Treponema Endosymbionts are the Dominant Bacterial Members with Ureolytic Potential in the Gut of the Wood-Feeding Termite, Reticulitermes Hesperus

Jonathan Y Lin (johlin@ucdavis.edu)
University of California Davis https://orcid.org/0000-0003-4977-2506
Laibin Huang
University of California Davis
Sung J Won
University of California Davis
Jorge L.M. Rodrigues
University of California Davis

Research Article

Keywords: Urease, ureC gene, gut microbial communities, termite gut, endosymbiont, Treponema

Posted Date: January 7th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1225749/v1

Treponema endosymbionts are the dominant bacterial members with ureolytic potential in the gut of the wood-feeding termite, *Reticulitermes hesperus*

Jonathan Y. Lin[1], Laibin Huang[1], Sung J. Won[1], and Jorge L.M. Rodrigues[1,2*]

[1]Department of Land, Air, and Water Resources, University of California, Davis, CA, USA 95616
[2]Environmental Genomics and Systems Biology Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA 94720

*Correspondence: jmrodrigues@ucdavis.edu

Running Title: Microbial urease genes in termite guts

Keywords: Urease, *ureC* gene, gut microbial communities, termite gut, endosymbiont, *Treponema*

ORCID
Jonathan Y. Lin: 0000-0003-4977-2506
Laibin Huang: 0000-0003-3885-1815
Jorge L.M. Rodrigues: 0000-0002-6446-6462

Abstract

Termites are remarkable for their ability to digest cellulose from wood as their main energy source, but the extremely low nitrogen (N) content of their diet presents a major challenge for N acquisition. Besides the activity of N_2-fixing bacteria in the gut, the recycling of N from waste products by symbiotic microbes as a complementary N-provisioning mechanism in termites remains poorly understood. In this study, we used a combination of high-throughput amplicon sequencing, quantitative PCR, and cultivation to characterize the microbial community capable of degrading urea, a common waste product, into ammonia in the guts of termites (*Reticulitermes hesperus*) from a wild and laboratory-reared colony. Taxonomic analysis indicated that a majority of the urease (*ureC*) genes in the termite gut (53.0%) matched with a *Treponema* endosymbiont of gut protists previously found in several other termites, suggesting an important contribution to the nutrition of essential cellulolytic protists. Furthermore, analysis of both the 16S rRNA and *ureC* amplicons revealed that the laboratory colony had decreased diversity and altered community composition for both prokaryotic and ureolytic microbial communities in the termite gut. Estimation by quantitative PCR showed that microbial *ureC* genes decreased in abundance in the laboratory-reared colony compared to the wild colony. In addition, most of our cultivated isolates appeared to originate from non-gut environments. Together, our results underscore a more important role for ureolysis by endosymbionts within protists than by free-swimming bacteria in the gut lumen of *R. hesperus*.

Introduction

Nitrogen (N) is an essential element and a major component of nucleic acids and proteins. Although this element is abundant as an atmospheric gas (N_2), the amount of N in bioavailable forms is scarce and can be a limiting factor for growth, reproduction, and survival [1]. N-limitation is frequently a challenge for herbivorous insects, whose plant-based diets often lack sufficient

74 quantities of essential N-metabolites that they cannot produce themselves to build cells and
75 tissues [2, 3]. To meet their nutritional demands, insects have developed various methods to
76 acquire or conserve N, including mechanisms developed from beneficial partnerships with their
77 microbiomes [1, 3]. One symbiont-mediated strategy adopted by insects is microbial N recycling,
78 where bacterial or fungal symbionts located within the gut lumen, hemolymph, or specialized
79 host cells convert waste products excreted by the host into compounds that can be re-absorbed
80 by the insect to support N conservation [1].

81 Symbiotic N recycling has been characterized in several lineages of ants and
82 cockroaches [1]. In carpenter ants (*Camponotus* spp.), *Blochmania* endosymbionts housed
83 within host bacteriocytes are capable of degrading urea ($CO(NH_2)_2$), a common waste product,
84 into ammonia (NH_3) to synthesize essential amino acids for host assimilation [4, 5]. In contrast,
85 arboreal turtle ants (*Cephalotes* spp.) possess a diverse, conserved community of extracellular
86 bacterial symbionts in the gut [6], of which many strains have urease enzyme activity [7] and are
87 capable of converting urea into essential amino acids for the host [8]. Furthermore, the
88 endosymbiont *Blattabacterium*, found in bacteriocytes of cockroaches, also contain genes for
89 urease [9] to degrade and recycle urea converted from uric acid by the host [10]. However, aside
90 from these seminal examples, the extent that symbiotic microbial communities contribute to N
91 recycling in insects remains largely unexplored [1].

92 Termites are social insects descended from wood-feeding cockroaches [11] and have
93 long been studied for their ability to thrive on a diet of lignocellulose, the principal component of
94 woody plant material [12, 13]. Their ability to efficiently digest lignocellulose is driven by essential
95 contributions from deeply evolved, mutualistic symbionts found in their hindguts consisting of
96 archaea and bacteria in the 'higher termites' (Family *Termitidae*) and a tripartite community of
97 free-swimming archaea, bacteria, and cellulolytic protists with their ecto- and endosymbionts in
98 the evolutionarily basal 'lower termites' [12, 13]. Although biological N_2 fixation by bacterial

symbionts is a prominent route of N acquisition in termites [11, 13, 14], the high variation in nitrogen fixation rates between species suggests these insects have devised different methods to manage N limitation [15]. In particular, the high total nitrogen doubling times estimated for some termites such as the *Reticulitermes* make it difficult for them to rely on biological N_2 fixation as the sole process to satisfy all host N requirements [15]. As *Reticulitermes* species typically feed on decaying wood containing fungal growth, the relative enrichment of N from fungi [16] likely minimizes the need for biological N_2 fixation and increases the contribution of N recycling of waste to their total N economy [1, 12].

Termites release most nitrogenous waste as uric acid, which is stored in their fat pads and has been shown to accumulate with laboratory maintenance [17]. A previous study showed that gut bacteria in wood-feeding *Reticulitermes flavipes* termites are capable of recycling N derived from uric acid for re-absorption into host tissue [18]. This process was confirmed in several bacterial strains isolated from termites that can ferment uric acid to produce ammonia [19, 20] and is estimated to provide up to 30% of the total N annually for an average termite colony (comprising from 60,000 up to 1,000,000 individuals) [18]. However, during waste recycling, uric acid can be converted into urea by host or symbiont enzymes [1]. In addition, urea can be excreted as a waste product by protist cells [21] or other symbionts. At this step, whether there are symbionts that can recycle urea N by producing urease enzymes to catalyze the breakdown of urea to ammonia for assimilation in the termite gut is unknown. Furthermore, whether prolonged laboratory maintenance and the associated buildup of uric acid in termites will affect the dynamics of ureolytic microbes is unclear. Thus, this represents a knowledge gap in termite-microbial symbiosis and nutrient cycling processes important for host insect fitness. In this study, we aimed to 1) uncover the taxonomic diversity and abundance of symbionts with ureolytic potential inhabiting the termite gut, and to 2) compare microbial community dynamics between a colony of wild and laboratory-reared termites. We focused on the lower termites as

they are the most well-studied and hypothesized that wild and laboratory-reared populations of termites will harbor distinct prokaryotic and ureolytic microbial communities in their guts.

Methods

Termite Collection and Maintenance

Termites (*Reticulitermes hesperus*) were collected from the University of California Davis (UC Davis) Putah Creek Riparian Reserve (38.524° N 121.783° W) in March 2017. Approximately 200-300 termites, including members from worker, soldier, and reproductive castes were collected from a fallen log (colony TH1) and reared in the laboratory for seven months as described previously [22]. Briefly, termites were maintained in plastic boxes with autoclaved sand and fed with oak wood collected from the same area. Boxes were kept at room temperature, ventilated, and periodically remoistened with sterile distilled water. Wild termites (colony WTH1) belonging to the same species (see termite identification below) were collected at the UC Davis Stebbins Cold Canyon Reserve (38.507° N 122.097° W) in November 2017 and held for less than 36 hours before immediate degutting. Only termites from the worker caste were used for experiments.

Gut Dissection, Library Preparation, and Amplicon Sequencing

Hindguts (11-12 per group) were removed by first surface-washing each worker in sterile urea isolation broth (UIB, per liter: 5 g of NaCl, 2 g of peptone, 9.5 g of K_2HPO_4, 9.1 g of KH_2PO_4, 1 mL of 1000x trace elements solution (**Table S4**), and 10 g of urea added aseptically as a solution after cooling) before pulling the thorax and anus apart using sterile forceps [23]. For DNA extraction, single, whole hindguts were placed into bead-beating tubes containing Powerbead solution and solution C1 from the DNeasy Powersoil DNA extraction kit (Qiagen, Germantown, MD, USA) and stored at -20°C until extraction. DNA was extracted using a vortex adaptor

149 according to the manufacturer's instructions and was quantified using the Qubit dsDNA HS
150 assay kit (Life Technologies, Carlsbad, CA, USA). To characterize the prokaryotic community,
151 the V4 hypervariable region of the 16S rRNA gene was amplified from each sample in triplicate
152 using the primer pair 515F/806R [24], which was designed to include Illumina adaptors and 12
153 bp barcode sequences. All primer sequences and PCR cycling conditions are listed in **Table S1**.
154 Amplification were carried out in 20 µL volume reactions containing 10 µL Phusion Hot-Start II
155 High-Fidelity Master Mix (ThermoFisher, Waltham, MA, USA), 0.5 µM each primer, 10 ng sample
156 DNA, and 4 µL sterile ddH_2O. The resulting amplicons were inspected by gel electrophoresis on
157 a 1% agarose gel, pooled in equimolar concentrations, and sequenced (paired-end 250 bp) on
158 a MiSeq platform (Illumina, San Diego, CA, USA) at the UC Davis DNA Technologies core facility
159 using the custom sequencing and indexing primers described in Caporaso et. al [24].
160 To characterize the ureolytic microbial community, the *ureC* gene, which encodes for the
161 catalytic subunit of urease containing several conserved regions, was amplified from each
162 sample in triplicate using modified versions of the primers UreC-F and UreC-R [25, 26]. A unique
163 12 bp barcode sequence was added to the reverse primer for each sample (**Table S2**). The
164 resulting amplicons were inspected by gel electrophoresis as above, pooled in equimolar
165 concentrations, and sequenced on a MiSeq platform as above using standard sequencing
166 primers following enzymatic ligation of Illumina sequencing adaptors.
167
168 ***Termite Identification***
169 Termite heads removed simultaneously during degutting were pooled (15-20 heads per
170 extraction) and DNA was extracted from wild and laboratory-reared workers using the DNeasy
171 Powersoil Kit as described above. Purified DNA was used for amplification and sequencing of
172 the mitochondrial cytochrome oxidase II (*COII*) gene using the primers A-tLEU and B-tLYS [27].
173 PCR products were then purified using the Ultraclean PCR clean up kit (MO-BIO Laboratories,

Carlsbad, CA, USA) and sent to the UC Davis College of Biological Sciences DNA sequencing facility (Davis, CA, USA) for Sanger sequencing using an ABI 3730 platform (Applied Biosystems, Foster City, CA, USA) to ensure the termites were *R. hesperus* (**Figure S1**).

Quantitative PCR

To determine the abundance of the prokaryotic and ureolytic microbial community in termite guts, quantitative PCR (qPCR) was performed on each DNA sample using non-barcoded versions of the degenerate primers 515F and 806R for the 16S rRNA gene [28] and UreC-F and UreC-R for the *ureC* gene [25]. qPCR for each target gene was performed in a 20 µL reaction mixture containing 10 µL SsoAdvanced Universal SYBR Green Supermix (Bio-Rad Laboratories, Hercules, CA, USA), 0.5 µM each primer, 10 ng template DNA, and 4 µL sterile ddH_2O. Reactions were performed on a Bio-Rad CFX Connect System (Bio-Rad Laboratories, Hercules, CA, USA) and amplification of the 16S rRNA gene consisted of an initial denaturation of 95°C for 3 min, followed by 39 cycles of 95°C for 10 s and 60°C for 30 s. Amplification of the *ureC* gene consisted of an initial denaturation of 95°C for 3 min, followed by 39 cycles of 95°C for 10 s and 52°C for 30 s. Quantification was performed by comparing the C_t values of unknown samples to a standard curve (with a detection range of 10^1-10^9 copies) generated with the pCR Blunt II-TOPO vector (Invitrogen, Carlsbad, CA, USA) containing a PCR-amplified fragment of each target. Coefficient of determination (R^2) values and amplification efficiency percentages for the standard curves were 0.993 and 108.8% for the 16S rRNA gene and 0.965 and 91.3% for *ureC* gene, respectively. Triplicate reactions were performed for each gene per sample and a melting curve analysis was performed after each assay to ensure specificity of the amplified products.

Sequence Data Processing

198 Raw reads for the 16S rRNA gene sequences were processed using DADA2 v.1.6 [29]
199 implemented in R v.3.4.4. Briefly, paired-end fastq files were processed by quality-trimming
200 forward and reverse reads to 200 and 150 bp lengths, respectively. After sequence dereplication,
201 merging, error correction, and chimera removal, Exact Sequence Variants (ESVs) were inferred
202 and taxonomic identification was performed using the SILVA 16S database v.128.
203 Raw reads for the *ureC* gene sequences were first quality checked using FastQC [30]
204 before paired-end reads were merged with FLASH [31] using default parameters. Merged
205 sequences were demultiplexed in QIIME v.1.9 [32] using the 'split_libraries_fastq.py' and
206 'split_sequence_file_on_sample_ids.py' scripts. Then, the forward and reverse primer
207 sequences were trimmed from each file using BBDuk [33] before chimera detection and removal
208 with the 'identify_chimeric_sequences.py' and 'filter_fasta.py' scripts in QIIME using USEARCH
209 v.6.1. To identify the taxonomy of representative sequences, microbial *ureC* gene sequences
210 were downloaded from the FunGene repository [34]. In addition, *ureC* gene sequences from
211 bacteria originating from the termite gut were downloaded from NCBI and both sequence
212 datasets were compiled into a custom gene package using the '-create' command implemented
213 in GraftM [35] for protein sequence alignment, hidden Markov model (HMM) construction, and
214 phylogenetic tree building. Taxonomy was assigned by using the '-graft' command against the
215 compiled *ureC* gene package in GraftM, which places query sequences onto the *ureC* reference
216 tree with pplacer [36] using a default likelihood cut off value of 0.75. Representative sequences
217 for both genes were aligned using MAFFT [37] and maximum-likelihood trees were constructed
218 using FastTree [38] with default parameters. Phylogenetic trees were visualized using the
219 'phyloseq' and 'ggtree' packages in R [39, 40].
220
221 ***Cultivation and Screening Procedures***

To confirm whether *ureC* gene sequences obtained from termites originated from viable microbial symbionts potentially capable of degrading urea, we cultivated bacteria from hindgut samples and screened for the presence of the *ureC* gene. Termite hindguts (10-15) were removed from each colony and homogenized in sterile UIB using a pestle. Afterwards, the homogenate was serially diluted and spread onto plates containing Urea Isolation Agar (UIA), which contains the same components as UIB but with the addition of agar (15 g per L). Since the termite hindgut is spatially stratified with respect to oxygen concentration [41], replicate extraction and plating procedures were performed under an atmospheric O_2 concentration (20.9% at 1 atm) and inside a 2% O_2 atmosphere-controlled glove box fitted with an oxygen sensor and automated controller (Coy Labs, Grass Lake, MI, USA) using a gas mixture of 5% CO_2 and 95% N_2 at room temperature for approximately 1 month. Individual colonies were streaked onto fresh UIA plates as they appeared and reinoculated onto plates three subsequent times to confirm isolation. Genomic DNA was extracted from cells by using a QIAamp DNA mini kit (Qiagen, Germantown, MD, USA) according to the manufacturer's instructions and used for repetitive element palindromic (rep) PCR fingerprinting analysis to identify unique isolates [42]. Briefly, approximately 10-20 ng of DNA from each isolate was used for PCR using the BOXA1R primer [42] and the resulting fingerprint patterns were inspected on a 1% agarose gel. For each unique isolate, the 16S rRNA gene was amplified using the universal primers 63F and 1389R [43]. The PCR products were purified using ExoSAP-IT (USB Corporation, Cleveland, OH, USA) and Sanger sequencing was performed at the UC Berkeley DNA sequencing facility using BigDye terminator chemistry. Resulting sequences were quality-trimmed using 4Peaks v.1.8 (Nucleobytes, Amsterdam, NL) and BLAST searches in the NCBI database [44] were performed to identify the closest matches to our isolates. To determine whether the strains possessed the genomic potential for ureolysis, PCR using the UreC-F and UreC-R primers was performed on DNA extracted from each isolate as described above and visualized by gel electrophoresis to

247 determine the presence of a *ureC* gene fragment. DNA from *Proteus vulgaris* strain ATCC 6380
248 (Microbiologics, St. Cloud, MN, USA) and *Escherichia coli* strain BL21 were used as the positive
249 and negative controls, respectively. All strains were archived by suspending cultures in UIB
250 supplemented with 20% (v/v) glycerol and storing them in cryovials at -80°C.
251
252 ***Statistical Analysis***
253 All data were first tested for assumptions of normality and homogeneity of variance
254 before comparison using a one-way Analysis of Variance (ANOVA) test and a Tukey Honestly
255 Significant Difference post-hoc test to identify significant differences between wild and
256 laboratory-reared termites. Data that failed to meet assumptions for ANOVA were compared by
257 using a nonparametric Kruskal-Wallis test followed by a Dunn's post-hoc test.
258 Bray-Curtis distance matrices were calculated for the 16S rRNA and *ureC* gene datasets
259 for Non-Metric Multidimensional Scaling (NMDS) analysis to determine overall differences in the
260 composition (presence and relative abundance) of prokaryotic and ureolytic microbial
261 communities between wild and laboratory-reared termites. Permutational Multivariate Analysis
262 of Variance (PERMANOVA) tests were performed using 999 permutations to assess differences
263 in community composition using the 'adonis' function in the 'vegan' package [45]. Within-group
264 dispersions were calculated and their differences between termite colonies were tested using
265 the 'betadisper' and 'permutest' functions [45]. DESeq2 [46] was used to identify significantly
266 enriched taxa between colonies by estimating standard errors, transforming read counts to
267 obtain log fold change, performing Wald tests, and applying Benjamini-Hochberg corrections on
268 reported *P*-values. Differentially abundant taxa were visualized using heatmaps produced with
269 the 'pheatmap' package [47]. All statistical analyses were performed using R and a value of $P <$
270 0.05 was considered statistically significant.
271

Data Availability

The 16S rRNA and *ureC* gene sequencing data were deposited to the NCBI sequence read archive (SRA) under BioProject PRJNA660442 with the following BioSample accession numbers: SAMN15949798 (16S rRNA) and SAMN15949813 (*ureC*). The custom GraftM gene package used for taxonomic classification of *ureC* gene sequences is available at https://github.com/jonathanylin/Termite_gut_urease.

Results

Sequencing results and normalization

After quality filtering, error correction, and chimera removal, the total number of paired-end reads was 675,612 for the 16S rRNA gene and 417,234 for the *ureC* gene across 23 samples. The number of 16S rRNA gene sequences per sample ranged from 16,533 to 41,428 with an average of 29,374, whereas the number of *ureC* gene sequences per sample ranged from 11,308 to 32,520 with an average of 20,749. The 16S rRNA and *ureC* gene datasets were rarefied to 16,000 and 11,000 sequences per sample, respectively, to ensure equal depth for statistical analysis. After rarefaction, the total number of unique phylotypes was 1,358 and 286 for the 16S rRNA and *ureC* genes, respectively (**Figure 1E & 2E**). Overall, the termite gut classified by the 16S rRNA gene was dominated by the phyla *Spirochaetes* (34.2%), *Proteobacteria* (20.9%), *Elusimicrobia* (14.9%), *Firmicutes* (15.1%), *Bacteroidetes* (8.9%), *Planctomycetes* (2.0%), and *Actinobacteria* (1.4%), with *Synergistetes* (0.8%), *Verrucomicrobia* (0.6%), and *Euryarcheota* (0.4%) comprising the main phyla at under 1%. (**Figure 1F**). The most abundant ESVs were from the genera *Treponema* (32.3%), Candidatus *Endomicrobium* (19.5%), *Treponema* 2 (10.4%), *Desulfovibrio* (6.2%), and Candidatus *Armantifilum* (4.5%) (**Figure S2**). Across all samples, the taxonomic composition of *ureC* gene comprised the phyla *Proteobacteria* (36.7%), *Spirochaetes*

297 (27.4%), and *Firmicutes* (21.7%), with most of the remaining sequences not yielding any
298 classification (12.0%) (**Figure 2F**). At the genus level, the most abundant classified phylotype
299 based on the *ureC* gene was from the genus *Treponema* (53.0%), followed by *Pseudomonas*
300 (8.3%) and *Bacillus* (5.7%) (**Figure S3**).
301
302 **Differences in 16S rRNA and *ureC* gut microbial community**
303 Prokaryotic alpha diversity was significantly different between members of the two
304 colonies, with a higher average number of ESVs observed in termites from colony WTH1
305 compared to termites from colony TH1 (**Figure 1A**, ANOVA, $F = 12.2$, $P = 0.002$). Similar results
306 were observed by using the Shannon diversity (**Figure 1B**, Kruskal-Wallis, $\chi^2 = 9.1$, $P = 0.001$)
307 and Shannon evenness indexes (**Figure 1C**, Kruskal-Wallis, $\chi^2 = 6.7$, $P = 0.005$). Termites from
308 both colonies had a substantially different microbial community composition based on the 16S
309 rRNA gene (PERMANOVA, $R^2 = 0.395$, $F = 13.7$, $P = 0.001$). These results are reflected in the
310 clear separation of microbial communities by colony type observed in the NMDS ordination
311 (**Figure 1D**). A permutational test of the multivariate homogeneity of variances revealed no
312 differences in dispersion between groups (BETADISPER, $F = 0.18$, $P = 0.71$). Colony WTH1 had
313 a significantly higher number of unique ESVs than colony TH1, while both shared a core
314 community of 423 ESVs (**Figure 1E**).
315 Similar to the 16S rRNA gene results, the *ureC* gene richness was significantly decreased
316 in the laboratory-reared colony compared to the wild colony (**Figure 2A**, ANOVA, $F = 10.8$, $P =$
317 0.0036), and these results were consistent when our analyses were performed using two
318 additional non-parametric measurements of diversity: the Shannon diversity (**Figure 2B**, ANOVA,
319 $F = 18.1$, $P = 0.00036$) and Shannon evenness index (**Figure 2C**, ANOVA, $F = 13.7$, $P = 0.001$).
320 Classification of the *ureC* gene also revealed a distinct community composition of ureolytic

microbes between colonies (PERMANOVA, R^2 = 0.270, F = 7.8, P = 0.001) with separation apparent in the NMDS ordination (**Figure 2D**). No differences in *ureC* gene dispersion between groups were detected (BETADISPER, F = 0.024, P = 0.91). Termites from both colonies shared a core community of 161 *ureC* species, while termites from colony WTH1 had a higher number of unique *ureC* species compared to termites from colony TH1 (**Figure 2E**).

Differences in relative abundance of prokaryotic and ureolytic taxa between colonies

Based on the 16S rRNA gene, the relative abundance of the phyla *Spirochaetes* (38.0% vs 30.4%), *Proteobacteria* (24.6% vs 17.1%), *Bacteroidetes* (9.8% vs 8.0%), and *Actinobacteria* (1.7% vs 0.99%) were higher in termites from colony WTH1 compared to the termites from colony TH1 (**Figure 1F**). The relative abundance of the *Elusimicrobia* and *Planctomycetes* increased from 8.6% to 21.2% and 0.6% to 4.9%, respectively, in the termites from colony TH1 compared to termites from colony WTH1 (**Figure 1F**). Differentially abundant taxa between both colonies that were significant at the order level were illustrated using a heatmap and by calculating $\log_2$-transformed abundances (**Figure 1G**).

Consistent with the 16S rRNA gene results, the relative abundance of *ureC* genes classified as *Proteobacteria* decreased from 46.5% in colony WTH1 to 27.0% in colony TH1 (**Figure 2F**). The proportion of *ureC* genes that could not be assigned to any known phylum also decreased from 13.2% in colony WTH1 to 10.8% in colony TH1 (**Figure 2F**). In contrast, the relative abundance of *ureC* gene sequences belonging to the phyla *Spirochaetes* and *Firmicutes* increased from 22.4% and 15.3% in the termites from colony WTH1 to 32.4% and 28.0% in termites from colony TH1, respectively (**Figure 2F**). Significant responders identified at the order level corroborated trends observed at the phyla level, with *ureC* genes classified as

Pseudomonadales, *Campylobacterales*, and *Burkholderiales* decreased and the *Spirochaetales*, *Clostridiales*, and *Bacillales* increased in colony TH1 compared to colony WTH1 (**Figure 2G**).

Notably, all *ureC* gene sequences annotated as *Spirochaetes* were mapped to a single phylotype, Urec_98, which was classified as a species within the genus *Treponema*. Urec_98 was the single most abundant phylotype among all *ureC* genes at the genus level for both colonies (**Figure S3**). Additional phylogenetic analysis placed Urec_98 in a clade of *ureC* genes from a novel *Treponema* species previously identified as an endosymbiont of *Eucomonympha* protists in the termite gut [48] (**Figure S4**). Urec_98 shared a 96.61% identity in the protein coding sequence with the *ureC* genes from these endosymbionts, and this sequence was noticeably distinct from the *ureC* gene of *Treponema bryantii* and *Treponema ruminis* as well as from other bacteria for which data are available (**Figure S4**). These results indicate that Urec_98 is evolutionarily distant from other *Treponema* species with ureolytic potential and likely an endosymbiont of protists in the termite gut.

16S rRNA and *ureC* gene abundance

16S rRNA and *ureC* gene copy numbers were quantified by qPCR as a proxy for absolute abundance. The *ureC* gene copy number was higher in the hindguts of termites from colony WTH1 than in termites from colony TH1 (**Figure 3A**, ANOVA, $F = 21.95$, $P = 0.0001$). In contrast, the 16S rRNA gene copy number did not differ between colonies (**Figure 3B**, Kruskal-Wallis, $\chi^2 = 0.034$, $P = 0.43$). Consequently, the overall proportion of *ureC* gene copies, calculated as the ratio of *ureC* gene copies to total 16S rRNA gene copies, decreased from an average of 2.1% in colony WTH1 to 0.77% of the total microbial community in colony TH1 (**Figure 3C**, Kruskal-Wallis, $\chi^2 = 16.0$, $P < 0.0001$).

368 **Cultivation analyses**

369 A total of 192 isolates were retrieved from our cultivation procedures, with 86 isolates
370 cultivated from colony WTH1 and 106 isolates cultivated from colony TH1 (**Table S4**). Overall,
371 the majority of identified strains were from the phyla *Proteobacteria* (148 isolates), followed by
372 the *Bacteroidetes* (11 isolates) and *Firmicutes* (6 isolates) (**Figure S5A**). At the genus level, most
373 strains cultured from colony WTH1 were identified as *Pseudomonas*, *Acinetobacter*, and
374 *Chryseobacterium* species, while a substantial number of strains cultivated from termites in
375 colony TH1 were identified as *Citrobacter* species (**Figure S5B**). Many of the identified isolates
376 were also detected from amongst the most abundant *ureC* gene amplicon sequences (**Figure**
377 **S3**). The termites from colony WTH1 had a higher proportion of cultivated isolates with genomic
378 potential for ureolysis than termites from colony TH1 based on diagnostic PCR of the *ureC* gene
379 (**Figure S5C**).
380
381 **Discussion**

382 Conditions associated with laboratory-rearing can affect insect physiology and their gut
383 microbiota. However, the impact of prolonged laboratory maintenance on the gut microbiome
384 has rarely been documented in termites. One previous study reported significant differences in
385 gut bacterial composition between field-collected and laboratory-reared *Coptotermes*
386 *formosanus* termites using cultivation-based methods [49], and another observed a gradual
387 decrease in alpha diversity across several *R. flavipes* colonies [27]. In this study, we showed that
388 wild and laboratory-reared colonies of *Reticulitermes hesperus* workers have substantially
389 distinct gut microbial communities. Compared to termites retrieved from colony WTH1, the
390 laboratory-reared termites from colony TH1 had significantly decreased alpha diversity,
391 evenness, and significantly altered community compositions for both the 16S rRNA and *ureC*

genes (**Figures 1 & 2**). Together, these results show that common laboratory maintenance has the potential to affect both the prokaryotic and potential ureolytic microbial community over the course of 7 months. It is important to highlight that while our findings match general trends of decreased diversity and altered microbial community composition with laboratory maintenance observed in other insects [50–52], the mechanisms driving these changes may vary considerably based on host physiology, behavior, and rearing strategies. Recently, Tasaki et. al [53] reported that reactive oxygen species (ROS) are important sources of stress for termites during laboratory maintenance, demonstrating that the protective effects against ROS by endogenous host enzymes became decreased in *R. speratus* workers after 5 weeks in the laboratory. While ROS are generated during normal metabolic activities, their concentrations can be exacerbated by UV irradiation and high oxygen levels leading to damaged DNA and proteins. Intense UV irradiation leads to phototoxicity in termites [54]; hyperoxia can also kill termites [55] and stimulate oxygen-guarding behaviors, presumably to protect oxygen-sensitive microbial symbionts [56]. Although not lethal, the light and oxygen levels experienced by *Reticulitermes* workers under laboratory conditions are likely comparatively higher than in their natural subterranean habitats where colonies can establish down to 5-6 meters [57]. We surmise that consistent, long-term exposure to increased light and oxygen under normal laboratory conditions eventually overwhelms host and microbial protective mechanisms to oxidative stress, thereby inducing shifts in the gut microbial community. However, the relative importance of these factors, particularly at intensities relevant to laboratory environments on termite fitness and their gut microbial community has not yet been determined.

Termites void most nitrogenous waste as uric acid, which is synthesized and stored in the fat bodies [17]. It has been shown that uric acid levels in termites increase with laboratory maintenance, accumulating up to 45.4% of termite dry weight after 15 months [17]. Uric acid synthesis increases immediately following laboratory storage [58] and its deposition in fat tissue

causes an intense white coloration visible on laboratory-reared termites [19, 58]. Uric acid can be broken down via the uricolytic pathway where uric acid is transformed into several forms, including urea, before conversion into ammonia by microbial ureases for re-assimilation [1]. Observing that tissue from *R. flavipes* workers lacked uricase enzymes, Potrikus and Breznak [18] performed isotope tracer experiments to confirm that uric acid-degrading gut bacteria contribute to N-recycling and provision uric acid N for host re-absorption. The authors proposed that uric acid is transported by the termite into the gut through the Malpighian tubules to stimulate microbial N-recycling activity [18]. However, the authors did not determine whether strains possessing urease genes were involved in this N-recycling pathway. In this study, we asked whether bacteria with the potential to degrade urea originating from either uric acid or other waste products appear to play an important role to the N-recycling economy. Furthermore, we reasoned that increased uric acid levels with laboratory maintenance may reflect an increase in its mobilization and availability, potentially increasing downstream urea concentrations and altering the ureolytic community. Although we detected ureolytic bacteria in the gut (**Figure 2**) and isolated bacteria with ureolytic potential (**Figure S5**), the overall abundance of *ureC* genes in the termite gut was relatively low, averaging 835 copies per ng gut DNA in all samples and constituting 0.4% to 2.9% of the total prokaryotic community estimated by qPCR (**Figure 3A & C**). These results demonstrate that ureolytic potential is present in the gut, albeit likely representing a minor contribution to waste N-recycling. In addition, our finding of decreased *ureC* gene abundance (**Figure 3**) and reduced *ureC* gene diversity (**Figure 2A**) in the laboratory-reared termites in colony TH1 may indicate that uric acid mobilization is not increased with laboratory-maintenance or not converted to urea by uricolytic gut microbes. To date, there has been no direct evidence that uric acid is mobilized from termite fat body stores [1, 59]. Chappell and Slaytor [58] proposed that an elevated proportion of uric acid over the total N content in older laboratory termites indicates that no mobilization occurs *in situ*. Instead, necrophagy by

other colony members has been hypothesized as a mechanism for uric acid delivery for N-recycling by gut symbionts, but this has not been further explored [1, 58]. Several uric acid-degrading bacteria have been cultivated from termite guts and characterized [19, 20]. In one instance, urea was not produced as an intermediate by a *Streptococcus* strain during uric acid degradation [19]. In addition, several of our ureolytic isolates identified as *Citrobacter farmeri* (**Table S4**) appeared to be similar to a few uric acid degrading *Citrobacter farmeri* strains isolated previously [20]. This suggests that uricolytic bacteria may directly assimilate uric acid N or are also capable of simultaneously using urea *in situ*, and hence may not contribute to the urea pool available for use by other bacteria in the gut lumen. However, whether there are other strains capable of producing urea and other intermediates from uric acid in the gut requires further study.

The *Proteobacteria* and *Firmicutes* constituted 58.4% of *ureC* gene sequences in all samples at the phylum level (**Figure 2D**). *Pseudomonas* and *Bacillus* species were identified as the most abundant taxa at the genus level within these groups, respectively (**Figure 2E & S3**). These bacteria are not known to be abundant in the gut [27] or part of the core microbiome of lower termites [13]. Indeed, in our 16S dataset, *Pseudomonas* species only represented 2.6% of all ESVs, whereas *Bacillus* species were detected at less than 0.01%. Furthermore, the 16S rRNA gene sequences from our *Pseudomonas* and *Bacillus* isolates matched closely with strains previously cultivated from soil and plant-associated habitats (**Table S4**), suggesting that they may be of non-gut origin. Transient microbes from soil, plant, or other sources may be ingested with food and pass through the termite gut with limited host interactions and are expected to be detected at low frequencies [60]. Our detection of bacteria that are putatively from non-gut sources, paired with our findings of low *ureC* gene copy numbers suggests that many resident microbes in the termite gut do not possess urease genes in their genomes. By contrast, *ureC* genes are in soil are dominated by members of the *Proteobacteria* and *Firmicutes* [61], which

likely explains the *Bacillus* and *Pseudomonas* species we found at high proportions in our *ureC* dataset but detected at low proportions in our 16S rRNA gene dataset. Overall, these results indicate that a majority of the ureolytic microorganisms in *R. hesperus* represent a small proportion of the total gut prokaryotic community, and that many of these abundant *ureC* genes appear to originate from bacteria that are transient rather than co-evolved, core members of the gut microbial community.

We observed that a single partial gene sequence, classified as Urec_98 and identified as belonging to the genus *Treponema*, was the most abundant in the *ureC* dataset, representing 53.0% of *ureC* phylotypes at the genus level across all samples (**Figure S3**). Further analysis showed that Urec_98 is closely related to "*Candidatus* Treponema intracellularis," (**Figure S4**), an endosymbiont of *Eucomonympha* protists in wood-feeding *Hodotermopsis sjoestedti* termites [48]. The genome of "*Candidatus* T. intracellularis" contains genes encoding for urease as well as a membrane-bound urea channel, indicating its ability to both transport and use urea excreted by its host [48]. "*Candidatus* T. intracellularis" falls within the termite *Treponema* cluster II [48], a defined clade of *Treponem*a ectosymbionts attached to the cell surface of termite gut protists [62, 63]. This clade, along with a group of free-swimming termite *Treponema* species (cluster I) comprise an abundant and highly co-evolved community of *Spirochaetes* within the termite gut [64]. In our 16S rRNA gene dataset, we found that both *Treponema* clusters represented 42.8% of all ESVs and together their relative abundances did not differ significantly between termites from either colony (**Figure 1D**). Thus, in contrast to the *Pseudomonas* and *Bacillus* species identified in the *ureC* gene sequences, Urec_98 is likely a member of the autochthonous gut microbial community in termites.

To date, the only other members of the *Treponema* for which genomes are available on NCBI that possess genes for urease and urea transporters are *T. bryantii* and *T. ruminis*, two *Spirochaetes* originally isolated from the bovine rumen [65, 66], an environment where urea

492 enters from the bloodstream as a major source of waste N in ruminant animals [67]. By
493 comparison, the genomes of *T. primitia* and *T. azonutricium* [68], two free-swimming species
494 (representing termite *Treponema* cluster I) isolated from the termite gut [69], do not contain genes
495 encoding for any urease subunits or urea transporters. This suggests that unlike the rumen
496 environment, the termite gut lumen likely does not have a significant flux of urea which may
497 underscore a lack of selective pressure for free-swimming *Spirochaetes* to possess urease
498 genes. Thus, it appears that the endosymbiotic *Treponema* are the only known *Treponema*
499 species to date in the termite gut that contain urease genes. Ohkuma and colleagues [48] found
500 that a close relative of *Eucomonympha* protists, *Teranympha mirabilis*, also contains
501 endosymbionts related to "*Candidatus* T. intracellularis" in the guts of *R. speratus* termites.
502 Further phylogenetic work revealed that the *Treponema* endosymbiont of *T. mirabilis* is a different
503 species from "*Candidatus* T. intracellularis" in *Eucomonympha*, indicating strong cospeciating
504 relationships between the protist host and endosymbiont [70]. Our study provides evidence that
505 *Treponema ureC* genes from *R. hesperus* termites share high homology with *ureC* genes from
506 "*Candidatus* T. intracellularis" (**Figure S4**), suggesting that ureolysis by *Treponema*
507 endosymbionts within protist hosts is persistent across several lineages of wood-feeding
508 termites.

509 Besides the *Treponema* endosymbionts, several extracellular bacteria previously isolated
510 from termite guts have been shown to have urease enzyme activity, such as *Comamonas*
511 *odontotermitis* [71]; or encode operons for urease and their transporters in their genomes
512 including a *Citrobacter* strain [72], *Sporomusa termitida* [73], *Stenoxybacter acetivorans* [74], and
513 several *Verrucomicrobia* strains [75–79]. Yet, owing to their low abundance relative to the total
514 prokaryotic community and the fact that we did not detect any of these taxa at proportions
515 greater than 0.5% in our *ureC* gene dataset suggests that ureolysis in the gut lumen by free-
516 swimming bacteria likely does not produce a significant quantity of recycled N for *Reticulitermes*

termites. By contrast, we reasoned that our detection of *Treponema* endosymbionts at much higher proportions from our *ureC* sequences suggests a more important role for urea recycling inside protists. This is substantiated by a previous finding of another termite endosymbiont, "*Candidatus* Azobacteroides pseudotrichonymphae," a *Bacteroidales* strain that, like "*Candidatus* T. intracellularis," also possesses a gene cluster encoding a urease and urea transporter [80]. This endosymbiont was found to be abundant in *Pseudotrichonympha* protists, a sister lineage to the *Eucomonympha-Teranympha* protists, which suggests that phylogenetically diverse bacterial species may have convergently established similar functional niches for N recycling within protist hosts [48]. In addition to possessing a complete operon for urease and its transporter, both *"Candidatus* T. intracellularis" and "*Candidatus* A. pseudotrichonymphae" have genes for nitrogen fixation, implying that depending on conditions within the host cytoplasm, these endosymbionts can transport and recycle external urea or fix N_2 for biosynthesis of nitrogenous compounds to benefit the protist [80, 81]. This degree of versatility for N metabolism is expected to enable the protist to grow efficiently and remain stable during nutrient fluctuations in the gut, thereby allowing the termite to maintain cellulolytic protists essential for host nutrition [48].

It is important to note that as with any study, our work contains caveats resulting from our methods. First, approximately 12% of all *ureC* gene sequences we obtained could not be assigned to any known phylum, indicating that the termite gut may contain several hitherto unknown urease genes that require further characterization. Second, our cultivation efforts focused solely on culturing bacteria under oxic (~21% O_2) and hypoxic (2% O_2) conditions. Despite persistent oxygen flux in the hindgut (34), many termite symbionts are anaerobic with low tolerances for oxygen [13]. Thus, whether there is a significant component of viable ureolytic bacteria occupying anoxic niches in the termite gut warrants future investigation. Finally, the two colonies of termites we collected in this study were genetically distinct (**Figure S1**) and were not

542 derived from the same geographic population. Hence, the differences in the gut microbial
543 communities we observed between these two colonies cannot be attributed solely to the effect
544 of laboratory maintenance. Hongoh and colleagues [82] found that although *Reticulitermes*
545 termites from across Japan shared a core gut bacterial microbiome, differences in taxonomic
546 composition were found between sampling sites and among different termite species. Similarly,
547 a characterization of *R. flavipes* termites collected throughout Connecticut and Massachusetts
548 revealed the presence of a core gut microbiome with site-specific differences, possibly due to
549 fluctuations in low abundance taxa between samples [27]. At the genus level, we observed
550 similar fluctuations in the low abundance taxa between colonies from both our 16S rRNA and
551 *ureC* gene datasets (**Figures S2 & S3**). These differences could be due the genetic divergence
552 between hosts or from site-specific diets prior to our collection and laboratory rearing. While it
553 is likely that laboratory maintenance contributed to the differences in gut microbial communities
554 due to our observations of 1) large effects with little overlap between both colonies and 2)
555 decreased richness and evenness in the laboratory-reared colony consistent with a previous
556 report [27], the specific effects of laboratory-rearing compared to colony specificity cannot be
557 resolved and are beyond the scope of this study. Therefore, future studies are required to clarify
558 the impact of laboratory rearing on termite gut microbial communities.
559
560 **Conclusion**
561 In summary, we investigated the prokaryotic and ureolytic microbial communities in the
562 hindguts of *R. hesperus* workers from two colonies. We found that termites from colony TH1,
563 which was reared in the laboratory for 7 months, had a shifted community composition and
564 decreased diversity for both the 16S rRNA and *ureC* genes in the gut. Quantification with qPCR
565 revealed that *ureC* genes represented a relatively low proportion of the overall gut prokaryotic
566 community, and we documented a decrease in *ureC* gene abundance in colony TH1 compared

to colony WTH1. Our taxonomic characterization of *ureC* gene sequences showed a significant number of *ureC* genes matching to a *Treponema* endosymbiont of gut protists previously found in several other termites, suggesting that urea-recycling within protist hosts may be conserved across the lower termites. Together with our cultivation findings, where a majority of our isolates appeared to originate from non-gut environments, our results highlight a more important role for ureolysis by endosymbionts within protists than by free-swimming bacteria in the gut lumen of *R. hesperus*. Thus, ureolytic endosymbionts are likely important for maintaining the stability of essential cellulolytic protists within the tripartite microbial community in the guts of the lower termites. Our study raises several questions for future work on the distribution of ureolytic microbes in other termites with differing diets and across the phylogenetic tree. For instance, it has not yet been explored whether litter-feeding *Rhynchotermes* termites, which have lower biological N_2-fixation rates than wood-feeding termites [83] and several *Firmicutes* strains implicated in potential uric acid degradation [84] have a significant ureolytic component in their guts. In addition, it is unknown whether ureolytic microbes in the guts of soil-feeding termites, which are likely to be more abundant than in wood-feeders due to relatively higher densities of *ureC* genes in soil [61] contribute significantly to symbiotic N-recycling.

Acknowledgements

We thank the members of the Soil Ecogenomics lab, Eloi Parladé, Albert Barberán, and Michael Messina for assistance with sampling. We acknowledge Yang Ouyang for uploading *ureC* seed sequences to the Fungene Repository. This research was supported in part by a UC Davis Henry A. Jastro Research Award (to JYL) and a UC Davis Provost's Undergraduate Fellowship (to SJW).

Declarations

Funding

This research was supported in part by a UC Davis Henry A. Jastro Research Award and a UC Davis Provost's Undergraduate Fellowship.

Competing Interests

All authors declare that they have no competing interests.

Ethics Approval

No approval of research ethics committees was required to accomplish the goals of this study because experimental work was conducted with an unregulated invertebrate species.

Data Availability

The 16S rRNA and *ureC* gene sequencing data were deposited to the NCBI sequence read archive (SRA) under BioProject PRJNA660442 with the following BioSample accession numbers: SAMN15949798 (16S rRNA) and SAMN15949813 (*ureC*). The custom GraftM gene package used for taxonomic classification of *ureC* gene sequences is available at https://github.com/jonathanylin/Termite_gut_urease.

Author Contributions

JYL and JLMR contributed to the study conception and design. Material preparation, data collection, and analysis were performed by JYL, LH, and SJW. JYL wrote the manuscript and all authors commented on previous versions of the manuscript. All authors read and approved the final version of the manuscript.

References

1. Hansen AK, Pers D, Russell JA (2020) Symbiotic solutions to nitrogen limitation and amino acid imbalance in insect diets. In: Oliver KM, Russell JA (eds) Advances in Insect Physiology. Academic Press, pp 161–205

2. Mattson WJ (1980) Herbivory in relation to plant nitrogen content. Annu Rev Ecol Syst 11:119–161

3. Douglas AE (2009) The microbial dimension in insect nutritional ecology. Funct Ecol 23:38–47

4. Gil R, Silva FJ, Zientz E, et al (2003) The genome sequence of *Blochmannia* floridanus: Comparative analysis of reduced genomes. Proc Natl Acad Sci USA 100:9388–9393

5. Feldhaar H, Straka J, Krischke M, et al (2007) Nutritional upgrading for omnivorous carpenter ants by the endosymbiont *Blochmannia*. BMC Biol 5:48

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

7. Lin JY, Russell JA, Sanders JG, Wertz JT (2016) *Cephaloticoccu*s gen. nov., a new genus of '*Verrucomicrobia'* containing two novel species isolated from *Cephalotes* ant guts. Int J Syst Evol Microbiol 66:3034–3040

8. Hu Y, Sanders JG, Łukasik P, et al (2018) Herbivorous turtle ants obtain essential nutrients from a conserved nitrogen-recycling gut microbiome. Nat Commun 9:964

9. Sabree ZL, Kambhampati S, Moran NA (2009) Nitrogen recycling and nutritional provisioning by *Blattabacterium*, the cockroach endosymbiont. Proc Natl Acad Sci USA 106:19521–19526

10. Patiño-Navarrete R, Piulachs M-D, Belles X, et al (2014) The *cockroach Blattella germanic*a obtains nitrogen from uric acid through a metabolic pathway shared with its bacterial endosymbiont. Biol Lett 10:20140407

11. 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

12. Wertz JT, Béchade B (2020) Symbiont-mediated degradation of dietary carbon sources in social herbivorous insects. In: Oliver KM, Russell JA (eds) Advances in Insect Physiology. Academic Press, pp 63–109

13. Brune A (2014) Symbiotic digestion of lignocellulose in termite guts. Nat Rev Micro 12:168–180

14. Breznak JA, Brill WJ, Mertins JW, Coppel HC (1973) Nitrogen fixation in termites. Nature 244:577–580

15. Breznak JA (2000) Ecology of prokaryotic microbes in the guts of wood- and litter-feeding termites. In: Abe T, Bignell DE, Higashi M (eds) Termites: Evolution, Sociality, Symbioses, Ecology. Springer Netherlands, Dordrecht, pp 209–231

16. Collins NM (1983) The utilization of nitrogen resources by termites (Isoptera). In: Lee JA, McNeill S, Rorison IH (eds) Nitrogen as an Ecological Factor. Blackwell Scientific Publications, Oxford, pp 381–412

17. Potrikus CJ, Breznak JA (1980) Uric acid in wood-eating termites. Insect Biochem 10:19–27

18. Potrikus CJ, Breznak JA (1981) Gut bacteria recycle uric acid nitrogen in termites: A strategy for nutrient conservation. Proc Natl Acad Sci USA 78:4601–4605

19. Potrikus CJ, Breznak JA (1980) Anaerobic degradation of uric acid by gut bacteria of termites. Appl Environ Microbiol 40:125–132

20. Thong-On A, Suzuki K, Noda S, et al (2012) Isolation and characterization of anaerobic bacteria for symbiotic recycling of uric acid nitrogen in the gut of various termites. Microbes Environ 27:186–192

21. Weatherby JH (1929) Excretion of nitrogenous substances in Protozoa. Physiol Zool 2:375–394

22. Odelson DA, Breznak JA (1983) Volatile fatty acid production by the hindgut microbiota of xylophagous termites. Appl Environ Microbiol 45:1602–1613

23. Matson E, Ottesen E, Leadbetter J (2007) Extracting DNA from the gut microbes of the termite (*Zootermopsis nevadensis*). J Vis Exp e195

24. Caporaso JG, Lauber CL, Walters WA, et al (2012) Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. ISME J 6:1621–1624

25. Reed KE (2001) Restriction enzyme mapping of bacterial urease genes: using degenerate primers to expand experimental outcomes. Biochem Mol Biol Edu 29:239–244

26. Jin D, Zhao S, Zheng N, et al (2017) Differences in ureolytic bacterial composition between the rumen digesta and rumen wall based on *ureC* gene classification. Front Microbiol 8:385

27. Benjamino J, Graf J (2016) Characterization of the core and caste-specific microbiota in the termite, *Reticulitermes flavipes*. Front Microbiol 7:171

28. Rubin BER, Sanders JG, Hampton-Marcell J, et al (2014) DNA extraction protocols cause differences in 16S rRNA amplicon sequencing efficiency but not in community profile composition or structure. Microbiologyopen 3:910–921

29. Callahan BJ, McMurdie PJ, Rosen MJ, et al (2016) DADA2: High-resolution sample inference from Illumina amplicon data. Nat Methods 13:581–583

30. Andrews S (2010) FastQC: A quality control tool for high throughput sequence data.

31. Magoč T, Salzberg SL (2011) FLASH: fast length adjustment of short reads to improve genome assemblies. Bioinformatics 27:2957–2963

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

33. Bushnell B BBMap. In: SourceForge. https://sourceforge.net/projects/bbmap/

34. Fish JA, Chai B, Wang Q, et al (2013) FunGene: The functional gene pipeline and repository. Front Microbiol 4:291

35. Boyd JA, Woodcroft BJ, Tyson GW (2018) GraftM: A tool for scalable, phylogenetically informed classification of genes within metagenomes. Nucleic Acids Res 46:e59–e59

36. Matsen FA, Kodner RB, Armbrust EV (2010) pplacer: Linear time maximum-likelihood and Bayesian phylogenetic placement of sequences onto a fixed reference tree. BMC Bioinform 11:538

37. Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. Mol Biol Evol 30:772–780

38. Price MN, Dehal PS, Arkin AP (2010) FastTree 2 – approximately maximum-likelihood trees for large alignments. PLoS One 5:e9490

39. McMurdie PJ, Holmes S (2013) phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. PLoS One 8:e61217

40. Yu G, Smith DK, Zhu H, et al (2017) ggtree: An R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. Methods Mol Biol 8:28–36

41. Brune A, Emerson D, Breznak JA (1995) The termite gut microflora as an oxygen sink: Microelectrode determination of oxygen and pH gradients in guts of lower and higher termites. Appl Environ Microbiol 61:2681–2687

42. Dombek PE, Johnson LK, Zimmerley ST, Sadowsky MJ (2000) Use of repetitive DNA sequences and the PCR To differentiate *Escherichia coli* isolates from human and animal sources. Appl Environ Microbiol 66:2572–2577

43. Baker GC, Smith JJ, Cowan DA (2003) Review and re-analysis of domain-specific 16S primers. J Microbiol Methods 55:541–555

44. Altschul SF, Gish W, Miller W, et al (1990) Basic local alignment search tool. J Mol Biol 215:403–410

45. Oksanen J, Blanchet FG, Friendly M, et al (2019) vegan: Community ecology package

721 46. Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion
722 for RNA-seq data with DESeq2. Genome Biol 15:550

723 47. Kolde R (2019) pheatmap: Pretty heatmaps

724 48. Ohkuma M, Noda S, Hattori S, et al (2015) Acetogenesis from H_2 plus CO_2 and nitrogen
725 fixation by an endosymbiotic spirochete of a termite-gut cellulolytic protist. Proc Natl
726 Acad USA 112:10224–10230

727 49. Husseneder C, Berestecky JM, Grace JK (2009) Changes in composition of culturable
728 bacteria community in the gut of the Formosan subterranean termite depending on
729 rearing conditions of the host. Ann Entomol Soc Am 102:498–507

730 50. Kakumanu ML, Maritz JM, Carlton JM, Schal C (2018) Overlapping community
731 compositions of gut and fecal microbiomes in lab-reared and field-collected German
732 cockroaches. Appl Environ Microbiol 84:e01037-18

733 51. Waltmann A, Willcox AC, Balasubramanian S, et al (2019) Hindgut microbiota in
734 laboratory-reared and wild *Triatoma infestans*. PLoS Negl Trop Dis 13:e0007383

735 52. Smith CC, Srygley RB, Healy F, et al (2017) Spatial structure of the Mormon cricket gut
736 microbiome and its predicted contribution to nutrition and immune function. Front
737 Microbiol 8:801

738 53. Tasaki E, Sakurai H, Nitao M, et al (2017) Uric acid, an important antioxidant contributing
739 to survival in termites. PLoS One 12:e0179426

740 54. Siderhurst M, James D, Bjostad L (2006) Ultraviolet light induced autophototoxicity and
741 negative phototaxis in *Reticulitermes* termites (*Isoptera*: *Rhinotermitidae*). Sociobiology
742 48:27–49

743 55. Bready JK, Friedman S (1963) Oxygen poisoning of the termite, *Reticulitermes flavipes*
744 (Kollar), and protection by carbon dioxide. J Insect Physiol 9:337–347

745 56. Lighton JRB, Ottesen EA (2005) To DGC or not to DGC: Oxygen guarding in the termite
746 *Zootermopsis nevadensis* (*Isoptera*: *Termopsidae*). J Exp Biol 208:4671–4678

747 57. Kamble S (1991) Subterranean termites and their control. Historical Materials from
748 University of Nebraska-Lincoln Extension

749 58. Chappell DJ, Slaytor M (1993) Uric acid synthesis in freshly collected and laboratory-
750 maintained *Nasutitermes walkeri* hill. Insect Biochem Mol Biol 23:499–506

751 59. Slaytor M, Chappell DJ (1994) Nitrogen metabolism in termites. Comparative Biochemistry
752 and Physiology Part B: Comparative Biochemistry 107:1–10

753 60. Waidele L, Korb J, Voolstra CR, et al (2017) Differential ecological specificity of protist and
754 bacterial microbiomes across a set of termite species. Front Microbiol 8:2518

61. Oshiki M, Araki M, Hirakata Y, et al (2018) Ureolytic prokaryotes in soil: Community abundance and diversity. Microbes Environ 33:230–233

62. Ohkuma M, Iida T, Kudo T (1999) Phylogenetic relationships of symbiotic spirochetes in the gut of diverse termites. FEMS Microbiol Lett 181:123–129

63. Noda S, Ohkuma M, Yamada A, et al (2003) Phylogenetic position and in situ Identification of ectosymbiotic *Spirochetes* on protists in the termite gut. Appl Environ Microbiol 69:625–633

64. Brune A (2006) Symbiotic associations between termites and prokaryotes. In: Dr MDP, Falkow S, Rosenberg E, et al (eds) The Prokaryotes. Springer New York, pp 439–474

65. Stanton TB, Canale-Parola E (1980) *Treponema bryantii* sp. nov., a rumen spirochete that interacts with cellulolytic bacteria. Arch Microbiol 127:145–156

66. Newbrook K, Staton GJ, Clegg SR, et al (2017) *Treponema ruminis* sp. nov., a spirochaete isolated from the bovine rumen. Int J Syst Evol Microbiol 67:1349–1354

67. Mobley HL, Hausinger RP (1989) Microbial ureases: significance, regulation, and molecular characterization. Microbiol Rev 53:85–108

68. Graber JR, Leadbetter JR, Breznak JA (2004) Description of *Treponema azotonutricium* sp. nov. and *Treponema primitia* sp. nov., the first *Spirochetes* isolated from termite guts. Appl Environ Microbiol 70:1315–1320

69. Leadbetter JR, Schmidt TM, Graber JR, Breznak JA (1999) Acetogenesis from H_2 plus CO_2 by Spirochetes from termite guts. Science 283:686–689.

70. Noda S, Shimizu D, Yuki M, et al (2018) Host-symbiont cospeciation of termite-gut cellulolytic protists of the genera *Teranympha* and *Eucomonympha* and their *Treponema* endosymbionts. Microbes Environ 33:26–33

71. Chou J-H, Sheu S-Y, Lin K-Y, et al (2007) *Comamonas odontotermitis* sp. nov., isolated from the gut of the termite *Odontotermes formosanus*. Int J Syst Evol Microbiol 57:887–891

72. Fontes-Perez H, Olvera-García M, Chávez-Martínez A, et al (2015) Genome sequence of Citrobacter sp. CtB7.12, isolated from the gut of the desert subterranean termite *Heterotermes aureus*. Genome Announc 3:e01290-15

73. Breznak JA, Switzer JM, Seitz H-J (1988) *Sporomusa termitida* sp. nov., an H_2/CO_2-utilizing acetogen isolated from termites. Arch Microbiol 150:282–288

74. Wertz JT, Breznak JA (2007) *Stenoxybacter acetivorans* gen. nov., sp. nov., an acetate-oxidizing obligate microaerophile among diverse O_2-consuming bacteria from termite guts. Appl Environ Microbiol 73:6819–6828

75. Wertz JT, Kim E, Breznak JA, et al (2012) Genomic and physiological characterization of the *Verrucomicrobia* isolate *Diplosphaera colitermitum* gen. nov., sp. nov., reveals microaerophily and nitrogen fixation genes. Appl Environ Microbiol 78:1544–1555

76. Lin JY, Rodrigues JLM (2020) *Geminisphaera*. In: Bergey's Manual of Systematics of Archaea and Bacteria. Wiley, pp 1–7

77. Kotak M, Lin JY, Isanapong J, Rodrigues JLM (2020) Draft genome sequences of strains TAV3 and TAV4 (*Verrucomicrobia*: *Opitutaceae*), isolated from a wood-feeding termite, and in silico analysis of their polysaccharide-degrading enzymes. Microbiol Resour Announc 9:e01192-19

78. Kotak M, Isanapong J, Goodwin L, et al (2015) Complete genome sequence of the Opitutaceae bacterium strain TAV5, a potential facultative methylotroph of the wood-feeding termite *Reticulitermes flavipes*. Genome Announc 3:e00060-15

79. Isanapong J, Goodwin L, Bruce D, et al (2012) High-quality draft genome sequence of the Opitutaceae bacterium strain TAV1, a symbiont of the wood-feeding termite *Reticulitermes flavipes*. J Bacteriol 194:2744–2745

80. Hongoh Y, Sharma VK, Prakash T, et al (2008) Genome of an endosymbiont coupling N_2 fixation to cellulolysis within protist cells in termite gut. Science 322:1108–1109

81. Hongoh Y (2010) Diversity and genomes of uncultured microbial symbionts in the termite gut. Biosci Biotech Biochem 74:1145–1151

82. Hongoh Y, Deevong P, Inoue T, et al (2005) Intra- and interspecific comparisons of bacterial diversity and community structure support coevolution of gut microbiota and termite host. Appl Environ Microbiol 71:6590–6599

83. Prestwich GD, Bentley BL, Carpenter EJ (1980) Nitrogen sources for neotropical nasute termites: Fixation and selective foraging. Oecologia 46:397–401

84. Ottesen EA, Leadbetter JR (2011) Formyltetrahydrofolate synthetase gene diversity in the guts of higher termites with different diets and lifestyles. Appl Environ Microbiol 77:3461–3467

Figure Legends

Figure 1. Gut prokaryotic communities in termites from colony WTH1 and TH1 based on the 16S rRNA gene. Differences in (**A**) richness, (**B**) Shannon diversity, (**C**) Shannon evenness, and (**D**) community composition based on Bray-Curtis distances between colonies. (**E**) Venn diagram representing shared ESVs, (**F**) bar plot showing average relative abundances of the 10 most abundant phyla, and (**G**) heatmap indicating differentially abundant ($P < 0.05$) taxa at the order level between colonies.

Figure 2. Gut ureolytic microbial communities in termites from colony WTH1 and TH1 based on the *ureC* gene. Differences in (**A**) richness, (**B**) Shannon diversity, (**C**) Shannon evenness, and (**D**) community composition based on Bray-Curtis distances between colonies. (**E**) Venn diagram representing shared species, (**F**) bar plot showing average relative abundances of the 10 most abundant phyla, and (**G**) heatmap indicating differentially abundant ($P < 0.05$) taxa at the order level between colonies.

Figure 3. Changes in the abundance of *ureC* and 16S rRNA genes between wild and laboratory-reared termites. Copy numbers of the (**A**) *ureC* and (**B**) 16S rRNA gene, and (**C**) the proportion of *ureC* gene copies (calculated as the ratio of *ureC* gene copies to total 16S rRNA gene copies) between colonies.

Figures

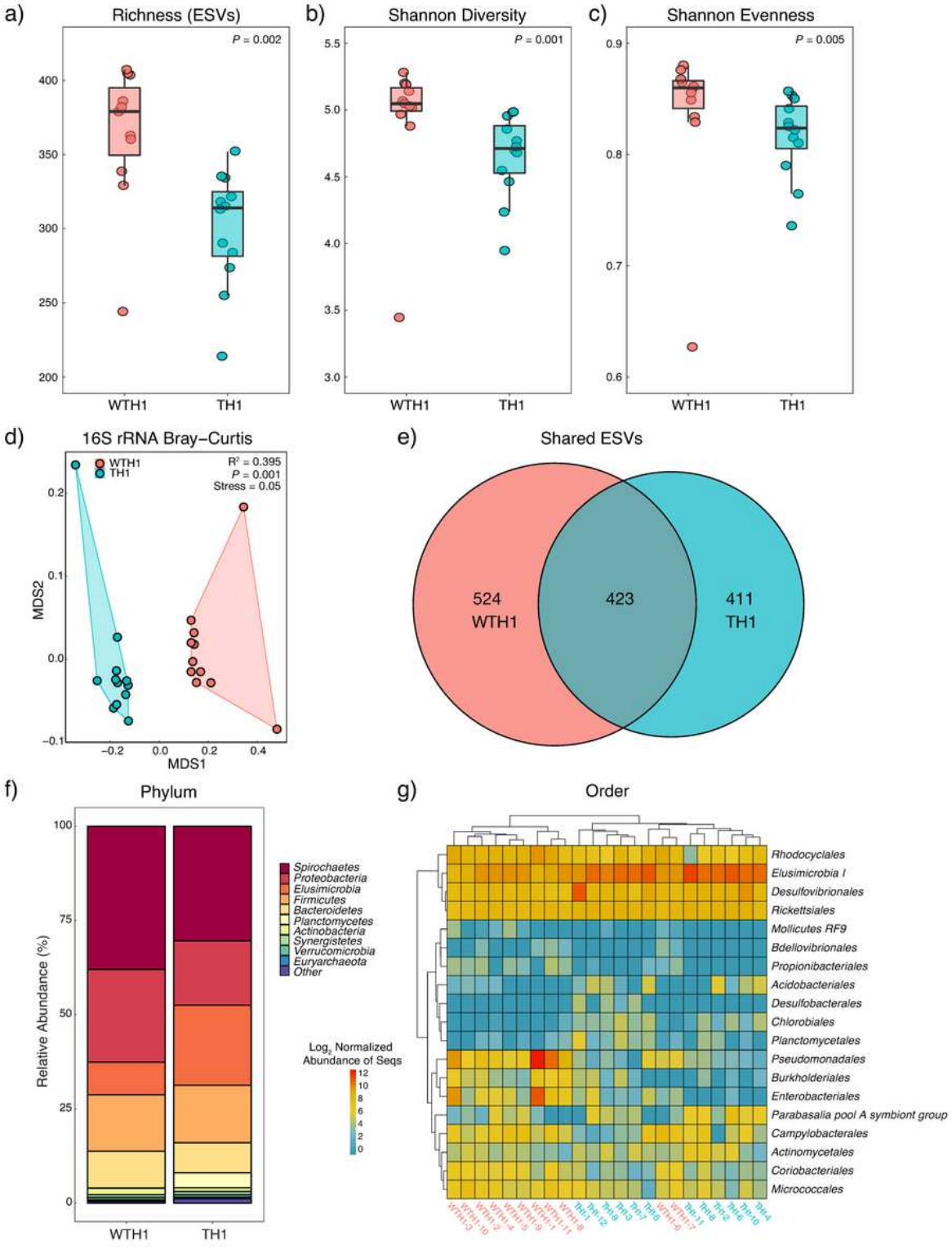

Figure 1

Gut prokaryotic communities in termites from colony WTH1 and TH1 based on the 16S rRNA gene. Differences in **(A)** richness, **(B)** Shannon diversity, **(C)** Shannon evenness, and **(D)** community composition based on Bray-Curtis distances between colonies. **(E)** Venn diagram representing shared

ESVs, **(F)** bar plot showing average relative abundances of the 10 most abundant phyla, and **(G)** heatmap indicating differentially abundant (P < 0.05) taxa at the order 828 level between colonies.

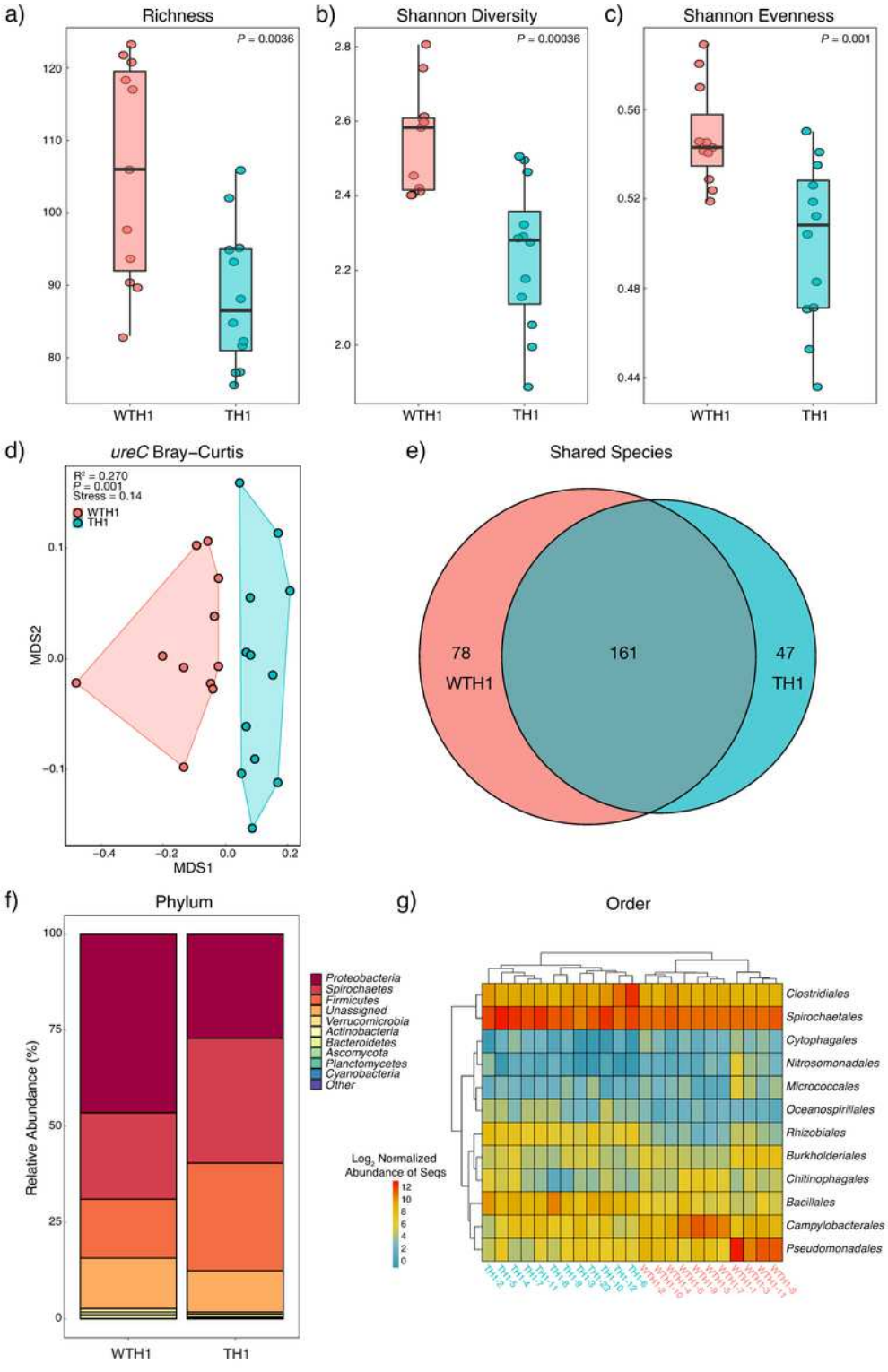

Figure 2

Gut ureolytic microbial communities in termites from colony WTH1 and TH1 based on the ureC gene. Differences in **(A)** richness, **(B)** Shannon diversity, **(C)** Shannon evenness, and **(D)** community

composition based on Bray-Curtis distances between colonies. **(E)** Venn diagram representing shared species, **(F)** bar plot showing average relative abundances of the 10 most abundant phyla, and **(G)** heatmap indicating differentially abundant (P < 0.05) taxa at the order level between colonies.

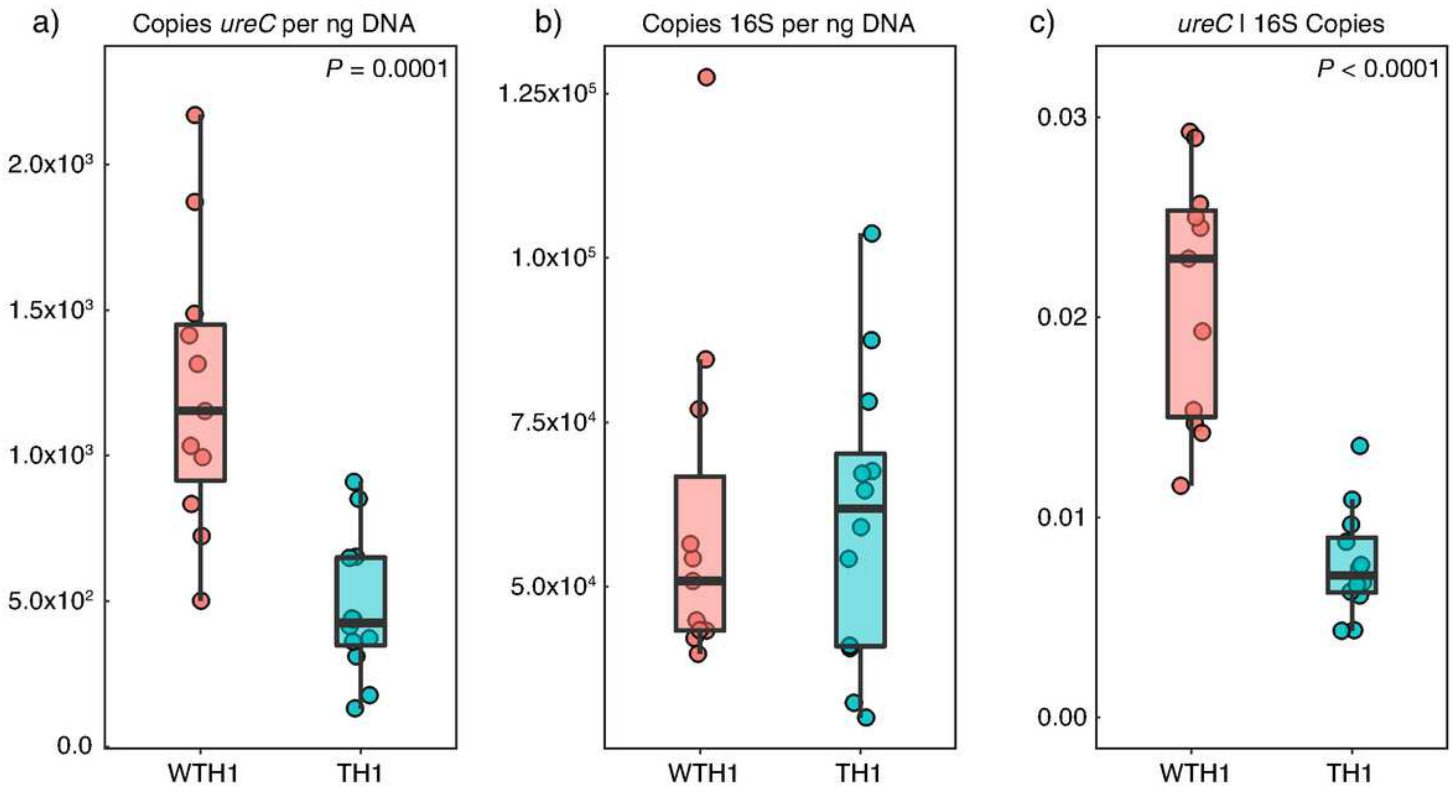

Figure 3

Changes in the abundance of *ureC* and 16S rRNA genes between wild and laboratory reared termites. Copy numbers of the **(A)** *ureC* and **(B)** 16S rRNA gene, and **(C)** the proportion of *ureC* gene copies (calculated as the ratio of *ureC* gene copies to total 16S rRNA gene copies) between colonies

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- TermiteUreaseManuscriptSupplementREVISEDFINAL.pdf