Biological Journal of the Linnean Society, 2018, **123**, 728–738. With 3 figures.

Cycad-feeding insects share a core gut microbiome

SHAYLA SALZMAN*, MELISSA WHITAKER and NAOMI E. PIERCE

Department of Organismic and Evolutionary Biology, Harvard University, 26 Oxford Street, Cambridge, MA 02149, USA

Received 29 September 2017; revised 31 January 2018; accepted for publication 31 January 2018

Five insect species including three species of weevils (Coleoptera) and two species of lycaenid butterflies (Lepidoptera) that feed exclusively on the carcinogenic and neurotoxic tissues of cycads were found to share a core set of bacterial phylotypes, including the bacterium *Raoultella ornithinolytica*, which has known anti-cancer and nitrogen-fixing capabilities. The other shared bacteria belong to lineages that include insect-associated and extremophilic taxa. The presence of *Raoultella ornithinolytica* and an unknown Enterobacteriaceae was in contrast to a set of non-cycad-feeding relatives of these insects, none of which contained this same set of shared bacterial phylotypes. Given the considerable phylogenetic distance between the cycadivorous insect species as well as the fact that shared microbiota are not found in their non-cycad-feeding relatives, our data suggest that this core set of shared bacteria are important in helping cycad feeders detoxify their poisonous host plants.

ADDITIONAL KEYWORDS: BMAA – β-methylamino-L-alanine – coevolution – Cycadales – gut microbiome – herbivory – lycaenid – methylazoxymethanol.

INTRODUCTION

The plant order Cycadales comprises ten genera with ~350 species found across the tropics (Calonje, Stevenson & Stanberg, 2017). These dioecious gymnosperms are often obligately insect-pollinated and in many cases provide a brood site for pollinators that feed and develop on their tissue (Norstog, Stevenson & Niklas, 1989; Stevenson, Norstog & Fawcett, 1998; Terry *et al.*, 2012; Brookes *et al.*, 2015; Valencia-Montoya *et al.*, 2017). Cycads are among the most ancient lineages of seed plants with a fossil record extending back over 250 My (Mamay, 1969; Gao & Thomas, 1989) and while they are currently the most endangered plant order in the world (IUCN, 2017), they were once a dominant component of the Mesozoic flora (Friis, Chaloner & Crane, 1987; Thomas & Spicer, 1987). Cycads produce many secondary metabolites (De Luca *et al.*, 1982; Khabazian *et al.*, 2002), including two highly toxic compounds found in species throughout the order: methylazoxymethanol (MAM) (De Luca *et al.*, 1980; Moretti, Sabato, Gigliano, 1983) and β-methylamino-L-alanine (BMAA) (Vega & Bell, 1967). Whereas non-cycadivorous insect herbivores do not encounter these plant compounds in their diets, insects that have specialized on cycads must be capable of contending with both toxins concurrently, and each compound acts in very different ways.

MAM has both carcinogenic and neurotoxic effects (Laqueur & Spatz, 1968; Morgan & Hoffmann, 1983). This compound occurs in the plants in a non-toxic form in which the toxic agent, MAM, is attached to a glycoside. While MAM-glycosides are found in all cycad genera (De Luca *et al.*, 1980; Moretti, Sabato & Gigliano, 1981, 1983), the non-toxic storage form may differ depending on the glycoside attached to MAM. The two most common MAM-glycosides are cycasin, in which the glycoside is β-D-glucose (Nishida, Kobayashi & Nagahama, 1955), and macrozamin, in which the glycoside is a disaccharide of glucose and xylose (Lythgoe & Riggs, 1949). In both cases, toxicity results from cleavage of the glycoside from MAM by the activity of endogenous glucosidase enzymes (Laqueur & Spatz, 1968; Schneider *et al.*, 2002) that are produced in the digestive tracts of mammals and insects (Conchie & MacDonald, 1959; Terra & Ferreira, 1994). Once cleaved, MAM spontaneously degrades into formaldehyde and methyl-diazonium, a potent methylating agent (Morgan & Hoffmann, 1983). MAM-induced genetic alterations have been described in plants, mammals, yeast, bacteria and insects (Morgan &

*Corresponding author. E-mail: shaylasalzman@fas.harvard.edu

Hoffmann, 1983). MAM-glycosides have been found in all cycad genera and in all plant tissues that have been tested, including seeds, leaves, pollen and ovulate cones, roots and stems (Cooper, 1941; Riggs, 1954; De Luca *et al.*, 1980; Moretti *et al.*, 1981, 1983; Blagrove, Lilley & Higgins, 1984; Rothschild, Nash & Bell, 1986; Yagi & Tadera, 1987; Bowers & Larin, 1989; Lindblad, Tadera & Yagi, 1990; Nash, Bell & Ackery, 1992; Vovides *et al.*, 1993; Castillo-Guevara & Rico-Gray, 2003; Yagi, 2004; Prado, 2011; Nair & Staden, 2012; Prado *et al.*, 2014, 2016).

The second class of cycad toxins, BMAA, has received considerable attention by virtue of its implication as the causative agent of amyotrophic lateral sclerosis-parkinsonism-dementia, a human neurobiological disease that is endemic to the island of Guam and is also referred to as Guam's dementia (Cox, Banack & Murch, 2003). BMAA was first isolated from cycads by Vega & Bell (1967) in response to an extremely detailed account of cycad consumption and toxic effects in humans and cattle (Whiting, 1963). BMAA's toxicity arises from its disruption of glutamate receptor function. Glutamate receptors have deep homology and are found across plants and animals (Lam *et al.*, 1998; Chiu *et al.*, 1999; Lacombe *et al.*, 2001). BMAA works as a neurotoxin in insects (Goto, Koenig & Ikeda, 2012) and mammals, causing convulsions and neural degeneration (Spencer *et al.*, 1987) as well as abnormalities in brain development (Kisby, Moore & Spencer, 2013). In *Arabidopsis*, BMAA-induced glutamate receptor blockage affects signal transduction causing hypocotyl elongation and inhibiting cotyledon opening (Brenner *et al.*, 2000). Moreover, as a non-protein amino acid, BMAA can be incorporated into proteins, fundamentally altering their structure and function (Dunlap *et al.*, 2013). It is unknown how the cycad protects itself from this endogenous toxin. BMAA has been found in all cycad genera and in all tested tissues, including leaves, pollen and ovulate cones, seeds, pollen and roots (Dossaji & Bell, 1973; Duncan, Kopin & Crowley, 1989; Norstog & Fawcett, 1989; Vovides *et al.*, 1993; Pan *et al.*, 1997a, b; Banack & Cox, 2003).

Given the highly toxic nature of these compounds and their presence throughout plant tissues, only specialized insects are able to utilize cycads as a food source. Many of these are pollinating herbivores, including species from several genera of beetles that feed on pollen cone parenchyma tissue and coevolve with their host cycads (Donaldson, Nänni & Bösenberg, 1995; Stevenson *et al.*, 1998; Terry *et al.*, 2012; Suinyuy, Donaldson & Johnson, 2015). In addition to pollinating herbivores, several folivorous Lepidoptera are obligate cycad feeders for either their entire larval development or for their first few instars (e.g. Sihvonen, Staude & Mutanen, 2015).

Relatively little is known about potential toxin tolerance mechanisms in cycadivorous insects, which could include methods to detoxify, sequester and/or avoid plant defensive chemicals. One study focused on BMAA avoidance in the pollinating weevil, *Rhopalotria furfuracea* (previously *R. mollis*), which feeds on pollen cone parenchyma tissue of *Zamia furfuraceae* (Norstog & Fawcett, 1989). In this case, staining experiments demonstrated that plants sequester BMAA in specialized plant cells (idioblasts) that are able to pass through the insect gut intact. Interestingly, these idioblast cells were found intact in the pollen cones on which the weevils feed, but burst open in ovulate cones, which the weevils visit but never eat. The authors suggested that this plant mechanism restricts pollinator herbivory to the expendable pollen cone tissue. No known or proposed mechanism exists for BMAA tolerance or avoidance in leaf or ovulate cone feeders where the toxin is not sequestered in plant idioblasts.

The only other investigation into cycad toxin tolerance mechanisms focused on MAM tolerance in a leaf-feeding moth. A β-glucosidase enzyme was found to be localized in the guts of the larvae of the 'echo moth', *Sierarctia echo*, feeding on leaves of *Zamia integrifolia*, and this insect was shown to produce cycasin when fed the toxic MAM (Teas, 1967). Echo moths are aposematically coloured as both larvae and adults, and are thought to sequester and utilize the host plant's cycasin for protection. Teas (1967) hypothesized that this endogenous β-glucosidase enzyme rehydrolyses the toxic MAM with a glycoside, reverting the compound into its non-toxic form, cycasin. Laqueur & Spatz (1968) suggested that this enzyme could be of microbial origin, yet this remains untested, and it is unknown whether the enzyme is present in the guts of other cycadivorous insects.

The suggestion that microbes may play a role in toxin tolerance in *Sierarctia echo* moths is particularly interesting given the growing evidence that many herbivorous insects rely on symbiotic gut bacteria to mediate the challenges associated with plant-based diets (Douglas, 2013), including degrading plant secondary metabolites (Boone *et al.*, 2013; Ceja-Navarro *et al.*, 2015) and countering specialized plant defences (Chu *et al.*, 2013). It is possible that cycadivorous insects rely on gut bacteria to ameliorate their highly toxic diets, and given the similar plant defensive chemistry that all cycad herbivores are exposed to, even distantly related insects may have converged upon similar bacterial associations. To investigate these possibilities, we used 16S amplicon sequencing to characterize and compare the gut bacterial communities of five species of cycadivorous insects from two different orders, and to identify and investigate bacterial phylotypes that are shared across these phylogenetically distinct insect species.

MATERIAL AND METHODS

We collected three Curculionidae (Coleoptera) weevil species and two Lycaenidae (Lepidoptera) butterfly species that feed on a variety of cycad species and tissues (Table 1, Fig. 1). Insects were either immediately flash frozen whole in liquid nitrogen or dissected and the gut preserved in ethanol, and all samples were stored at −80 °C. Whole insect samples were surface sterilized for 5 s in 10% bleach and rinsed in PBS prior to DNA extraction using the Powersoil DNA isolation kit and protocol (MoBio Laboratories, Carlsbad, CA, USA) with the addition of 60 µg proteinase K to the lysis buffer. DNA concentration was assessed using a Qubit Fluorometer (Invitrogen, Carlsbad, CA, USA) and samples with low DNA yields were concentrated using the MoBio protocol.

Extracted DNA was sent to Argonne National Laboratories (Lemont, IL, USA) for library preparation and sequencing of the V4 region of the 16S rRNA gene. Library preparation used barcoded primers 515F (5′-GTGYCAGCMGCCGCGGTAA-3′) and 806R (5′-GGACTACNVGGGTWTCTAAT-3′) and the methods from Caporaso *et al.* (2012). Libraries were pooled, and 150-bp paired end reads were sequenced on an Illumina MiSeq sequencer.

Raw sequences were preprocessed using previously published methods (Whitaker *et al.*, 2016). Because the 16S universal primers we used are known to also amplify organellar DNA, we compared chloroplast abundance across samples before removing non-target sequences, which included chloroplasts and mitochondria as well as the common laboratory contaminants, Staphylococcaceae and *Escherichia*. We then applied a filtering method requiring bacterial operational taxonomic units (OTUs) to be represented by at least ten sequences in the data set and at a minimum relative abundance of 0.05% per sample in order to be included in the analysis.

Core microbiomes were calculated using the *filterfun* function in the *phyloseq* package in R, requiring bacterial OTUs to be present in all replicates within a species (species cores) or all cycad herbivore replicates (cycad herbivore core, hereafter 'shared OTUs'). The taxonomic assignments of all core OTUs were further checked using NCBI BLAST and Seqmatch from the Ribosomal Database project (Cole *et al.*, 2014). Shared bacterial OTUs were further analysed using the Oligotyping pipeline v2.1 (Eren *et al.*, 2013) with a minimum substantive abundance of 10 and the smallest number of entropy positions needed to properly decompose oligotypes. For comparison, we searched for these shared OTUs in published surveys of the gut microbiomes of six species of non-cycad-ivorous 'outgroup' insects: two Lycaenidae (Lepidoptera) and four Curculionidae (Coleoptera) (details in Supporting Information S1). For these comparisons, we

Table 1. Cycad herbivore sampling included two cycad genera and multiple insect species, families and orders, feeding tissues and localities; sample numbers consider only samples that passed the filtering requirements and were included in analyses

Herbivore	Order	Life stage	Number	Preservation	Host plant	Feeding tissue	Locality
Rhopalotria furfuracea (O'Brian & Tang)	Coleoptera: Belidae	Adult	6	Whole flash frozen	*Zamia furfuracea* (Aiton)	Pollen cone microsporophyll	Miami, FL, USA
Pharaxanotha floridana (Casey)	Coleoptera: Erotylidae	Late instar	6	Whole flash frozen	*Zamia integrifolia* (Aiton)	Pollen cone peduncle	Naples, FL, USA
Eubulus sp. (Kirsh)	Coleoptera: Curculionidae	Late instar	6	Whole flash frozen	*Zamia aff. portoricensis* (Urban)	Subterranean stem	Miami, FL, USA
Eumaeus atala (Poey)	Lepidoptera: Lycaenidae	Late instar	5	Whole flash frozen	*Zamia integrifolia* (Aiton)	Leaf	Miami, FL, USA
Chilades pandava (Horsfield)	Lepidoptera: Lycaenidae	Late instar	7	Gut dissected in ethanol	*Cycas revoluta* (Thunberg)	Leaf	Singapore

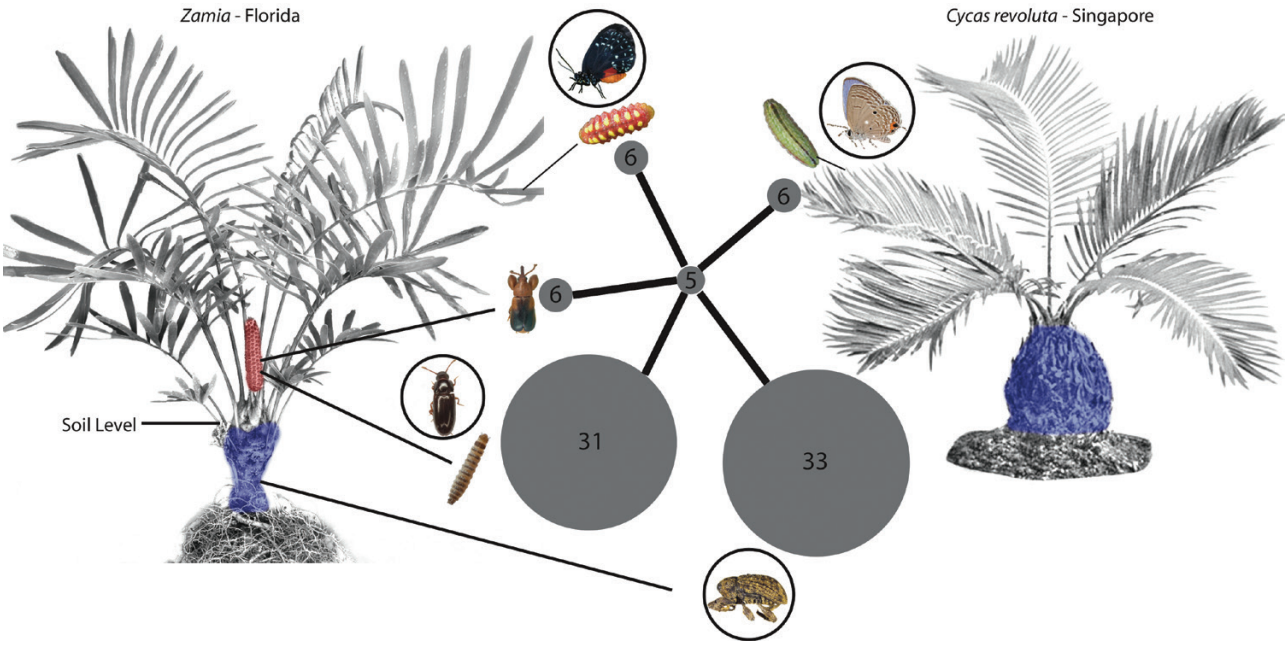

Figure 1. Cycad herbivores share five bacterial OTUs. Herbivore species are shown corresponding to the plant tissue on which they feed. Circles show core microbiome counts. Insets of adult insects are included in cases where we sampled the larval stage. Plant tissues are coloured to highlight the pollen cone (red) and stem (blue – underground for *Zamia integrifolia*). Herbivores of *Zamia* are (top to bottom): *Eumaeus atala* (note the aposematic colouring of both life stages), *Rhopalotria furfuracea*, *Pharaxanotha floridana* and *Eubulus* sp. The sole herbivore of *Cycas* is *Chilades pandava*.

processed and analysed raw 16S sequences of gut bacteria in the same way that we did with the cycadivorous insects.

For diversity analyses, libraries were first rarified to 10 000 sequences, retaining any library with at least 5000 sequences. The similarity of each sample's bacterial community composition was compared using the Bray–Curtis dissimilarity metric and hierarchical clustering based on weighted UniFrac distances where OTU abundances were averaged across replicates within a species. All sequence data are deposited in the EMBL-EBI database.

RESULTS

Raw sequences clustered into 1789 unique OTUs across the 31 samples. As expected, chloroplast 16S sequences were found in high abundance in caterpillars of both leaf-feeding Lepidoptera, *Eumaeus atala* and *Chilades pandava*, but were unexpectedly prevalent in cone-feeding weevils, *R. furfuracea*, as well (median 60% of the total sequences, 18% inter-quartile range). After abundance filtering and removing non-target OTUs, the filtered dataset lost one *Eum. atala* sample due to small library size but retained 177 taxonomic OTUs across the remaining 30 samples. This filtered dataset was used in all subsequent analysis.

Two additional samples were omitted following rarefaction for diversity analyses due to small library sizes, one *R. furfuracea* and one *Eum. atala*. The remaining samples clustered in a characteristic pattern in non-metric multidimensional scaling (NMDS) ordinations and hierarchical clustering (Fig. 2B, C). The gut microbiota of all five cycad-feeding insects were remarkably conserved, clustering mainly by species except for an overlap between *R. furfuracea* and *C. pandava* in NMDS ordinations of Bray–Curtis distances (stress 0.13; Fig. 2B). *Rhopalotria furfuracea* and *C. pandava* also grouped by similarity in hierarchical clustering (Fig. 2C). Taxonomy plots demonstrate a compositional similarity between these two species, driven by the dominance of Alicyclobacillaceae and Comamonadaceae in the gut bacterial communities of both insects (Fig. 2A).

The five most abundant bacterial families in our dataset were Alicyclobacillaceae, Enterobacteriaceae, Moraxellaceae, Comamonadaceae and Enterococcaceae (Fig. 2A). Bacterial OTUs found in all samples of a species are reported briefly in Table 2 and in detail in Supporting Information S2. Most significantly, the microbiota of all five cycad-feeding species showed significant overlap for five OTUs that were present in all replicates (Table 2). The *greengenes* taxonomic assignments, however, did not match NCBI BLAST results for any of these five shared OTUs. Only one of

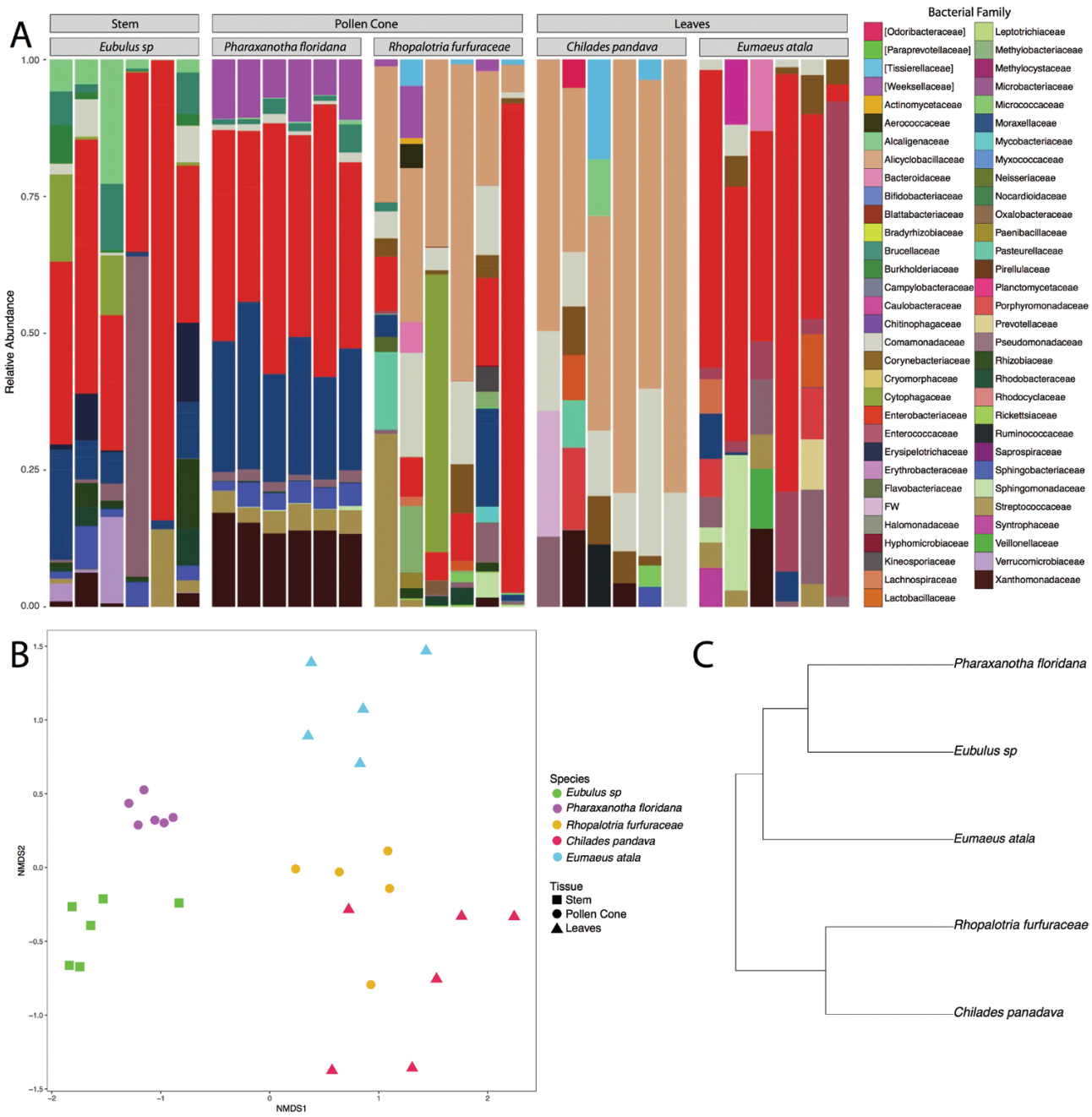

Figure 2. (A) The relative abundance of the 20 most abundant bacterial families per sample shows that Enterobacteriaceae and Alicyclobacillaceae are dominant in the guts of several species. (B) In ordinations based on Bray–Curtis distances (stress 0.13), gut bacterial communities generally cluster according to species, except for a striking similarity in the gut communities of *Rhopalotria furfuracea* and *Chilades pandava*. (C) Hierarchical clustering of insect species by bacterial community composition also highlights the similarity between *R. furfuracea* and *C. pandava*.

the shared OTUs could be identified to species using NCBI BLAST (OTU 5: *Raoutella ornithinolytica* in the Enterobacteriaceae). The four remaining OTUs could only be identified to the family level using NCBI BLAST search (OTUs 4 and 249 in the Enterobacteriaceae, OTU 3 in the Alicyclobacillaceae and OTU 85 in the Rickettsiaceae; Table 2). Unidentified bacteria are not surprising when exploring novel environmental samples, and the four OTUs that could be identified only to family belong to bacterial families that include insect-associated and extremophilic bacteria. Of these five shared OTUs, OTU 4 (Enterobacteriaceae) and OTU 5

Table 2. Five bacterial OTUs are shared across five species of herbivores feeding on cycads

	Chilades pandava	*Eumaeus atala*	*Rhopalotria furfuracea*	*Pharaxanotha floridana*	*Eubulus* sp.
Unknown Alicyclobacillaceae	3	3	3	3	3
Unknown Rickettsiaceae	85	85	85	85	85
Unknown Enterobacteriaceae	249	249	249	249	249
Enterobacteriaceae *Raoultella ornithinolytica*	5	5	5	5	5
Unknown Enterobacteriaceae	4	4	4	4	4
				596	596
Comamonadaceae *Curvibacter*	13		13	13	
Enterococcaceae *Enterococcus*		15		15	15
Comamonadaceae *Lampropedia*				36	36
Comamonadaceae *Comamonas*				432	432
Streptococcaceae *Lactococcus*				10	10
Alcaligenaceae *Achromobacter*				24	24
Unknown Blattabacteriaceae				240	240
Sphingobacteriaceae *Sphingobacterium*				17	17
Moraxellaceae *Acinetobacter*				763	187
				2	2
					91
					37
					21
					47

OTUs that were found to be part of more than one species' core microbiome are shown and highlighted in grey. Full species microbiota are listed in Supporting Information S2. OTUs were considered to be part of the species core microbiota when present in 100% of the samples for that species. They are designated here by their OTU number.

(*Raoutella ornithinolytica*) were found solely in cycad herbivores. None was found in any of the outgroup weevils. However, OTUs 3, 85 and 249 were found in both outgroup butterflies.

Oligotyping results for shared OTUs are presented in Figure 3 (data in Supporting Information S3). Twenty-one unique oligotypes were found for *Raoutella ornithinolytica* (OTU 5). *Pharaxanotha floridana* had the most consistent composition of *Raoutella ornithinolytica* oligotypes, with *Eubulus* sp. showing a similar although less constrained pattern (Fig. 3). In contrast, *Eum. atala* was dominated (> 93%) by one *Raoutella ornithinolytica* oligotype across all replicates. Twenty unique oligotypes were found for OTU 4 (Unknown Enterobactereaceae). Again, *P. floridana* and *Eub.* sp. showed consistent patterns of OTU 4 oligotype composition, whereas OTU 4 oligotype composition was highly variable in the remaining species (Fig. 3). Only three oligotypes were recovered for OTU 249 (Unknown Enterobactereaceae), one of which dominated all samples. Finally, oligotyping analysis found no entropy peaks for either OTU 85 (Unidentified) or OTU 3 (Unknown Alicyclobacillaceae).

DISCUSSION

Our results show that five cycad-feeding insect species from two orders share a core set of bacterial OTUs in their gut microbiota, despite being generally distinct in overall bacterial community composition. While the functions of these OTUs remain unknown, our results are consistent with the hypothesis that gut bacteria may mediate herbivory of cycads, and they identify specific bacterial phylotypes as candidates for future functional study.

In terms of the entire community of gut microbiota, we found that each insect harbours a distinctive species-specific bacterial assemblage, with the exception of *R. furfuracea* weevils and *C. pandava* butterflies, whose communities of microbiota were surprisingly similar. For example, chloroplasts were found in high abundance in all Lepidoptera, as well as *R. furfuracea*, but were not found in the other Coleoptera. In ordination plots based on community dissimilarity metrics, the gut bacterial communities of *R. furfuracea* weevils clustered more closely with *C. pandava* than with *P. floridana*, the other cone-feeding coleopteran (Fig. 2B). Larvae and adults of *R. furfuracea* feed only

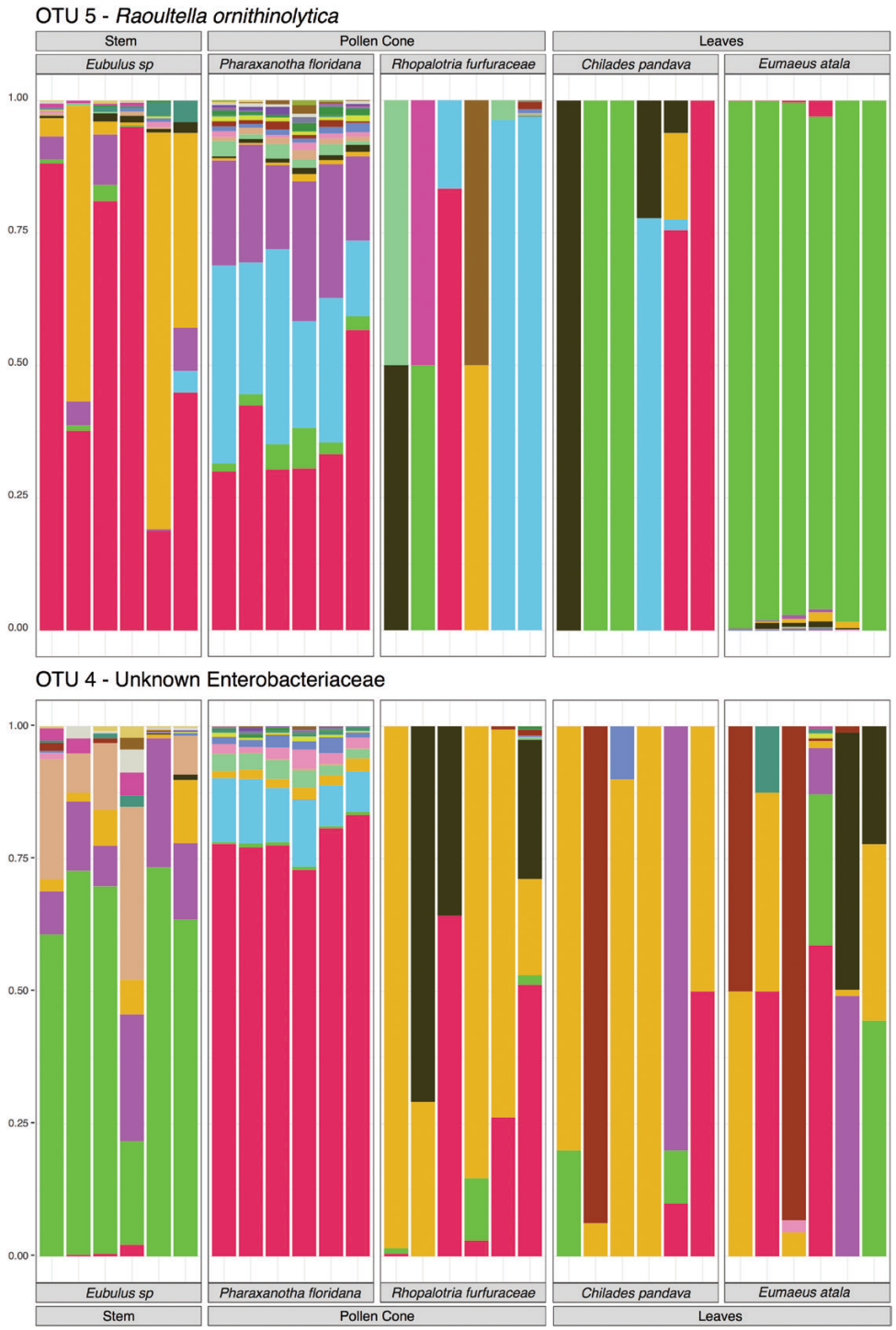

Figure 3. Oligotyping analysis of the two OTUs that were unique to cycad herbivores. *Eumaeus atala* was dominated by one oligotype for OTU 5 (*Raoultella ornithinolytica*). OTU 4 (Unknown Enterobacteriaceae) displayed highly constrained compositional patterns within *Eubulus* sp. and *Pharaxanotha floridana*. OTU 85 (Unidentified) and OTU 3 (Unknown Alicyclobacillaceae) were each represented by a single oligotype across all samples, and OTU 249 (Unknown Enterobacteriaceae) was dominated by a single oligotype across all species, so oligoptyping results for those three OTUs are not shown here.

on the microsporophyll of the pollen cone (Fawcett & Norstog, 1993), which is developmentally a modified leaf (Gifford & Foster, 1989), whereas *P. floridana* larvae feed on cone peduncle tissue (Fawcett & Norstog, 1993), which is developmentally a modified stem (Gifford & Foster, 1989). It is possible that the bacterial compositional similarities we observe between *R. furfuraceae* and *C. pandava* reflect developmentally related chemical or nutritional similarities in the plant tissues on which they feed. Further research on the chemical composition of various cycad tissues will be necessary to test this hypothesis.

Ordination plots demonstrated that the microbiota of *P. floridana*, and to some extent those of *Eub.* sp., were more similar to each other than they were to those of the remaining three species. In fact, whether looking at the OTU (Fig. 2A) or sub-OTU level (oligotype; Fig. 3), *P. floridana* samples exhibited highly conserved bacterial community compositions across replicates. It is unclear if these conserved assemblages are a product of limited environmental exposure, vertical transmission or selection by the host. *Pharaxanotha floridana* beetles feed on pollen for their first two instars, and then spend the remainder of their larval development within the peduncle of the pollen cone, a fairly closed environment (Fawcett & Norstog, 1993). Further sampling of *Pharaxanotha* adults, which feed exclusively on cycad pollen (Fawcett & Norstog, 1993), would help to elucidate whether these insects harbour consistent bacterial communities during all life stages. This would enable us to assess the relative contributions of host selection versus environmental exposure in determining gut bacterial community assemblages in these beetles.

Four of the five shared OTUs found in the guts of cycad-feeding insects were unidentifiable to genus. The one that could be identified, however, offers some insight into potential activities of the gut bacterial community. OTU 5 was identified as *Raoultella ornithinolytica*, a bacterium that has been shown to fix nitrogen in the guts of wild *Ceratitis capitata* fruit flies (Behar, Yuval & Jurkevitch, 2005) and to elicit cytotoxicity and apoptotic and necrotic death of mammalian cancer cells due to the activity of a protein–polysaccharide complex (Fiołka *et al.*, 2013). Oligotyping of *Raoultella ornithinolytica* revealed consistent strain-level compositional patterns within *Eum. atala*, *P. floridana*, and *Eub.* sp. (Fig. 3). The leaf- and cone-feeding *Eumaeus* larvae are more vagile than the cone- and stem-feeding *Pharaxanotha* and *Eubulus* larvae, such that we might expect *Eum. atala* to be exposed to a greater diversity of bacteria in the environment. It is therefore somewhat unexpected that the *Eum. atala* caterpillars were dominated by only one oligotype. *Raoultella ornithinolytica* warrants further functional research to assess whether this bacterium provides critical benefit to cycadivorous insects, such as detoxification of their poisonous cycad host plants.

OTU 4 was identified with equal confidence to two bacterial genera, *Pantoea* and *Klebsiella*, and in our analysis it remains an unidentified member of the family Enterobacteriaceae. Oligotyping revealed conserved strain-level compositional patterns in *P. floridana* and *Eub.* sp. for this OTU (Fig. 3). As these are the two least mobile insects in our dataset, it is unclear if this pattern arises from limited exposure to environmental bacteria, or from selection on the part of the host or gut environment. While it is unknown whether this bacterium might contribute to host nutrition or fitness, we can make some inferences based on its similarity to *Pantoea* and *Klebsiella*. *Pantoea* is a known gut symbiont in many insects, and it has been shown to confer nutritional benefits including nitrogen fixation, toxin degradation and hydrolysis of proteins (Sood & Nath, 2002; MacCollom *et al.*, 2009). Additionally, ingested *Pantoea* and *Klebsiella* have both been shown to colonize the guts of insects and be subsequently vertically transmitted (Lauzon *et al.*, 2009). Future work should focus on isolating and identifying this OTU and on characterizing its metabolic capabilities and symbiotic potential.

Of the remaining shared OTUs, two were initially identified by *greengenes* taxonomic assignment as *Buchnera* (OTU 249) and *Wolbachia* (OTU 85), both known insect symbionts. However, these assignments were not supported by the NCBI BLAST database, and both OTUs displayed little oligotypic variation. OTU 3 (Unidentified Alicyclobacillaceae) showed no oligotypic variation across samples, despite being represented by a large number of reads (16383). Bacteria in this family are found in extreme environments, often growing at extreme temperatures or pH (Vos *et al.*, 2009). With no information about the function of these bacteria or their symbiotic potential as well as their presence in the 'outgroup' butterfly samples, it is impossible to determine whether they have a functional relationship with their cycadivorous insects.

To our knowledge, this survey represents the first report of gut bacterial associations that are conserved among a phylogenetically and geographically disparate set of herbivores that share a common toxic host plant. By comparing the composition of gut bacterial communities of five species of cycad-feeding insects, we found that each insect harbours a distinctive species-specific bacterial assemblage, with the exception of *R. furfuraceaea* weevil larvae and *C. pandava* caterpillars that host similar communities of microbiota. Most importantly, we found that all five of the insect species share a core set of bacterial OTUs, which we believe warrant future functional study. We identified

one of these shared bacteria as *Raoultella ornithinolytica*, a species with documented anti-tumour and nitrogen fixation capabilities. Future comparative surveys of the microbiota of cycad herbivores should include a broader range of insect taxa, developmental stages and feeding ecologies, as well as an investigation into the functional profiles of core bacterial OTUs. Such studies would be invaluable in elucidating the role of microbial symbionts in mediating cycadivorous diets.

ACKNOWLEDGEMENTS

We thank Rory Maher for the use of his photographs of *Eumaeus atala*, *Pharaxanotha floridana* and *Rhopalotria furfuracea* and William Tang and Horace Tan for the use of their photographs of *Eubulus* sp. and *Chilades pandava*, respectively. We are grateful to the Montgomery Botanical Center for its helpful and knowledgeable staff including Patrick Griffith, Michael Calonje, Claudia Calonje, Stella Cuestas and Vickie Murphy and the use of their collections, as well as to Dennis Stevenson for his generosity and deep knowledge of cycad biology. The manuscript was greatly improved by the comments of three anonymous reviewers. The authors are aware of no conflicts of interest to report. SS was supported by a National Science Foundation Graduate Research Fellowship and MW was supported by an NSF Postdoctoral Research Fellowship in Biology (1309425) and grants from the Wildlife Reserves Singapore Conservation Fund and the British Ecological Society. SS and NP conceived the project. SS and MW conducted the fieldwork, laboratory work and analysed the data. All authors contributed to the preparation of the manuscript.

REFERENCES

Banack SA, Cox PA. 2003. Distribution of the neurotoxic non-protein amino acid BMAA in *Cycas micronesica*. *Botanical Journal of the Linnean Society* **143:** 165–168.

Behar A, Yuval B, Jurkevitch E. 2005. Enterobacteria-mediated nitrogen fixation in natural populations of the fruit fly *Ceratitis capitata*. *Molecular Ecology* **14:** 263702643.

Blagrove RJ, Lilley GG, Higgins TJV. 1984. Macrozin, a novel storage globulin from seeds of *Macrozamia communis* L. Johnson. *Functional Plant Biology* **11:** 69–77.

Boone CK, Keefover-Ring K, Mapes AC, Adams AS, Bohlmann J, Raffa KF. 2013. Bacteria associated with a tree-killing insect reduce concentrations of plant defense compounds. *Journal of Chemical Ecology* **39:** 1003–1006.

Bowers MD, Larin Z. 1989. Aquired chemical defense in the lycaenid butterfly, *Eumaeus atala*. *Journal of Chemical Ecology* **15:** 1133–1146.

Brenner ED, Martinez-Barboza N, Clark AP, Liang QS, Stevenson DW, Coruzzi GM. 2000. Arabidopsis mutants resistant to S(+)-β-methyl-α-β-diaminopropionic acid, a cycad-derived glutamate receptor agonist. *Plant Physiology* **124:** 1615–1624.

Brookes DR, Hereward JP, Terry LI, Walter GH. 2015. Evolutionary dynamics of a cycad obligate pollination mutualism–Pattern and process in extant *Macrozamia cycads* and their specialist thrips pollinators. *Molecular Phylogenetics and Evolution* **93:** 83–93.

Calonje M, Stevenson DW, Stanberg L. 2017. *The world list of cycads, online edition*. Available at: http://www.cycadlist.org.

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

Castillo-Guevara C, Rico-Gray V. 2003. The role of macrozamin and cycasin in cycads (Cycadales) as antiherbivore defenses. *Journal of the Torrey Botanical Society* **130:** 206–217.

Ceja-Navarro JA, Vega FE, Karaoz U, Hao Z, Jenkins S, Lim HC, Kosina P, Infante F, Northen TR, Brodie EL. 2015. Gut microbiota mediate caffeine detoxification in the primary insect pest of coffee. *Nature Communications* **6:** 7618.

Chiu J, DeSalle R, Lam HM, Meisel L, Coruzzi G. 1999. Molecular evolution of glutamate receptors: a primitive signaling mechanism that existed before plants and animals diverged. *Molecular Biology and Evolution* **16:** 826–838.

Chu CC, Spencer JL, Curzi MJ, Zavala JA, Seufferheld MJ. 2013. Gut bacteria facilitate adaptation to crop rotation in the western corn rootworm. *Proceedings of the National Academy of Sciences of the United States of America* **110:** 11917–11922.

Cole JR, Wang Q, Fish JA, Chai B, McGarrell DM, Sun Y, Brown CT, Porras-Alfaro A, Kuske CR, Tiedje JM. 2014. Ribosomal Database Project: data and tools for high throughput rRNA analysis. *Nucleic Acids Research* **42:** D633–D642.

Conchie J, MacDonald DC. 1959. Glycosidases in the mammalian alimentary tract. *Nature* **184:** 1233–1235.

Cooper JM. 1941. Isolation of a toxic principle from the seeds of *Macrozamia spiralis*. *Proceedings of the Royal Society of New South Wales* **74:** 450–454.

Cox PA, Banack SA, Murch SJ. 2003. Biomagnification of cyanobacterial neurotoxins and neurodegenerative disease among the Chamorro people of Guam. *Proceedings of the National Academy of Sciences of the United States of America* **100:** 13380–13383.

Donaldson JS, Nänni I, Bösenberg DW. 1995. The role of insects in the pollination of *Encephalartos altensteinii*. In: Vorster P, ed. *Proceedings of the Third International Conference on Cycad Biology*. Stellenbosch: Cycad Society of South Africa, 423–434.

Douglas AE. 2013. Microbial brokers of insect-plant interactions revisited. *Journal of Chemical Ecology* **39:** 952–961.

De Luca P, Moretti A, Sabato S, Gigliano GS. 1980. The ubiquity of cycasin in cycads. *Phytochemistry* **19:** 2230–2231.

De Luca P, Moretti A, Sabato S, Gigliano GS. 1982. A comparative study of cycad mucilages. *Phytochemistry* **21:** 1609–1611.

Dossaji SF, Bell EA. 1973. Distribution of α-amino-β-methylaminopropionic acid in *Cycas*. *Phytochemistry* **12:** 143–144.

Duncan MW, Kopin IJ, Crowley JS. 1989. Quantification of the putative neurotoxin 2-amino-3-(methylamino) propanoic acid (BMAA) in cycadales: analysis of the seeds of some members of the family Cycadaceae. *Journal of Analytical Toxicology* **13:** 169–175

Dunlap RA, Cox PA, Banack SA, Rodgers KJ. 2013. The non-protein amino acid BMAA is misincorporated into human proteins in place of L-serine causing protein misfolding and aggregation. *Plos One* **8:** e75376.

Eren AM, Maignien L, Sul WJ, Murphy LG, Grim SL, Morrison HG, Sogin ML. 2013. Oligotyping: differentiating between closely related microbial taxa using 16S rRNA gene data. *Methods in Ecology and Evolution* **4:** 1111–1119.

Fawcett P, Norstog K. 1993. *Zamia pumila* in South Florida: a preliminary report on its pollinators *R. Slossoni*, and snout weevil and *P. zamiae*, a clavicorn beetle. In: Stevenson DW, Norstog KJ, eds. *The biology, structure and systematics of the cycadales*. Milton: Palm & Cycad Societies of Australia Ltd, 109–119.

Fiołka MJ, Lewtak K, Rzymowska J, Grzywnowicz K, Hulas-Stasiak M, Sofinska-Chmiel W, Skrzypiec K. 2013. Antifungal and anticancer effects of a polysaccharide-protein complex from the gut bacterium *Raoultella ornithinolytica* isolated from the earthworm *Dendrobaena veneta*. *Pathogens and Disease* **69:** 49–61.

Friis EM, Chaloner WG, Crane PR, eds. 1987. *The origins of angiosperms and their biological consequences*. Cambridge: Cambridge University Press.

Gao Z, Thomas BA. 1989. A review of fossil cycad mega-sporophylls, with new evidence of *Crossozamia pomel* and its associated leaves from the lower Permian of Taiyuan, China. *Review of Palaeobotany and Palynology* **60:** 205–223.

Gifford EM, Foster AS. 1989. Pteridospermophyta (seed ferns), Cycadeoidophyta (Cycadeoids), and Cycadophyta (Cycads).In: Gifford EM, Foster AS, eds. *Morphology and evolution of vascular plants*, *3rd edn*. San Francisco: W.H. Freeman, 345–384.

Goto JJ, Koenig JH, Ikeda K. 2012. The physiological effect of ingested β-N-methylamino-L-alanine on a glutamatergic synapse in an in vivo preparation. *Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology* **156:** 171–177.

IUCN. 2017. *IUCN summary statistics*. Available at: http://www.iucnredlist.org/about/summary-statistics.

Khabazian I, Bains JS, Williams DE, Cheung J, Wilson JM, Pasqualotto BA, Pelech SL, Anderson RJ, Wang YT, Liu L, Nagai A, Kim SU, Craig UK, Shaw CA. 2002. Isolation of various forms of sterol beta-D-glucoside from the seed of *Cycad circinalis*: neurotoxicity and implications for ALS-parkinsonism dementia complex. *Journal of Neurochemistry* **82:** 516–528

Kisby GE, Moore H, Spencer PS. 2013. Animal models of brain maldevelopment induced by cycad plant genotoxins. *Birth Defect Research C* **99:** 247–255.

Lacombe B, Becker D, Hedrich R, DeSalle R. 2001. The identity of plant glutamate receptors. *Science* **292:** 1486–1487.

Lam HM, Chiu J, Hsieh MH, Meisel L, Oliveira IC. 1998. Glutamate-receptor genes in plants. *Nature* **396:** 125–126.

Laqueur GL, Spatz M. 1968. Toxicology of cycasin. *Cancer Research* **28:** 2262–2267.

Lauzon CR, McCombs SD, Potter SE, Peabody NC. 2009. Establishment and vertical passage of *Enterobacter* (*Pantoea*) *agglomerans* and *Klebsiella pneumoniae* through all life stages of the Mediterranean fruit fly (Diptera: Tephritidae). *Annals of the Entomological Society of America* **102:** 85–95.

Lindblad P, Tadera K, Yagi F. 1990. Occurrence of azoxyglycosides in *Macrozamia riedlei* and their effects on the free-living isolated *Nostoc* PCC 73102. *Environmental and Experimental Botany* **30:** 429–434.

Lythgoe B, Riggs NV. 1949. Macrozamin. Part I. The identity of the carbohydrate component. *Journal of the Chemical Society* **4:** 2716–2718.

MacCollom GB, Lauzon CR, Sjogren RE, Meyer WL, Olday F. 2009. Association and attraction of blueberry maggot fly curran (Diptera: Tephritidae) to *Pantoea* (*Enterobacter*) *agglomerans*. *Environmental Entomology* **38:** 116–120.

Mamay SH. 1969. Cycads: fossil evidence of late Paleozoic origin. *Science* **164:** 295–296

Moretti A, Sabato S, Gigliano GS. 1981. Distribution of macrozamin in Australasian cycads. *Phytochemistry* **20:** 1415–1416.

Moretti A, Sabato S, Gigliano GS. 1983. Taxonomic significance of methylazoxymethanol glycosides in the cycads. *Phytochemistry* **22:** 115–117.

Morgan RW, Hoffmann GR. 1983. Cycasin and its mutagenic metabolites. *Mutation Research* **114:** 19–58.

Nair JJ, van Staden J. 2012. Isolation and quantification of the toxic methylazoxymethanol glycoside macrozamin in selected South African cycad species. *South African Journal of Botany* **82:** 108–112.

Nash RJ, Bell EA, Ackery PR. 1992. The protective role of cycasin in cycad-feeding lepidoptera. *Phytochemistry* **31:** 1955–1957.

Nishida K, Kobayashi A, Nagahama T. 1955. Studies on cycasin, a new toxic glycoside of *Cycas revolute* Thunb. *Agricultural Chemical Society Bulletin* **19:** 77–83.

Norstog KJ, Stevenson DW, Niklas KJ. 1989. The role of beetles in the pollination of *Zamia furfuracea* L. fil (Zamiaceae). *Biotropica* **18:** 300–306.

Norstog KJ, Fawcett P. 1989. Insect–cycad symbiosis and its relation to the pollination of *Zamia furfuracea* (Zamiaceae) by *Rhopalotria mollis* (Curculionidae). *American Journal of Botany* **76:** 1380–1394.

Pan M, Mabry TJ, Beale JM, Mamiya BM. 1997a. Nonprotein amino acids from *Cycas revolute*. *Phytochemistry* **45:** 517–519.

Pan M, Mabry TJ, Cao P, Moini M. 1997b. Identification of nonprotein amino acids from cycad seeds as N-ethoxycarbonyl ethyl ester derivatives by positive chemical-ionization gas chromatography-mass spectrometry. *Journal of Chromotography A* **787:** 288–294.

Prado A. 2011. The cycad herbivores. *Bulletin de la Société d'entomologie du Québec* **18:** 3–6.

Prado A, Sierra A, Windsor D, Bede JC. 2014. Leaf traits and herbivory levels in a tropical gymnosperm, *Zamia stevensonii* (Zamiaceae). *American Journal of Botany* **101:** 437–447.

Prado A, Rubio-Mendez G, Yañez-Espinosa L, Bede JC. 2016. Ontogenetic changes in azoxyglycoside levels in the leaves of *Dioon edule* Lindl. *Journal of Chemical Ecology* **42:** 1142–1150.

Riggs NV. 1954. The occurrence of macrozamin in the seeds of cycads. *Australian Journal of Chemistry* **7:** 123–124.

Rothschild M, Nash RJ, Bell EA. 1986. Cycasin in the endangered butterfly *Eumaeus atala* florida. *Phytochemistry* **25:** 1853–1854.

Schneider D, Wink M, Sporer F, Lounibos P. 2002. Cycads: their evolution, toxins, herbivores and insect pollinators. *Naturwissenschaften* **89:** 281–294.

Sihvonen P, Staude HS, Mutanen M. 2015. Systematic position of the enigmatic African cycad moths: an integrative approach to a nearly century old problem (Lepidoptera: Geometridae, Diptychini). *Systematic Entomology* **40:** 606–627.

Sood P, Nath A. 2002. Bacteria associated with *Bactrocera* sp. (Diptera: Tephritidae)–isolation and identification. *Pest Management and Economic Zoology* **10:** 1–9.

Spencer PS, Nunn PB, Hugon J, Ludolph AC, Ross SM, Roy DN, Robertson RC. 1987. Guam amyotrophic lateral sclerosis-parkinsonism-dementia linked to a plant excitant neurotoxin. *Science*. **237:** 517–522.

Stevenson DW, Norstog KJ, Fawcett PKS. 1998. Pollination biology of cycads. In: Owens SJ, Rudall PJ, eds. *Reproductive biology*. Kew: Royal Botanic Gardens, 277–294.

Suinyuy TN, Donaldson JS, Johnson SD. 2015. Geographical matching of volatile signals and pollinator olfactory responses in a cycad brood-site mutualism. *Proceedings of the Royal Society B* **282:** 2015–2053.

Teas HJ. 1967. Cyasin synthesis in *Seirarctia echo* (lepidoptera) larvae fed methylazoxymethanol. *Biochemical and Biophysical Research Communications* **26:** 686–690.

Terra WR, Ferreira C. 1994. Insect digestive enzymes: properties, compartmentalization and function. *Comparative Biochemistry and Physiology Part B: Comparative Biochemistry* **109:** 1–62.

Terry I, Tang W, Blake AST, Donaldson JS, Singh R, Vovides AP, Jaramillo AC. 2012. An overview of cycad pollination studies. *Memoirs of the New York Botanical Garden* **106:** 352–394.

Thomas BA, Spicer RA. 1987. *The evolution and palaeobiology of land plants*. London: Croom Helm.

Valencia-Montoya WA, Tuberquia D, Guzman PA, Cardona-Duque J. 2017. Pollination of the cycad *Zamia incognita* A. Lindstr. & Idárraga by *Pharaxonotha* beetles in the Magdalena Medio Valley, Colombia: a mutualism dependent on a specific pollinator and its significance for conservation. *Arthropod-Plant Interactions* **2017:** 1–17.

Vega A, Bell EA. 1967. α-Amino-β-methylaminopropionic acid, a new amino acid from seeds of *Cycas circinalis*. *Phytochemistry* **6:** 759–762.

Vos P, Garrity G, Jones D, Krieg NR, Ludwig W, Rainey FA, Schleifer KH, Whitman W, eds. 2009. *Bergey's manual of systematic bacteriology Volume 3: The Firmicutes*. New York: Springer, 229–243.

Vovides AP, Norstog KJ, Fawcett P, Duncan MW, Nash RJ, Molsen DV. 1993. Histological changes during maturation in male and female cones of the cycad *Zamia furfuracea* and their significance in relation to pollination biology. *Botanical Journal of the Linnean Society* **111:** 241–252.

Whitaker MR, Salzman S, Sanders J, Kaltenpoth M, Pierce NE. 2016. Microbial communities of lycaenid butterflies do not correlate with larval diet. *Frontiers in Microbiology* **7:** 1920.

Whiting MG. 1963. Toxicity of cycads. *Economic Botany* **17:** 270–302.

Yagi F, Tadera K. 1987. Azoxyglycoside contents in seeds of several cycad species and various parts of Japanese cycad. *Agricultural and Biological Chemistry* **51:** 1719–1721.

Yagi F. 2004. Azoxyglycoside content and β-glycosidase activities in leaves of various cycads. *Phytochemistry* **65:** 3243–3247.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

S1. Non-cycadivorous lycaenid and curculionid samples that passed filtering requirements and were analysed for the presence of shared bacterial OTUs found in cycad herbivores.
S2. Species core microbiome OTU and assigned taxonomy are listed with OTUs present in more than one species highlighted in grey. Species' microbiome cores were determined by OTU presence in 100% of samples of that species after filtering and are represented by their OTU number.
S3. Oligotyping results