ISME

www.nature.com/ismej

ARTICLE

Physiological and evolutionary contexts of a new symbiotic species from the nitrogen-recycling gut community of turtle ants

Benoît Béchade[1✉], Christian S. Cabuslay[1], Yi Hu[1,2], Caroll M. Mendonca[3], Bahareh Hassanpour[3], Jonathan Y. Lin[4], Yangzhou Su[4], Valerie J. Fiers[1], Dharman Anandarajan[1], Richard Lu[1], Chandler J. Olson[1,5], Christophe Duplais[6], Gail L. Rosen[7], Corrie S. Moreau[6,8], Ludmilla Aristilde[3], John T. Wertz[4] and Jacob A. Russell[1]

While genome sequencing has expanded our knowledge of symbiosis, role assignment within multi-species microbiomes remains challenging due to genomic redundancy and the uncertainties of in vivo impacts. We address such questions, here, for a specialized nitrogen (N) recycling microbiome of turtle ants, describing a new genus and species of gut symbiont—*Ischyrobacter davidsoniae* (*Betaproteobacteria*: *Burkholderiales*: *Alcaligenaceae*)—and its in vivo physiological context. A re-analysis of amplicon sequencing data, with precisely assigned *Ischyrobacter* reads, revealed a seemingly ubiquitous distribution across the turtle ant genus *Cephalotes*, suggesting ≥50 million years since domestication. Through new genome sequencing, we also show that divergent *I. davidsoniae* lineages are conserved in their uricolytic and urea-generating capacities. With phylogenetically refined definitions of *Ischyrobacter* and separately domesticated *Burkholderiales* symbionts, our FISH microscopy revealed a distinct niche for *I. davidsoniae*, with dense populations at the anterior ileum. Being positioned at the site of host N-waste delivery, in vivo metatranscriptomics and metabolomics further implicate *I. davidsoniae* within a symbiont-autonomous N-recycling pathway. While encoding much of this pathway, *I. davidsoniae* expressed only a subset of the requisite steps in mature adult workers, including the penultimate step deriving urea from allantoate. The remaining steps were expressed by other specialized gut symbionts. Collectively, this assemblage converts inosine, made from midgut symbionts, into urea and ammonia in the hindgut. With urea supporting host amino acid budgets and cuticle synthesis, and with the ancient nature of other active N-recyclers discovered here, *I. davidsoniae* emerges as a central player in a conserved and impactful, multipartite symbiosis.

The ISME Journal; https://doi.org/10.1038/s41396-023-01490-1

INTRODUCTION

Symbiotic bacteria have played integral roles in the evolution of insects, shaping critical physiological processes while impacting host nutrition [1]. The nature of symbiotic associations varies across and within insect families, showing substantial taxonomic diversity [2–4]. Variation at the intra-family level is notable among the ants (*Hymenoptera*: *Formicidae*), where bacteriocyte-symbioses and high biomass symbiotic gut communities appear rare [5–7]. When present, however, they are impactful, with several such symbionts upgrading ant nutrition in the face of herbivorous diets. Key to these symbioses are nitrogen (N) recycling metabolisms, which fuel the synthesis of amino acids (AAs) ([8, 9], see also [10–14]). Such symbiont-mediated N-recycling appears common among termites and cockroaches, and more variable among hemipterans, true fruit flies, and beetles [15]. Uncertainties remain in each group, with a lack of detailed insight into the identities and origins of N-waste molecules, and the metabolic coordination required among N-recycling community members [16].

An ideal system for investigating N-recycling symbioses is comprised by the turtle ants from the sister genera *Cephalotes* and *Procryptocerus*, and their high biomass gut microbiomes [7, 17]. Localizing to the midgut, ileum, and rectum [18], extracellular bacterial symbionts of turtle ants are accommodated within enlarged, sometimes invaginated, gut compartments presumably modified to house a high biomass community [19–21]. Trophallactic symbiont transfer among adult nestmates [22, 23], and vertical transfer through queens, have perpetuated these symbioses, enforcing partner fidelity across a ≥50 million year timespan [24, 25].

[1]Department of Biology, Drexel University, 3245 Chestnut St., Philadelphia, PA 19104, USA. [2]State Key Laboratory of Earth Surface Processes and Resource Ecology and Ministry of Education Key Laboratory for Biodiversity Science and Ecological Engineering, College of Life Sciences, Beijing Normal University, 100875 Beijing, China. [3]Department of Civil and Environmental Engineering, McCormick School of Engineering and Applied Science, Northwestern University, Evanston, IL 60208, USA. [4]Department of Biology, Calvin University, 1726 Knollcrest Circle SE, Grand Rapids, MI 49546-4402, USA. [5]Department of Biological Sciences, University of Alabama, 1325 Hackberry Ln, Tuscaloosa, AL 35487, USA. [6]Department of Entomology, Cornell University, Cornell AgriTech, Geneva, NY 14456, USA. [7]Ecological and Evolutionary Signal-Processing and Informatics Laboratory, Department of Electrical and Computer Engineering, Drexel University, 3141 Chestnut St., Philadelphia, PA 19104, USA. [8]Department of Ecology and Evolutionary Biology, Cornell University, Ithaca, NY 14853, USA. ✉email: ben.bechade@gmail.com

Received: 21 March 2023 Revised: 21 July 2023 Accepted: 27 July 2023
Published online: 09 August 2023

The contemporary impacts of this symbiosis include: (1) recycling of N [9], thought to be relatively low in turtle ant diets [26], (2) essential AA acquisition [9, 27], and (3) construction of the ant cuticle [27]. Demonstrated through a combination of genomics, in vitro assays, antibiotic treatments, and isotope labeling experiments, urea is a crucial ingredient for each of the above three functions [9, 27, 28]. Since insects cannot catabolize this compound, the use of urea-derived N must be facilitated by bacterial metabolism [15]. Urease function is encoded by two or more symbionts in most turtle ant hosts. Included among these are the ubiquitous, cospeciating *Cephaloticoccus* symbionts [9, 24].

Genomic analyses suggest other symbionts are N-recyclers as well, deriving urea through the upstream metabolism of purines, including urate (i.e., uric acid) [9, 28]. However, conservation of such roles remains uncertain outside of *Cephaloticoccus*, stemming in part from current challenges with symbiont phylogenetics. In addition, despite extensive (meta)genomic investigation and the suggested trend of divided N-recycling labor, realized gene expression and the in vivo contexts of N-recycling genes remain unclear. The mystery deepens, further, amidst the complexities of a multi-generational, multi-caste, social insect society with trophallaxis-mediated nutrient transfer [29, 30].

Among the eight bacterial orders found commonly in turtle ant guts [17, 31–33], the *Burkholderiales* includes symbionts with N-recycling potential [9, 28]. Representation within this order appears diverse, with several operational taxonomic units (OTUs) from the *Alcaligenaceae* and one from the *Comamonadaceae*, being common in turtle ant adults and more sporadic among their larvae [24]. Past phylogenetics on turtle ant-associated taxa affiliated with the *Burkholderiales* suggested several distinct, yet host-specific lineages. But inferences have relied on sparse sampling of Sanger-sequenced 16S rRNA genes or on short amplicon sequence reads, yielding variable results across studies [31, 32].

Among insects, the *Alcaligenaceae* family has served as a common source of symbiosis, with representatives in coleopteran, hymenopteran, and lepidopteran hosts [34–36]. In turtle ants, symbionts from the *Alcaligenaceae* encode conserved abilities to synthesize B-vitamins, AAs [28], and a diverse array of specialized metabolites (e.g. terpenes, polyketides type I, siderophores) that could potentially serve as inter-symbiont signal molecules [37]. These turtle ant-associated symbionts may also be involved in CoA-related metabolism and the decomposition of phenol or lignin [28]. Of greatest pertinence here, is their genomic capacity for N-recycling involving the lysis of urate (i.e., uricolysis), terminating with the production of urea [9, 28].

Testing the contributions of one such *Alcaligenaceae* symbiont toward N-recycling pathways, we combine microscopy, phylogenetics, in vitro metabolic assays, and in vivo multi-omic screenings to describe a new lineage of insect-associated symbionts, with broad distributions across turtle ants. Members of this new lineage encode a diverse suite of intriguing biosynthetic and catabolic metabolisms. Of central relevance are our findings that these symbionts and other co-resident symbionts express complementary portions of the purine biosynthesis and uricolytic pathways. Through its assisted production of urea this new symbiont emerges as a key, cooperative N-recycler, with long-standing impacts on the use of relatively poor diets amidst the ancestrally predaceous and omnivorous ants [38].

METHODS

Detailed protocols for in vitro assays, phylogenetics, transmission electron microscopy, genomics, analyses of amplicon sequencing data, fluorescent in situ hybridization (FISH) microscopy, metatranscriptomics, and metabolomics are available as Supplementary Methods. Highlights from each are presented below.

Bacterial cultivation and identification

Bacterial symbionts, including strain Cv33a from the newly described symbiont species, were isolated, cultured, and characterized in vitro [9, 28]. Isolates were routinely cultured in trypticase soy broth or trypticase soy agar supplemented with 5% newborn calf serum (Sigma-Aldrich, St. Louis, MO, USA). The Cv33a isolate grew without the addition of calf serum when the medium was supplemented with 20 mg/l L-asparagine, 1 ml/l vitamin supplement MD-VS, and trace mineral supplement MD-TMS (ATCC, Manassas, VA, USA).

All isolates were identified through 16S rRNA gene sequencing with the universal primers 63F (5′-CAGGCCTAACACATGCAAGTC-3′) and 1389R (5′-ACGGGCGGTGTGTACAAG-3′) [39]. The resulting amplicons were sequenced at the Michigan State University Genomics Core Facilities using Sanger sequencing. For isolates assigning to the *Burkholderiales*, we used BLAST to identify relatives within NCBI, downloading their 16S rRNA gene sequences for use in a multiple sequence alignment. This input was used for phylogenetic inference using RAxML HPC2 [40] on the XSEDE computing cluster [41].

Genomics

The Cv33a genome was previously sequenced from a cultured isolate from *Cephalotes varians* turtle ants [9]. Following prior methods [28], we screened this genome for genes encoding metabolic pathways with central relevance in other symbioses—i.e., biosynthesis of AAs and B-vitamins, N-recycling, sulfur and central carbon metabolism, and catabolism of plant cell wall molecules [42–49]. In brief, we performed function searches on the IMG/M-ER website [50], searching first with KEGG Orthology identifiers (KOs) from KEGG metabolic and biosynthetic pathways. Enzyme commission numbers (EC), Pfam identifiers, and COG identifiers were also used in these searches, yielding broad lists of encoded enzymes impacting the above functions. Expression of the encoded pathways, and of key metabolite transporters, was subsequently examined in our transcriptomic analyses, as detailed below.

To complement the Cv33a genome analysis, we sequenced the genome of a cultured relative from the *Alcaligenaceae* hosted by *Procryptocerus* nr. *pictipes* (strain CSM3487_49), which is an ant species from the sister genus to *Cephalotes*. Although *Procryptocerus* and *Cephalotes* diverged over 50 million years ago [51], species from both genera nest similarly in pre-existing tree cavities and mainly feed on a cryptic, N-limited diet [52]. Ants from both genera also harbor gut symbionts from the same lineages of turtle ant specialists [17, 18, 31], sharing these through (proctodeal) trophallaxis among nestmates [23, 52]. Phylogenetics placed the *P.* nr. *pictipes* CSM3487_49 isolate into a lineage with Cv33a and a bacterium encoding a "binned" draft genome, identified from the metagenome of *C. grandinosus* [9]. These new Cv33a-related genomes were screened for genes in the aforementioned functional pathways, using the Function Profile tool in IMG/M-ER [50].

Symbiont localization via amplicon sequencing analyses and FISH microscopy

Analyses of amplicon sequencing data. With improved phylogenetic definitions of turtle ant-specialized taxa affiliated with the *Burkholderiales*, we used amplicon sequence data from two prior papers to study the relative abundance, cross-species incidence, and localization of the Cv33a/CSM3487_49 lineage. To achieve this, we first mapped previously generated 97%-OTUs from 13 *Cephalotes* species [24] to the 16S rRNA gene sequences from the Cv33a and CSM3487_49 genomes, using BLASTn [53] on the high-performance computer cluster from the University Research Computing Facility at Drexel University. We then performed similar BLASTn searches of these Cv33a and CSM3487_49 sequences against the NCBI nucleotide (nt) database. A single OTU (OTU011) was the top hit of both Cv33a or CSM3487_49, and was tentatively assigned to this newly described lineage. Two additional BLASTs were then performed using the V4 16S rRNA sequence of OTU011 as the query. In the first, 16S rRNA sequences from all turtle ant symbionts with sequenced genomes were used as subjects. For the second, subject sequences came from the NCBI nt database.

Having established OTU011 from Cv33a/CSM3487_49 as reciprocal best hits, we next studied 16S rRNA gene sequences from the amplicon sequence variants (ASVs) of Flynn et al. [18], who studied turtle ant microbiomes across the digestive tract, separating the crop, midgut, ileal, and rectal compartments. Using a similar BLAST approach to that above, we first used the ASV sequences from Flynn et al. as queries, BLAST'ing these against the OTU sequences of Hu et al. [24]. In a separate search, we

used the representative sequence of OTU011 in a BLASTn search against all ASV 16S rRNA sequences from the Flynn et al. study. All ASVs that were reciprocal best hits to OTU011 were assigned to this Cv33a/CSM3487_49 symbiont lineage. Using the summed abundances of Cv33a/CSM3487_49-assigning ASVs within each gut compartment, along with 16S rRNA qPCR data, we gained insight into variable abundance of this symbiont across the gut and host ant species.

FISH microscopy. To further our localization efforts, we used whole-mount fluorescence microscopy for more fine-scale Cv33a localization within the gut of adult *C. varians*. Whole guts were dissected from workers of two colonies collected in the Florida Keys in August 2019 and March 2022. These ants were euthanized in ethanol prior to fixation in PBS-4% formaldehyde. We used a combination of fluorescent rRNA probes specific to the Cv33a strain (Table S1) and to the *Betaproteobacteria* [54]. The Cv33a-specific probe was designed with the webtool DECIPHER [55] using full length 16S rRNA genes from the genomes of turtle ant symbionts [9]. The specificity of this probe was tested in a series of hybridization experiments (Fig. S1).

The whole-mount FISH protocol was carried out as described elsewhere [56]. To reduce tissue autofluorescence, specimens were bleached in 6% hydrogen peroxide solution and in 80% ethanol for 7 days. Fresh solution exchanges and ethanol rinses were performed daily for the first 3 days of this protocol. Specimens were then rehydrated in PBSTx (PBS with 3 ml of Triton X-100 added to each liter) three times for 10 min each, followed by a rinse in hybridization buffer (0.02 M Tris-HCl, 0.9 M NaCl, 0.01% SDS, and 35 or 45% formamide in distilled water), and the final addition of hybridization buffer with probes and DAPI staining agent (Table S1). Samples were kept overnight at room temperature then washed in PBSTx before being suspended in CitiFluor Mountant Solution (Electron Microscopy Sciences, Hatfield, PA, USA) on a microscope slide for imaging on an Olympus (Tokyo, Japan) FV1000 confocal microscope.

Metatranscriptomics

Dietary experiment and RNA sequencing. To study the in vivo roles of gut symbionts, we performed RNAseq on gut tissues from field caught *C. varians* and those from lab specimens fed diets containing natural foods (pollen), or constituents thereof (urea from mammalian urine, cellobiose from pollen) [9, 28, 32]. Through this approach, we hoped to learn whether symbionts of mature adult workers are responsive to diet and to increase our chances of detecting gene expression through a range of dietary contexts.

For this research, *C. varians* turtle ants were collected in the Florida Keys in October 2014. One day after arrival in the lab, 15 adult workers from each of four colonies were dissected, with their guts being subsequently preserved in RNAlater and stored at −80 °C. The rest were reared in a warm room and fed with a nutrient-defined holidic diet [57] for 1 week. These four colonies were then, each, evenly divided into four groups of 16–45 adult workers and fed one of four experimental diets, twice weekly for 3 weeks. All dietary treatments included the aforementioned holidic diet, administered in liquid form. This was the sole food source for the "control" diet. For the remaining three treatments, we used this holidic diet to dissolve the supplementary food sources, creating diets of 6% honeybee pollen (Y.S. Organic Bee Farms, Sheridan, IL, USA), 4% cellobiose, or 1% urea (Sigma-Aldrich).

Following the dietary experiment, ants were surface sterilized using 6% bleach and handled with bleach-sterilized tweezers. The midgut, ileum, and rectum from each of these ants, and from the field-caught *C. varians*, were dissected and ground in 500 µl of TRIzol Reagent (Invitrogen, Waltham, MA, USA) with sterile plastic pestles. Samples were then preserved at −80 °C until RNA extraction. A protocol involving TRIzol lysis coupled with the RNeasy mini kit (Qiagen Ltd., Hilden, Germany) procedure and DNase I treatment was used for RNA extraction on pools of 15–45 guts obtained for each colony × treatment. We assessed RNA quality with a 2100 Bioanalyzer (Agilent, Santa Clara, CA, USA) and depleted rRNA using Ribo-Zero Magnetic Gold Epidemiology Kits (Epicentre Technologies Corporation, Madison, WI, USA). RNA quantity from one sample was insufficient, leading to a design of 19 pooled RNA samples for our metatranscriptomics: four colonies × four lab treatments; plus three out of these same four colonies from the field. Each RNA sample was assessed for contaminating DNA via PCR with universal 16S rRNA primers 9Fa (5′-GAGTTTGATCITIGCTCAG-3′) and 1513R (5′-TACIGITACCTTGTTACGACTT-3′) [58]. The contaminant-free samples were then sent to the UC Berkeley Genomics Facility for cDNA synthesis, library preparation/multiplexing, and sequencing in one flow cell lane on a HiSeq2500 System (Illumina, San Diego, CA, USA). The sequencing machine was set to rapid run mode, with paired-end sequencing runs of 250 bp in length.

Assembly and transcript count estimation. Work reported here was run on hardware supported by Drexel's University Research Computing Facility. We first checked the quality of the FASTq raw reads using MultiQC v1.9 [59]. Raw read pre-treatment was performed using software compiled in the SAMSA2 pipeline v2.2.0 [60]. Sequencing adapters were trimmed using Trimmomatic v0.36 [61], and paired-end reads were merged using Pear v0.9.10 [62]. We removed residual rRNAs using SortMeRna v2.1 [63], before reverifying the quality of pre-processed metatranscriptomics with MultiQC.

The taxonomic representation in each metatranscriptome was estimated by calculating the metatranscriptomic read coverage on previously sequenced host and symbiont genomes [9], using Bowtie2 v2.4.1, SAMtools v1.2, and Qualimap v2.2.1 [64–66]. Metatranscriptomic reads that mapped to the host *C. varians* genome [67] were removed from downstream analyses using Bowtie2 and SAMtools.

Using default settings of Trinity v2.11.0 [68], we performed a de novo assembly of pre-processed reads from all 19 metatranscriptomes, creating a single assembled metatranscriptome. The quality of the assembly was verified using MetaQUAST, from QUAST v5.0.2 [69], with *Cephalotes*-derived symbiont sequenced genomes [9] as references to estimate the proportion of symbiont genomes being expressed. Candidate protein-coding regions were then detected using TransDecoder v5.5.0 [70] and clustered into unique genes ("unigenes") using CD-HIT v4.7 [71], following default parameters.

For each of the 19 metatranscriptomes, we estimated unigene representation by first mapping raw reads to assembled unigenes using Salmon v1.3.0 [72] in an alignment-based mode, with Bowtie2 as the aligner. Abundance was obtained as counts of mapping reads using the script align_and_estimate_abundance.pl from Trinity [73]. Recent comparative studies suggest that coupling within-species normalization with adjustments for total RNA estimates in metatranscriptomes can yield robust quantifications [74]. Therefore, we first normalized unigene counts within individual symbionts by the counts of symbiont-specific housekeeping genes using the *DESeq2* R package [75]. We then performed a taxon-specific scaling [76] across metatranscriptomes to adjust for symbiont-wide differential RNA quantities. Finally, we standardized these normalized unigene counts by effective unigene length in our transcript per million (TPM) calculations. Such normalized, relative gene expression units were referred to as "normalized TPM" hereafter.

Annotation and pathway reconstruction. Using *Cephalotes*-derived symbiont genomes [9], a custom database was compiled by concatenating encoded AA sequences from (1) *Cephalotes*-derived symbiont genomes sequenced from cultured isolates, (2) a *C. varians* gut metagenome (PL010-W), and (3) the *C. varians* host genome [67].

Metatranscriptome-assembled unigenes were first annotated with three different databases including our custom database, the NCBI RefSeq database [77] from July 2020, and the SEED Subsystem database [78] from September 2017. Toward this end, we used a SAMSA2 script [60] to run DIAMOND v0.8.38 BLASTx [79] with default settings. The metatranscriptome assembly was also submitted to BlastKOALA in January 2022 to obtain additional unigene KEGG KO functional annotations. Symbiont taxonomic assignments from unigene annotation with our custom database were used to extract all annotated unigenes from single symbionts separately. Through these combined processes, we compiled symbiont-specific annotated draft transcriptome files.

To reconstruct symbiont-expressed metabolic pathways, unigene normalized TPMs, averaged or summed across all 19 metatranscriptome libraries, were grouped by KEGG KOs. Lists of expressed KEGG KOs were uploaded to the online KEGG Mapper—Color [80] to infer expressed pathways. These lists of expressed KEGG KOs were also uploaded to MicrobiomeAnalyst [81], and a "Shotgun Data Profiling" was generated to obtain symbiont expression grouped into KEGG KO categories.

Metabolomics

Insect sampling and antibiotic experiment. To understand how gut symbionts shape the identities and abundances of purine/N-waste metabolites within adult turtle ants, we performed antibiotic manipulations and metabolomic analyses on tissues from lab-reared workers. Live ants from six *C. varians* colonies (Florida Keys, August 2019) were reared in a Percival environmental chamber, with daily day:dusk:night:dawn lighting

cycles of full-lights:half-lights:no-lights:half-lights, across 10:2:10:2-h alternations (with 26:25:24:25 °C, temperature cycling). Before the experiment, the ants were fed three times per week with a customized holidic diet [57], stained with 0.1% of methylene blue, and diluted in half in 20% sucrose water. In January 2020, 60 mature adult workers were randomly selected from each colony to be used in the dietary experiment, when possible. The lowest number of ants used in the experiment for a colony was 28. Half of the ants from each colony (i.e., "antibiotic-treated group") were kept in a separate box and fed holidic diet enriched with 5% dietary fibers (Fiber Blend from the Vitamin Shoppe, Secaucus, NJ, USA), and 0.01% antibiotic cocktail (1:1:1 kanamycin:rifampicin:tetracycline, [9]). The remaining half was reared separately and fed on the same diet without antibiotics (i.e., "control group"). The experiment was conducted for 6–10 weeks.

To verify the efficacy of antibiotics, we performed qPCR on bacterial 16S rRNA genes from three ants per experimental group, per colony (Dataset S1). The remaining ants were processed with sterilized instruments (i.e., forceps, capillary tubes) for use in metabolomic assays. These ants were, first, surface sterilized using 6% bleach. After harvesting hemolymph, we dissected their infrabuccal pocket, crop, midgut, ileal, and rectal compartments in cold, sterile PCR water. With intended use of infrabuccal pocket and hemolymph data in a future publication, we include data from only the crop, midgut, ileum, and rectum here.

Individual gut compartments from several ants of the same colony and treatment were pooled in a single pre-chilled low-protein binding 1.5 ml tube. In total, six groups of ants (three colony replicates × two conditions) were dissected. Each group had their own dissection blank sample, consisting of an adult leg left in a droplet of sterile water near the zone where the ants were dissected, to pick up contaminants that might infiltrate the dissected tissues. Following dissection, all tissue samples were homogenized by grinding them in 15 µl cold sterile water using sterile plastic pestles. They were then stored at −80 °C until further processing.

Metabolite extraction, identification, and quantification. Following antibiotic treatment, tissues from the three colonies that had the highest bacterial load reduction (Dataset S1) were used in metabolite extraction. Of the 15 µl of homogenized ant tissues, 10 µl were used in metabolite extraction while 5 µl were used in protein quantification (see below). In addition to ant tissues, an extraction blank sample was included, consisting of sterile water instead of ant tissue. Aliquots of the synthetic diets that were fed to antibiotic-treated and experimental control ants were also included in metabolite extraction and analysis.

Solvents were kept cold on ice or dry ice over the extraction process, and tubes and glassware were pre-chilled in a freezer before extraction. Metabolites from the diluted samples were extracted in a 1 ml methanol:acetonitrile:water (40:40:20) extraction solvent through a series of centrifugations. The tubes containing metabolite extracts were then sealed with paraffin and shipped in dry ice to the Aristilde lab at Northwestern University. Upon reception, the samples were dried using a nitrogen blowdown evaporator. Metabolites were analyzed using ultra-high-performance liquid chromatography (LC) (Thermo Scientific Dionex UltiMate 3000) coupled to high-resolution/accurate mass spectrometry (MS) (Thermo Scientific Q Exactive Hybrid Quadrupole-Orbitrap) controlled with XCalibur software (Thermo Scientific, Waltham, MA, USA), as described previously [82, 83].

Blank correction and normalization. To correct for potential contamination that may have occurred during metabolite extraction, the concentration of each metabolite in each sample was subtracted by its concentration from the extraction blank sample. We then corrected for potential contaminants stemming from dissection using dissection blank samples (Dataset S2).

To avoid biases in metabolite quantities being disproportionally higher in heavier tissues and in colonies with larger ant individuals, we standardized blank-corrected metabolite concentrations by protein concentration from the same samples, using this protein measure as a proxy for insect tissue biomass [83]. To quantify proteins, we used a DC Protein Assay Kit (Bio-Rad Laboratories, Hercules, CA, US) with standard dilutions of known concentrations of bovine serum albumin protein, following the manufacturer's standard procedure for 96-well plates. Absorbance was read at 750 nm using a plate reader. Blank-corrected absolute metabolite concentrations were then divided by sample-specific protein quantities, which had also been blank-corrected (protein concentration from an ant group minus protein concentration in dissection blank sample from the same ant group). Metabolite concentrations were then quantified as micromole per gram of protein ("µmol/g of protein"). We acknowledge that, with this normalization method, we possibly underestimated metabolite concentrations from midguts, since protein concentrations in these tissues were remarkably high (midguts: 4.19 mg/ml ± 1.10, Dataset S2).

Data analysis. To fit concentrations across metabolites to a normal distribution, we performed generalized logarithmic base 10 (glog) transformation, a method that stabilizes the variance [84], implementing these calculations in MetaboAnalyst [85]. The glog-transformed data were then loaded on R v3.6.3 [86] for statistical analyses. Due to our small sample size and the zero-inflation (i.e., high number of zero values) inherent to metabolomic studies, glog-transformed data did not always approximate normal distributions. Nevertheless, in metabolomic studies that use few replicates, it has been recommended to apply parametric tests [87]. We therefore tested whether the metabolite concentrations were significantly different between antibiotic-treated and control ants using linear mixed models implemented in the *lme4* R package [88]. The models were as follows: "Metabolite concentration ~ Condition * Tissue + (1|Colony of origin)". The last term of the model, included as a random effect, allowed us to control for ants originating from the same colony.

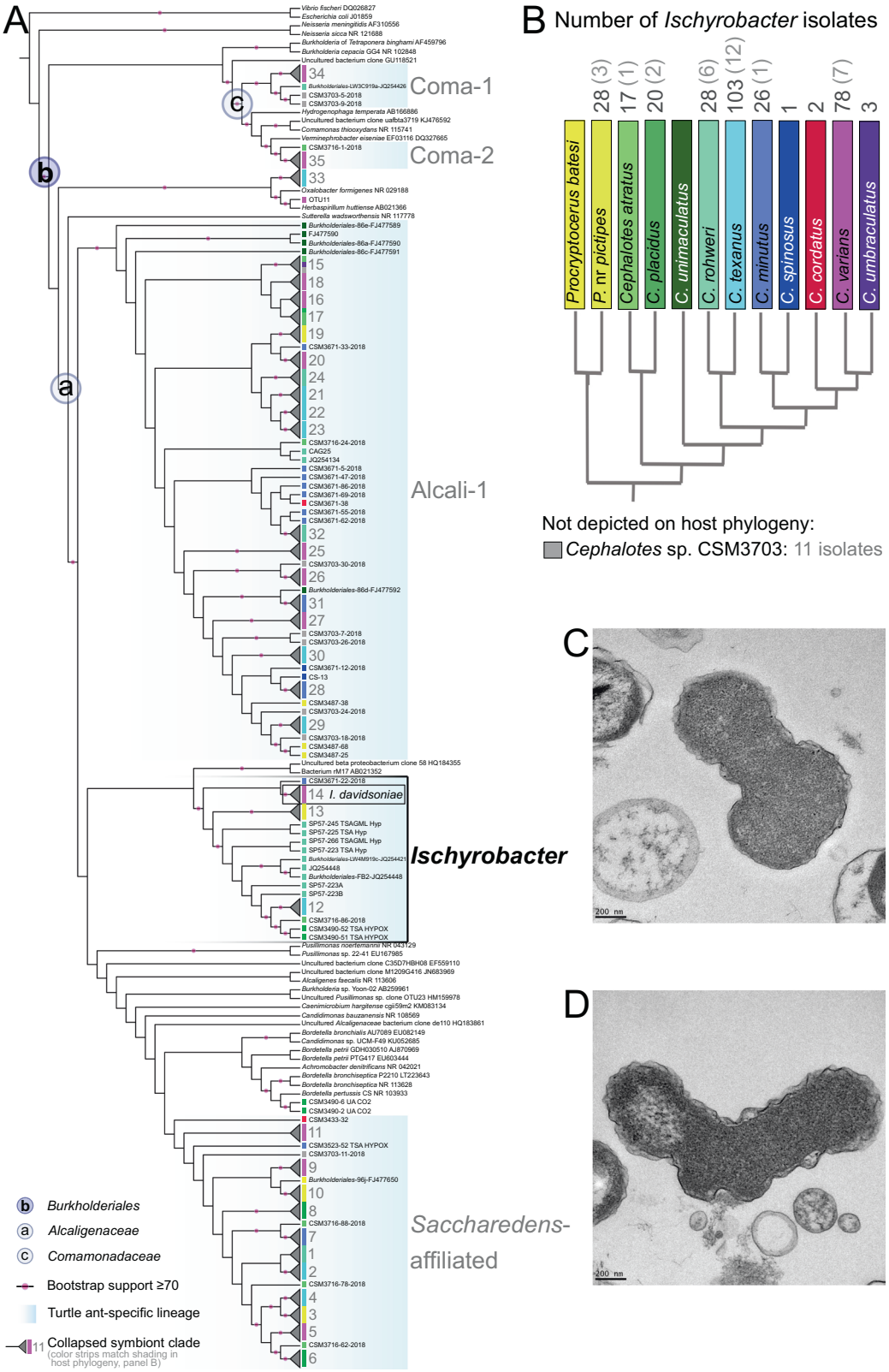
We next performed pairwise post hoc tests to identify tissues where antibiotic treatment significantly altered metabolite concentrations. We used the *emmeans* R package [89] to compute the estimated margin means [90] on the interaction Condition:Tissue terms of our linear mixed models. In doing so we applied an asymptotic degrees-of-freedom method. All p values from these tests were adjusted using the false discovery rate method [91] with the "p.adjust" function from the R *stats* package, focusing only on purine metabolites ($n = 9$ purine concentrations).

RESULTS/DISCUSSION

Ischyrobacter davidsoniae—a newly described symbiont genus and species from turtle ants

Through a multi-year effort at cultivation and molecular barcoding aimed at symbionts from 11 turtle ant species, we identified hundreds of bacterial isolates assigning to the order *Burkholderiales* (Fig. 1A). Phylogenetics of 16S rRNA genes revealed that nearly all of 317 isolates from the *Burkholderiales* fell into one of five turtle ant-specific clades. A protein phylogeny, focused on members of the *Alcaligenaceae* and built with three genes conserved among bacteria (SmpB, RpoB, and UvrB), further supports the existence of separately domesticated clades affiliated with the *Burkholderiales* exclusive to turtle ants (Fig. S2). One of these clades found on both trees included the Cv33a isolate from *C. varians*, which clustered with other *C. varians*-derived isolates (Figs. 1A and S2). Bootstrap values for this clade reached 97%, and 100% for the three-protein phylogenetic tree. Clustering of cultured isolates into this same clade on the 16S rRNA tree suggested that the Cv33a lineage had representation in seven of the 11 turtle ant species targeted with culturing. Absence was found in only those species with lower sampling effort (Fig. 1B). This cohesive group thus appears widespread across turtle ants. We return to this possibility, in our analysis of amplicon sequence data, below.

Providing insight into divergence from non-symbiotic relatives, we found that the Cv33a 16S rRNA gene shared ~93% nucleotide identity with homologs from *Bordetella hinzi*, *B. holmesii*, and *Pusillimonas noertemannii*. Under a range of molecular clock speeds [92, 93], this suggests that their last common ancestor was relatively ancient. Percent identity to the closest relatives in our 16S rRNA phylogeny (NCBI accession numbers HQ184355 and ABO21352, Fig. 1A) ranged from 92.7–94.4%, falling below the 95% genus-level threshold adopted in prior studies (e.g. [94]). While these unstudied, bacteria may prove the closest relatives of the Cv33a lineage, a sister-taxon status to *Pusillimonas* was supported in our three-protein phylogenetic tree (Fig. S2). Genome-wide pairwise average nucleotide identity (ANI) between Cv33a and this *Pusillimonas* genus fell at 72% (Table S2). These ANIs, in turn, fell just below the mean (74.0%) and median (73.1%) demarcation values among genera [95].



We cultivated and characterized strain Cv33a, isolated from the gut of a *C. varians* ant worker. This strain was vibrioid in shape, in culture, and ranged from 1.4–1.8 µm in length and 0.4–0.6 µm in width (Fig. 1C, D). Its morphology was in line with a morphotype observed and described by Roche and Wheeler (1997) [21] from *C. rohweri* hindguts: a short rod "with wavy outer membranes". Strain Cv33a had a Gram-negative type outer membrane and thin cell wall, as confirmed with light microscopy and Gram staining. We further characterized the biochemical and metabolic traits of this bacterium through additional in vitro assays (see Fig. S3 and Table S3).

Fig. 1 Genetic relationships and morphology of the newly described *Ischyrobacter* clade. A 16S rRNA maximum likelihood phylogeny illustrating the diversity of cultured isolates affiliated with the *Burkholderiales* from turtle ants. Symbiont phylogeny includes 317 cultured isolates with partially sequenced 16S rRNA assigning to the *Burkholderiales* order. These isolates came from 20 colonies spanning 10 species of *Cephalotes* and one species of *Procryptocerus*. Host species of the isolates present at the tips are identified by the strip colors, with correspondence shown in (**B**). Turtle ant-specific clades are highlighted with light blue shaded squares. Most clades that included three or more members of one strain from one host ant species were collapsed to economize space. The identifiers of the strains in collapsed clades can be found in our data repository. These collapsed clades were numbered to ease referencing. The *Ischyrobacter* clade is highlighted with a black border. It includes genome-sequenced isolates Cv33a in clade 14 and CSM3487_49 in clade 13 (see Fig. S2). In addition to cultured isolates, this analysis included closely related sequences from previous culture-independent studies using cloning (after universal primer amplification) or direct sequencing (after amplification with diagnostic PCR)—each of these is associated with an NCBI accession number. Included further were sequences from cultured isolates with sequenced genomes (e.g. "Cv33", "Cag32"), and a subset of the relevant reference sequences affiliated with the *Burkholderiales* from a prior 16S rRNA amplicon sequencing study ([32]; names starting with "OTU"). Sequences that were not derived from turtle ants were identified through BLASTn searches. The tree was rooted using two members of the *Gammaproteobacteria*. Names for turtle ant-derived cultured isolates were after their colony collection codes, and began with "Cs", "CSM", "JDR", "JR" or "SP". **B** Phylogeny of host turtle ants, based off of Price et al. [126]. Color shading given to turtle ant taxa is used for the color strips and circles in (**A**) to identify the host species, when possible (see gray box below phylogeny for exception). Indicated above phylogeny are the numbers of cultured isolates from each species of interest included in the phylogenetic analysis of the *Burkholderiales* (**A**), and, in parentheses, the numbers falling within the *Ischyrobacter* clade. **C**, **D** Transmission electron microscopy images of the Cv33a isolate strain revealing size and shape of *I. davidsoniae* bacterial cells.

Combined genetic and phenotypic divergence from neighboring taxa warrant the classification of strain Cv33a as a new species and genus within the *Burkholderiales* family *Alcaligenaceae*, proposed here as *Ischyrobacter* gen. nov., *davidsoniae* sp. nov. *Ischyrobacter* gen. nov. (Is.chy.ro.bac'ter. Gr. masc. adj. *ischyros*, strong; N.L. masc. n. *bacter*, rod or bacterium; N.L. masc. n. *Ischyrobacter*, a strong bacterium). *Ischyrobacter davidsoniae* sp. nov. (da.vid.so'ni.ae. N.L. gen. n. *davidsoniae*, named in honor of Dr. Diane Davidson). These terms, together, represent the strong contributions made by this bacterium to key aspects of herbivorous ant ecology, and Davidson's innovative research within the fields of ant ecology and symbiosis [22, 26].

Prior sequencing of the *I. davidsoniae* Cv33a isolate genome had revealed its capacity to participate in N-recycling by converting purine metabolites (guanine, hypoxanthine, xanthine, and urate) into urea [9]. Our additional genomic annotations, here, uncovered further metabolic features (Dataset S3), while assessment of two additional genomes from strains in the *Ischyrobacter* clade (*grandinosus*-Bin-5 from *C. grandinosus* and CSM3487_49 from *P.* nr *pictipes*, Fig. S2) revealed that several of these properties are conserved (Dataset S4). We note, further, that the 82.8% ANI between CSM3487_49 and Cv33a (Table S2) is comparable to 83% ANI among divergent *Buchnera* strains, the primary endosymbionts of most aphids, with 50–70 million years of separation [96, 97]. N-recycling conservation in *Ischyrobacter* genomes (Dataset S5) has thus clearly spanned a substantial period of evolutionary history, which is discussed below.

Broad distributions and gut localization among Ischyrobacter symbionts

Using prior 16S rRNA amplicon data, we assessed the prevalence of the monophyletic *I. davidsoniae* lineage across turtle ant species. Toward this end, we first mapped 16S rRNA V4 region reads from a previous study [24] onto our genome-derived 16S rRNA sequences for *I. davidsoniae*. In doing so we identified one OTU (OTU011) showing higher relatedness to our *Ischyrobacter* clade than to any other turtle ant symbiont or any bacterium from NCBI (Dataset S6). While not dominant, this *Ischyrobacter*-matching OTU011 ranked among the top 10 most abundant symbionts in whole worker microbiomes across 13 *Cephalotes* species surveyed in a prior study [24]. This OTU was also recorded in immature (pre-mated) queens and in some middle-to-late stage larvae [24]. Through additional mappings, this *Ischyrobacter* OTU011 was found to be highly similar to 16 amplicon sequence variants (ASVs), sampled from gut tissues of 11 *Cephalotes* species, including six unique to this second amplicon sequence study [18]. Through similar means, *Ischyrobacter* was found in one studied species of *Procryptocerus* [18]. Across the two assessed amplicon sequencing studies, *Ischyrobacter* was, thus, present in all 20 surveyed turtle ant species.

Roughly matching prior patterns at the *Burkholderiales* order level [18], we found the 16 *Ischyrobacter*-mapping ASVs from Flynn et al. [18] to be at low abundance in the crop and midgut, but at high relative abundance in the ileum and rectum (Fig. 2A). Accounts of total symbiont titer across each gut compartment, ascertained through analysis of amplicon sequencing and qPCR, suggested the ileum was the primary site of *Ischyrobacter* colonization (Dataset S7).

FISH microscopy added new precision to these results across gut compartments. Most intriguing was the high concentration of *I. davidsoniae* at the anterior ileum of *C. varians* workers, with close proximity to the pyloric valve (Figs. 2B–D and S4). High abundance at the site of Malpighian tubule-mediated, insect N-waste delivery fit some expectations for this presumed N-recycler [98]. Fluorescent signal for *I. davidsoniae* remained high in the enlarged portion of this ileal gut compartment, found immediately downstream of the pyloric valve, before diminishing toward the medial and posterior segments and then, reappearing in the rectum (Figs. 2B–D and S4). Signal with our generic betaproteobacterial probe revealed some partially distinct localizations, with non-*Ischyrobacter* members of this group being uniquely abundant within the medial ileum (Fig. 2B). Through our efforts here, *Ischyrobacter* becomes the second turtle ant-associated symbiont genus with a well-defined spatial ecology, joining *Cephaloticoccus* (*Verrucomicrobia*: *Opitutales*) [18, 33, 99]. Localization inferences had been more straightforward for this latter urea-recycling symbiont, due to lower diversity, with just a single domestication within the *Opitutales* lineage [24].

In vivo I. davidsoniae metabolism inferred from transcriptomics—amino acids and precursors

To provide insight into the realized function of *I. davidsoniae*, we mapped RNAseq reads from 19 *C. varians* gut metatranscriptome libraries against reference genomes of turtle ant gut symbionts. As described earlier, RNAseq libraries came from adult workers preserved directly from the field or after dietary experiments in the lab. Through these combined efforts, we gained insights into separate symbiont contributions under a range of dietary conditions. A total of 4500 assembled contigs mapped back to the *I. davidsoniae* isolate genome. These contained 1530 unique genes ("unigenes", Dataset S8), suggesting expression of 45.3% of the genome, with an average of 8.76x coverage per gene. Change in diet had minimal detectable impact on *I. davidsoniae* gene expression, since only one unigene was significantly differentially expressed, i.e., between ants on urea-rich versus control diets (Fig. S5).

While we return later for discussion of symbiont-mediated N-recycling gene expression, we first focus on gene expression in amino acid (AA) metabolism, a critical function of the turtle ant

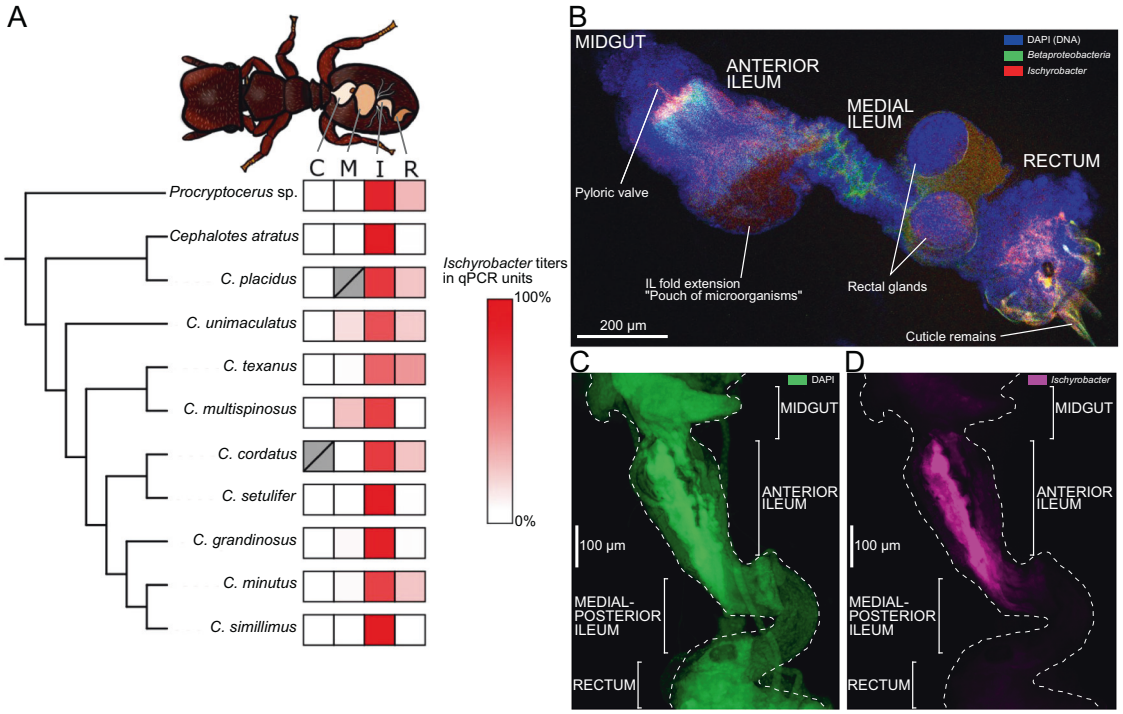


Fig. 2 ***Ischyrobacter*** **localization in turtle ant gut sections. A** Proportion of *Ischyrobacter* populations found in separate gut sections for different turtle ant species across two recent studies [18, 24]. Results combine amplicon sequencing data with qPCR for abundance comparisons, weighting the gut section *Ischyrobacter* relative abundance (amplicon sequencing) by total bacterial absolute abundance recorded in turtle ant species (qPCR). Gray crossed out squares correspond to gut sections with no qPCR data from studies used. For the color-coded squares, "C" is for crop, "M" is for midgut, "I" is for ileum, and "R" is for rectum. **B** FISH microscopy image showing colonization of *I. davidsoniae* in hindgut sections of a *C. varians* worker. Note that the posterior portion of the ileum is hidden by rectal glands. Blue color depicts DNA, green is for *Betaproteobacteria* (including *Ischyrobacter* sp.), and red shows *I. davidsoniae* cells. Uppercase labels correspond to anatomical regions as in Bution and Caetano [19]. **C**, **D** Single color images representing gut colonization by *I. davidsoniae* in another *C. varians* worker. Both panels are the same picture but colorized via distinct fluorescent signal: green for DNA and pink for *I. davidsoniae* cells.

gut microbe symbiosis [27]. Our transcriptomic analyses revealed "Amino Acid Metabolism" to be the second most highly expressed KEGG functional category by *I. davidsoniae* in vivo (9.8% of normalized TPMs on average, Fig. S6). While past genomic studies have emphasized the importance of symbiont-encoded AA biosynthesis [9, 28], we found that *I. davidsoniae* expressed genes not only to synthesize these compounds, but to also catabolize them. In vitro metabolic assays showed that four polar AAs, including one essential AA (L-histidine) and three non-essential AAs (L-aspartate, L-glutamate, and L-serine), supported Cv33a strain growth, among the twelve AAs examined (Table S3). In vivo gene expression, accordingly, suggested use of these four polar AAs to synthesize several other proteinogenic AAs, including two essential AAs (L-methionine and L-threonine) and six non-essential AAs (glycine, L-alanine, L-arginine, L-cysteine, L-glutamine, L-methionine, and L-threonine) (Figs. 3 and S7). *I. davidsoniae* also expressed genes to export branched-chain AAs or their precursors (Fig. S8). Despite encoding a shikimate pathway (Figs. 3 and S7), this symbiont did not express this pathway in full under the studied conditions, raising questions on its importance at earlier life stages [100], such as the callow worker stage, when cuticle formation is an active, intensive process [23].

In vivo I. davidsoniae metabolism inferred from transcriptomics—cofactor, vitamin, and energy metabolism

Rivaling the expression of AA metabolism genes was the expression of genes shaping energy, cofactor, and B-vitamin metabolism. Indeed, 42.1% of total normalized TPM *I. davidsoniae* expression, on average, was assigned to the KEGG category "Metabolism of Cofactors and Vitamins", and 7.4% was assigned to "Energy Metabolism" (Fig. S6). Genes for eight subunits of the large NADH-quinone oxidoreductase complex were expressed by *I. davidsoniae* in vivo (Figs. 3 and S9), indicating a likely functional complex to oxidize NADH to NAD^+ in the generation of ATP. The gene encoding NuoN, a subunit of this complex, was the third most highly expressed KO encoded by this symbiont (5039 normalized TPMs total for K00343, Dataset S8), and the only one to be significantly differentially expressed between ants fed control and urea-rich diets (Fig. S5).

Also related to NADH/NAD^+ is the enzyme nicotinate phosphoribosyltransferase, a function encoded by the *pncB* gene (K00763), which aids in converting nicotinate to NAD^+. This gene was the most highly expressed KO function in vivo, for *I. davidsoniae* (33163 normalized TPMs total for this KO). Another highly expressed gene, *ribF* (K11753), is involved in the utilization of a B-vitamin to produce an enzymatic co-factor. Specifically, this gene codes for riboflavin kinase/FMN adenylyltransferase, which transforms riboflavin into FAD. The assigned *ribF* KO function ranked fourth among those expressed by *I. davidsoniae* (4381 normalized TPMs).

Only portions of the TCA cycle were expressed in vivo (Figs. 3 and S9), despite representation of the complete TCA cycle in the *I. davidsoniae* genome. However, expression results suggested that several key metabolites in the TCA cycle could be produced through other pathways, including the catabolism of aromatic AAs (Fig. S7), other aromatic compounds (e.g. benzoate, vanillate derivatives), or secondary metabolites (e.g. 3-oxopimeloyl-CoA, salicylate) (Fig. S9). Genes coding for transport of aromatic compounds may also have been expressed by this symbiont because some of the *liv* unigenes (presumably for branched-chain AA transport) were annotated as benzoate transporter genes by alternative databases (Fig. S8 and Dataset S8). These combined patterns imply the use of aromatic compounds as a carbon source,

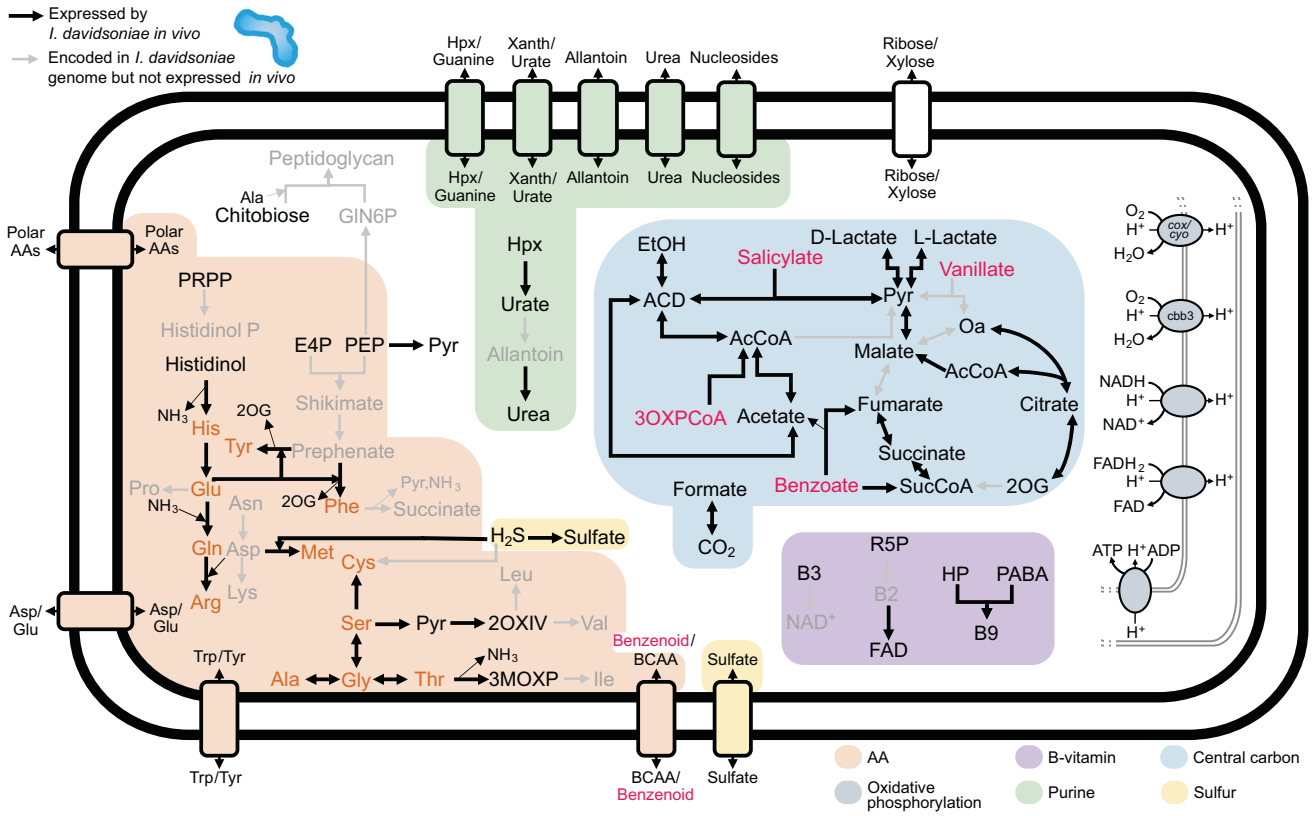


Fig. 3 Summary of *Ischyrobacter davidsoniae* metabolism. Black arrows depict pathways expressed by *I. davidsoniae* in vivo, while gray arrows show pathways that are encoded in its genome but not fully expressed by this symbiont in any of our 19 metatranscriptomic libraries. These metatranscriptomes derived from *C. varians* adult workers that were either directly preserved from the field or fed with different diets enriched in pollen, cellobiose, or urea. Metabolites with gray font indicate that the genes to complete their synthesis are not expressed in vivo. Orange font corresponds to amino acids that could be synthesized through *I. davidsoniae* expressed genes, whereas pink font represents aromatics (i.e., "benzenoids") that could be degraded by *I. davidsoniae*. Orange background is for amino acid metabolism, purple background is for B-vitamin metabolism, blue background is for central carbon metabolism, blue-gray background is for oxidative phosphorylation (depicted with mitochondrial membranes), green background is for purine metabolism and N-recycling, and yellow background is for sulfur metabolism. "2OG" = 2-oxoglutarate, "2OXIV" = 2-oxoisovalerate, "3MOXP" = 3-methyl-2-oxopentanoate, "3OXPCoA" = 3-oxopimeloyl-coenzyme A, "AA" = amino acid, "AcCoA" = acetyl coenzyme A, "ACD" = acetaldehyde, "Ala" = L-alanine, "Arg" = L-arginine, "Asn" = L-asparagine, "Asp" = L-aspartate, "B2" = riboflavin, "B6" = nicotinate, "B9" = folate, "BCAA" = branched-chain amino acid, "Cys" = L-cysteine, "E4P" = D-erythrose 4-phosphate, "EtOH" = ethanol, "FAD" = flavin adenine dinucleotide, "Gln" = L-glutamine, "GlN6P" = N-acetylglucosamine 6-phosphate, "Glu" = L-glutamate, "Gly" = glycine, "H_2S" = hydrogen sulfide, "His" = L-histidine, "Histidinol P" = histidinol phosphate, "HP" = 6-hydroxymethyl-7,8-dihydropterin, "Hpx" = hypoxanthine, "Ile" = L-isoleucine, "Leu" = L-leucine, "Lys" = L-lysine, "Met" = L-methionine, "NH_3" = ammonia, "Oa" = oxaloacetate, "PABA" = 4-amino benzoate, "PEP" = phosphoenolpyruvate, "Phe" = L-phenylalanine, "PRPP" = 5-phosphoribosyl diphosphate, "Pyr" = pyruvate, "R5P" = ribulose 5-phosphate, "Ser" = L-serine, "SucCoA" = succinyl coenzyme A, "Thr" = L-threonine, "Trp" = L- tryptophan, "Tyr" = L-tyrosine, "Val" = L-valine, "Xanth" = xanthine.

with influx into energy-generating pathways. Such aromatic molecules may be obtained from plant-derived lignins, flavonoids, or plant defensive compounds [101, 102]. The abilities of *I. davidsoniae* to scavenge such aromatics is unique among ant-associated symbionts, though similar to findings from termite and desert locust symbioses [103, 104].

Our final focus outside of the N-recycling category includes two components of sulfur metabolism—an overlooked element of terrestrial animal symbioses (but see [42, 105])—with joint implications for AA and energy metabolism. First among these were assimilatory capacities to import and reduce sulfate to hydrogen sulfide. Being encoded by the *Blochmannia* symbionts of carpenter ants [106], nearly all steps required for such capacities were expressed by *I. davidsoniae* (Fig. S9). Beyond this, *I. davidsoniae* expressed genes involved in hydrogen sulfide oxidation, producing sulfate in two energy-generating steps. Biological sulfide oxidation is typically conducted in oxic-anoxic transition zones (high-to-zero O_2 concentrations) with high levels of hydrogen sulfide [107]. While it is not yet known whether anoxic conditions are common in turtle ant guts, our in vitro assays showed that *I. davidsoniae* Cv33a could only be grown under oxic to microoxic conditions (Fig. S3), raising questions on the context for sulfide-oxidizing gene expression.

Ischyrobacter N-recycling inferred from genomics and transcriptomics

Prior work has demonstrated that the urea-recycling activities of gut symbionts support free AA pools and cuticle synthesis of adult *Cephalotes* ants [27]. These symbionts gain further cuticular and N economy relevance from their localization at the site of host N waste delivery (Fig. 2A, B) and their genome-encoded urea-generating abilities [9]. Our genomic analyses revealed here that urea-generating metabolisms are encoded in the two *Ischyrobacter* isolates with sequenced genomes: Cv33a and CSM3487_49 (Fig. 4A and Dataset S5). The partial, binned, *Ischyrobacter* genome from *C. grandinosus* encodes at least a portion of this same pathway, as well. Thus, while symbiont roles can change after domestication (e.g. [108–111]), we see little evidence for such change to longstanding capacities for *Ischyrobacter* N-recycling. Future studies on this possibility are of clear interest.

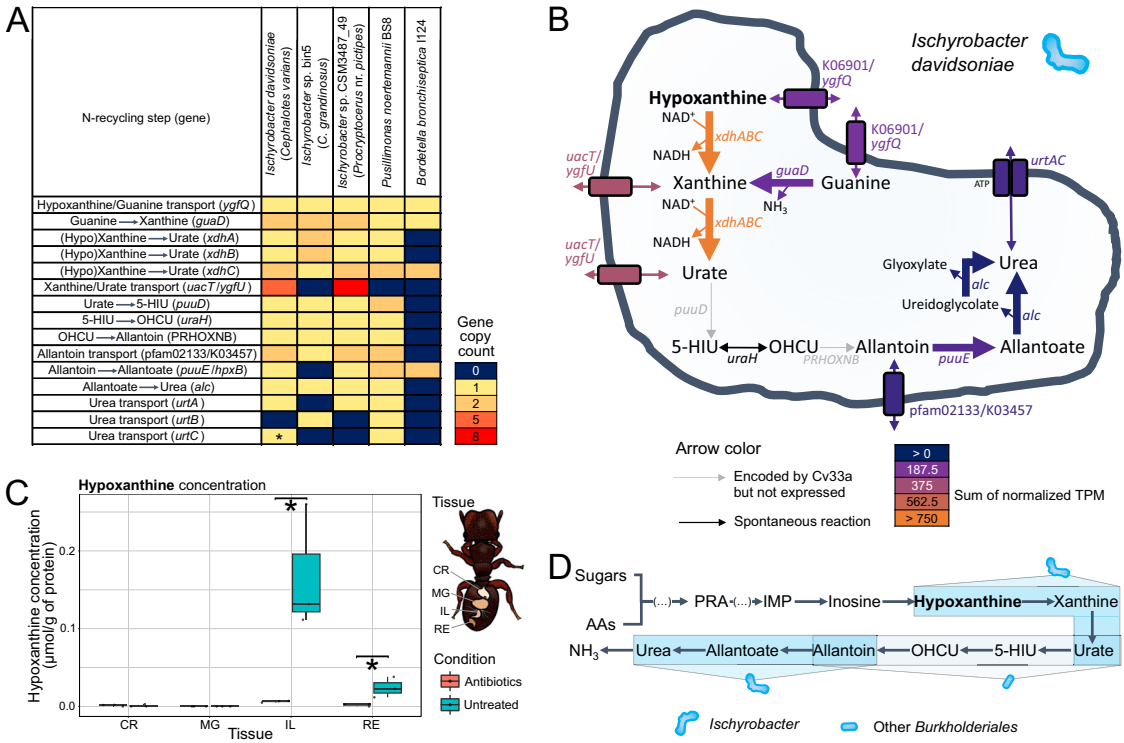


Fig. 4 Nitrogen (N)-recycling by *Ischyrobacter* symbionts inferred from multi-omic data. A N-recycling steps encoded in the genome of *Ischyrobacter* symbionts and the closest relative bacterial species. Values in the heatmap correspond to the number of genes coding for specific steps of the N-recycling pathway. For each column, host species is indicated between parentheses. The asterisk refers to a gene that was not found in the Cv33a isolate genome, but in the metagenome-assembled genome Bin3.1 from the *C. varians* PL010 gut metagenome, which is almost identical to the Cv33a isolate genome [28]. **B** Expression level of *I. davidsoniae* genes that code for N-recycling pathway steps. Arrow color depicts the sum of normalized transcript per million (TPM) for individual enzyme/transporter subunits across the 19 metatranscriptome libraries. For each function, the most highly expressed subunits are shown. "5-HIU" = 5-hydroxyisourate, "OHCU" = 2-oxo-4-hydroxy-4-carboxy-5-ureidoimidazoline. **C** Concentration of hypoxanthine, as measured through in vivo metabolomics across gut sections in antibiotic-treated (red) and control (blue) ants. Gut sections are as follows: "CR" = crop, "MG" = midgut, "IL" = ileum, and "RE" = rectum. Asterisks indicate significant differences between the two conditions for metabolite concentrations in given gut sections, at FDR-adjusted $p < 0.1$, tested via emmeans of the interaction Conditions:Tissue based on a linear mixed model. **D** Overview of the focal N-recycling pathway. Vibrioid-shaped bacterial cells correspond to *I. davidsoniae* and its expression, while the rod is for another turtle ant-associated taxon affiliated with the *Burkholderiales*. "AAs" = amino acids, "IMP" = inosine 5'-monophosphate, "NH_3" = ammonia, "PRA" = 5-phosphoribosylamine.

To test whether the urea-generating, purine-catabolizing genes of *Ischyrobacter* are expressed in mature adult workers, we examined transcripts from this 7-step pathway across our 19 RNAseq libraries. We first note that *I. davidsoniae* was one of four or more symbiont lineages expressing transporters for hypoxanthine or guanine (K06901/*ygfQ*, Fig. S10)—purine precursors within the urea-generating pathway. This symbiont was unique in its expression of the *guaD* gene converting guanine to xanthine, and of the *xdh* genes encoding the multi-subunit xanthine dehydrogenase known to yield urate from hypoxanthine or xanthine (step 3, Fig. S10). The *xdh* genes encoding these three subunits were among the most highly expressed *I. davidsoniae* unigenes (Fig. 4B), with a sum of normalized TPMs >45 for each (Dataset S8). At the top of this list was a single unigene encoding the XdhB subunit, which was the sixth most abundant in the *I. davidsoniae* Cv33a draft transcriptome (sum of normalized TPM = 3160).

Despite the above evidence that *I. davidsoniae* produces urate (i.e., uric acid) in vivo, we found no evidence for expression of this symbiont's *puuD* uricase (Fig. 4B). Assays performed on gut-isolated *I. davidsoniae* Cv33a further suggested that urate was not used under the studied in vitro conditions. Not only was uricase activity negligible in these cultures, but urate did not increase *I. davidsoniae* growth when added to growth media (Fig. S11). Expression of other N-recycling genes, but not *puuD* uricase, is consistent with the N-recycling gene architecture within the Cv33a genome. Specifically, while this symbiont possesses most such genes in proximity on a single gene cluster, *puuD* was found on the antisense DNA strand [9], suggesting separate gene regulation and distinct contexts for expression. It is thus plausible that this *puuD* uricase gene is transcribed by *I. davidsoniae* when ants consume foods with high levels of urate or similar molecules (e.g. reptile urine [112]). It is also possible that expression occurs when *I. davidsoniae* colonizes larvae, in young queens or soldiers, or during early stages of colony lifespan.

Fitting with its lack of in vivo uricase-encoding gene expression, *I. davidsoniae* expresses genes coding for a transporter enabling cross-membrane passage of urate or xanthine (UacT/YgfU, Fig. 4B). Here we propose that it exports the former molecule into the gut lumen, where it is acquired by other symbionts for further metabolism. Supporting this latter supposition was our finding that three or more non-*Ischyrobacter* symbionts from the *Burkholderiales* expressed *puuD* uricase (step 4, Fig. S10) and PRHOXNB OHCU decarboxylase genes, filling this "gap" within the urea-generating pathway to make allantoin from urate. Then, *I. davidsoniae* appears likely to import allantoin, given its expression of the pfam02133/K03457 transporter across the studied adult workers (Fig. 4B). This symbiont further transcribed the genes to produce urea (Figs. 4B and S10), matching in vitro observations of allantoin-based urea production ([9], Fig. S12).

With evidence that *I. davidsoniae* and other symbionts collectively express all necessary steps for the uricolytic, urea-

generating pathway, our data suggest N-recycling to be a multi-symbiont, collaborative activity. Given the further expression of urea transporter genes (*urtA* and *urtC* transcripts) by *I. davidsoniae*, and of urease genes by *Cephaloticoccus* and symbionts from the *Rhizobiales* order (Fig. S10), we see a fulfilled potential for collaborative symbiont derivation of NH_3. *I. davidsoniae* expressed genes for NH_3 assimilation into AAs via glutamine synthetase ("GS"; *glnA*, Fig. S6), while other symbionts encode GS/GOGAT genes as well [9], hinting at combined N-assimilation efforts with further impact on ant AA budgets.

Multi-omic insights into the N-recycling collaboration—symbiont origins of I. davidsoniae-utilized hypoxanthine

With transcriptional evidence elucidating symbiont roles in turtle ant N-recycling, our data provide further support for Davidson's hypothesis on the enrichment of N-provisioning symbioses among herbivorous ants [26]. We next turned to an in vivo metabolomic analysis to further study host- and symbiont-impacted N-recycling, assessing whether antibiotic treatment, that induced drops in symbiont titers (Dataset S1), impacts the concentration of several purines in different turtle ant gut compartments. We note first that absolute measures of urate, allantoin, and allantoate, which constitute common insect N-wastes [113], were not possible due to the absence of standards for these three compounds in our LC-MS runs. Relative measures obtained here (Fig. S14), instead provide an interesting start for future study. However, absolute measures were obtained for several purine precursors of urate and its derivatives. Among these were guanine and xanthosine, found sporadically, at low concentrations, across the crop, midgut, ileum, and rectum of *C. varians* workers (Fig. S13). More abundant, with intriguing differences across gut compartments and under antibiotic treatment, were adenine, adenosine monophosphate (AMP), guanosine, hypoxanthine, inosine, inosine monophosphate (IMP), and xanthine (Figs. 4C, 5A and S13). Across the detected molecules, only adenine and guanosine were found in the provided ant diet, although, at very low concentrations (<0.2 µM), suggesting purines in *C. varians* guts came from ant or bacterial metabolism.

Of central importance to *Ischyrobacter* metabolism (Fig. 4A, B) were findings that xanthine and hypoxanthine were the most abundant purines across the ileal and rectal gut compartments of symbiont-housing, control ants (Figs. 4C and S13). But in symbiont-suppressed, antibiotic-treated-individuals, hypoxanthine was virtually absent (antibiotic versus control in ileum: emmeans, estimate = −1.41, *z*-ratio = −3.19, *p* adjusted <0.05; antibiotic versus control in rectum: emmeans, estimate = −1.17, *z*-ratio = −2.66, *p* adjusted <0.05, Fig. 4C). Xanthine levels were also strongly reduced under antibiotic treatment (Fig. S13 and Dataset S9). These data suggest that symbiotic gut bacteria are a major source of both metabolites.

Given the high abundance of *I. davidsoniae* in both the ileum and rectum (Fig. 2A, B), and its expression of genes to import and convert hypoxanthine into urate (Fig. 4B), symbiont-produced hypoxanthine may be a major fuel for this part of *I. davidsoniae* metabolism. To identify hypoxanthine producers, we broadened our study of the purine metabolic pathway, noting first that hypoxanthine can be directly produced from inosine in one enzymatic step. Our metabolomic data showed inosine to be the most abundant midgut purine within symbiont-housing ants, after AMP and IMP (Figs. 5A and S13). But antibiotic treatment eliminated nearly all midgut inosine (emmeans, estimate = −1.54, *z*-ratio = −6.71, *p* adjusted <0.001, Fig. 5A and Dataset S9). In contrast to these results were findings from the ileum where inosine levels were elevated by antibiotic treatment (emmeans, estimate = 0.66, *z*-ratio = 2.87, *p* adjusted <0.05, Fig. 5A and Dataset S9). Symbionts would thus appear to enact net inosine production within the midgut, and net inosine consumption

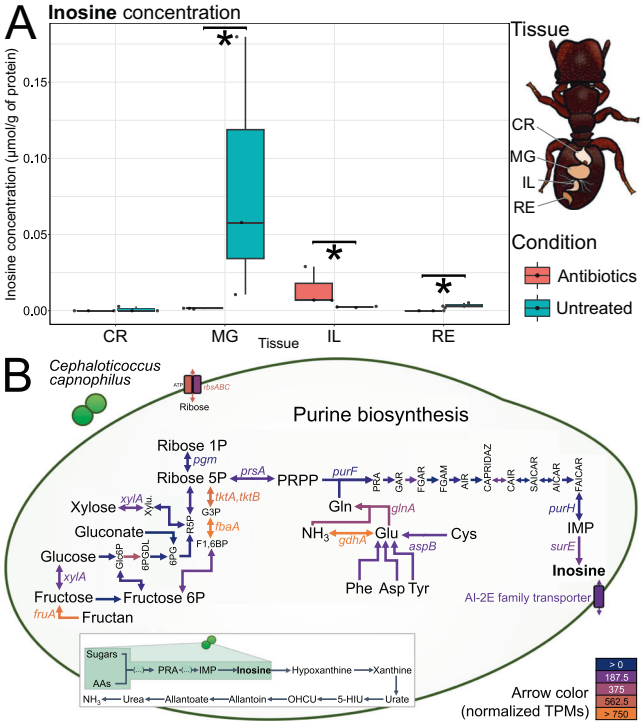


Fig. 5 Inosine production by *Cephaloticoccus capnophilus* in the midgut. A Concentration of inosine, as measured through in vivo metabolomics across gut sections in antibiotic-treated (red) and control (blue) ants. Gut sections are as follows: "CR" = crop, "MG" = midgut, "IL" = ileum, and "RE" = rectum. Asterisks indicate significant differences between the two conditions for metabolite concentrations in given gut sections, at FDR-adjusted $p < 0.1$, tested via emmeans of the interaction Conditions:Tissue based on a linear mixed model. **B** Purine biosynthesis pathway and associated transporters encoded and expressed by the midgut dominant *Cephaloticoccus capnophilus* in vivo. Arrow color shows, for the highest *Cephaloticoccus* expressed genes, their transcript per million (TPM), averaged across the 19 metatranscriptome libraries. "P" is the abbreviation for phosphate, "6PG" = 6-phospho-D-gluconate, "6PGDL" = 6-phospho-D-glucono-1,5-lactone, "AI-2E" = Autoinducer-2 exporter, "AICAR" = 5-aminoimidazole-4-carboxamide ribotide, "AIR" = aminoimidazole ribotide, "Asp" = L-aspartate, "CAIR" = 1-(5-phospho-D-ribosyl)-5-amino-4-imidazolecarboxylate, "CAPRIDAZ" = 5-carboxyamino-1-(5-phospho-D-ribosyl)imidazole, "Cys" = L-cysteine, "F1,6BP" = β-D-fructose 1,6-bisphosphate, "FAICAR" = 5′-phosphoribosyl-5-formamido-4-imidazolecarboxamide, "FGAM" = 5′-phosphoribosyl-N-formylglycinamidine, "FGAR" = N-formylglycinamide ribonucleotide, "G3P" = glyceraldehyde 3-phosphate, "GAR" = glycinamide ribonucleotide, "Glc6P" = β-D-glucose 6-phosphate, "Gln" = L-glutamine, "Glu" = L-glutamate, "IMP" = inosine 5′-monophosphate, "NH_3" = ammonia, "Phe" = L-phenylalanine, "PRA" = 5-phosphoribosylamine, "PRPP" = 5-phosphoribosyl diphosphate, "R5P" = ribulose 5-phosphate, "SAICAR" = 1-(5′-phosphoribosyl)-5-amino-4-(N-succinocarboxamide)-imidazole, "Tyr" = L-tyrosine, "Xylu." = xylulose.

downstream within the ileum, where *Ischyrobacter* would utilize the derived hypoxanthine (Fig. 6).

Metatranscriptomics allowed an independent test of these metabolome-inferred hypotheses, helping to identify inosine producers and the inosine consumers presumably using this molecule to make hypoxanthine. We first note that several symbionts with known abundance in the ileum expressed the 5′-nucleotidase *surE* gene to convert IMP into inosine (Fig. S10). This contingent includes *Cephaloticoccus*, cospeciating symbionts living within both the midgut and ileum [18, 33, 99]. Transcripts mapping to the *Cephaloticoccus capnophilus* reference genome (strain Cv41) revealed full pathway expression for de novo inosine

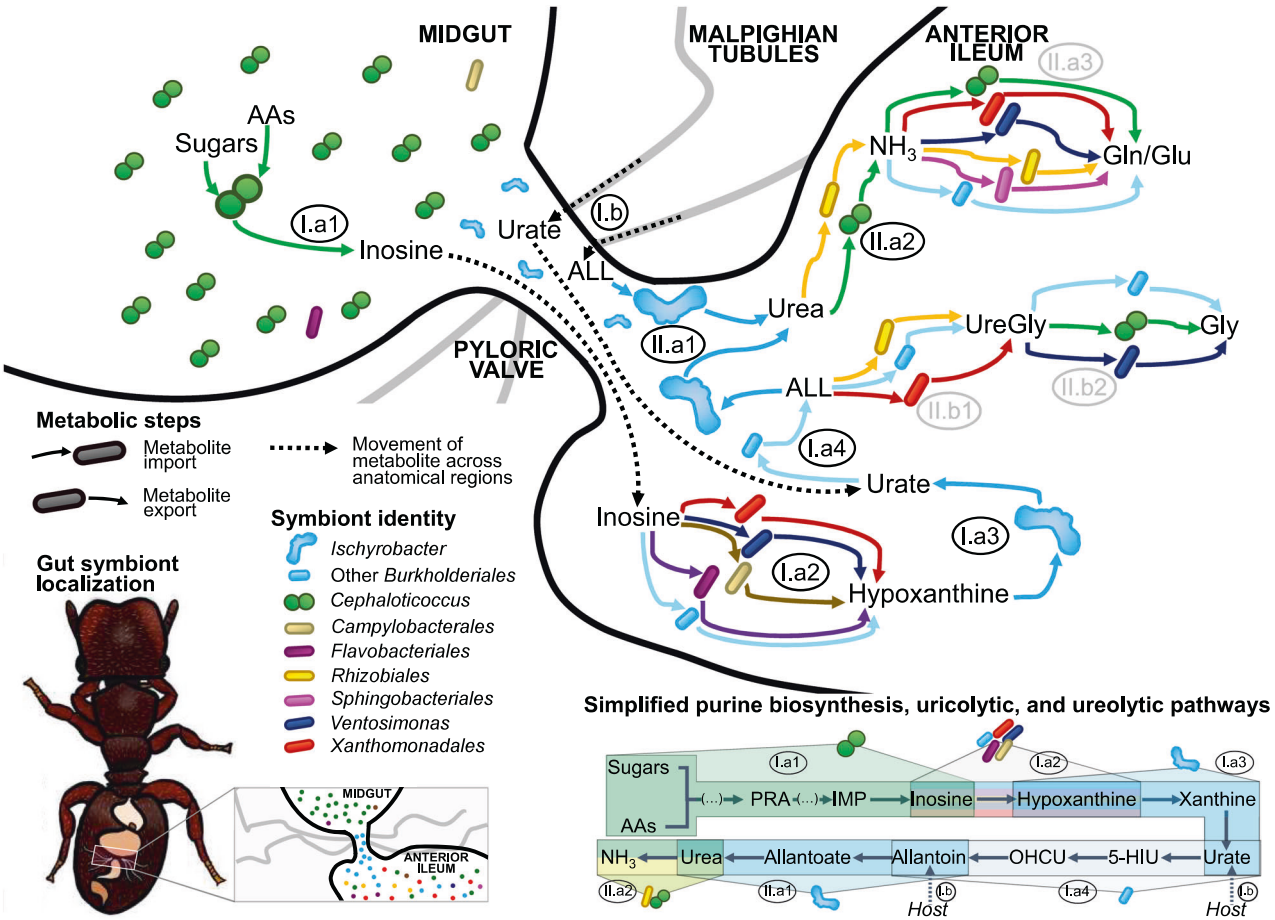


Fig. 6 Proposed model on purine metabolism and N-recycling in symbiont communities of turtle ant mature workers' gut. Metabolic pathways shown here take place at the junction between the midgut and the anterior ileum of mature turtle ant workers, but note that the metabolic functions depicted in the anterior ileum may also be performed by the same symbionts in other regions of the hindgut (e.g. medial ileum, posterior ileum, rectum). The pyloric valve connects midgut and anterior ileum, and is the site of host-derived N-waste release from the Malpighian tubules to the gut lumen. Dashed arrows represent movements of metabolites from one anatomical region to another. Colored arrows correspond to symbionts expressing metabolic steps in vivo. The taxonomic order or genus of bacterial symbionts and their gut position are shown in the key on the left. Metabolic steps are labeled as follows: "I" is for the production of allantoin, performed either by symbionts ("I.a") or by the host ("I.b"). "II" is for the utilization of allantoin to make new amino acids via two different pathways: through urea ("II.a": Gln/Glu urease pathway), depicted in the pathway diagram at the bottom right, or through ureidoglycine ("II.b": Gly ureidoglycine pathway, see [28]). "AAs" = amino acids, "ALL" = allantoin, "Gly" = glycine, "NH_3" = ammonia, "UreGly" = ureidoglycine.

synthesis from sugars and AAs (Fig. 5B). A role for this symbiont in inosine provisioning was further supported by its expression of genes encoding an autoinducer-2 exporter family protein with known roles in purine biosynthesis [114, 115], and therefore, a potential role in inosine export. With this symbiont being, by far, the most abundant microbe within the midgut [18], we argue that it is likely the primary source of midgut inosine (Figs. 5B and 6).

Following inosine synthesis and release into the midgut, this purine might then transit to the ileum where it could further be processed into hypoxanthine, through the single, aforementioned step encoded by the *punA* purine phosphorylase gene. Our metatranscriptomic analyses support this possibility, with *punA* expression by symbionts known from the ileum, assigning to the orders *Burkholderiales* (strain Cv52), *Campylobacterales*, *Flavobacteriales*, *Pseudomonadales* (*Ventosimonas gracilis*), and *Xanthomonadales* (Fig. S10) [18]. Hence, with hypoxanthine being a plausible major input for *Ischyrobacter*-assisted urea-generation (Fig. 4D), these collective results suggest a multipartite, symbiont-autonomous means for turtle ant N-recycling (Fig. 6). The influence of these activities on hosts, and the relative importance of symbiont- versus host- or diet- derived purine wastes in the turtle ant N-economies remain to be studied.

CONCLUSION

Through microscopy, in vitro assays, phylogenetics, and multi-omics we characterize an N-recycling symbiont lineage from the newly described genus *Ischyrobacter* and describe its functional integration within a diverse gut symbiont community. While less abundant than some members of the turtle ant gut community [32], *Ischyrobacter* symbionts appear ubiquitous within the ileum of adult workers from these neotropical ants. Such high prevalence is likely supported by vertical transfer through colony founding queens, based on recent findings from pre-mated queen microbiomes [24]. While we cannot rule out some horizontal transfer, sequence divergence among the Cv33a and CSM3487_49 *Ischyrobacter* strains suggests fairly ancient common ancestry. This raises the possibility that fidelity-enforcing mechanisms, like queen-based vertical transmission, have assisted the long-term maintenance of this symbiont, resembling findings from other socially transmitted symbioses, like those in termites and honeybees [29, 116].

N-recycling is a key component of several symbioses within detritus- and plant-feeding insects, including coleopterans, dipterans, hemipterans, hymenopterans, and isopterans. Achieved through single-species mechanisms in some of these insects,

Ischyrobacter's contributions to N-recycling appear partitioned, adding to growing evidence for multi-species N-recycling systems [15]. Specifically: (1) the positioning of *I. davidsoniae* at a site of host-waste N delivery, and (2) the findings that bacteria in the *C. varians* ileum synthesize the hypoxanthine and allantoin used for *Ischyrobacter*-based urea generation, place this symbiont as a central team player in a collaborative N-recycling symbiosis (Fig. 6). Adding to this is the apparent midgut, *Cephaloticoccus*-derived origin of the inosine precursor for hypoxanthine. While some N-recycling functions of *Ischyrobacter* (and *Cephaloticoccus*) appear conserved, such conservation, and the broader impacts of *Ischyrobacter* on colony-wide N-economies require further study.

Altogether, our results reveal that urea production by ileum- or rectum-dwelling *I. davidsoniae* need not fully rely on diet or host N-waste delivery [113], but can instead (or additionally) extend from the complementary purine metabolism by other gut symbionts. Microbes often evolve in a community context. Selection to alleviate metabolic burden [117, 118] may combine with deletional mutation biases [119] to yield metabolic complementarity among recurring community members [120], or among such members and their hosts. Such patterns and processes would conform the Black Queen Hypothesis [121], and our expression data have bearing on the fit of this hypothesis for the turtle ant system. Specifically, while prior metagenome data suggested some metabolic complementarity within the turtle ant gut microbiome, partial N-recycling redundancies [9] failed to fully support the Black Queen Hypothesis, raising questions on confounding factors of spatial symbiont ecology or unforeseen tendencies toward competition [122, 123]. The reasons for redundancy and non-redundancy in this system remain unclear [124], and it will be crucial to understand the evolution of gene loss and differential gene expression across the history of this symbiosis.

Retention of genes going unexpressed in our studied context raises further questions of particular intrigue. In social insect colonies, symbiont function may be partitioned across caste, development, and colony life cycle [24]. It is still unknown whether *Ischyrobacter* functions found silent in mature adult workers play some role in callow workers or larvae, and whether they matter in soldiers, queens, or other stages of colony development. Recognizing these other contexts, we might further ask how symbiont activities in mature adult workers shape colony-level fitness, and how this extends from recycling-mediated impacts on such ants' AA budgets. Adding to the complexity of multipartite symbioses existing already in non-social insects [125], these questions suggest important targets for future study across the N-recycling pathway and beyond.

DATA AVAILABILITY

The *Ischyrobacter davidsoniae* Cv33a genome (Project ID: Gp0110144) and the *Ischyrobacter* sp. CSM3487_49 genome (Project ID: Gp0624441) are available from the IMG/M-ER website (https://img.jgi.doe.gov/) by searching the project IDs. The *Ischyrobacter* sp. bin5 draft genome from *C. grandinosus* is accessible on the IMG/M-ER *C. grandinosus* metagenome page (Project ID: Gp0125967), by selecting "Metagenome Bins" in the "Statistics" tab. The *Cephaloticoccus capnophilus* genome (Project ID: Gp0110136) was downloaded from the IMG/M-ER website by searching the project ID. All 16S amplicon sequencing data are from Flynn et al. [18], https://github.com/peterjflynn/cephalotes_gut_localization and Hu et al. [24], https://doi.org/10.5061/dryad.kwh70rz5d, and have been deposited in the GenBank Short Read Archive under BioProject PRJNA668764 at https://www.ncbi.nlm.nih.gov/bioproject/668764 and BioProject PRJNA767930 at https://www.ncbi.nlm.nih.gov/bioproject/767930, respectively. Supplementary data can be found in our data repository at: https://doi.org/10.6084/m9.figshare.21989537.

REFERENCES

1. Klepzig KD, Adams AS, Handelsman J, Raffa KF. Symbioses: a key driver of insect physiological processes, ecological interactions, evolutionary diversification, and impacts on humans. Environ Entomol. 2009;38:67–77.
2. Dale C, Moran NA. Molecular interactions between bacterial symbionts and their hosts. Cell. 2006;126:453–65.
3. Salem H, Kaltenpoth M. Beetle–bacterial symbioses: endless forms most functional. Annu Rev Entomol. 2022;67:201–19.
4. Sudakaran S, Kost C, Kaltenpoth M. Symbiont acquisition and replacement as a source of ecological innovation. Trends Microbiol. 2017;25:375–90.
5. Moreau CS. Symbioses among ants and microbes. Curr Opin Insect Sci. 2020;39:1–5.
6. Russell JA, Sanders JG, Moreau CS. Hotspots for symbiosis: function, evolution, and specificity of ant-microbe associations from trunk to tips of the ant phylogeny (Hymenoptera: Formicidae). Myrmecol News. 2017;24:43–69.
7. Sanders JG, Lukasik P, Frederickson ME, Russell JA, Koga R, Knight R, et al. Dramatic differences in gut bacterial densities correlate with diet and habitat in rainforest ants. Integr Comp Biol. 2017;57:705–22.
8. Feldhaar H, Straka J, Krischke M, Berthold K, Stoll S, Mueller MJ, et al. Nutritional upgrading for omnivorous carpenter ants by the endosymbiont *Blochmannia*. BMC Biol. 2007;5:48.
9. Hu Y, Sanders JG, Łukasik P, D'Amelio CL, Millar JS, Vann DR, et al. Herbivorous turtle ants obtain essential nutrients from a conserved nitrogen-recycling gut microbiome. Nat Commun. 2018;9:964.
10. Bisch G, Neuvonen M-M, Pierce NE, Russell JA, Koga R, Sanders JG, et al. Genome evolution of Bartonellaceae symbionts of ants at the opposite ends of the trophic scale. Genome Biol Evol. 2018;10:1687–704.
11. Jackson R, Monnin D, Patapiou PA, Golding G, Helanterä H, Oettler J, et al. Convergent evolution of a labile nutritional symbiosis in ants. ISME J. 2022;16:2114–22.
12. Klein A, Schrader L, Gil R, Manzano-Marín A, Flórez L, Wheeler D, et al. A novel intracellular mutualistic bacterium in the invasive ant *Cardiocondyla obscurior*. ISME J. 2016;10:376–88.
13. Neuvonen M-M, Tamarit D, Näslund K, Liebig J, Feldhaar H, Moran NA, et al. The genome of Rhizobiales bacteria in predatory ants reveals urease gene functions but no genes for nitrogen fixation. Sci Rep. 2016;6:39197.
14. Rubin BER, Kautz S, Wray BD, Moreau CS. Dietary specialization in mutualistic acacia-ants affects relative abundance but not identity of host-associated bacteria. Mol Ecol. 2019;28:900–16.
15. Hansen AK, Pers D, Russell JA. Chapter Five—Symbiotic solutions to nitrogen limitation and amino acid imbalance in insect diets. In: Oliver KM, Russell JA (eds). Advances in insect physiology, vol. 58. 1st ed. Academic Press, Cambridge, 2020. pp 161–205.
16. Estes AM, Hearn DJ, Agrawal S, Pierson EA, Dunning Hotopp JC. Comparative genomics of the *Erwinia* and *Enterobacter* olive fly endosymbionts. Sci Rep. 2018;8:15936.
17. Sanders JG, Powell S, Kronauer DJC, Vasconcelos HL, Frederickson ME, Pierce NE. Stability and phylogenetic correlation in gut microbiota: lessons from ants and apes. Mol Ecol. 2014;23:1268–83.
18. Flynn PJ, D'Amelio CL, Sanders JG, Russell JA, Moreau CS. Localization of bacterial communities within gut compartments across *Cephalotes* turtle ants. Appl Environ Microbiol. 2021;87:e02803–20.
19. Bution ML, Caetano FH. Ileum of the *Cephalotes* ants: a specialized structure to harbor symbionts microorganisms. Micron. 2008;39:897–909.
20. Bution ML, Caetano FH. The midgut of *Cephalotes* ants (Formicidae: Myrmicinae): ultrastructure of the epithelium and symbiotic bacteria. Micron. 2010;41:448–54.
21. Roche RK, Wheeler DE. Morphological specializations of the digestive tract of *Zacryptocerus rohweri* (Hymenoptera: Formicidae). J Morphol. 1997;234:253–62.
22. Cook SC, Davidson DW. Nutritional and functional biology of exudate-feeding ants. Entomol Exp Appl. 2006;118:1–10.
23. Lanan MC, Rodrigues PADP, Agellon A, Jansma P, Wheeler DE. A bacterial filter protects and structures the gut microbiome of an insect. ISME J. 2016;10:1866.
24. Hu Y, D'Amelio CL, Béchade B, Cabuslay CS, Łukasik P, Sanders JG, et al. Partner fidelity and environmental filtering preserve stage-specific turtle ant gut symbioses for over 40 million years. Ecol Monogr. 2023;93:e1560.
25. Nalepa CA. Origin of mutualism between termites and flagellated gut protists: transition from horizontal to vertical transmission. Front Ecol Evol. 2020;8:14.
26. Davidson DW, Cook SC, Snelling RR, Chua TH. Explaining the abundance of ants in lowland tropical rainforest canopies. Science. 2003;300:969–72.
27. Duplais C, Sarou-Kanian V, Massiot D, Hassan A, Perrone B, Estevez Y, et al. Gut bacteria are essential for normal cuticle development in herbivorous turtle ants. Nat Commun. 2021;12:676.
28. Béchade B, Hu Y, Sanders JG, Cabuslay CS, Łukasik P, Williams BR, et al. Turtle ants harbor metabolically versatile microbiomes with conserved functions across development and phylogeny. FEMS Microbiol Ecol. 2022;98:fiac068.
29. Nalepa CA. Origin of termite eusociality: trophallaxis integrates the social, nutritional, and microbial environments. Ecol Entomol. 2015;40:323–35.

30. Meurville MP, LeBoeuf AC. Trophallaxis: the functions and evolution of social fluid exchange in ant colonies (Hymenoptera: Formicidae). Myrmecol News. 2021;31:1–30.
31. Anderson KE, Russell JA, Moreau CS, Kautz S, Sullam KE, Hu Y, et al. Highly similar microbial communities are shared among related and trophically similar ant species. Mol Ecol. 2012;21:2282–96.
32. Hu Y, Lukasik P, Moreau CS, Russell JA. Correlates of gut community composition across an ant species (*Cephalotes varians*) elucidate causes and consequences of symbiotic variability. Mol Ecol. 2014;23:1284–300.
33. Kautz S, Rubin BER, Russell JA, Moreau CS. Surveying the microbiome of ants: comparing 454 pyrosequencing with traditional methods to uncover bacterial diversity. Appl Environ Microbiol. 2013;79:525–34.
34. Meyer JM, Hoy MA. Molecular survey of endosymbionts in Florida populations of *Diaphorina citri* (Hemiptera: Psyllidae) and its parasitoids *Tamarixia radiata* (Hymenoptera: Eulophidae) and *Diaphorencyrtus aligarhensis* (Hymenoptera: Encyrtidae). Fla Entomol. 2008;91:294–304.
35. Paniagua Voirol LR, Frago E, Kaltenpoth M, Hilker M, Fatouros NE. Bacterial symbionts in Lepidoptera: their diversity, transmission, and impact on the host. Front Microbiol. 2018;9:556.
36. Alonso-Pernas P, Arias-Cordero E, Novoselov A, Ebert C, Rybak J, Kaltenpoth M, et al. Bacterial community and PHB-accumulating bacteria associated with the wall and specialized niches of the hindgut of the forest cockchafer (*Melolontha hippocastani*). Front Microbiol. 2017;8:291.
37. Chanson A, Moreau CS, Duplais C. Assessing biosynthetic gene cluster diversity of specialized metabolites in the conserved gut symbionts of herbivorous turtle ants. Front Microbiol. 2021;12:1640.
38. Nelsen MP, Ree RH, Moreau CS. Ant–plant interactions evolved through increasing interdependence. Proc Natl Acad Sci USA. 2018;115:12253–8.
39. Osborn AM, Moore ERB, Timmis KN. An evaluation of terminal-restriction fragment length polymorphism (T-RFLP) analysis for the study of microbial community structure and dynamics. Environ Microbiol. 2000;2:39–50.
40. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 2014;30:1312–3.
41. Miller MA, Pfeiffer W, Schwartz T. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In: Proc. of the Gateway Computing Environments Workshop (GCE). New Orleans: IEEE; 2010. pp 1–8.
42. Russell CW, Bouvaine S, Newell PD, Douglas AE. Shared metabolic pathways in a coevolved insect-bacterial symbiosis. Appl Environ Microbiol. 2013;79:6117–23.
43. Zheng H, Powell JE, Steele MI, Dietrich C, Moran NA. Honeybee gut microbiota promotes host weight gain via bacterial metabolism and hormonal signaling. Proc Natl Acad Sci USA. 2017;114:4775–80.
44. Sanders JG, Beinart RA, Stewart FJ, Delong EF, Girguis PR. Metatranscriptomics reveal differences in in situ energy and nitrogen metabolism among hydrothermal vent snail symbionts. ISME J. 2013;7:1556–67.
45. Bauer E, Kaltenpoth M, Salem H. Minimal fermentative metabolism fuels extracellular symbiont in a leaf beetle. ISME J. 2020;14:866–70.
46. Duron O, Morel O, Noël V, Buysse M, Binetruy F, Lancelot R, et al. Tick-bacteria mutualism depends on B vitamin synthesis pathways. Curr Biol. 2018;28:1896–902.e5.
47. Calusinska M, Marynowska M, Bertucci M, Untereiner B, Klimek D, Goux X, et al. Integrative omics analysis of the termite gut system adaptation to *Miscanthus* diet identifies lignocellulose degradation enzymes. Commun Biol. 2020;3:275.
48. Sabree ZL, Kambhampati S, Moran NA. Nitrogen recycling and nutritional provisioning by *Blattabacterium*, the cockroach endosymbiont. Proc Natl Acad Sci USA. 2009;106:19521–6.
49. Shigenobu S, Watanabe H, Hattori M, Sakaki Y, Ishikawa H. Genome sequence of the endocellular bacterial symbiont of aphids *Buchnera* sp. APS. Nature. 2000;407:81–6.
50. Chen I-MA, Chu K, Palaniappan K, Pillay M, Ratner A, Huang J, et al. IMGs/M v. 5.0: an integrated data management and comparative analysis system for microbial genomes and microbiomes. Nucleic Acids Res. 2019;47:D666–77.
51. Price SL, Blanchard BD, Powell S, Blaimer BB, Moreau CS. Phylogenomics and fossil data inform the systematics and geographic range evolution of a diverse neotropical ant lineage. Insect Syst Div. 2022;6:9.
52. Wheeler DE. Behavior of the ant, *Procryptocerus scabriusculus* (Hymenoptera: Formicidae), with comparisons to other Cephalotines. Psyche. 1984;91:171–92.
53. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST+: architecture and applications. BMC Bioinform. 2009;10:421.
54. Manz W, Amann R, Ludwig W, Wagner M, Schleifer K-H. Phylogenetic oligodeoxynucleotide probes for the major subclasses of Proteobacteria: problems and solutions. Syst Appl Microbiol. 1992;15:593–600.
55. Wright ES, Yilmaz LS, Corcoran AM, Ökten HE, Noguera DR. Automated design of probes for rRNA-targeted fluorescence in situ hybridization reveals the advantages of using dual probes for accurate identification. Appl Environ Microbiol. 2014;80:5124–33.
56. Łukasik P, Newton JA, Sanders JG, Hu Y, Moreau CS, Kronauer DJC, et al. The structured diversity of specialized gut symbionts of the New World army ants. Mol Ecol. 2017;26:3808–25.
57. Straka J, Feldhaar H. Development of a chemically defined diet for ants. Insectes Soc. 2007;54:100–4.
58. Russell JA, Moreau CS, Goldman-Huertas B, Fujiwara M, Lohman DJ, Pierce NE. Bacterial gut symbionts are tightly linked with the evolution of herbivory in ants. Proc Natl Acad Sci USA. 2009;106:21236–41.
59. Ewels P, Magnusson M, Lundin S, Käller M. MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics. 2016;32:3047–8.
60. Westreich ST, Treiber ML, Mills DA, Korf I, Lemay DG. SAMSA2: a standalone metatranscriptome analysis pipeline. BMC Bioinform. 2018;19:175.
61. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30:2114–20.
62. Zhang J, Kobert K, Flouri T, Stamatakis A. PEAR: a fast and accurate Illumina paired-end read merger. Bioinformatics. 2014;30:614–20.
63. Kopylova E, Noe L, Touzet H. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. Bioinformatics. 2012;28:3211–7.
64. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012;9:357–9.
65. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The sequence alignment/map format and SAMtools. Bioinformatics. 2009;25:2078–9.
66. Okonechnikov K, Conesa A, García-Alcalde F. Qualimap 2: advanced multi-sample quality control for high-throughput sequencing data. Bioinformatics. 2016;32:292–4.
67. Barkdull M, Moreau CS. Worker reproduction and caste polymorphism impact genome evolution and social genes across the ants. Genome Biol Evol. 2023;15:evad095.
68. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol. 2011;29:644–52.
69. Mikheenko A, Saveliev V, Gurevich A. MetaQUAST: evaluation of metagenome assemblies. Bioinformatics. 2016;32:1088–90.
70. Haas B. TransDecoder. 2022. https://github.com/TransDecoder/TransDecoder.
71. Fu L, Niu B, Zhu Z, Wu S, Li W. CD-HIT: accelerated for clustering the next-generation sequencing data. Bioinformatics. 2012;28:3150–2.
72. Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. Salmon provides fast and bias-aware quantification of transcript expression. Nat Methods. 2017;14:417–9.
73. Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, et al. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat Protoc. 2013;8:1494–512.
74. Zhang Y, Thompson KN, Huttenhower C, Franzosa EA. Statistical approaches for differential expression analysis in metatranscriptomics. Bioinformatics. 2021;37:i34–41.
75. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15:550.
76. Klingenberg H, Meinicke P. How to normalize metatranscriptomic count data for differential expression analysis. PeerJ. 2017;5:e3859.
77. O'Leary NA, Wright MW, Brister JR, Ciufo S, Haddad D, McVeigh R, et al. Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. Nucleic Acids Res. 2016;44:D733–45.
78. Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, et al. The SEED and the rapid annotation of microbial genomes using subsystems technology (RAST). Nucleic Acids Res. 2014;42:D206–14.
79. Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. Nat Methods. 2015;12:50–60.
80. Kanehisa M, Sato Y, Kawashima M. KEGG mapping tools for uncovering hidden features in biological data. Protein Sci. 2022;31:47–53.
81. Chong J, Liu P, Zhou G, Xia J. Using MicrobiomeAnalyst for comprehensive statistical, functional, and meta-analysis of microbiome data. Nat Protoc. 2020;15:799–21.
82. Lu W, Clasquin MF, Melamud E, Amador-Noguez D, Caudy AA, Rabinowitz JD. Metabolomic analysis via reversed-phase ion-pairing liquid chromatography coupled to a stand alone orbitrap mass spectrometer. Anal Chem. 2010;82:3212–21.
83. Ankrah NYD, Wilkes RA, Zhang FQ, Aristilde L, Douglas AE. The metabolome of associations between xylem-feeding insects and their bacterial symbionts. J Chem Ecol. 2020;46:735–44.
84. Durbin BP, Hardin JS, Hawkins DM, Rocke DM. A variance-stabilizing transformation for gene-expression microarray data. Bioinformatics. 2002;18:S105–10.
85. Pang Z, Zhou G, Ewald J, Chang L, Hacariz O, Basu N, et al. Using MetaboAnalyst 5.0 for LC–HRMS spectra processing, multi-omics integration and covariate adjustment of global metabolomics data. Nat Protoc. 2022;17:1735–61.
86. R Core Team. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2020.

87. Huang Z, Wang C. A review on differential abundance analysis methods for mass spectrometry-based metabolomic data. Metabolites. 2022;12:305.
88. Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. J Stat Softw. 2015;67:1–48.
89. Lenth RV. emmeans: Estimated marginal means, aka least-squares means. 2022.
90. Searle SR, Speed FM, Milliken GA. Population marginal means in the linear model: an alternative to least squares means. Am Stat. 1980;34:216–21.
91. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Ser B Stat Methodol. 1995;57:289–300.
92. Ochman H, Elwyn S, Moran NA. Calibrating bacterial evolution. Proc Natl Acad Sci USA. 1999;96:12638–43.
93. Moran NA, Munson MA, Baumann P, Ishikawa H. A molecular clock in endosymbiotic bacteria is calibrated using the insect hosts. Proc R Soc Lond B Biol Sci. 1993;253:167–71.
94. Srinivasan R, Karaoz U, Volegova M, MacKichan J, Kato-Maeda M, Miller S, et al. Use of 16S rRNA gene for identification of a broad range of clinically relevant bacterial pathogens. PLoS ONE. 2015;10:e0117617.
95. Barco RA, Garrity GM, Scott JJ, Amend JP, Nealson KH, Emerson D. A Genus definition for bacteria and archaea based on a standard genome relatedness index. mBio. 2020;11:02475–19.
96. Richter M, Rosselló-Móra R. Shifting the genomic gold standard for the prokaryotic species definition. Proc Natl Acad Sci USA. 2009;106:19126–31.
97. Tamas I, Klasson L, Canbäck B, Näslund AK, Eriksson A-S, Wernegreen JJ, et al. 50 million years of genomic stasis in endosymbiotic bacteria. Science. 2002;296:2376–9.
98. Engel P, Moran NA. The gut microbiota of insects—diversity in structure and function. FEMS Microbiol Rev. 2013;37:699–35.
99. Lin JY, Russell JA, Sanders JG, Wertz JT. *Cephaloticoccus* gen. nov., a new genus of 'Verrucomicrobia' containing two novel species isolated from *Cephalotes* ant guts. Int J Syst Evol Microbiol. 2016;66:3034–40.
100. Stoll S, Feldhaar H, Gross R. Transcriptional profiling of the endosymbiont *Blochmannia floridanus* during different developmental stages of its holometabolous ant host. Environ Microbiol. 2009;11:877–88.
101. Margesin R, Volgger G, Wagner AO, Zhang D, Poyntner C. Biodegradation of lignin monomers and bioconversion of ferulic acid to vanillic acid by *Paraburkholderia aromaticivorans* AR20-38 isolated from Alpine forest soil. Appl Microbiol Biotechnol. 2021;105:2967–77.
102. Singh S, Kaur I, Kariyat R. The multifunctional roles of polyphenols in plant-herbivore interactions. Int J Mol Sci. 2021;22:1442.
103. Dillon R, Charnley K. Mutualism between the desert locust *Schistocerca gregaria* and its gut microbiota. Res Microbiol. 2002;153:503–9.
104. Ngugi DK, Tsanuo MK, Boga HI. Benzoic acid-degrading bacteria from the intestinal tract of *Macrotermes michaelseni* Sjöstedt. J Basic Microbiol. 2007;47:87–92.
105. Santos-Garcia D, Latorre A, Moya A, Gibbs G, Hartung V, Dettner K, et al. Small but powerful, the primary endosymbiont of moss bugs, *Candidatus* Evansia muelleri, holds a reduced genome with large biosynthetic capabilities. Genome Biol Evol. 2014;6:1875–93.
106. Gil R, Silva FJ, Zientz E, Delmotte F, Gonzalez-Candelas F, Latorre A, et al. The genome sequence of *Blochmannia floridanus*: comparative analysis of reduced genomes. Proc Natl Acad Sci USA. 2003;100:9388–93.
107. Jørgensen BB, Postgate JR, Postgate JR, Kelly DP. Ecology of the bacteria of the sulphur cycle with special reference to anoxic—oxic interface environments. Philos Trans R Soc Lond B Biol Sci. 1982;298:543–61.
108. Meseguer AS, Manzano-Marín A, Coeur d'Acier A, Clamens A-L, Godefroid M, Jousselin E. *Buchnera* has changed flatmate but the repeated replacement of co-obligate symbionts is not associated with the ecological expansions of their aphid hosts. Mol Ecol. 2017;26:2363–78.
109. Zheng H, Nishida A, Kwong WK, Koch H, Engel P, Steele MI, et al. Metabolism of toxic sugars by strains of the bee gut symbiont *Gilliamella apicola*. mBio. 2016;7:e01326–16.
110. Salem H, Kirsch R, Pauchet Y, Berasategui A, Fukumori K, Moriyama M, et al. Symbiont digestive range reflects host plant breadth in herbivorous beetles. Curr Biol. 2020;30:2875–86.
111. Li Y, Leonard SP, Powell JE, Moran NA. Species divergence in gut-restricted bacteria of social bees. Proc Natl Acad Sci USA. 2022;119:e2115013119.
112. Shoemaker V, Nagy KA. Osmoregulation in amphibians and reptiles. Annu Rev Physiol. 1977;39:449–71.
113. O'Donnell M. Insect excretory mechanisms. In: Simpson SJ (ed). Advances in insect physiology, vol. 35. 1st ed. Academic Press, Cambridge, 2008. pp 1–22.
114. Cho B-K, Federowicz SA, Embree M, Park Y-S, Kim D, Palsson BØ. The PurR regulon in *Escherichia coli* K-12 MG1655. Nucleic Acids Res. 2011;39:6456–64.
115. Rettner RE, Saier MH Jr. The autoinducer-2 exporter superfamily. Micro Physiol. 2010;18:195–205.
116. Powell JE, Martinson VG, Urban-Mead K, Moran NA. Routes of acquisition of the gut microbiota of the honey bee *Apis mellifera*. Appl Environ Microbiol. 2014;80:7378–87.
117. Mee MT, Collins JJ, Church GM, Wang HH. Syntrophic exchange in synthetic microbial communities. Proc Natl Acad Sci USA. 2014;111:E2149–56.
118. Pande S, Kost C. Bacterial unculturability and the formation of intercellular metabolic networks. Trends Microbiol. 2017;25:349–61.
119. Mira A, Ochman H, Moran NA. Deletional bias and the evolution of bacterial genomes. Trends Genet. 2001;17:589–96.
120. McCutcheon JP, von Dohlen CD. An interdependent metabolic patchwork in the nested symbiosis of mealybugs. Curr Biol. 2011;21:1366–72.
121. Morris JJ, Lenski RE, Zinser ER. The Black Queen hypothesis: evolution of dependencies through adaptive gene loss. mBio. 2012;3:e00036–12.
122. Hillesland KL, Stahl DA. Rapid evolution of stability and productivity at the origin of a microbial mutualism. Proc Natl Acad Sci USA. 2010;107:2124–9.
123. Oliveira NM, Niehus R, Foster KR. Evolutionary limits to cooperation in microbial communities. Proc Natl Acad Sci USA. 2014;111:17941–6.
124. Louca S, Polz MF, Mazel F, Albright MBN, Huber JA, O'Connor MI, et al. Function and functional redundancy in microbial systems. Nat Ecol Evol. 2018;2:936–43.
125. Hussa EA, Goodrich-Blair H. It takes a village: ecological and fitness impacts of multipartite mutualism. Annu Rev Microbiol. 2013;67:161–78.
126. Price SL, Etienne RS, Powell S. Tightly congruent bursts of lineage and phenotypic diversification identified in a continental ant radiation. Evolution. 2016;70:903–12.

ACKNOWLEDGEMENTS

We thank Scott Powell and Jignasha Rana who collected specimens used in this study. We are grateful to Benjamin Rubin who participated in *C. varians* genome sequencing and assembly. Thanks also go to Rebecca Wilkes who worked on metabolomic sample preparation and Dan Freeman who helped in genomic analyses. We further thank Alicia Pastor at Michigan State University for the work on TEM images, and Aharon Oren and Bernhard Schink for their help naming *Ischyrobacter davidsoniae*. Finally, we thank Katrina Terry, Azad Ahmed, Rachel Ehrlich, and Steven Lang from the Drexel Genomic Core Facility who worked on the PacBio sequencing preparation and processing.

AUTHOR CONTRIBUTIONS

BB, CSC, YH, JTW, and JAR designed research; BB, CSC, YH, CMM, BH, JYL, CD, LA, JTW, and JAR contributed to conceptualization; BB, CSC, YH, CMM, BH, JYL, YS, VJF, DA, RL, CJO, CSM, JTW, and JAR performed research; BB, CSC, JTW, and JAR analyzed data; BB and JAR wrote the manuscript; BB, YH, GLR, LA, JTW, and JAR reviewed and edited the manuscript; CSC, GLR, CSM, LA, JTW, and JAR contributed to funding; all authors approved the final version.

FUNDING

This work was supported by National Science Foundation CBET-1653092 to LA, DEB-1442156 to JTW, DEB-1900357 to CSM, DOB-1442144 to JAR, GRF-2041772 to CSC, and TUES-1245632 to GLR and JAR.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41396-023-01490-1.

Correspondence and requests for materials should be addressed to Benoît Béchade.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.