



Universiteit  
Leiden  
The Netherlands

## **Gut microbiomes of cycad-feeding insects tolerant to $\beta$ -methylamino-L-alanine (BMAA) are rich in siderophore biosynthesis**

Gutiérrez-García, K.; Whitaker, M.R.L.; Bustos Díaz, E.D.; Salzman, S.; Ramos-Aboites, H.E.; Reitz, Z.L.; ... ; Barona-Gómez, F.

### **Citation**

Gutiérrez-García, K., Whitaker, M. R. L., Bustos Díaz, E. D., Salzman, S., Ramos-Aboites, H. E., Reitz, Z. L., ... Barona-Gómez, F. (2023). Gut microbiomes of cycad-feeding insects tolerant to  $\beta$ -methylamino-L-alanine (BMAA) are rich in siderophore biosynthesis. *Isme Communications*, 3(1). doi:10.1038/s43705-023-00323-8

Version: Publisher's Version

License: [Creative Commons CC BY 4.0 license](https://creativecommons.org/licenses/by/4.0/)

Downloaded from: <https://hdl.handle.net/1887/3718717>

**Note:** To cite this publication please use the final published version (if applicable).

## ARTICLE OPEN



# Gut microbiomes of cycad-feeding insects tolerant to $\beta$ -methylamino-L-alanine (BMAA) are rich in siderophore biosynthesis

Karina Gutiérrez-García<sup>1,2,10</sup>, Melissa R. L. Whitaker<sup>3,4,10</sup>, Edder D. Bustos-Díaz<sup>1,5</sup>, Shayla Salzman<sup>6,7</sup>, Hilda E. Ramos-Aboites<sup>1</sup>, Zachary L. Reitz<sup>7</sup>, Naomi E. Pierce<sup>8</sup>, Angélica Cibrián-Jaramillo<sup>8,9</sup> and Francisco Barona-Gómez<sup>1,5</sup>

© The Author(s) 2023

Ingestion of the cycad toxins  $\beta$ -methylamino-L-alanine (BMAA) and azoxyglycosides is harmful to diverse organisms. However, some insects are specialized to feed on toxin-rich cycads with apparent immunity. Some cycad-feeding insects possess a common set of gut bacteria, which might play a role in detoxifying cycad toxins. Here, we investigated the composition of gut microbiota from a worldwide sample of cycadivorous insects and characterized the biosynthetic potential of selected bacteria. Cycadivorous insects shared a core gut microbiome consisting of six bacterial taxa, mainly belonging to the Proteobacteria, which we were able to isolate. To further investigate selected taxa from diverging lineages, we performed shotgun metagenomic sequencing of co-cultured bacterial sub-communities. We characterized the biosynthetic potential of four bacteria from *Serratia*, *Pantoea*, and two different *Stenotrophomonas* lineages, and discovered a suite of biosynthetic gene clusters notably rich in siderophores. Siderophore semi-untargeted metabolomics revealed a broad range of chemically related yet diverse iron-chelating metabolites, including desferrioxamine B, suggesting the occurrence of an unprecedented desferrioxamine-like biosynthetic pathway that remains to be identified. These results provide a foundation for future investigations into how cycadivorous insects tolerate diets rich in azoxyglycosides, BMAA, and other cycad toxins, including a possible role for bacterial siderophores.

ISME Communications; <https://doi.org/10.1038/s43705-023-00323-8>

## INTRODUCTION

Cycads are among the oldest seed plants, with a lineage that traces back to the Permian [1]. The genome of one cycad species in the genus *Cycas* was recently published, and with more than 10 Gbp, suggests a series of gene expansions as well as a whole genome duplication prior to divergence from its ginkgo sister taxa [2]. These tropical gymnosperms harbor defensive chemicals in their leaves and other tissues, including the azoxyglycoside methylazoxymethanol (MAM), or cycasin, and the non-proteinogenic amino acid  $\beta$ -Methylamino-L-Alanine (BMAA), both of which are toxic to diverse organisms [3, 4]. Previous studies have demonstrated that several insect herbivores are able to detoxify and even sequester the cycad-derived MAM glycoside, cycasin [4–7], whereas mechanisms of BMAA resistance, just as its biosynthesis, remain elusive [8, 9]. BMAA is capable of blocking glutamate receptors [3], causing neurotoxic damage in insects [10, 11] and mammals [12] as well as developmental problems in *Arabidopsis thaliana* [13]. This contrasts with the recent

observation that BMAA quantities in cycad leaves may not be acutely toxic or deterrent to insects [14].

In addition to MAM and BMAA, the genomes of a few species of *Cycas* cycads encode for a functional FitD insect toxin that was acquired by horizontal gene transfer from a microbial source, and then expanded into multiple copies [2]. Many bacteria also have *fitD* genes, as well as the homologous genes *mcf* ('makes caterpillar floppy'), which have been shown to be broadly distributed and to contribute to insecticidal activity in plant- and nematode-associated bacteria [15, 16]. These observations, together with BMAA and MAM toxicity, suggest that cycads' diverse specialized metabolites exert significant evolutionary pressure on cycad-specialized insects and other organisms. Even so, several specialized insect species are able to feed on cycad tissues without adverse effects, including some beetles, thrips, moths, and butterflies [17, 18], consistent with the so-called "gut microbial facilitation hypothesis" [19]. At least two species of cycad-feeding Lepidoptera sequester BMAA in their larval and adult

<sup>1</sup>Evolution of Metabolic Diversity Laboratory, Unidad de Genómica Avanzada (Langebio), Cinvestav-IPN, Km 9.6 Libramiento Irapuato - León, Irapuato, Guanajuato 36824, México. <sup>2</sup>Department of Embryology, Carnegie Institution for Science, 3520 San Martin Drive, Baltimore, MD 21218, USA. <sup>3</sup>Museum of Comparative Zoology, Department of Organismic and Evolutionary Biology, Harvard University, 26 Oxford Street, Cambridge, MA 02138, USA. <sup>4</sup>Department of Biological Sciences, East Tennessee State University, Johnson City, TN 37614, USA. <sup>5</sup>Institute of Biology, Leiden University, Sylviusweg 72, Leiden 2333 BE, The Netherlands. <sup>6</sup>University of Georgia, Entomology Department, Athens, GA 30602, USA. <sup>7</sup>Bioinformatics Group, Wageningen University, Droevendaalsesteeg 1, 6708PB Wageningen, The Netherlands. <sup>8</sup>Ecological and Evolutionary Genomics Laboratory, Unidad de Genómica Avanzada (Langebio), Cinvestav-IPN, Km 9.6 Libramiento Irapuato - León, Irapuato, Guanajuato 36824, México. <sup>9</sup>Naturalis Biodiversity Center, Darwinweg 2, 2333 CR Leiden, The Netherlands. <sup>10</sup>These authors contributed equally: Karina Gutiérrez-García, Melissa R. L. Whitaker. <sup>✉</sup>email: [melliwhitaker@gmail.com](mailto:melliwhitaker@gmail.com); [f.barona.gomez@biology.leidenuniv.nl](mailto:f.barona.gomez@biology.leidenuniv.nl)

Received: 2 May 2023 Revised: 17 October 2023 Accepted: 25 October 2023

Published online: 22 November 2023

tissues [18], and for at least one group of cycad specialists, *Eumaeus* butterflies, toxin tolerance appears to be a key innovation in the evolution of these insects [20].

The gut bacterial communities of many insects are essential for nutrient acquisition, digestion, and detoxification [21]. Many microorganisms are capable of enzymatically degrading plant-specialized metabolites, and growing evidence suggests that symbiotic gut bacteria can help herbivorous insects cope with diets rich in plant defensive chemicals [22], and can play a role in resistance to insecticides [23, 24]. However, despite the overall effect of plant secondary metabolites on the plant, insect, and soil microbiomes, which have been broadly documented [25, 26], examples of phytochemicals specifically affecting the metabolomes of these microbiomes are scarce or nonexistent [27, 28]. An earlier study of the microbiota of cycadivorous insects identified a small core of bacterial taxa that are shared across phylogenetically and geographically distinctive insects [29]. The present study aimed to analyze the gut bacteria of a larger sampling of cycad feeding insects, and to explore possible functions by assessing specialized metabolites biosynthesized by sub-communities and selected bacterial taxa isolated from the guts of these insects.

To do so, we characterized the bacterial communities of cycadivorous insects from around the world using 16 S sequencing, and then applied *EcoMining* [30], a method of capturing bacterial groups and metabolites of interest by co-culturing bacterial sub-communities for shotgun metagenomics, phylogenomics, and integrated omics analyses. The advantage of this method is that it allows for the distillation of functionally relevant fractions of the entire microbiome based on a priori biological hypotheses. We then investigated the metabolic functions of selected bacterial taxa by mining for biosynthetic gene clusters (BGCs) predicted to direct the synthesis of specialized metabolites, and in some cases, confirming the expression of diverse yet related siderophores through semi-targeted siderophore metabolomics. Our results provide insight into the functional capabilities of bacteria found in the guts of cycadivorous insects, and their potential importance for chemically-mediated cycad-insect interactions.

## MATERIALS AND METHODS

### Microbiome global sampling of cycadivorous insects

With the aim of capturing as much ecological and trophic diversity as possible, 12 insect species from 3 orders were collected from the USA, South Africa, China, Australia, Singapore, and Thailand between 2014 and 2017 (Fig. 1). Over 90 insect samples were included in the analysis, with detailed collecting information in Table S1. All species included in the study are obligate cycad specialists with two exceptions: the neotropical moth, *Seirarctia echo* (Erebidae), which feeds on cycads and plants from multiple families of angiosperms [17], although the specimens we analyzed were collected from *Zamia* cycads; and the larvae of the African moth *Zerenopsis lepida*, which are obligate cycad folivores in early instars, although some individuals switch to feeding on angiosperms in the fourth instar [31]. We were able to collect late instar *Z. lepida* feeding on *Encephalartos* and *Stangeria* cycads as well as an angiosperm (*Maesa lanceolata*).

Insects were processed differently depending on collector, insect type, and size. Lepidopteran larvae in their third instar or older were collected with sterile forceps, submerged in ethanol, and rinsed in sterile PBS. Their guts were then excised using flame-sterilized dissecting tools. Dissected guts were preserved in individual vials containing 97% ethanol. The guts of *Zerenopsis lepida* were not dissected due to their small size; instead, they were surface-sterilized using a 10% bleach solution and preserved whole. Coleopteran specimens were either immediately flash frozen in liquid nitrogen or preserved in ethanol. Whole specimens were then submerged for 5 s in a 10% bleach solution and rinsed in PBS for surface sterilization prior to DNA extraction. Finally, specimens from the thysanopteran species *Cycadothrips chadwicki* were neither surface sterilized nor dissected due to their small size; instead, approximately 150 individual thrips were pooled together into a single sample for DNA extraction.

### Amplicon metagenomics of the cycadivorous insect's gut microbiome

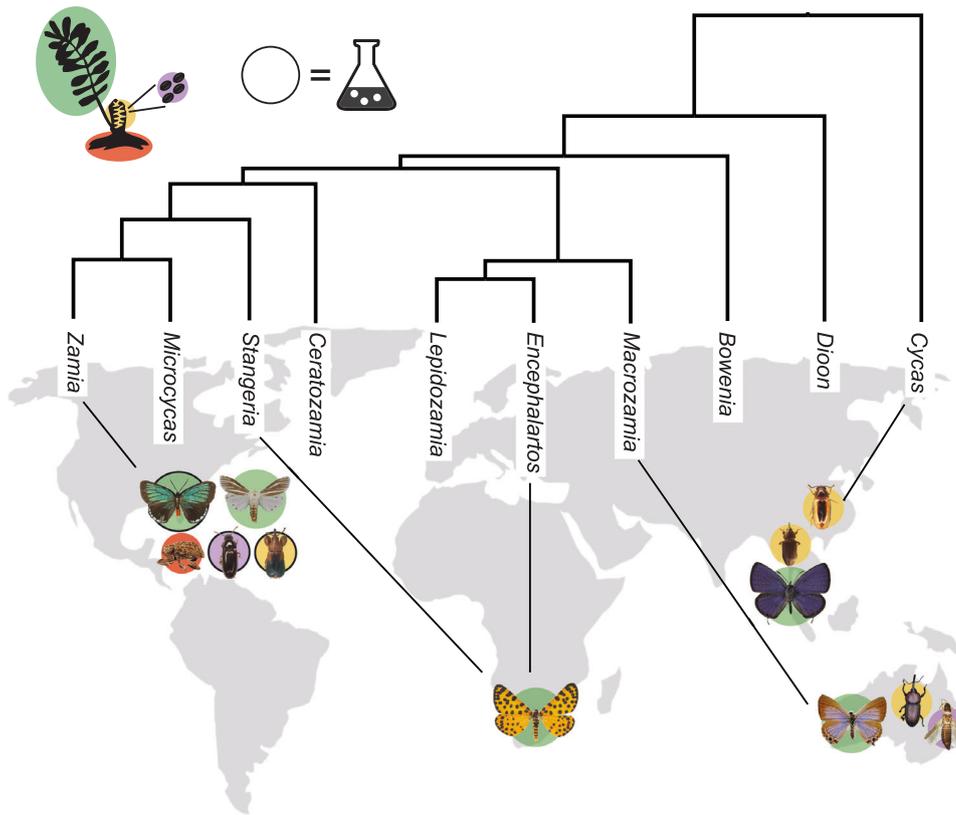
DNA extractions were performed using the Powersoil DNA isolation kit (MoBio Laboratories, Carlsbad, CA, USA) according to the manufacturer's instructions with the addition of 60 µg of proteinase K to the lysis buffer prior to bead-beating. DNA concentration was assessed using a Qubit Fluorometer (Invitrogen, Carlsbad, CA, USA), and samples with low DNA yields were concentrated as per the DNA extraction kit's recommendations. DNA extracts were sent to Argonne National Laboratories (Lemont, IL, USA) for library preparation and 16 S rRNA gene amplification and sequencing. Amplification of the 16 S rRNA was done using barcoded primers 515 F (5'-GTGTCAGCMGCCGCGGTAA-3') and 806 R (5'-GGAC-TACNVGGGTWCTAAT-3') of the V4 region [32], and libraries were pooled and sequenced with Illumina MiSeq instrument in 2 × 150 paired-end mode. Raw sequences were processed following previously published methods [33]. Briefly, sequences were demultiplexed and quality filtered in QIIME [34] then de novo clustered at 97% identity using UPARSE [35]. QIIME was used to assign taxonomy and to construct a phylogenetic tree of representative sequences. The resulting BIOM table and phylogenetic tree were exported to R and analyzed with the phyloseq [36], the phyloseq-extended [37], and the microbiome packages. Non-target sequences (mitochondria, chloroplasts, *Wolbachia*, and the common lab contaminants *Escherichia* and *Staphylococcus*) were removed from the dataset and spurious sequences were identified and removed using the PERfect package's permutation filtering method implemented with the "fast" algorithm [38]. We defined a species' core microbiome as the set of bacterial operational taxonomic units (OTUs) present in at least 83% of replicates within a species (corresponding to 5 out of 6 replicates, for example). For species represented by multiple life stages (e.g., larva & adult) or, in the case of *Zerenopsis lepida*, multiple host plants, cores were calculated separately for each grouping. For samples with only a single replicate (e.g., the pooled *Cycadothrips* sample), a within-species core could not be obtained, but these samples were included in an across-species core, defined as OTUs present in 98% of all samples.

### Co-culture inoculation for shotgun metagenomics and bacterial isolation

Three insect species were collected and processed to provide inocula for co-cultures: *Rhopalotria slossoni* (Coleoptera), *Pharaxanatha floridana* (Coleoptera), and *Eumaeus atala* (Lepidoptera) (Table S1). Twenty-two *R. slossoni* adults, fourteen *P. floridana* adults, and two *E. atala* larvae were surface sterilized by successive washes in 70% ethanol and sterile dd-MilliQ water. The guts of *E. atala* larvae were excised using sterile dissecting tools and then suspended, individually, in sterilized dd-MilliQ water and homogenized with a sterile micropestle. Coleopteran specimens were pooled by species, suspended in sterilized dd-MilliQ water, and homogenized with a sterile micropestle. Resulting biomass from each sample was divided to inoculate two 100 ml co-cultures per sample of liquid Brain Heart Infusion (BHI) medium (Becton Dickinson, Germany) at 0.05x, one without BMAA and one with 20 µM of BMAA, a concentration used in previous studies showing the toxic effects on insects [39, 40]. Co-cultures were incubated at 23 °C with agitation at 150 r.p.m for 48 h, after which each co-culture was used to inoculate petri dishes containing different bacterial media described below. DNA was extracted from the eight co-cultures for shotgun metagenomics. Two additional co-cultures (BMAA +/-) were generated from the pooled *P. floridana* biomass and were used for genomic isolation but not shotgun metagenomics.

### Shotgun metagenome sequencing and taxonomic analysis

DNA was extracted from co-cultures using a standard phenol-chloroform protocol [41], and DNA quality was checked by Nanodrop 2000/2000c (Thermo Scientific, USA). Libraries were prepared with the Truseq nano kit and sequenced at UGA, Cinvestav (Irapuato, Mexico) using the NextSeq Illumina platform 2 × 150 paired-end reads format. Obtained reads were checked with FastQC v0.11.9 [42], and low-quality bases were trimmed with Trimmomatic v0.32 [43]. Metagenomic profiling of high-quality paired reads was performed using kraken2 [44] with the maxikraken database (September 2018 release; available at [https://lomanlab.github.io/mockcommunity/mc\\_databases.html](https://lomanlab.github.io/mockcommunity/mc_databases.html)). Abundances, taxonomic lineage tables, and metadata were exported to R using the kraken-biom software [45] to be analyzed with the phyloseq [36], the taxize [46], and the microbiome packages.



**Fig. 1 Cycadivorous insect sampling strategy.** Coleoptera, Lepidoptera, and Thysanoptera were collected from four continents. Black lines connect each insect species to the cycad genus on which it feeds, and background colors represent the plant tissue on which each insect is specialized: green = leaves, yellow = cones, purple = pollen, and red = roots. All cycad herbivores encounter cycad toxins, regardless of the plant tissue on which they feed. Dark black outlines indicate the three insect genera that were subjected to the *EcoMining* approach.

### Shotgun metagenome assembly, mining, and binning

Assembly of the six metagenomes obtained was performed with metaSPAdes v3.10 using default parameters [47]. The resulting scaffolds were annotated using RAST [48], taxonomically classified using kraken2 with the same database and parameters used for read classification, and mined using antiSMASH v5.0 [49], BiG-SCAPE and CORASON [50]. To obtain the core biosynthetic gene clusters (BGCs), each complete and non-redundant BGC obtained from antiSMASH was used to construct a presence/absence matrix. A core BGCs plot was then created using the UpSetR R package v1.4.0 [51], with all the set intersections in this matrix. Scaffolds were binned with the PATRIC Metagenomic Binning Service [52] using default parameters, from which nine metagenome-assembled genomes (MAGs) were obtained: 4 for *Stenotrophomonas*, 4 for *Serratia*, and 1 for *Pantoea* (Table S2). To obtain 16S sequences from the metagenomes we used the 16S sequence recovery pipeline of Anvi'o [53].

### Phylogenomic analysis of selected bacteria

To establish the relationships between MAGs and genome sequences of bacterial isolates, phylogenies of *Stenotrophomonas*, *Pantoea*, and *Serratia* were constructed using representative genomes from each genus, along with our newly generated genome sequences and recovered MAGs (Tables S3, S4, and S5). For each phylogeny, a protein core was obtained using Get-Phylomarkers [54], computed with Get-Homologues using the algorithms bidirectional best-hit (BDBH), Clusters of Orthologous Groups-triangles (COGtriangles), and OrthoMCL (Markov Clustering of orthologs, OMCL) with default parameters (Tables S6, S7, and S8). Resulting matrices, containing 39, 64, and 712 proteins, respectively, were used to construct phylogenetic trees with MrBayes v3.2 [55] with a mixed substitution model based on posterior probabilities (aamodel[Wag]1.000) for proteins for 100,000 generations. The trees were visualized with FigTree v1.4.2 [56].

### Isolation and identification of bacteria from co-cultures in semi-selective media

Semi-selective media were chosen to target specific bacterial phyla that are known to be represented in cycadivorous insects' microbiomes [29]. These were: 1) SFM media, (mannitol: 20 g/L; soya flour 20 g/L; and agar 17 g/L for solid media); 2) ISP4 and ISP4N<sup>-</sup>, (starch: 10.0 g/L; dipotassium phosphate: 1 g/L; magnesium sulfate: 1 g/L; sodium chloride: 1 g/L; ammonium sulfate: 2 g/L for ISP4 media, none for ISP4N<sup>-</sup> media; calcium carbonate: 2 g/L; ferrous sulfate: 1 mg/L; magnesium chloride: 1 mg/L; zinc sulfate: 1 mg/L; final pH 7.0; and agar 17 g/L for solid media); 3) BHI media at concentrations 0.02x and 0.05x (Becton Dickinson, Germany); 4) FLA media (Fortified lipid agar) (tryptic soy broth: 16 g/L; vegetable oil: 10 mL/L; nutrient broth 12 g/L; yeast extract: 5 g/L; and agar 17 g/L for solid media); and 5) NTBA media (peptone: 5 g/L; beef extract: 3 g/L; bromothymol blue: 0.025 g/L; 2,3,5-triphenyl tetrazolium chloride: 0.04 g/L; and agar 17 g/L for solid media).

The Petri dishes were incubated at 22 °C for 72 h. After this, only unique bacterial morphotypes were selected for further isolation and purification in their corresponding media. Once bacterial morphology was homogeneous on each isolation plate, isolates were grown on 100 ml of their corresponding media for 48 h. DNA extraction and quality measurements were done as before. gDNA was used for PCR amplification of the 16S rRNA region using the universal bacterial primers 27 F and 1492 R [57]. The amplification yielded PCR fragments of 1.4 Kbp in length, which were purified with the Qiagen QIAquick PCR purification kit (Hilden, Germany) and sequenced using Sanger chemistry at UGA, Cinvestav (Irapuato, Mexico). Taxonomic identity of the sequences was determined using Blastn against the SILVA rRNA database v132 [58]. Each sequence was classified according to the best hit (>97% sequence identity) (Table S9). An alignment of all of the obtained 16S rRNA sequences was constructed with ClustalW [59] and edited with Gblocks v0.91 to remove ambiguous positions [60]. A phylogenetic tree was constructed using MrBayes v3.2 [55], with a 4by4 nucleotide substitution model for 100,000 generations,

with sampling every 100 generations. The graphical representation of the phylogenetic trees was obtained with Iroki Tree Viewer [61].

### Genomic sequencing and natural products mining of isolated strains

We selected six strains of interest from the isolates based on two criteria: 1) their prevalence in the gut microbiomes of *R. slossoni*, *P. floridana*, and *E. atala*; and 2) their biosynthetic potential to produce conserved natural products. For genomic characterization of these strains, we grew monocultures of the selected *Serratia*, *Pantoea*, and *Stenotrophomonas* strains, and the obtained biomass was used for DNA extraction and quality measurements as before. Libraries were prepared using the Truseq nano kit and sequenced at UGA, Cinvestav (Irapuato, Mexico) using the NextSeq Illumina platform with 2 × 150 paired-end reads format. Raw reads were checked with FastQC v0.11.9 and low-quality bases were trimmed using Trimmomatic v0.32 [43]. Multiple de novo genomes were assembled with SPAdes v3.15.3 using default parameters [62] and annotated with RAST [48] (Table S2). CORASON [50] was used to mine turnerbactin-like siderophores in the different MAGs and genomes within a phylogenetic context. CORASON hits that contained homologs of the turnerbactin-like non-ribosomal peptide synthetase (NRPS) were obtained using a BlastP (e-value and cut-off of 0.001).

### Siderophore genome mining

To further characterize the putative siderophore BGCs detected by antiSMASH v5.0 [49], we functionally annotated genes in the BGCs that were previously implicated in iron metabolism. This included enzymes annotated as acyl CoA-dependent acyltransferases, FAD-dependent amine monooxygenases, and iron reductases, as well as TonB-dependent receptors and ABC transporters. The presence of conserved DNA motifs corresponding to iron boxes, known to be involved in iron regulation in bacteria [63], was also determined and used to mine the genomes. An iron box database was constructed using DNA motifs previously described for Gram-positive bacteria (TTAGGTTAGGCTCACCTAA and TGATAATNATTATCA) [64, 65], and Gram-negative bacteria (GATAATGATAATCATTATC, GATAAAATTAATCAGCCTC, and ATTAATAAAAACCATTGTC, GATAATGAGAATCATTATT, GATAATTGTTATCGTTTGC, TATAATGATACGCATTATC, TGTAATGATAACCATTCTC, GAATATGATTATCATTTTC, GAAAATGATAATCATATTC, ATAAATGATAATCATTATT, GATAATCATTTCATATC) [64, 66, 67]. The presence of iron box motifs was analyzed with PREDetector v3.1 [68] using the iron box motifs dataset.

### Identification of bacterial siderophores through HPLC and MS

The *Pantoea*, *Serratia*, and *Stenotrophomonas* bacteria were subjected to semi-untargeted siderophore extraction and characterization. Each strain was cultivated in two separate Erlenmeyer flasks containing LB medium, with and without 200 μM of the iron-chelator compound 2,2'-dipyridyl (DIPY) for four days, as previously [65]. Culture growth was checked by OD<sub>600</sub> measurements every 12 h and cultures were harvested upon reaching saturation. Cultures were then centrifuged at 8000 rpm and the supernatant was collected and lyophilized. Dry extracts were resuspended in 10 ml of sterile dd-MilliQ water and separated in two Falcon tubes, each with 5 ml of the resuspended extracts. Iron-chelating metabolites were converted to their ferric complexes by addition of 1 M FeCl<sub>3</sub> to one of the tubes of each condition. All samples were then concentrated by filtering the supernatants with Millipore Millex-GN filters with a nylon membrane of 0.22 μm. Samples were analyzed on a Thermo Ultimate 3000-uHPLC instrument equipped with a quaternary pump, a diode array detector, and a ZORBAX Eclipse XDB-C18 (Agilent, USA) analytical column (4.6 × 150 mm, 5 μm). Analytical conditions were as follows: The mobile phase comprised a binary system of eluent A, 0.1% trifluoroacetic acid, and eluent B, 100% acetonitrile. The run consisted of a gradient from 0 to 100% B for 25 min. Tris-hydroxamate-Fe<sup>3+</sup> complexes, formed after the addition of iron chloride to bacterial extracts, were detected at a wavelength of 435 nm. Fractions were collected and pooled by bacterial genus for mass spectrometry (MS) analysis in an ion trap LTQ Velos (Thermo Scientific, Waltham, USA). MS/MS analysis of selected ions was performed with a collision energy of 20 eV, and spectra were analyzed using the XCalibur software v3.0 (Thermo Scientific, USA). Metal-chelating metabolites were identified by comparing fragmentation patterns with those of siderophores previously reported [65] or included in the Bertrand's siderophore database ([http://bertrandsamuel.free.fr/siderophore\\_base/index.php](http://bertrandsamuel.free.fr/siderophore_base/index.php)) and the GNPS database used to generate GNPS feature-based molecular networks [69, 70].

## RESULTS

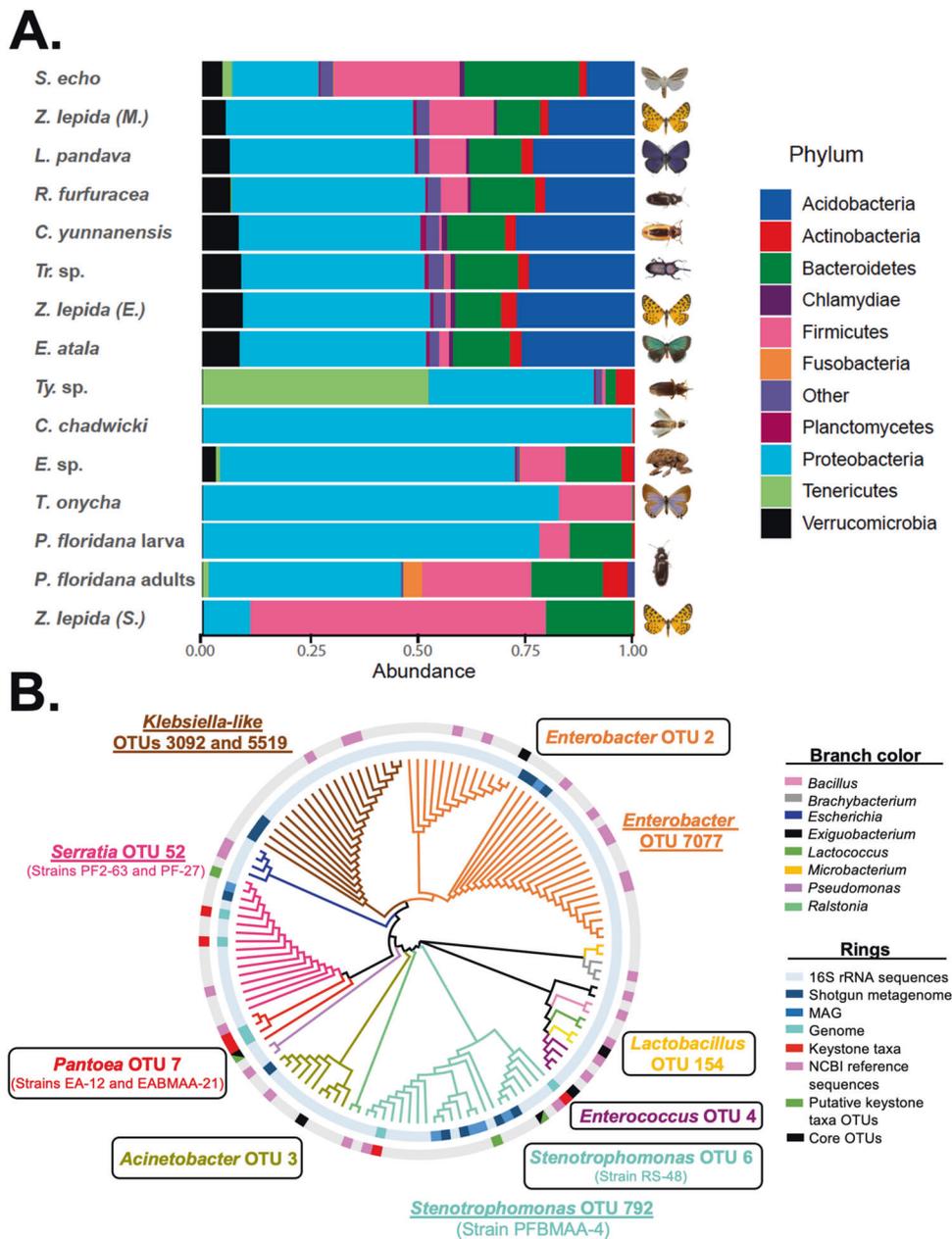
### 16S taxonomic profiling of cycadivorous insects' guts identifies a global microbiome core

16S amplicon sequencing of DNA extracted from the guts of a global sampling of cycadivorous insect species yielded 6,317,577 bacterial reads. After removing non-target and spurious OTUs, we were left with 5,027,153 reads across 92 samples. Observed bacterial species richness ranged from 41 to >700 OTUs within a single insect sample. A total of 26 bacterial phyla were detected (Fig. 2A), with Proteobacteria comprising moderate to dominant portions of the microbiomes of most insects with a few exceptions: *Zerenopsis lepida* larvae feeding on *Stangeria* cycads (but not other plants) were dominated by Firmicutes, *Tychiodes* sp. adults were dominated by Tenericutes, and *Seirarctia echo* larval microbiomes were composed of relatively equal proportions of Proteobacteria, Bacteroidetes, and Firmicutes. Insects' core microbiomes ranged widely in richness, from 17 core OTUs in *Luthrodes pandava* larvae ( $n = 7$ ) to 185 OTUs in *Eubulus* sp. larvae ( $n = 7$ ). In general, the core microbiomes of Coleoptera tended to be larger than in Lepidoptera. A cross-species core included 6 OTUs found in all insect species, with highly variable relative abundances across samples. Some core OTUs differ slightly from previously reported core bacteria [29], likely due to methodological differences in how the core was defined between studies. In the present study, core OTUs belong to the genera *Acinetobacter* (OTU 3), *Pantoea* (OTU 7), *Enterobacter* (OTU 2), *Enterococcus* (OTU 4), *Lactobacillus* (OTU 154), and *Stenotrophomonas* (OTU 6). These taxa were then placed within a phylogenetic context (Fig. 2B) following isolation and analysis of selected bacterial species, as described in the following sections.

### Isolation of bacteria from the microbiomes of cycadivorous insects' guts

To direct bacterial isolation towards ecologically relevant strains from the broad diversity of cycadivorous insects investigated, we adopted the *EcoMining* approach, which we have successfully employed to investigate the microbiome of cycad's coralloid roots and its biosynthetic potential [71]. We obtained bacterial isolates from the sub-community co-cultures of the dissected guts of *Eumaeus atala*, *Pharaxanatha floridana*, and *Rhopalotria slossoni* (Fig. 1, circles). This approach led to two types of cultures for further analyses: (i) bacterial isolates obtained in selective and semi-selective media, which were characterized after amplification and sequencing of their 16S rRNA gene, and eventually their whole genome for selected strains; and (ii) sub-community co-cultures grown under functionally relevant conditions, such as the presence of the cycad toxin BMAA, which were shotgun sequenced to identify core taxonomic and functional genes, both at the single-gene level and as part of metagenome-assembled genomes (MAGs) (Table 1).

To compare the metagenomics datasets generated from cultures (i.e., *EcoMining*) and culture-independent methods (16S profiling), we extracted a subset of the 16S profiling data from *Rhopalotria furfuracea* adults ( $n = 6$ ), *Pharaxanatha floridana* adults ( $n = 7$ ), and *Eumaeus atala* larvae ( $n = 5$ ), for comparison to the taxonomic information generated by co-culture shotgun metagenomic sequencing. Alpha diversity analysis showed that shotgun metagenomes recovered more OTUs than 16S amplicon sequencing (Fig. S1A) and that these OTUs belonged to fewer phyla, with Proteobacteria being the most dominant phylum, reflecting a partial but marginal growth bias of the co-cultures (Fig. S1B). These analyses also showed that addition of BMAA to the co-cultures does not alter the taxonomic composition of bacterial communities; instead, host species identity was the biggest factor contributing to taxonomic dissimilarities among sub-communities (Fig. S1A). The fact that BMAA did not alter the microbial diversity of the co-cultures is congruent with the fact



**Fig. 2** Composition and distribution of cycadivorous insect's gut microbiome. **A** Community composition among a global sampling of cycad-feeding insects. Proteobacteria dominate the gut microbiomes of most cycadivorous insects, followed by Firmicutes and Bacteroidetes. **B** 16 S rRNA phylogenetic tree including 79 bacterial strains isolated from the guts of *Eumaeus atala*, *Pharaxanotia floridana*, and *Rhopalotria slossoni*. The 16 S rRNA gene from the six genomes of the genera *Serratia*, *Pantoea*, and *Stenotrophomonas*, plus all the metagenomes data and the 16 S amplicon data were included. Core taxa identified after the integration of different datasets are highlighted: core taxa are shown within boxes and semi-conserved taxa shown underlined. Specific strains selected for further experimental characterization are shown between parentheses below their cognate OTUs.

that the isolated sub-communities are adapted to this toxin present in the insect's diet.

Having assessed the validity of the co-culture metagenomes to recover core taxa identified by the 16 S metagenomic profiling, we sought to identify bacteria present in all of the sub-community co-cultures. We found that the Proteobacteria diversity recovered from the shotgun metagenomes and their corresponding 16 S sequences included the genera *Serratia*, *Acinetobacter*, *Stenotrophomonas*, and *Enterobacter* across all three insect species, albeit with different patterns of prevalence (Fig. S2). Given that Proteobacteria was also the most abundant phylum in the 16 S amplicon dataset, it is reasonable to conclude that this is a

constitutive part of the insects' gut microbiomes. 16 S metagenomic profiling identified several bacteria from the Enterobacteriaceae family that were not found in cultures but these taxa were only present in the metagenomes obtained from *Pharaxanotia floridana* with and without BMAA (PF and PFBMAA, respectively), and *Rhopalotria slossoni* (RS). Based on the 16 S phylogeny (Fig. 2B), the closest match to these previously uncultured bacteria belongs to the genus *Klebsiella* (OTUs 3092 and 5519), with sequence identity scores between 90% and 93% depending on the metagenome.

Using this integrated dataset as a guide, we then selected for further experimental characterization the following strains that we

**Table 1.** Metagenomes and genomes generated in this study.

(Meta)genome sequence/strain	Type	Bacterial genus	Origin <sup>a</sup>	Accession number <sup>b</sup>
EA	Metagenome	NA	EA	SRR24296969
EABMAA	Metagenome	NA	EA + BMAA	SRR24296968
PF	Metagenome	NA	PF	SRR24296967
PFBMAA	Metagenome	NA	PF + BMAA	SRR24296966
RS	Metagenome	NA	RS	SRR24296965
RSBMAA	Metagenome	NA	RS + BMAA	SRR24296964
Pa-EAmG	MAG	<i>Pantoea</i>	EA	SAMN34374683
Se-PFmG	MAG	<i>Serratia</i>	PF	SAMN34374684
St-PFmG	MAG	<i>Stenotrophomonas</i>	PFBMAA	SAMN34374685
Se-PFBMAAmG	MAG	<i>Serratia</i>	PFBMAA	SAMN34374686
St-PFBMAAmG	MAG	<i>Stenotrophomonas</i>	PFBMAA	SAMN34374687
Se-RSmG	MAG	<i>Serratia</i>	RS	SAMN34374688
St-RSmG	MAG	<i>Stenotrophomonas</i>	RS	SAMN34374689
Se-RSBMAAmG	MAG	<i>Serratia</i>	RSBMAA	SAMN34374690
St-RSBMAAmG	MAG	<i>Stenotrophomonas</i>	RSBMAA	SAMN34374691
EA-12	Genome	<i>Pantoea</i>	EA	SAMN34373067
EABMAA-21	Genome	<i>Pantoea</i>	EABMAA	SAMN34373068
PFBMAA-4	Genome	<i>Stenotrophomonas</i>	PFBMAA	SAMN34373069
RS-48	Genome	<i>Stenotrophomonas</i>	RS	SAMN34373070
PF2-63	Genome	<i>Serratia</i>	PF2	SAMN34373071
PF-27	Genome	<i>Serratia</i>	PF	SAMN34373072

NA not applicable, EA *Eumaeus atala*, PF *Pharaxanota floridana*, RS *Rhopalotria slossoni*.

<sup>a</sup>The origin refers to the co-culture from which the metagenome DNA was extracted and sequenced or the bacterial strain isolated and subsequently sequenced.

<sup>b</sup>NCBI accession number.

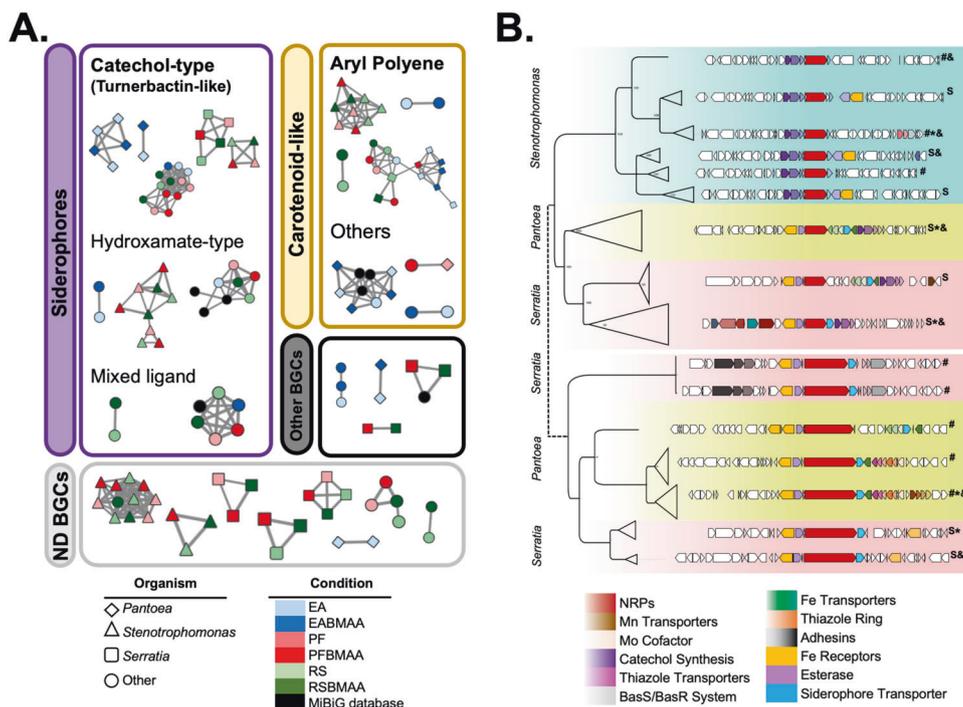
could isolate and confirm by direct sequencing of their rRNA 16 S gene: (i) EA-12 and EABMAA-21 as representatives of the *Pantoea* 'OTU 7' from *E. atala*; (ii) strains PF2-63 and PF-27 as representatives of the *Serratia* 'OTU 52' from *P. floridana*; and (iii) the strains PFBMAA-4 and RS-48, as representative of the *Stenotrophomonas* 'OTU 792' and 'OTU 6' from *P. floridana* and *R. slossoni*, respectively. Interestingly, these latter strains represent two distinct lineages within the *Stenotrophomonas* genus. These OTUs are either present or ubiquitous in the microbiomes of all insect species from the global sampling (Fig. S2, Table S10). We also identified and isolated strains of the previously mentioned uncultured core *Klebsiella*-like genus, but further analysis of these isolates was not pursued due to our inability to obtain representative MAGs and the uncertainty about their taxonomic classification needed for phylogenomics analyses.

### Phylogenomics of natural products highlights bacterial siderophore BGCs

Metagenomes from the six sub-community co-cultures were analyzed with antiSMASH and found to contain a total of 262 predicted BGCs, of which 85 were complete and non-redundant. These included 16 different types of BGCs, the most abundant of which were biosynthetic systems for highly conserved aryl polyenes and divergent non-ribosomal peptides (NRPs), with 20 and 19 BGCs, respectively (20% of the total). The remaining BGCs were bacteriocins (8), terpenes (11), NRPS-independent siderophores (5), hybrid PKS-NRPs (3), thiopeptides (3), resorcinols (2), butyrolactone (2), Hserlactone (1), type III PKS (1), lasso peptide (1), acyl amino acids (1), type II PKS (1), ectoine (1), and others (6) (Table S11). Of these, two core BGCs were present in all *EcoMining* metagenomes from the three insects: one turnerbactin-like divergent BGC from the catechol-type siderophore category, and

one highly conserved carotenoid-like aryl polyene (Fig. S3, Table S12). Further analysis of these two core BGCs using BiG-SCAPE revealed that the turnerbactin-like and aryl polyene BGCs formed five and four clans, respectively, mostly composed of scaffolds assigned to *Pantoea* "OTU 7" mainly from the *E. atala* metagenomes, and *Serratia* "OTU 52" and *Stenotrophomonas* "OTU 6" and "OTU 792" mainly from the *R. slossoni* metagenomes. While the aryl polyenes are highly conserved the siderophore BGCs showed signs of diversification. For instance, other siderophore BGCs of the hydroxamate and mixed ligands types were found in *Stenotrophomonas* and other bacterial genera, such as *Acinetobacter*, *Klebsiella*, and *Enterobacter* (Fig. 3A).

To explore the evolutionary dynamics of the biosynthetic potential of these bacterial taxa, whole BGC CORASON phylogenies and core proteome species phylogenies were obtained. The *Pantoea* phylogeny indicates that *Pantoea* sp. EA-12 and *Pantoea* sp. EABMAA-21 are part of a monophyletic clade that includes *Pantoea* species previously isolated from plants and rivers, while *Pantoea* sp. Pa-EAmG belongs to a monophyletic clade of *Pantoea* isolated from insects, soil, and plants (Fig. S4). In the *Serratia* phylogeny, *Serratia* sp. PF2-63 and *Serratia* sp. PF-27 formed a monophyletic clade with MAGs *Serratia* sp. Se-RSBMAAmG, *Serratia* sp. Se-PFmG, *Serratia* sp. Se-PFBMAAmG, and *Serratia* sp. Se-RSmG, and *Serratia* sp. OMLW3, an isolate from the insect, *Orius majusculus* [72] (Fig. S5). The *Stenotrophomonas* phylogeny was composed of two monophyletic clades, one containing the genomes and MAGs from *Stenotrophomonas* sp. RS-48, *Stenotrophomonas* sp. St-RSmG, and *Stenotrophomonas* sp. St-RSBMAAmG, plus isolates from soils and rivers; and the other composed of *Stenotrophomonas* sp. PFBMAA-4, *Stenotrophomonas* sp. St-PFmG, and *Stenotrophomonas* sp. St-PFBMAAmG (Fig. S6). These phylogenies place the strains within an evolutionary context specific to



**Fig. 3 Conserved BGCs of cycadivorous core gut microbiome.** **A** BIG-SCAPE networks containing 85 complete and non-redundant BGCs identified in the six metagenomes obtained from the *EcoMining* experiments. Turnerbactin-like, catechol-type siderophore, and aryl polyene BGCs formed five and four clans, respectively. **B** CORASON analysis with 230 turnerbactin-like BGCs from *Serratia*, *Pantoea*, and *Stenotrophomonas* genomes were used to reconstruct a BGC phylogeny. Turnerbactin-like BGCs can be separated into two groups based on changes in the NRPs' protein sequences and their genomic context. The first group was found in all three bacterial genera, while the second was found only in *Serratia* and *Pantoea*. Two types of divergently related BGCs were found: (i) *bona fide* catechol-type siderophore BGCs (marked with S), and (ii) siderophore-like BGCs, no Fur box detected (#). BGCs found in the isolated strains from co-cultures are denoted with an asterisk (\*), and BGCs found in MAGs with ampersands (&). EA, *Eumaeus atala* (BMAA +/-); PF, *Pharaxonotha floridana* (BMAA +/-); RS, *Rhopalotria slossoni* (BMAA +/-). Expanded versions of the trees supporting this general phylogeny are provided as supplementary information (Figs. S8, S9, S10, and S11).

the insect's gut, and provide a framework to look into BGC distribution.

A contrasting evolutionary dynamic was found for these two conserved BGCs. On one hand, the carotenoid-like aryl polyene BGC was highly conserved across the phylogenies of the core bacteria *Pantoea* "OTU 7", *Serratia* "OTU 52", and *Stenotrophomonas* "OTU 6 and OTU 792", with changes limited to taxonomic distance (Fig. S7). In contrast, the turnerbactin-like BGCs, for which in-depth functional annotation of each locus was performed, were composed of two types of divergently related BGCs: (i) *bona fide* catechol-type siderophore BGCs and (ii) siderophore-like BGCs for which some of the expected key functional elements for a siderophore BGC, e.g., regulatory iron boxes could not be detected (Figs. S8, S9, S10, and S11). This analysis also showed that turnerbactin-like BGCs could be separated into two groups based on changes in the NRP protein sequences and their genomic vicinity. The first group was found in all three bacterial genera, while the second was found only in *Serratia* and *Pantoea* (Fig. 3B). The gene composition of the BGCs—i.e., the presence or absence of genes for recognition of certain metals such as Fe, Mn, and Mo—further separated the turnerbactin-like BGCs found in each genus.

#### Semi-untargeted metabolomics of bacterial siderophores: evidence of an unprecedented desferrioxamine pathway

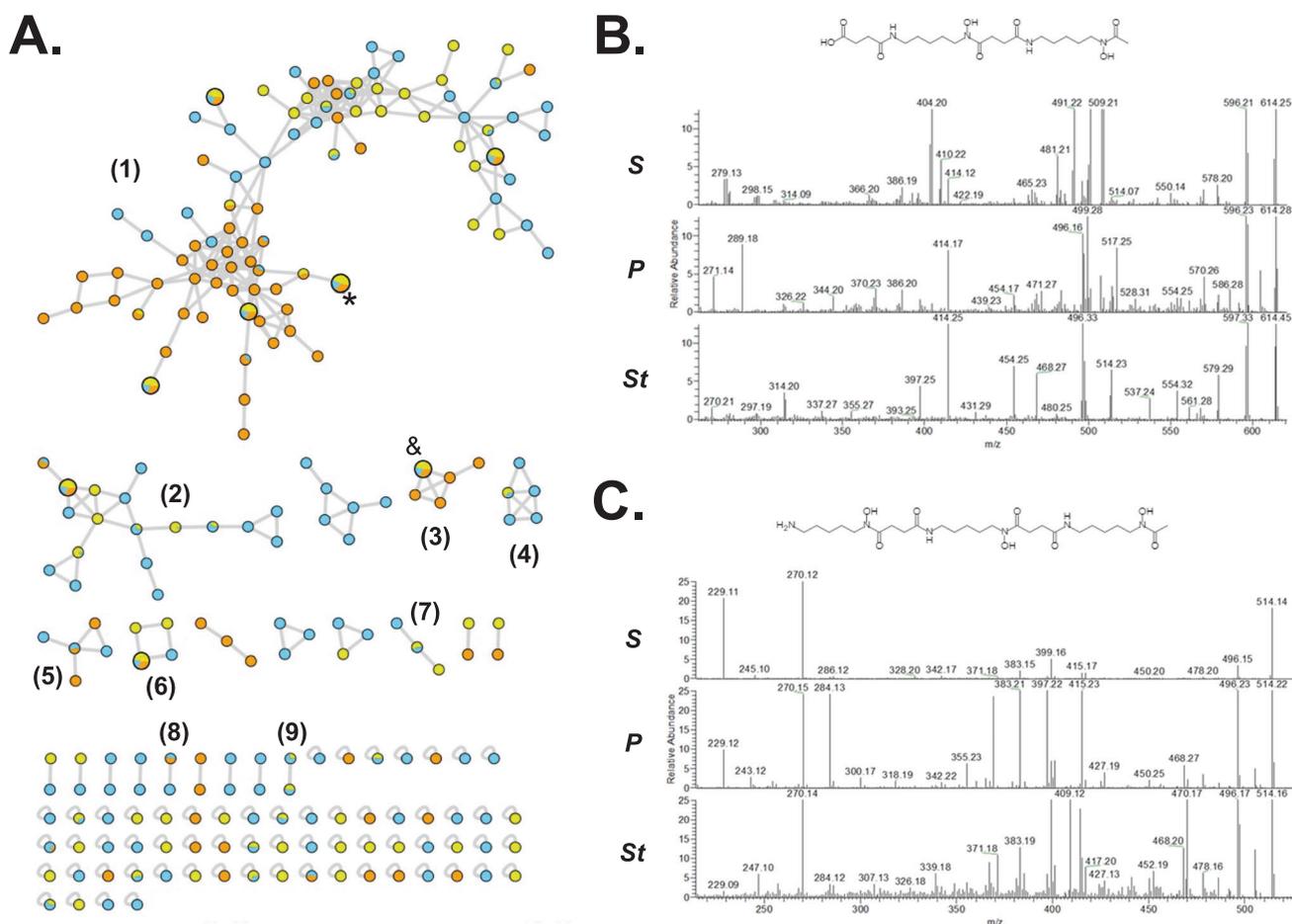
To explore the chemical diversity of siderophores in the isolated strains, Global Natural Products Social (GNPS) molecular networks [70] were constructed using MS data obtained from pooled peaks collected after HPLC separation (Fig. S12). To ensure that metal-chelating metabolites were targeted, we adopted a differential growth condition and analytical sample preparation with +/- ferric

iron, as we have previously done in other microbial ecology studies [65]. Identification of metal chelating metabolites revealed 9 clusters of related compounds, with 7 of these metabolites conserved between the three bacterial core genera *Pantoea*, *Serratia*, and *Stenotrophomonas* (Fig. 4A, Table S13). Among these, we identified as conserved the siderophores ferrioxamine H ( $[\text{M}-2\text{H}+\text{Fe}]^+ = 514.1717 \text{ m/z}$ , cluster 1, Fig. 4B) and B ( $[\text{M}-2\text{H}+\text{Fe}]^+ = 614.2715 \text{ m/z}$ , cluster 3, Fig. 4C), which are the succinylated dimer and the acetyl-capped trimer, respectively.

Unexpectedly, however, a putative desferrioxamine (*des*) BGC was not detected in any of the assembled metagenomes of the co-cultures, nor in the genomes of the isolated taxa that were sequenced independently. Our unsuccessful genome mining efforts included antiSMASH analyses; direct mapping of the reads of all six genome sequencing libraries against the *des* BGC from a close relative known to produce desferrioxamines, *Pantoea agglomerans* (MiBiG ID: BGC0001572); and identification of the Fur-dependent iron boxes implicated in iron metabolism regulation and previously exploited for siderophores genome mining [73]. The results of the latter analysis identified loci potentially involved in iron metabolism, of which some included biosynthetic enzymes and/or receptors, but were dissimilar to the DesABCD enzymes or the DesE receptor (Table S14). Together, these results provide evidence that diverse metal-chelating metabolites are present in the gut microbiome of cycad-feeding insects.

#### DISCUSSION

According to the 16S community surveys, the core gut microbiomes of Coleoptera tended to be richer than those of



**Fig. 4 Conserved chemically divergent siderophores present in core taxa. A** GNPS networks revealed 9 conserved clusters of metal-chelating metabolites produced by core bacteria of the genera *Pantoea* (Blue), *Serratia* (Yellow), and *Stenotrophomonas* (Orange). Four of these clusters included at least one metabolite shared amongst all three genera, but cluster 1 had four of these metabolites. Ferrioxamines H (\*) and B (&) were identified in clusters 1 and 3, respectively, after a combined manual and GNPS annotation. Selected metabolites of each group, and their MS features, are provided in Table S13. **B** Ferrioxamine H and **C** Ferrioxamine B. MS/MS spectra of these known siderophores was manually analyzed and compared with standards and databases to confirm the validity of the GNPS annotation. *S* *Serratia*, *P* *Pantoea*, *St* *Stenotrophomonas*.

Lepidoptera, which is consistent with previous reports that lepidopteran larvae typically do not host large resident bacterial communities [33, 74]. However, we found that a cross-species core microbiome is widespread among cycadivorous insects—including several lepidopterans—regardless of geographic location, insect order, life stage, host use, or feeding guild. Several of these core bacteria belong to clades that engage in beneficial and pathogenic associations with plants and insects, such as *Pantoea* and *Stenotrophomonas*. Although it is possible that similar microbiome convergence may be a more widespread feature of insect faunas that feed on chemically distinctive host plants, as postulated by the so-called “gut microbial facilitation hypothesis” [19], most insect microbiome surveys focus on a single species or group rather than on the microbiota derived from a community of herbivores, making it difficult to compare our results to those from similarly specialized insect faunas. For example, it would be interesting to compare our results with pine-feeding insects and their apparently convergent gut microbial associates that aid in the detoxification of terpenes [75–77] if complete and comparable data sets become available.

Co-culturing of microbial sub-communities, the key feature of the *EcoMining* approach, revealed that BMAA does not affect bacterial community composition. This is to be expected, as the insects’ gut lumens should contain BMAA from the insects’ food

plants, and any bacteria that cannot survive in BMAA-rich environments are likely naturally excluded from these communities. Genome and metagenome mining revealed a rich biosynthetic potential, including a set of core BGCs found in all three insect species. Although only three insect species were included in the BGC analysis, all 12 insect species from the global 16S survey contained the same bacterial OTUs from which the core BGCs were found. It may therefore be the case that the specific biosynthetic potential observed in the gut microbiomes of *Eumaeus*, *Rhopolotria*, and *Pharaxanotia* is widespread among cycadivorous insects, and could potentially contribute into the ability of these microorganisms to tolerate BMAA. Further BGC analysis predicted a high diversity of metal-chelating metabolites, e.g., siderophores, which are low molecular weight metabolites secreted by bacteria to chelate metals, often in response to iron deficiency, but which also play many ecologically relevant roles within bacterial communities [78]. As for the microbiomes of cycad-feeding insects, it remains to be seen if these siderophores are specific to core bacterial taxa or are common features of other bacteria within insects’ gut microbiomes.

Several catechol-type siderophores were predicted from the genomes of bacteria known to be enterobactin-like siderophore producers, such as *Pantoea* and *Stenotrophomonas* [79, 80]. Based on this evidence, the availability of iron (and potentially also other

metals) may drive the diversification of these specialized metabolites within the insects' gut microbiomes, as was recently shown in *Serratia plymuthica* with its combinatorial siderophore biosynthetic potential from polyamines [81]. Given that previous research into insect-siderophore interactions is limited, it is unclear whether the rich metal-chelating diversity observed within the microbiomes of cycadivorous is a widespread feature of insect gut microbiomes, or if it is limited to insects with specialized diets. Semi-untargeted metabolomics and molecular networking of metal-chelating metabolites produced by selected bacteria taxa showed signs of chemical diversification, consistent with the genetic diversity revealed by (meta)genome mining. Carotenoid-like aryl polyene biosynthetic systems, on the other hand, were found to be conserved and remain to be investigated, but it is interesting to note that they have been implicated in honeybee resistance to pathogenic fungi [82].

It is unknown how siderophores in the guts of cycadivorous insects might impact host fitness. A beneficial role of siderophores has been observed in plants, whereby siderophores produced by commensal or mutualistic bacteria can reduce growth of plant pathogens [83]. While research into siderophore-insect interactions is limited, previous studies have detected bacterial siderophores in the guts of mosquitoes, honey bees, grasshoppers, and moths [84–87]. In each of these cases, however, siderophore production was implicated in bacterial pathogenicity with clear costs for host fitness. Indeed, for many phyto- and entomopathogenic bacteria, siderophore production contributes to host colonization and microbial virulence, for example by competing with the host for iron and other essential micronutrients that are typically limited in plants. Without detailed assessments of the availability of iron and other metals to cycadivorous insects, it is unclear whether siderophore-producing bacteria may similarly compete with their hosts for iron. It also remains to be determined whether the siderophore chemical diversity measured here is different to that seen in taxonomically-related bacteria inhabiting different niches, a tantalizing challenge just starting to become accessible [88].

An interesting possibility is that siderophore-producing bacteria could provide protective benefits for cycadivorous insects by mediating chemical interactions between metal ions and BMAA inside the gut lumen. BMAA is also capable of chelating iron and other metals [89], and the formation of BMAA-metal ion complexes has been proposed as one possible cause of neurotoxicity in humans [90–92]. The chelation of divalent metal ions alters the equilibrium between BMAA and its corresponding carbamate adducts, which are required for binding to glutamate receptors and exerting neurotoxic effects [91]. Within the insect gut, the affinity of siderophores, such as desferrioxamines, for metal ions is likely to be higher than that of BMAA, such that bacterial siderophores may compete with BMAA for limited iron (and/or other metal ions serving as enzyme co-factors) within the insect gut lumen. In turn, this would alter the equilibrium between BMAA and its carbamate adducts affecting BMAA's toxicity within the gut. This hypothesis warrants further investigation. A somewhat analogous scenario is found in the plant coumarins, e.g., the highly toxic and biologically active scopoletin, implicated in shaping the root microbiota by a 'secondary' siderophore activity [93, 94] that acts in concert with the bacterial redox-active phenazines produced by *Pseudomonas* [95].

Chemical diversification of bacterial specialized metabolites in the guts of cycadivorous insects is likely the result of complex ecological and evolutionary forces, including the biomolecular and physicochemical features of natural products synthesized by the insect and/or plant hosts, and may lead to novel and unexpected metabolic functions [96]. Within this context, iron-limitation is likely not the only factor driving siderophore production within the insects' bacterial communities. Instead, we hypothesize that

metal-chelating metabolites, e.g., siderophores, serve other functions besides chelating metals as a means for nutrient acquisition, as recently suggested in other systems [97]. That this genetic and chemical diversity is present in bacterial taxa found across all 12 of the insects we surveyed suggests that these specialized metabolites are a key feature of cycadivorous insects' gut microbiomes. Future studies that take a plant-centered approach to surveying insect microbiomes, in combination with functional investigations, may ultimately yield more insight into the role of insects' gut microbiomes in facilitating convergent adaptation to specific hosts.

## DATA AVAILABILITY

Raw sequence data and genome assemblies used in this study are deposited in the NCBI archive under accession number PRJNA961367. 16 S analysis code, biom file, phylogenetic tree, and metadata files are available on the Harvard Dataverse at <https://doi.org/10.7910/DVN/OUSYHA>

## REFERENCES

1. Condamine FL, Nagalingum NS, Marshall CR, Morlon H. Origin and diversification of living cycads: a cautionary tale on the impact of the branching process prior in Bayesian molecular dating. *BMC Evol Biol.* 2015;15:65.
2. Liu Y, Wang S, Li L, Yang T, Dong S, Wei T, et al. The Cycas genome and the early evolution of seed plants. *Nature Plants.* 2022;8:389–401.
3. Liu X, Rush T, Zapata J, Lobner D.  $\beta$ -N-methylamino-L-alanine induces oxidative stress and glutamate release through action on system Xc<sup>-</sup>. *Exp Neurol.* 2009;217:429–33.
4. Schneider D, Wink M, Sporer F, Lounibos P. Cycads: their evolution, toxins, herbivores and insect pollinators. *Naturwissenschaften.* 2002;89:281–94.
5. Nash RJ, Bell EA, Ackery PR. The protective role of cycasin in cycad-feeding lepidoptera. *Phytochemistry.* 1992;31:1955–7.
6. Rothschild M, Nash RJ, Bell EA. Cycasin in the endangered butterfly *Eumaeus atala florida*. *Phytochemistry.* 1986;25:1853–4.
7. Teas HJ. Cycasin synthesis in *Seirarctia echo* (Lepidoptera) larvae fed methylazoxymethanol. *Biochem Biophys Res Commun.* 1967;26:686–90.
8. Koksharova OA, Safronova NA. Non-Proteinogenic Amino Acid  $\beta$ -N-Methylamino-L-Alanine (BMAA): Bioactivity and Ecological Significance. *Toxins.* 2022;14:539.
9. Mantas MJ, Nunn PB, Codd GA, Barker D. Genomic insights into the biosynthesis and physiology of the cyanobacterial neurotoxin 3-N-methyl-2,3-diaminopropanoic acid (BMAA). *Phytochemistry.* 2022;200:113198.
10. Okle O, Rath L, Galizia CG, Dietrich DR. The cyanobacterial neurotoxin beta-N-methylamino-L-alanine (BMAA) induces neuronal and behavioral changes in honeybees. *Toxicol Appl Pharmacol.* 2013;270:9–15.
11. Zhou X, Escala W, Papapetropoulos S, Bradley WG, Zhai RG. BMAA neurotoxicity in *Drosophila*. *Amyotroph Lateral Scler.* 2009;10:61–6.
12. Delcourt N, Claudepierre T, Maignien T, Arnich N, Mattei C. Cellular and Molecular Aspects of the  $\beta$ -N-Methylamino-L-alanine (BMAA) Mode of Action within the Neurodegenerative Pathway: Facts and Controversy. *Toxins.* 2017;10:6.
13. Brenner ED, Feinberg P, Runko S, Coruzzi GM. A mutation in the Proteasomal Regulatory Particle AAA-ATPase-3 in *Arabidopsis* impairs the light-specific hypocotyl elongation response elicited by a glutamate receptor agonist, BMAA. *Plant Mol Biol.* 2009;70:523–33.
14. Whitaker MR, Gilliéron F, Skirgaila C, Mescher MC, De Moraes CM. Experimental evidence challenges the presumed defensive function of a "slow toxin" in cycads. *Sci Rep.* 2022;12:6013.
15. Ruffner B, Péchy-Tarr M, Höfte M, Bloemberg G, Grunder J, Keel C, et al. Evolutionary patchwork of an insecticidal toxin shared between plant-associated pseudomonads and the insect pathogens *Photorhabdus* and *Xenorhabdus*. *BMC Genomics.* 2015;16:609.
16. Flury P, Aellen N, Ruffner B, Péchy-Tarr M, Fataar S, Metla Z, et al. Insect pathogenicity in plant-beneficial pseudomonads: phylogenetic distribution and comparative genomics. *ISME J.* 2016;10:2527–42.
17. Whitaker MR, Salzman S. Ecology and evolution of cycad-feeding Lepidoptera. *Ecol Lett.* 2020;23:1862–77.
18. Whitaker MR, Banack SA, Mescher MC, Cox PA, De Moraes CM. BMAA in cycad-feeding Lepidoptera: defensive sequestration or bioaccumulation? *Front Ecol Evol.* 2023;11:1114636.
19. Hammer TJ, Bowers MD. Gut microbes may facilitate insect herbivory of chemically defended plants. *Oecologia.* 2015;179:1–14.
20. Robbins RK, Cong Q, Zhang J, Shen J, Quer Riera J, Murray D, et al. A switch to feeding on cycads generates parallel accelerated evolution of toxin tolerance in

- two clades of *Eumaeus* caterpillars (Lepidoptera: Lycaenidae). *Proc Natl Acad Sci USA*. 2021;118:e2018965118.
21. Jing TZ, Qi FH, Wang ZY. Most dominant roles of insect gut bacteria: digestion, detoxification, or essential nutrient provision? *Microbiome*. 2020;8:38.
  22. van den Bosch TJ, Welte CU. Detoxifying symbionts in agriculturally important pest insects. *Microb Biotechnol*. 2017;10:531–40.
  23. Kikuchi Y, Hayatsu M, Hosokawa T, Nagayama A, Tago K, Fukatsu T. Symbiont-mediated insecticide resistance. *Proc Natl Acad Sci*. 2012;109:8618–22.
  24. Blanton AG, Peterson BF. Symbiont-mediated insecticide detoxification as an emerging problem in insect pests. *Front Microbiol*. 2020;11:547108.
  25. Jacoby RP, Koprivova A, Kopriva S. Pinpointing secondary metabolites that shape the composition and function of the plant microbiome. *J Exp Bot*. 2021;72:57–69.
  26. Shikano I, Rosa C, Tan CW, Felton GW. Trittrophic interactions: microbe-mediated plant effects on insect herbivores. *Ann Rev Phytopathol*. 2017;55:313–31.
  27. Pang Z, Chen J, Wang T, Gao C, Li Z, Guo L, et al. Linking plant secondary metabolites and plant microbiomes: a review. *Front Plant Sci*. 2021;12:621276.
  28. Koprivova A, Kopriva S. Plant secondary metabolites altering root microbiome composition and function. *Curr Opin Plant Biol*. 2022;67:102227.
  29. Salzman S, Whitaker M, Pierce NE. Cycad-feeding insects share a core gut microbiome. *Biol J Linn Soc*. 2018;123:728–38.
  30. Cibrián-Jaramillo A, Barona-Gómez F. Increasing Metagenomic Resolution of Microbiome Interactions Through Functional Phylogenomics and Bacterial Sub-Communities. *Front Genet*. 2016;7:4.
  31. Staude HS, Sihvonen P. Revision of the African geometrid genus *Zerenopsis* C. & R. Felder—moths with peculiar life histories and mating behaviours (Geometridae: Ennominae: Diptychini). *METAMORPHOSIS. J Lepidopter Soc Afr*. 2014;25:11–55.
  32. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N, et al. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J*. 2012;6:1621–4.
  33. Whitaker MR, Salzman S, Sanders J, Kaltenpoth M, Pierce NE. Microbial communities of lycaenid butterflies do not correlate with larval diet. *Front Microbiol*. 2016;7:1920.
  34. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, et al. QIIME allows analysis of high-throughput community sequencing data. *Nat Methods*. 2010;7:335–6.
  35. Edgar RC. UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat Methods*. 2013;10:996–8.
  36. McMurdie PJ, Holmes S. phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLoS ONE*. 2013;8:e61217.
  37. Mariadassou M. PhyloseqExtended. R package version 0.1.1.6. 2021. <https://github.com/mahendra-mariadassou/phyloseq-extended>
  38. Smirnova E, Huzurbazar S, Jafari F. PERfect: PERmutation Filtering test for microbiome data. *Biostatistics*. 2019;20:615–31.
  39. Berntzon L, Erasmie S, Celepli N, Eriksson J, Rasmussen U, Bergman B, et al. BMAA inhibits nitrogen fixation in the cyanobacterium *Nostoc* sp. PCC 7120. *Mar Drugs*. 2013;11:3091–108.
  40. Popova AA, Rasmussen U, Semashko TA, Govorun VM, Koksharova OA. Stress effects of cyanotoxin  $\beta$ -methylamino-L-alanine (BMAA) on cyanobacterial heterocyst formation and functionality. *Environ Microbiol Rep*. 2018;10:369–77.
  41. Sambrook J, Russell DW. Purification of nucleic acids by extraction with phenol: chloroform. *Cold Spring Harb Protoc*. 2006;2006:pdb-rot4455.
  42. Wingett SW, Andrews S. FastQ Screen: A tool for multi-genome mapping and quality control. *F1000Research*. 2018;7:1338.
  43. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 2014;30:2114–20.
  44. Wood DE, Lu J, Langmead B. Improved metagenomic analysis with Kraken 2. *Genome Biol*. 2019;20:3.
  45. Dabdoub SM. kraken-biom: Enabling interoperable format conversion for Kraken results (Version 1.2). 2016. <https://github.com/smdabdoub/kraken-biom>.
  46. Chamberlain SA, Szöcs E. taxize: taxonomic search and retrieval in R. *F1000Research*. 2013;2:191.
  47. Nurk S, Meleshko D, Korobeynikov A, Pevzner PA. metaSPAdes: a new versatile metagenomic assembler. *Genome Res*. 2017;27:824–34.
  48. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, et al. The RAST Server: rapid annotations using subsystems technology. *BMC Genomics*. 2008;9:75.
  49. Blin K, Shaw S, Steinke K, Villebro R, Ziemert N, Lee SY, et al. antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Res*. 2019;47:W81–7.
  50. Navarro-Muñoz JC, Selem-Mojica N, Mullowney MW, Kautsar SA, Tryon JH, Parkinson EL, et al. A computational framework to explore large-scale biosynthetic diversity. *Nat Chem Biol*. 2020;16:60–68.
  51. Lex A, Gehlenborg N, Strobel H, Vuillemot R, Pfister H. UpSet: Visualization of Intersecting Sets. *IEEE Trans Vis Comput Graph*. 2014;20:1983–92.
  52. Parrello B, Butler R, Chlenski P, Pusch GD, Overbeek R. Extraction of near-complete genomes from metagenomic samples: a new service in PATRIC. *Cold Spring Harb Lab*. 2019.12.13.874651.
  53. Eren AM, Esen ÖC, Quince C, Vineis JH, Morrison HG, Sogin ML, et al. Anvi'o: an advanced analysis and visualization platform for 'omics data. *PeerJ*. 2015;3:e1319.
  54. Vinuesa P, Ochoa-Sánchez LE, Contreras-Moreira B. GET\_PHYLOMARKERS, a software package to select optimal orthologous clusters for phylogenomics and inferring pan-genome phylogenies, used for a critical geno-taxonomic revision of the genus *Stenotrophomonas*. *Front Microbiol*. 2018;9:771.
  55. Ronquist F, Huelsenbeck JP. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*. 2003;19:1572–4.
  56. Rambaut A. FigTree v1. 4.2, a graphical viewer of phylogenetic trees. <http://tree.bio.ed.ac.uk/software/figtree/>.
  57. Lane D. 16S/23S rRNA sequencing. *Nucleic Acid Tech. Bac. Syst*. 1991;1:115–75.
  58. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res*. 2013;41:D590–D596.
  59. Thompson JD, Gibson TJ, Higgins DG. Multiple sequence alignment using ClustalW and ClustalX. *Curr Protoc Bioinform*. 2003;1:2–3.
  60. Castresana J. Selection of Conserved Blocks from Multiple Alignments for Their Use in Phylogenetic Analysis. *Mol Biol Evol*. 2000;17:540–52.
  61. Moore RM, Harrison AO, McAllister SM, Polson SW, Wommack KE. Iroki: automatic customization and visualization of phylogenetic trees. *PeerJ*. 2020;8:e8584.
  62. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol*. 2012;19:455–77.
  63. Hantke K. Iron and metal regulation in bacteria. *Curr Opin Microbiol*. 2001;4:172–7.
  64. Baichoo N, Helmann JD. Recognition of DNA by Fur: a Reinterpretation of the Fur Box Consensus Sequence. *J Bacteriol*. 2002;184:5826–32.
  65. Cruz-Morales P, Ramos-Aboites HE, Licona-Cassani C, Selem-Mojica N, Mejía-Ponce PM, Souza-Saldívar V, et al. Actinobacteria phylogenomics, selective isolation from an iron oligotrophic environment and siderophore functional characterization, unveil new desferrioxamine traits. *FEMS Microbiol Ecol*. 2017;93:1–12.
  66. Sauvage C, Franza T, Expert D. Analysis of the *Erwinia chrysanthemi* ferrichrysobactin receptor gene: resemblance to the *Escherichia coli* fepA-fes bidirectional promoter region and homology with hydroxamate receptors. *J Bacteriol*. 1996;178:1227–31.
  67. Panina EM, Mironov AA, Gelfand MS. Comparative analysis of FUR regulons in gamma-proteobacteria. *Nucleic Acids Res*. 2001;29:5195–206.
  68. Hiard S, Marée R, Colson S, Hoskisson PA, Titgemeyer F, van Wezel GP, et al. PREDetector: a new tool to identify regulatory elements in bacterial genomes. *Biochem Biophys Res Commun*. 2007;357:861–4.
  69. Wang M, Carver JJ, Phelan VV, Sanchez LM, Garg N, Peng Y, et al. Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. *Nature Biotechnol*. 2016;34:828–37.
  70. Aron AT, Gentry EC, McPhail KL, Nothias LF, Nothias-Esposito M, Bouslimani A, et al. Reproducible molecular networking of untargeted mass spectrometry data using GNPS. *Nat Protoc*. 2020;15:1954–91.
  71. Gutiérrez-García K, Bustos-Díaz ED, Corona-Gómez JA, Ramos-Aboites HE, Selem-Mojica N, Cruz-Morales P, et al. Cycad Coralloid Roots Contain Bacterial Communities Including Cyanobacteria and Caulobacter spp. That Encode Niche-Specific Biosynthetic Gene Clusters. *Genome Biol Evol*. 2019;11:319–34.
  72. Chen X, Hitchings MD, Mendoza JE, Balanza V, Facey PD, Dyson PJ, et al. Comparative Genomics of Facultative Bacterial Symbionts Isolated from European Orius Species Reveals an Ancestral Symbiotic Association. *Front Microbiol*. 2017;8:1969.
  73. Fillat MF. The FUR (ferric uptake regulator) superfamily: diversity and versatility of key transcriptional regulators. *Arch Biochem Biophys*. 2014;546:41–52.
  74. Hammer TJ, Janzen DH, Hallwachs W, Jaffe SP, Fierer N. Caterpillars lack a resident gut microbiome. *Proc Natl Acad Sci*. 2017;114:9641–6.
  75. Liu F, Ye F, Cheng C, Kang Z, Kou H, Sun J. Symbiotic microbes aid host adaptation by metabolizing a deterrent host pine carbohydrate d-pinitol in a beetle-fungus invasive complex. *Sci Adv*. 2022;8:eadd5051.
  76. Berasategui A, Axelsson K, Nordlander G, Schmidt A, Borg-Karlson AK, Gershenson J, et al. The gut microbiota of the pine weevil is similar across Europe and resembles that of other conifer-feeding beetles. *Mol Ecol*. 2016;25:4014–31.
  77. Berasategui A, Salem H, Paetz C, Santoro M, Gershenson J, Kaltenpoth M, et al. Gut microbiota of the pine weevil degrades conifer diterpenes and increases insect fitness. *Mol Ecol*. 2017;26:4099–110.
  78. Kramer J, Özkaya Ö, Kümmerli R. Bacterial siderophores in community and host interactions. *Nat Rev Microbiol*. 2020;18:152–63.

79. Hisatomi A, Shiwa Y, Fujita N, Koshino H, Tanaka N Identification and structural characterisation of a catechol-type siderophore produced by *Stenotrophomonas maltophilia* K279a. *Microbiology*. 2021;167:1071.
80. Reitz ZL, Sandy M, Butler A. Biosynthetic considerations of triscatechol siderophores framed on serine and threonine macrolactone scaffolds. *Metallomics*. 2017;9:824–39.
81. Cleto S, Haslinger K, Prather KLJ, Lu TK. Natural combinatorial genetics and prolific polyamine production enable siderophore diversification in *Serratia plymuthica*. *BMC Biol*. 2021;19:e46.
82. Miller DL, Smith EA, Newton ILG. A Bacterial Symbiont Protects Honey Bees from Fungal Disease. *mBio*. 2021;12:e0050321.
83. Kupferschmid P, Maurhofer M, Keel C. Promise for plant pest control: root-associated pseudomonads with insecticidal activities. *Front Plant Sci*. 2013;4:287.
84. Ganley JG, Pandey A, Sylvester K, Lu K, Toro-Moreno M, Rüttschlin S, et al. A Systematic Analysis of Mosquito-Microbiome Biosynthetic Gene Clusters Reveals Antimalarial Siderophores that Reduce Mosquito Reproduction Capacity. *Cell Chem Biol*. 2020;27:817–826.e5.
85. Hertlein G, Müller S, Garcia-Gonzalez E, Poppinga L, Süßmuth RD, Genersch E. Production of the catechol type siderophore bacillibactin by the honey bee pathogen *Paenibacillus* larvae. *PLoS One*. 2014;9:e108272.
86. Indiragandhi P, Anandham R, Madhaiyan M, Kim GH, Sa T. Cross-utilization and expression of outer membrane receptor proteins for siderophore uptake by Diamondback moth *Plutella xylostella* (Lepidoptera: Plutellidae) gut bacteria. *FEMS Microbiol Lett*. 2008;289:27–33.
87. Sonawane MS, Chaudhary RD, Shouche YS, Sayyed RZ. Insect gut bacteria: a novel source for siderophore production. *Proc Natl Acad Sci India Sec B*. 2018;88:567–72.
88. Elser D, Pflieger D, Villette C, Moegle B, Miesch L, Gaquerel E. Evolutionary metabolomics of specialized metabolism diversification in the genus *Nicotiana* highlights N-acylnicotine innovations. *Sci Adv*. 2023;9:eade8984.
89. Glover WB, Liberto CM, McNeil WS, Banack SA, Shipley PR, Murch SJ. Reactivity of  $\beta$ -methylamino-L-alanine in complex sample matrices complicating detection and quantification by mass spectrometry. *Anal Chem*. 2012;84:7946–53.
90. Diaz-Parga P, Goto JJ, Krishnan VV. Chemistry and chemical equilibrium dynamics of BMAA and its carbamate adducts. *Neurotox Res*. 2018;33:76–86.
91. Diaz-Parga P, Goto JJ, Krishnan VV. On the Differential Roles of Mg, Zn, and Cu in the Equilibrium of  $\beta$ -N-Methyl-Amino-L-Alanine (BMAA) and its Carbamates. *Neurotox Res*. 2021;39:6–16.
92. Nunn PB, O'Brien P, Pettit LD, Pyburn SI. Complexes of zinc, copper, and nickel with the nonprotein amino acid L- $\alpha$ -amino- $\beta$ -methylaminopropionic acid: A naturally occurring neurotoxin. *J Inorg biochem*. 1989;37:175–83.
93. Stringlis IA, De Jonge R, Pieterse CM. The age of coumarins in plant-microbe interactions. *Plant Cell Physiol*. 2019;60:1405–19.
94. Harbort, Hashimoto CJ, Inoue M, Niu H, Guan Y, Rombolà R, et al. Root-secreted coumarins and the microbiota interact to improve iron nutrition in *Arabidopsis*. *Cell Host Microbe*. 2020;28:825–37.
95. McRose DL, Li J, Newman DK. The chemical ecology of coumarins and phenazines affects iron acquisition by pseudomonads. *Proc Natl Acad Sci*. 2023;120:e2217951120.
96. Chevrette MG, Gutiérrez-García K, Selem-Mojica N, Aguilar-Martínez C, Yañez-Olvera A, Ramos-Aboites HE, et al. Evolutionary dynamics of natural product biosynthesis in bacteria. *Nat Prod Rep*. 2020;37:566–99.
97. Shepherdson EM, Elliot MA. Cryptic specialized metabolites drive *Streptomyces* exploration and provide a competitive advantage during growth with other microbes. *Proc Natl Acad Sci*. 2022;119:e2211052119.

## ACKNOWLEDGEMENTS

We are grateful to Jose Luis Steffani-Vallejo, Juan D. Camarena-Alba, and Paulina Mejía-Ponce for their help in laboratory work, and Nelly Selem-Mojica, Diego Garfias-Gallegos,

María Fernanda Contreras-González, Hannah Augustijn for bioinformatics support. Thanks also to Hermann Staude and Horace Tan for their invaluable contributions to fieldwork, to William Tang and Irene Terry for providing additional insect specimens, and Rory Maher for making data collection possible. This study was supported by CONACyT #169701 and FON.INST./265/2016 to ACJ. CONACyT #179290, #177568 and #285746 to FBG, also supported by a Newton Advanced Fellowship of The Royal Society, UK (NAF/R2/180631). National Science Foundation (NSF) of the US #1541560 to NEP. MRLW was supported by NSF of the US Postdoctoral Research Fellowship in Biology (1309425), with fieldwork funded by the Harvard Museum of Comparative Zoology's Putnam Expeditionary Fund, the Wildlife Reserves Singapore Conservation Fund, and the Explorer's Club. SS was supported by NSF of the US Graduate Research Fellowship and Postdoctoral Research Fellowship in Biology (1906333).

## AUTHOR CONTRIBUTIONS

KGG, EDBD, ACJ, and FBG designed and conducted the shotgun metagenomic sequencing of co-cultured bacterial sub-communities experiments, while MRLW, SS, and NEP conducted field work and designed and conducted the 16 rRNA amplicon sequencing experiments from a worldwide sample of cycadivorous insects. HERA performed the HPLC and MS for bacterial siderophores identification. ZLR helped with the siderophore identification. All the authors contributed to writing and editing the manuscript.

## 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/s43705-023-00323-8>.

**Correspondence** and requests for materials should be addressed to Melissa R. L. Whitaker or Francisco Barona-Gómez.

**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.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023