B).

**ASV Similarity Percentage Analyses**

Through the Similarity percentage (SIMPER) analysis run in Past3, the top ten bacterial
ASVs that were important in structuring the bacterial communities within the *Cephalotes* gut
compartments were identified (Table 2). Of these, two ASVs were associated with *C. texanus*
samples alone. Among the *C. texanus*-enriched ASVs, Opitutales ASV2 was only found in *C.
texanus*, exhibiting high read numbers in every individual, and highest abundance within the
midgut (Dataset S3). While not exclusive to *C. texanus*, Xanthomonadales ASV2 was enriched
in these hosts, being also found only in *C. minutus* across the eleven remaining species (Dataset
S3). Finally, Opitutales ASV1 (a sequence with just 1 bp difference vs. Opitutales ASV2) was
completely absent in *C. texanus* samples. It was, instead, found in generally high abundance
within the other sampled species, although at lower relative abundance in *C. grandinosus*, *C.
placidus*, and *Procryptocerus* (Dataset S3). Results of the pairwise comparisons with the
Wilcoxon rank-sum test with FDR correction for the ASVs discussed are found in Dataset S3.

**Co-Diversification Analyses**

The non-parametric Mantel test comparing the gaster (entire abdomen) Jaccard similarity
index distance matrix ASV table by species and the distance matrix by *Cephalotes* phylogeny
illustrated a statistically significant relationship (r=0.4017, p=0.006703). The Mantel tests comparing the crop and the midgut ASV samples Jaccard distance matrix to the host phylogeny did not find a statistically significant relationship, though possible trend for the midgut (r=0.0455, p=0.43822 for crop and r=0.3963, p=0.053524 for midgut). On the other hand, the mantel tests comparing the ileum and rectum ASV samples Jaccard distance matrix to the host phylogeny found a statistically significant relationship (r=0.3534, p=0.039836 for the ileum and r=0.2485, p = 0.037296 for the rectum). Of the six most abundant bacterial orders, the two significantly correlated with the *Cephalotes* phylogeny were Opitutales (r=0.4017, p=0.0356) and Burkholderiales (r=0.332, p=0.00325). Results for all mantel tests are found in Table 3.

Of the top six most abundant bacterial orders, Enterobacteriales and Rickettsiales were found to be negatively correlated based on the results of their mantel tests (r=-0.2352, p=0.61 for Rickettsiales and r=-0.1854, p=0.80 for Enterobacteriales). Neither of these bacterial orders co-speciate or are host specific within *Cephalotes* (Kautz et al., 2013b; Russell, 2012). Rickettsiales (*Wolbachia*) is often found at relatively high prevalence within *Cephalotes* species, but there is no evidence that Rickettsiales or Enterobacteriales is co-diversifying due to symbiosis with *Cephalotes* species.

**DISCUSSION**

Host-associated microbes perform a myriad of beneficial functions within their host. Across insects, microbes are involved in increasing tolerance to environmental perturbations (64, 65), priming the immune system (66–68), and aiding in digestion and nutrition (69–71). Within ants, gut-associated microbes contribute to a variety of nutritional adaptations. (9, 11, 72). Herbivorous turtle ants maintain a remarkably conserved symbiotic bacterial community, and the constituent microbes utilize recycled waste nitrogen to synthesize amino acids (12).
Since digestive function is not uniform across the alimentary canal, we expect the microbial community to be variable across the gut compartments within *Cephalotes*. While gut compartmentalization of microbes has been well studied in mammals and birds (73–75), and has been observed in insect groups such as termites (76) and beetles (77), few studies have documented specific microbial gut compartmentalization within ants. To understand if this pattern is common across Cephalotine ants, we examined the microbial community within each gut compartment in a variety of species spanning the *Cephalotes* phylogeny.

Gut compartment helps determine microbial community structure

Studies of a subsocial, wood-feeding beetle, *Odontotaenius disjunctus*, found their gut contains four discrete compartments, each with distinct microbial communities (77). There is evidence that this may be valuable in the aerobic and anaerobic mechanisms of energy extraction of woody material (78, 79). In addition, within both soil- and wood-feeding termites (*Cubitermes* and *Nasutitermes* respectively) morphological differentiation of their digestive tracts creates specialized environments which aid microbial metabolic processes that permit survival on hard to digest food sources (76, 80). Pollen is essential to the bee diet, however its cell wall is indigestible by the honeybee itself. Instead, bacteria found within the honey bee gut degrade pectin in the cell wall (81). The highly compartmentalized organization of bacteria within the honeybee gut may aid the bacteria in digesting recalcitrant forms of carbon from the pollen component of their diet.

The implication of gut compartmentalization with specialized microbes is that these microbes become increasingly specialized for their host, and over evolutionary time the host becomes increasingly reliant on these bacteria for survival. In addition, as microbes become specialized,
the gut of the insect becomes increasingly partitioned for these specialized symbionts (e.g. termite guts) (76). We found that gut compartment emerges as an important way to structure the microbial community within the Cephalotes digestive tract. In particular, crop and midgut microbiomes are distinct, both from each other and from the ileum/rectum (Fig. 2A; Fig. 5B; Fig. S2). The lack of uniformity across the digestive tract raises the possibility that symbionts perform distinct functions within each gut compartment (Fig. 7).

Some insects have a crop with abundant and diverse microbes (82, 83). However, more commonly the crop contains bacteria in low abundance or is almost entirely lacking microbes possibly due its routine evacuation (84, 85). In our study, these crop samples often hosted a different set of microbes than the rest of the compartments (Fig. 5B), although residents were found at low abundance. In general, bacterial community composition of crops varied widely across sibling ants, colonies, and species, exhibiting few consistent patterns across all samples (Fig. 2A; Fig. 3B). This suggests that these crop samples reflect a diverse suite of bacteria that are held at low density, which could be transient microbes from food sources or those acquired from oral-anal trophallaxis. In addition, qPCR values from the crop recorded uniformly negligible amounts of 16S rRNA gene copies (Fig. 4B) and the lowest number of raw reads of any gut compartment (Fig. 2B). The low quantity of bacteria within the crop samples along with a lack of bacterial uniformity is consistent with the function of the crop as the social stomach within the ant gut (35), and not a site for specialized bacterial colonization or for major nutritional/digestive function.

The microbial community of the midgut is distinct from other gut compartments and highly similar across colonies and species (Fig. 3C). Within all eleven Cephalotes species sampled, the midgut samples are dominated by a single bacterial order, Opitutales, comprised
almost exclusively of the specialized core symbiont genus *Cephaloticoccus* (Fig. 2A). This result is consistent with previous studies of *C. rohweri* midgut, which was completely dominated by what appeared to be *Cephaloticoccus* (26). Accordingly, the midgut had the lowest alpha diversity of all gut compartments based on every metric of alpha diversity we measured (Fig. 1).

Urea, major component of mammal urine, is often ingested by *Cephalotes* species (22, 23). The abundance of *Cephaloticoccus* found in the midgut, coupled with the discovery of urease-encoding capacities across this group suggests that they assist in urea metabolism (12). These bacteria also encode glutamate dehydrogenase (gdhA) and have complete pathways for synthesizing most essential amino acids (12, 86). This indicates that within the midgut, *Cephaloticoccus* may have the capacity to convert urea into ammonia, ammonia into glutamate, and to subsequently use glutamate as an N-donor in synthesizing several essential amino acids (Fig. 7). In addition, given the high abundance of *Cephaloticoccus* in the midgut, it likely has inherent properties that either allow it to outcompete other bacterial groups, or to uniquely thrive in the harsh midgut conditions.

The ileum contains a combination of both partially digested liquid food and nitrogenous waste. The uric acid in the ileum likely derives from two sources: ingested bird excrement and output from the Malpighian tubules (12, 32). Malpighian tubules spread into the *Cephalotes* body cavity and absorb nitrogenous waste, like uric acid, functioning as the main excretory organ of insects (3). Uric acid can be converted into urea with the aid of Burkholderiales bacteria (12). From there, Opitutales, some Rhizobiales, and occasionally Xanthomonadales or Burkholderiales could further break down urea into ammonia. Host bacteria could then assimilate the ammonia into glutamate, utilizing this molecule as an N-donor in the synthesis of different amino acids (Fig. 7). Burkholderiales ASV1 and ASV2 were most similar to Burkholderiales strain
POW0550W-166 (100% and 99.12% percent identity based on blastn, respectively; Accession Number MF441555). POW0550W-166 can produce urea from allantoin, which acts as a proxy for the latter portion of the uric acid degradation pathway (12). Since Burkholderiales ASVs make up the largest portion of the bacterial community within the ileum (26%), Burkholderiales bacteria may assist in uric acid degradation within the *Cephalotes* ileum. Microscopy studies have found large aggregations of bacterial cells along the entire ileum, with an especially high bacterial load at the midgut-ileum junction where the Malpighian tubules deposit nitrogenous waste (12, 32, 87). Hence, the localization of uric-acid degraders and symbionts likely to express urease within this chamber suggests the ileum as another important site for N-recycling and amino acid metabolism.

When comparing the bacterial community composition of the ileum and the rectum, the similarity is striking. The ileum and the rectum tightly overlap and cluster together in the Weighted uniFrac PCoA plot, which incorporates abundance of the ASVs as well as phylogenetic distance (Fig. 5B). Relative abundance of reads assigned to Burkholderiales, Opitutales, Rhizobiales, and Xanthomonadales were, furthermore, virtually identical across these gut compartments (Fig. 2A). These similarities are not surprising, since the ileum and the rectum together form the hindgut and both function in nutrient digestion and reabsorption. This raises the possibility that N-recycling and amino acid biosynthesis is further executed throughout the rectum. However, the rectum appears to harbor very small numbers of symbiotic bacteria, a finding derived from our qPCR assays and from microscopy studies of the *Cephalotes* gut (87).

We, thus, propose that rectal bacteria arrive with the passage of digestate from the distal digestive tract, which may enable successful microbial transmission through oral-anal trophallaxis.
503 **Caste may play a role in structuring microbial communities**

In addition to gut compartment, caste may play a minor role in structuring microbial gut communities. The three distinct *Cephalotes* castes sampled were worker, soldier, and queen. While microbes within the soldier caste samples were not significantly different from those in either the worker or the queen samples, worker samples exhibited slightly higher microbial Shannon diversity, ASV counts, Faith’s PD, and evenness than the queen samples (Fig. S1). Figure S4 highlights the taxonomic composition by caste type within the data. The queen, worker, and soldier samples all appear to primarily contain microbes associated with their core symbionts (Fig S4). This means that potentially differential caste exposure to environmental microbes most likely is not playing a major role in the structuring of the microbial gut community of any caste. In addition, there were no easily detectable differences in ASVs between workers, soldiers, and queens and therefore more fine-scale work would help tease apart these dissimilarities.

518 **Evolutionary history of the host structures gut bacterial communities**

Host phylogeny comprises another correlate of bacterial community composition within the *Cephalotes* gut. This study corroborates comparative studies that find *Cephalotes*-specific microbes comprising nearly all of the gut microbiome in a variety of *Cephalotes* species (17, 18, 21, 26). The Mantel test demonstrated a significant correlation between the *Cephalotes* phylogeny and a dendrogram constructed from beta diversity measures of bacterial community similarity in the gaster (r=0.547, p=0.0067). This supports work from a prior study focused mostly on a different set of *Cephalotes* species, which found evidence that the observed pattern
of microbial community correlation with host phylogeny was consistent with codiversification rather than diet or other selective conditions (21). In our study, among the six most abundant bacterial orders found within the ant gut, Opitutales and Burkholderiales were significantly correlated with the Cephalotes phylogenies. In addition, Rhizobiales and Xanthomonadales were also highly abundant, though not significantly correlated with the Cephalotes phylogeny. All four of these bacterial orders are part of the stable core microbiome within the Cephalotes digestive tract and are found at high prevalence (9, 14, 17). This could provide further evidence that these core microbial groups are co-diversifying within their Cephalotes host digestive tracts over evolutionary time.

The influence of the host phylogeny on the bacterial community varies by gut compartment

This study provides a finer context for understanding how the evolutionary history of Cephalotes is shaping the microbial community within the gut. We found that host-symbiont evolutionary correlation varied depending on the specific gut compartment within the digestive tract.

Unlike more posterior chambers, crop samples were highly variable within and between species, and were not colonized by a consistent set of symbionts or environmental bacteria (Fig. 3B). Only two species provided exceptions to this overall pattern: C. auricomus, with crops harboring large numbers of Rickettsiales, and C. simillimus, whose crop compartments were dominated by Rhizobiales. In accordance with the overall trends, a Mantel test (Table 3; r=0.0455 p=0.43822) suggested that host phylogeny does not correlate with community similarity of the crop microbiome. Any actual influence of host phylogeny may be hard to detect since the crop samples with the lowest bacterial density tended to have the highest abundance of...
non-specialized/core bacteria. Furthermore, a number of bacteria in the crop were not related to core symbionts, suggesting regular occurrence of environmental microbes or an inability to remove contaminant sequences that dominated the libraries of these microbially sparse samples.

Additionally, this study did not find a discernible influence of host phylogeny on the midgut bacterial community composition, despite a possible trend ($r=0.3963$, $p=0.053524$) While the crop samples had low abundance and variable bacterial communities, the midgut samples had generally high bacterial abundance and were almost completely uniform in bacterial composition i.e. dominated by a few *Cephaloticoccus* (Opitutales) ASVs. So even though an important core symbiont dominates the *Cephalotes* midgut, the uniformity of the midgut microbiome – and our reliance on slowly evolving 16S rRNA – may obscure finer-scale impacts of phylogeny (i.e. co-speciation) detectable through alternative methods.

Within the ileum and the rectum, the host’s evolutionary history appears to influence the bacterial community composition (Table 3). These gut compartments show a relatively consistent set of bacterial orders based on the specific gut compartment and host species sampled (Fig. 3D,E). While there was some variation in how bacterial communities mirror the *Cephalotes* host phylogeny here, there was still host-specific phylogenetic clustering (Fig. 6D,E). The ileum illustrated the greatest degree of visible clustering based on host phylogeny, as every species overlapped and clustered with their most closely related sister species, with the exception of *C. texanus/C. multispinosus* (Fig. 6D). The core bacteria found in the ileum may be driving this clustering based on host phylogeny. Since the ileum is the compartment where Burkholderiales, Opitutales, Rhizobiales, and Xanthomonadales are all found at high relative abundance, they may exhibit important nutritional functions. More in-depth studies specifically examining...
phylogenetic congruence on *Cephalotes* bacterial microbiome are needed to more definitively illustrate potential co-speciation based on gut compartment.

*Cephalotes texanus has a partially distinct gut microbial community*

Despite overall similarity in order-level composition (Fig. 3), the microbial community of *C. texanus* was significantly distinct from all other species sampled across the midgut, ileum, rectum, and gaster according to ASV-based measures (Fig. S3; Fig. 6A,C,D,E). Within the gaster samples, *Procryptocerus* and *C. texanus* both clustered away from the rest of *Cephalotes* samples when including *Procryptocerus*. Since *Procryptocerus* is the most distantly related group in the dataset, it is unsurprising that its bacterial community composition is distinct. However, *C. texanus* is not an early branching lineage within the *Cephalotes* genus. Within the sampled phylogenetic tree, *C. texanus* is a sister species to *C. multispinosus* (Fig. 6).

The majority of the *C. texanus* samples were live dissected from lab kept colonies instead of wild caught and preserved in the field (stored in 95-100% ethanol at -20°C until extraction) like other analyzed sample. This could explain why they clustered away from the other *Cephalotes* species (Fig. S3). Sanders et al. (2014) found that ethanol-preserved gasters recovered less Opitutales in 16S rRNA sequencing. However, we also included three wild caught and preserved *C. texanus* samples stored in 95% ethanol which clustered with the lab reared and live dissected *C. texanus* samples (Fig. S3), arguing against preservation method as the driver of this pattern.

A small number of ASVs unique to *C. texanus* drive the distinctive nature of their microbiome. Based on the SIMPER analysis, three of the top ten ASVs were associated specifically with (or lacking in) *C. texanus* (Opitutales ASV1, ASV2, and Xanthomonadales...
Opitutales ASV2 and Xanthomonadales ASV2 were found within *C. texanus*, whereas Opitutales ASV1 was found at high levels only within other *Cephalotes* species. (Dataset S3). Opitutales ASV2 was ubiquitous and at high abundance within every *C. texanus* sample, with especially high read numbers in the midgut. With a more detailed comparison, we found that Opitutales ASV2 was 100% identical (based on blastn) to the cultured symbiont bacterial strain JDR108-110A-112 (Accession Number MF945635). This strain is positive for urea degradation via urease, as in other *Cephaloticoccus* strains studied to date, and contributes to the N-recycling pathway (12). It is not yet clear, then, if this symbiont differs functionally from other symbionts in this lineage or whether its sequence differentiation is reflective of substitutions accruing within a specialized bacterium showing little capacity for horizontal transfer.

Gaster serves as a general proxy for digestive tract

Often in microbial studies of ant guts, DNA is isolated and sequenced from the entire gaster (or abdomen) as a proxy for the digestive tract (15, 87, 89). We tested this assumption by including both whole gaster and specific gut compartments. We compared the relative abundance and microbial community composition of the whole gasters to those of the specific gut compartments (Fig. 3A; Fig. 6A). The microbial community composition of the gaster samples was found to overlap with samples from all gut compartments: crop, midgut, rectum, and ileum (Fig. 2A). In addition, the qPCR values for the whole gasters were on average almost double any of the other gut compartments, and on average close to the sum of the averages of each of the individual compartments (Fig. 4B). The high microbial abundance within the gaster and the microbial community composition overlap provides evidence that gaster samples can serve as a rough proxy for the microbial community found within the digestive tract. Since dissecting each
minute gut compartment within the *Cephalotes* gut is difficult and time intensive, sampling whole gasters can be used for future studies examining gut symbionts within this genus. However, studies only sampling whole gasters would not be capable of discerning patterns of compartmentalization described in this study.

**Conclusions**

Our study sheds new light on the forces structuring the specialized, ancient microbial community within the digestive tract of *Cephalotes* turtle ants. Gut compartment plays a large role in shaping community structure within individual ants. Despite some taxonomic overlap between gut compartments, chambers of the *Cephalotes* digestive tract differ drastically in symbiont abundance, diversity, and relative abundance. Caste explains additional variation in microbiome composition, with queens harboring small differences in their microbiomes compared to those of workers. The evolutionary history of the *Cephalotes* species serves as the third, and final, factor found to shape the microbiome in our study, revealing a trend in which related ants harbor related microbiomes – quite possibly due to co-speciation. But remarkably, gut chamber-specific signatures of microbiome abundance, diversity, and composition have remained conserved across *Cephalotes* history, suggesting that this symbiosis has been fairly stable across a span of over 50 million years. Taken together, these findings are notable since they illustrate that specific gut compartments may be modifying and structuring their microbial community at a finer scale, possibly due to the varying physiological conditions in each gut chamber, but conceivably due to the influence of bacterial competition and more active measures by hosts to distinctly regulate their microbiome. Further studies including sequence isolates or metagenomic analysis should aim to elucidate the specific functional nature of these bacterial
communities within their respective gut compartments to better understand how they contribute
to digestion and nutrition within *Cephalotes*.

**ACKNOWLEDGEMENTS**

We thank the National Science Foundation for support for this project (NSF DEB 1900357 to CSM and 1442144 to JAR). We also thank Manuela Ramalho and Brooke Weigel for their help with the analysis of the microbial data as well as Maureen Coleman and Cathy Pfister for helpful discussions. We thank Milan Janda, Scott Powell, Shauna Price, and Sean O’Donnell for assistance with our collections and Steven Wang for the use of the *Cephalotes* photograph.

**REFERENCES**

1. Ley RE, Hamady M, Lozupone C, Turnbaugh PJ, Ramey RR, Bircher JS, Schlegel ML, Tucker TA, Schrenzel MD, Knight R, Gordon JI. 2008. Evolution of mammals and their gut microbes. Science 320:1647–51.

2. McFall-Ngai M, Hadfield MG, Bosch TCG, Carey H V., Domazet-Lošo T, Douglas AE, Dubilier N, Eberl G, Fukami T, Gilbert SF, Hentschel U, King N, Kjelleberg S, Knoll AH, Kremer N, Mazmanian SK, Metcalf JL, Nealson K, Pierce NE, Rawls JF, Reid A, Ruby EG, Rumpho M, Sanders JG, Tautz D, Wernegreen JJ. 2013. Animals in a bacterial world, a new imperative for the life sciences. Proc Natl Acad Sci 110:3229–3236.

3. Engel P, Moran NA. 2013. The gut microbiota of insects - diversity in structure and function. FEMS Microbiol Rev 37:699–735.

4. Kwong WK, Moran NA. 2015. Evolution of host specialization in gut microbes: The bee gut as a model. Gut Microbes 6:214–220.
Muegge BD, Kuczynski J, Knights D, Clemente JC, Gonzalez A, Fontana L, Henrissat B, Knight R, Gordon JI, Bain JR, Muehlbauer MJ, Ilkayeva O, Semenkovich CF, Funai K, Hayashi DK, Lyle BJ, Martini MC, Ursell LK, Clemente JC, Treuren W Van, Walters WA, Knight R, Newgard CB, Heath AC, Gordon JI. 2011. Diet Drives Convergence in Gut Microbiome Functions Across Mammalian Phylogeny and Within Humans. Science 332:970–974.

Hooper L V., Gordon JI. 2001. Commensal host-bacterial relationships in the gut. Science 292:1115–1118.

Cook SC, Davidson DW. 2006. Nutritional and functional biology of exudate-feeding ants. Entomol Exp Appl 118:1–10.

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

Russell JA, Moreau CS, Goldman-Huertas B, Fujiwara M, Lohman DJ, Pierce NE. 2009. Bacterial gut symbionts are tightly linked with the evolution of herbivory in ants. Proc Natl Acad Sci U S A 106:21236–21241.

Gil R, Silva FJ, Zientz E, Delmotte F, González-Candelas F, Latorre A, Rausell C, Kamerbeek J, Gadau J, Hölldobler B, Van Ham RCHJ, Gross R, Moya A. 2003. The genome sequence of Blochmannia floridanus: Comparative analysis of reduced genomes. Proc Natl Acad Sci U S A 100:9388–9393.

Feldhaar H, Straka J, Krischke M, Berthold K, Stoll S, Mueller MJ, Gross R. 2007. Nutritional upgrading for omnivorous carpenter ants by the endosymbiont Blochmannia. BMC Biol 5:1–11.

Hu Y, Sanders JG, Łukasik P, D’Amelio CL, Millar JS, Vann DR, Lan Y, Newton JA,
Schotanus M, Kronauer DJC, Pierce NE, Moreau CS, Wertz JT, Engel P, Russell JA. 2018. Herbivorous turtle ants obtain essential nutrients from a conserved nitrogen-recycling gut microbiome. Nat Commun 9:964.

13. Bisch G, Neuvonen MM, Pierce NE, Russell JA, Koga R, Sanders JG, Łukasik P, Andersson SGE. 2018. Genome evolution of Bartonellaceae symbionts of ants at the opposite ends of the trophic scale. Genome Biol Evol 10:1687–1704.

14. Anderson KE, Russell JA, Moreau CS, Kautz S, Sullam KE, Hu Y, Basinger U, Mott BM, Buck N, Wheeler DE. 2012. Highly similar microbial communities are shared among related and trophically similar ant species. Mol Ecol 21:2282–2296.

15. Łukasik P, Newton JA, Sanders JG, Hu Y, Moreau CS, Kronauer DJC, O’Donnell S, Koga R, Russell JA. 2017. The structured diversity of specialized gut symbionts of the New World army ants. Mol Ecol 26:3808–3825.

16. Wernegreen JJ, Kauppinen SN, Brady SG, Ward PS. 2009. One nutritional symbiosis begat another: Phylogenetic evidence that the ant tribe Camponotini acquired Blochmannia by tending sap-feeding insects. BMC Evol Biol 9:292.

17. Hu Y, Łukasik P, Moreau CS, Russell JA. 2014. Correlates of gut community composition across an ant species (Cephalotes varians) elucidate causes and consequences of symbiotic variability. Mol Ecol 23:1284–1300.

18. Kautz S, Rubin BER, Russell JA, Moreaua CS. 2013. Surveying the microbiome of ants: Comparing 454 pyrosequencing with traditional methods to uncover bacterial diversity. Appl Environ Microbiol 79:525–534.

19. Kelly M, Price SL, de Oliveira Ramalho M, Moreau CS. 2019. Diversity of Wolbachia Associated with the Giant Turtle Ant, Cephalotes atratus. Curr Microbiol 76:1330–1337.
20. Reeves DD, Price SL, Ramalho MO, Moreau CS. 2020. The Diversity and Distribution of Wolbachia, Rhizobiales, and Ophiocordyceps Within the Widespread Neotropical Turtle Ant, Cephalotes atratus (Hymenoptera: Formicidae). Neotrop Entomol 49:52-60.

21. Sanders JG, Powell S, Kronauer DJC, Vasconcelos HL, Frederickson ME, Pierce NE. 2014. Stability and phylogenetic correlation in gut microbiota: Lessons from ants and apes. Mol Ecol 23:1268–1283.

22. Baroni Urbani C, De Andrade ML. 1997. Pollen eating, storing, and spitting by ants. Naturwissenschaften 84:256–258.

23. Powell S. 2008. Ecological specialization and the evolution of a specialized caste in Cephalotes ants. Funct Ecol 22:902–911.

24. Eisner T. 1957. A comparative morphological study of the proventriculus of ants (Hymenoptera: Formicidae). Bull Mus Comp Zool 116:439-490.

25. Kwong WK, Medina LA, Koch H, Sing KW, Soh EJY, Ascher JS, Jaffé R, Moran NA. 2017. Dynamic microbiome evolution in social bees. Sci Adv 3:1–17.

26. Lanan MC, Rodrigues PAP, Agellon A, Jansma P, Wheeler DE. 2016. A bacterial filter protects and structures the gut microbiome of an insect. ISME J 10:1866–1876.

27. Wilson EO. 1976. A social ethogram of the neotropical arboreal ant Zacryptocerus varians (Fr. Smith). Anim Behav 24:354–363.

28. Hölldobler B, Wilson EO. 1990. The Ants. Harvard University Press, Cambridge, MA.

29. Wheeler DE. 1984. Psyche Behavior of the Ant, Procryptocerus Scabriusculus (Hymenoptera: Formicidae), with Comparisons to Other Cephalotines. Psyche (New York) 91:171–192.

30. Jaffe K, Caetano FH, Sánchez P, Hernández J V., Caraballo L, Vitelli-Flores J, Monsalve
W, Dorta B, Rodriguez Lemoine V. 2001. Sensitivity of ant (Cephalotes) colonies and individuals to antibiotics implies feeding symbiosis with gut microorganisms. Can J Zool 79:1120-1124s.

Roche RK, Wheeler DE. 1997. Morphological specializations of the digestive tract of Zacryptocerus rohweri (Hymenoptera: Formicidae). J Morphol 234:253–262.

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

Bution ML, Caetano FH. 2010. The midgut of Cephalotes ants (Formicidae: Myrmicinae): Ultrastructure of the epithelium and symbiotic bacteria. Micron 41:448–454.

Chapman RF. 1998. The Insects: Structure and Function, 4th Edition. Cambridge Univ Press, Cambridge, UK.

Davidson DW, Cook SC, Snelling RR. 2004. Liquid-feeding performances of ants (Formicidae): Ecological and evolutionary implications. Oecologia 139:255–266.

Eisner T, Wilson EO. 1952. The Morphology of the Proventriculus of a Formicine Ant. Psyche (New York) 59:47–60.

Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N, Owens SM, Betley J, Fraser L, Bauer M, Gormley N, Gilbert JA, Smith G, Knight R. 2012. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. ISME J 6:1621–1624.

Moreau CS, Rubin BER. 2017. Diversity and persistence of the gut microbiome of the giant neotropical bullet ant. Integr Comp Biol 57:682–689.

Ramalho MO, Bueno OC, Moreau CS, Forel C, Wheeler H. 2017. Microbial composition of spiny ants (Hymenoptera: Formicidae: Polyrhachis) across their geographic range.
Evol Bio 17:96.

40. Rubin BER, Sanders JG, Hampton-Marcell J, Owens SM, Gilbert JA, Moreau CS. 2014. DNA extraction protocols cause differences in 16S rRNA amplicon sequencing efficiency but not in community profile composition or structure. Microbiologyopen 3:910–921.

41. Wickham H. 2009. ggplot2 Elegant Graphics for Data Analysis. Springer, New York.

42. R Core Team. 2017. R: A language and environment for statistical computing. R Found Stat Comput Vienna, Austria.

43. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, Alexander H, Alm EJ, Arumugam M, Asnicar F, Bai Y, Bisanz JE, Bittinger K, Brejnrod A, Brislawn CJ, Brown CT, Callahan BJ, Caraballo-Rodríguez AM, Chase J, Cope EK, Da Silva R, Diener C, Dorrestein PC, Douglas GM, Durall DM, Duvall C, Edwarson CF, Ernst M, Estaki M, Fouquier J, Gauglitz JM, Gibbons SM, Gibson DL, Gonzalez A, Gorlick K, Guo J, Hillmann B, Holmes S, Holste H, Huttenhower C, Huttley GA, Janssen S, Jarmusch AK, Jiang L, Kaehler BD, Kang K Bin, Keefe CR, Keim P, Kelley ST, Knights D, Koester I, Kosciolek T, Kreps J, Langille MGI, Lee J, Ley R, Liu YX, Loftfield E, Lozupone C, Maher M, Marotz C, Martin BD, McDonald D, McIver LJ, Melnik A V., Metcalf JL, Morgan SC, Morton JT, Naimey AT, Navas-Molina JA, Nothias LF, Orchanian SB, Pearson T, Peoples SL, Petras D, Preuss ML, Pruesse E, Rasmussen LB, Rivers A, Robeson MS, Rosenthal P, Segata N, Shaffer M, Shiffer A, Sinha R, Song SJ, Spear JR, Swafford AD, Thompson LR, Torres PJ, Trinh P, Tripathi A, Turnbaugh PJ, Ul-Hasan S, van der Hooft JJJ, Vargas F, Vázquez-Baeza Y, Vogtmann E, von Hippel M, Walters W, Wan Y, Wang M, Warren J, Weber KC, Williamson CHD, Willis AD, Xu ZZ,
Zaneveld JR, Zhang Y, Zhu Q, Knight R, Caporaso JG. 2019. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat Biotechnol 37:852–857.

Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Põe AG, Goodrich JK, Gordon JI, Huttenhower C, Kelley ST, Knights D, Koenig JE, Ley RE, Lozupone CA, McDonald D, Muegge BD, Pirrung M, Reeder J, Sevinsky JR, Turnbaugh PJ, Walters WA, Widmann J, Yatsunenko T, Zaneveld J, Knight R. 2010. QIIME allows analysis of high-throughput community sequencing data. Nat Methods 7:335-336.

Kuczynski J, Stombaugh J, Walters WA, González A, Caporaso JG, Knight R. 2012. Using QIIME to analyze 16s rRNA gene sequences from microbial communities. Curr Protoc Microbiol Chapter 1: Unit 1E.5.

Katoh K, Misawa K, Kuma K, Miyata T. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res 30:3059–3066.

Stamatakis A. 2014. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30:1312–1313.

Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Prieels J, Glöckner FO. 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res 41:D590–D596.

Yilmaz P, Parfrey LW, Yarza P, Gerken J, Pruesse E, Quast C, Schweer T, Prieels J, Ludwig W, Glöckner FO. 2014. The SILVA and “All-species Living Tree Project (LTP)” taxonomic frameworks. Nucleic Acids Res 42:D643–D648.
50. Davis NM, Proctor DiM, Holmes SP, Relman DA, Callahan BJ. 2018. Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. Microbiome 6:226.

51. DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, Huber T, Dalevi D, Hu P, Andersen GL. 2006. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. Appl Environ Microbiol 72:5069-5072.

52. Faith DP. 1992. Conservation evaluation and phylogenetic diversity. Biol Conserv 61:1-10.

53. Pielou EC. 1966. The measurement of diversity in different types of biological collections. J Theor Biol 13:131-14.

54. Shannon CE. 1948. A Mathematical Theory of Communication. Bell Syst Tech J 27:379-423.

55. McMurdie PJ, Holmes S. 2013. Phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. PLoS One 8:e61217.

56. McArdle BH, Anderson MJ. 2001. Fitting Multivariate Models to Community Data: A Comment on Distance-Based Redundancy Analysis. Ecology 82:290-297.

57. Oksanen J, Guillaume BF, Kindt R, Legendre P, O’Hara RG, Simpson GL, Solymos P, Henry M, Stevens H, Wagner H. 2010. Vegan: community ecology package. R package version 1.17-0.

58. Goslee SC, Urban DL. 2007. The ecodist package for dissimilarity-based analysis of ecological data. J Stat Softw 22:1-19.

59. Ingala MR, Simmons NB, Wultsch C, Krampis K. 2018. Comparing Microbiome Sampling Methods in a Wild Mammal: Fecal and Intestinal Samples Record Different
60. Price SL, Powell S, Kronauer DJC, Tran LAP, Pierce NE, Wayne RK. 2014. Renewed diversification is associated with new ecological opportunity in the Neotropical turtle ants.

J Evol Biol 27:242–258.

61. Paradis E, Schliep K. 2019. Ape 5.0: An environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics 35:526-528.

62. Hammer Ø, Harper DAT, Ryan PD. 2001. Past: Paleontological statistics software package for education and data analysis. Palaeontol Electron 4:9.

63. K. R. Clarke. 1993. Non-parametric multivariate analyses of changes in community structure. Aust J Ecol 18:117-143.

64. Montllor C, Maxmen A, Purcell AH. 2002. Facultative bacterial endosymbionts benefit pea aphids Acyrthosiphon pisum under heat stress. Ecol Entomol 27:189-195.

65. Rio RVM, Wu YN, Filardo G, Aksoy S. 2006. Dynamics of multiple symbiont density regulation during host development: Tsetse fly and its microbial flora. Proc R Soc B Biol Sci 273:805–814.

66. Eleftherianos L, Atri J, Accetta J, Castillo JC. 2013. Endosymbiotic bacteria in insects: Guardians of the immune system? Front Physiol 4:46.

67. Feldhaaar H. 2011. Bacterial symbionts as mediators of ecologically important traits of insect hosts. Ecol Entomol 36:533–543.

68. Weiss B, Aksoy S. 2011. Microbiome influences on insect host vector competence.

Trends Parasitol 27:514–522.

69. Brune A, Dietrich C. 2015. The Gut Microbiota of Termites: Digesting the Diversity in the Light of Ecology and Evolution. Annu Rev Microbiol 69:145–166.
70. Douglas AE. 1998. Aphids and Their Symbiotic Bacteria Buchnera. Annu Rev Entomol 43:17–37.

71. Morales-Jiménez J, Vera-Ponce de León A, García-Domínguez A, Martínez-Romero E, Zúñiga G, Hernández-Rodríguez C. 2013. Nitrogen-Fixing and Uricolytic Bacteria Associated with the Gut of Dendroctonus rhizophagus and Dendroctonus valens (Curculionidae: Scolytinae). Microb Ecol 66:200–210.

72. Moreau CS. 2020. Symbioses among ants and microbes. Curr Opin Insect Sci 39:1–5.

73. Godoy-Vitorino F, Goldfarb KC, Karaoz U, Leal S, Garcia-Amado MA, Hugenholtz P, Tringe SG, Brodie EL, Domínguez-Bello MG. 2012. Comparative analyses of foregut and hindgut bacterial communities in hoatzins and cows. ISME J 6:531–541.

74. Hillman ET, Lu H, Yao T, Nakatsu CH. 2017. Microbial ecology along the gastrointestinal tract. Microbes Environ 32:300–313.

75. Bodawatta KH, Sam K, Jónsson KA, Poulsen M. 2018. Comparative analyses of the digestive tract microbiota of New Guinean Passerine birds. Front Microbiol 9:1830.

76. 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.). Appl Environ Microbiol 78:4691–4701.

77. Ceja-Navarro JA, Nguyen NH, Karaoz U, Gross SR, Herman DJ, Andersen GL, Bruns TD, Pett-Ridge J, Blackwell M, Brodie EL. 2014. Compartmentalized microbial composition, oxygen gradients and nitrogen fixation in the gut of Odontotaenius disjunctus. ISME J 8:6–18.

78. Ceja-Navarro JA, Karaoz U, Bill M, Hao Z, White RA, Arellano A, Ramanculova L,
Filley TR, Berry TD, Conrad ME, Blackwell M, Nicora CD, Kim YM, Reardon PN, Lipton MS, Adkins JN, Pett-Ridge J, Brodie EL. 2019. Gut anatomical properties and microbial functional assembly promote lignocellulose deconstruction and colony subsistence of a wood-feeding beetle. Nat Microbiol 4:864–875.

Nardi JB, Bee CM, Miller LA, Nguyen NH, Suh SO, Blackwell M. 2006. Communities of microbes that inhabit the changing hindgut landscape of a subsocial beetle. Arthropod Struct Dev 35:57–68.

Mikaelyan A, Meuser K, Brune A. 2017. Microenvironmental heterogeneity of gut compartments drives bacterial community structure in wood- and humus-feeding higher termites. FEMS Microbiol Ecol 93:1–11.

Kwong WK, Moran NA. 2016. Gut microbial communities of social bees. Nat Rev Microbiol 14:374–384.

Chen H, Hao D, Wei Z, Wang L, Lin T. 2020. Bacterial Communities Associated with the Pine Wilt Disease Insect Vector Monochamus alternatus (Coleoptera: Cerambycidae) during the Larvae and Pupae Stages. Insects 11:376.

Schauer C, Thompson CL, Brune A. 2012. The bacterial community in the gut of the cockroach Shelfordella lateralis reflects the close evolutionary relatedness of cockroaches and termites. Appl Environ Microbiol 78:2758–2767.

Douglas AE. 2015. Multiorganismal Insects: Diversity and Function of Resident Microorganisms. Annu Rev Entomol 60:17–34.

Martinson VG, Moy J, Moran NA. 2012. Establishment of characteristic gut bacteria during development of the honeybee worker. Appl Environ Microbiol 78:2830–2840.

Lin JY, Russell JA, Sanders JG, Wertz JT. 2016. Cephaloticoccus gen. Nov., a new genus
of ‘Verrucomicrobia’ containing two novel species isolated from Cephalotes ant guts. Int J Syst Evol Microbiol 66:3034–3040.

87. Sanders JG, Łukasik P, Frederickson ME, Russell JA, Koga R, Knight R, Pierce NE. 2017. Dramatic differences in gut bacterial densities correlate with diet and habitat in rainforest ants. Integr Comp Biol 57:705–722.

88. Powell S, Dornhaus A. 2013. Soldier-based defences dynamically track resource availability and quality in ants. Anim Behav 85:157–164.

89. Rubin BER, Kautz S, Wray BD, Moreau CS. 2019. Dietary specialization in mutualistic acacia-ants affects relative abundance but not identity of host-associated bacteria. Mol Ecol 28:900–916.

### Tables

#### Table 1. Pairwise PERMANOVA results of different gut compartments with the Jaccard distance matrix. Results highlighted in bold indicate p values less than 0.05.

| X1      | X2      | Sample size | pseudo-F   | p-value | pval-Bon |
|---------|---------|-------------|------------|---------|----------|
| Crop    | Gaster  | 100         | 45.746554  | 0.001   | 0.011    |
| Crop    | Ileum   | 150         | 128.470497 | 0.001   | 0.011    |
| Crop    | Midgut  | 129         | 152.16059  | 0.001   | 0.011    |
| Crop    | Rectum  | 131         | 90.7966672 | 0.001   | 0.011    |
| Gaster  | Ileum   | 154         | 22.354784  | 0.001   | 0.011    |
| Gaster  | Midgut  | 133         | 43.8965646 | 0.001   | 0.011    |
| Gaster  | Rectum  | 135         | 14.8418312 | 0.001   | 0.011    |
| Ileum   | Midgut  | 183         | 252.145455 | 0.001   | 0.011    |
| Ileum   | Rectum  | 185         | 2.19091684 | 0.034   | 0.34     |
| Midgut  | Rectum  | 164         | 166.291863 | 0.001   | 0.011    |
Table 2. SIMPER analysis table. This includes the top 10 main ASV taxa which contribute to the observed differences in community structure.

| Taxon                      | Av. dissim | % Contribution to Difference | Cumulative % | Mean Read Crop | Mean Read Midgut | Mean Read Ileum | Mean Read Rectum | Mean Read Gaster |
|----------------------------|------------|------------------------------|--------------|----------------|------------------|-----------------|------------------|-----------------|
| Opitutales ASV1            | 18.07      | 20.26                        | 20.26        | 1.23E+03       | 8.90E+03         | 1.40E+03        | 988              | 4.44E+03        |
| Opitutales ASV2            | 5.766      | 6.465                        | 26.72        | 2.12E+03       | 911              | 2.81            | 597              | 345             |
| Rhizobiales ASV1           | 2.954      | 3.313                        | 30.04        | 911            | 2.12E+03         | 281             | 1.02E+03         | 783             |
| Xanthomonadales ASV1       | 2.872      | 3.221                        | 33.26        | 911            | 2.12E+03         | 281             | 988              | 4.44E+03        |
| Alphaproteobacteria ASV1   | 2.628      | 2.946                        | 36.2         | 911            | 2.12E+03         | 281             | 988              | 4.44E+03        |
| Xanthomonadales ASV2       | 2.521      | 2.827                        | 39.03        | 911            | 2.12E+03         | 281             | 988              | 4.44E+03        |
| Rhizobiales ASV2           | 2.247      | 2.52                         | 41.55        | 911            | 2.12E+03         | 281             | 988              | 4.44E+03        |
| Xanthomonadales ASV3       | 1.605      | 1.799                        | 43.35        | 911            | 2.12E+03         | 281             | 988              | 4.44E+03        |
| Rhizobiales ASV3           | 1.438      | 1.613                        | 44.96        | 911            | 2.12E+03         | 281             | 988              | 4.44E+03        |
| Xanthomonadales ASV4       | 1.431      | 1.605                        | 46.57        | 911            | 2.12E+03         | 281             | 988              | 4.44E+03        |
Table 3. Mantel Tests by gut compartment and bacterial lineage tested against host phylogeny from Price et al. 2014. Based on 9999 permutations. Results highlighted in bold indicate p-values less than 0.05. The taxon-specific tests were performed with a subset of the 16S rRNA ASV data by excluding all other bacterial orders.

| Subject            | r statistic | p   |
|--------------------|-------------|-----|
| Overall Samples    | 0.2452      | 0.11188 |
| Crop               | 0.0455      | 0.43822 |
| Midgut             | 0.3963      | 0.053524 |
| Ileum              | **0.3534**  | **0.039836** |
| Rectum             | 0.2485      | 0.037296 |
| Gaster             | **0.547**   | **0.006703** |
| Opitutales         | **0.4017**  | **0.035567** |
| Rhizobiales        | 0.1339      | 0.25759 |
| Xanthomonadales    | 0.1893      | 0.19703 |
| Rickettsiales      | -0.02352    | 0.60958 |
| Enterobacteriales  | -0.1854     | 0.79937 |
| Burkholderiales    | **0.332**   | **0.00325** |
**Figure Legends**

**Figure 1.** Alpha Diversity Metrics by body compartment including (A) mean (±standard error) Shannon diversity, (B) mean (±standard error) Faith’s phylogenetic diversity, (C) mean (±standard error) Pielou’s evenness, (D) mean (±standard error) ASV count. Different letters at the top of the figure illustrate body compartments with significant differences (P<0.001) in this alpha diversity metric.

**Figure 2.** Taxa Bar Plots ordered by gut compartment. (A) Relative Abundance Plot of rarefied dataset excluding *Proxtercerus* samples colored by bacterial order of ASV. Sampling depth of 13814 reads. Each sample was a specific individual gut compartment sampled. (B) Abundance Plot, Raw dataset excluding *Proxtercerus* samples with no rarefaction, colored by bacterial order of ASV. Each sample was a specific individual gut compartment sampled. *Cephalotes* gut image credit: Corrie Moreau

**Figure 3.** Taxa Bar Plot based on Relative Abundance ordered by *Cephalotes* species (with a number and color corresponding to species) and the relative abundance bars are colored by percentage of bacterial order of ASV: (A) Gaster samples (B) Crop samples (C) Midgut samples (D) Ileum samples (E) Rectum samples. Sample names are coded by colony number, gut section, sample replicate number (i.e. 17M3 would be ant colony 17, midgut sample, replicate 3).

**Figure 4.** Mean qPCR values (copies bacterial 16S rRNA gene (rRNA/uL) across all samples by (A) gut compartment (colored by gut compartment) or by (B) gut compartment and species
(colored by species). All gut compartments had significantly different numbers of bacteria than each other, except for the crop and the rectum which were not statistically different. Significance was determined by a mixed effect model with gut compartment as a fixed effect and species as a random effect and Tukey’s HSD test for pairwise comparisons (p < 0.05). Different letters at the top of the specific compartment illustrate body compartments with significant differences.

**Figure 5.** Plots of Principal coordinates analysis (PCoA) with weighted UniFrac distance matrix with Proryptocerus samples excluded. Colors correspond to gut compartment. (A) whole gaster samples are included (B) whole gaster samples are excluded.

**Figure 6.** Principal coordinates analysis (PCoA) with Bray-Curtis dissimilarity distance matrix excluding *Proryptocerus* samples. Cephalotes phylogeny from Price et al. 2014 on the left of the PCoA plots. Colors correspond to Cephalotes species. (A) Gaster samples (B) Crop samples (C) Midgut samples (D) Ileum samples (E) Rectum samples.

**Figure 7.** Illustration of gut compartments within the Cephalotes digestive tract with main bacterial symbionts and N-waste recycling pathways. The crop considered the "social stomach" hosts a diverse collection of transient bacteria at low abundance. The midgut contains primarily Opitutales bacteria which have potential to assist in urea metabolism and amino acid biosynthesis. The uric acid found within the ileum could be degraded into urea with assistance from Burkholderiales bacteria, with this urea then being converted to amino acids with the aid of bacteria from the orders: Opitutales, Xanthomonadales, Rhizobiales, and Burkholderiales. The rectum contains a less abundant though similar suite of bacterial symbionts as the ileum, possibly
aiding in nitrogen recycling and amino acid biosynthesis but at a much smaller scale. (Cephalotes image courtesy of Steven Wang; reproduced with permission.)
Figure 3
A qPCR values across gut compartment

B qPCR values across gut compartment and species

Figure 4
Figure 5
Figure 6
Figure 7

N-waste recycling pathway

Urea → NH₃ → Glutamate → Essential amino acids (AAs) synthesis

Ingested by Cephalotes

Opitutales, Xanthomonadales, Burkholderiales, Rhizobiales

Urea → NH₃ → Glutamate → AAs

“Social Stomach”

Crop → Midgut → Ileum → Rectum → Malpighian Tubules

Urea recycling

Urea → NH₃ → Glutamate → AAs

Opitutales, Xanthomonadales, Burkholderiales, Rhizobiales

Urea → NH₃ → Glutamate → AAs

Ingested by Cephalotes

Opitutales, Xanthomonadales, Burkholderiales, Rhizobiales

Urea → NH₃ → Glutamate → AAs

Downloaded from http://aem.asm.org/ on April 27, 2021 by guest