New insights on cycad biology and evolution

Shayla Salzman¹, Melissa Whitaker¹, Adriel Sierra¹, Daniel E. Bustos Díaz¹, Angélica Cibrián-Jaramillo¹, Yang Liu¹, Francisco Barona Gómez¹, Shouzhou Zhang¹, Sunil Sahu¹, and Juan Carlos Villarreal A.¹

¹Affiliation not available

March 07, 2024

AUTHORS:
Salzman, Shayla¹, M.R.L. Whitaker², A.M. Sierra³, E.D Bustos-Díaz⁴, A. Cibrián-Jaramillo⁵, Y. Liu⁶, F. Barona-Gómez⁴, S.Z. Zhang⁶, S.K. Sahu⁷,⁹, J.C. Villarreal A.³,⁸,⁹

¹ University of Georgia, Entomology, Athens, GA, US
² East Tennessee State University, Biological Sciences, Johnson City, TN, US
³ Département de Biologie, Université Laval, G1V 0A6, Québec, Canada
⁴ Institute of Biology, University of Leiden, 2333 BE, Leiden, NL.
⁵ Naturalis Biodiversity Center, Darwinweg 2 2333 CR Leiden, The Netherlands
⁶ Shenzhen & Chinese Academy of Sciences, Fairy Lake Botanical Garden, Shenzhen, CN
⁷ State Key Laboratory of Agricultural Genomics, Key Laboratory of Genomics, Ministry of Agriculture, BGI Research, Shenzhen 518083, China
⁸ Smithsonian Tropical Research Institute, Ancón, Panamá
⁹ Corresponding authors

Plain Language Summary
Cycads are palm-like gymnosperms with a suite of distinctive characteristics such as specialized relationships with insects, unique toxicity, mobile sperm and specialized roots with nitrogen-fixing microbes. We summarize the last decade of cycad research highlighting insect mutualisms, functional microbial diversity in roots and leaves, genomic evidence of horizontal transfer between microbes and cycads and the evolutionary shift to non-motile sperm.

Summary
Cycads are an ancient lineage of gymnosperms that hold a special place in seed plant evolution as sister, along with Ginkgo, to all other gymnosperms. Cycads have myriad morphological, structural, physiological, chemical and behavioral adaptations that position them as a unique system to study genome evolution, symbioses, conservation and microevolutionary processes. To this end, we have provided an overview of cycad biology and recent advances in phylogenetics, symbiosis, metabolomics and genomics to highlight their potential to address key questions about the evolution of land plants.

Summary
Introduction

The great cycad biologist, Knut Norstog once suggested that we should use “the analogy of the Rosetta Stone for the fund of information stored within the living cycads and its importance to the interpretation of plant biology….the very ancient structures and developmental pathways of cycads enables us to make connections between the early origins of seed plants and their present-day counterparts” (Donaldson, 2003). Indeed, cycads maintain a plethora of traits that have influenced our understanding of land plant evolution, the origin of insect pollination, symbiosis biology and coevolution (Fig. 1). Some examples of these traits include: flagellated motile sperm, an ancestral land plant trait; thermogenesis in reproductive tissues; symbiotic brood-site pollination; specialized insect associates with diverse defensive ecologies; and morphologically distinct coralloid roots for housing nitrogen fixing microbiota (Fig. 1). For centuries, cycad research has touched every corner of plant biology including ecology, evolution, physiology, ethnobotany, phylogenetics, systematics, development, genomics, signaling, and paleobotany. Due to recent methodological and technical advances, cycad research is now, more than ever, uniquely positioned to address basic questions across the biological sciences. The rich paleobotanical and ethnobotanical cycad literature is beyond the scope of this review, interested readers should see Carrasco et al. (2022) or Coiro et al. (2023) for recent examples. Here we present the most recent advances in cycad research on land plant evolution, the evolution and mechanisms of insect pollination and herbivory, the biochemical basis of symbiosis, microbial symbionts, and plant genomics.
Cycad biology

Cycads are one of the largest extant gymnosperm lineages, comprising ten genera and ca. 375 species, two-thirds of which are included on the International Union for Conservation of Nature (IUCN) Red List of threatened species (Calonje et al., 2023). Cycads are placed at the base of the gymnosperm phylogeny and are amongst the most ancient extant seed plants, with likely origins in the Carboniferous. The biogeography of fossil cycad taxa has shown that cycads were globally widespread in the Jurassic, Cretaceous and early Palaeogene, with taxa found in the Arctic and Europe where no extant species occur (Coiro et al., 2023; Fig. 2). They appear to have undergone a period of extinction with recent radiation in the last ~20 Mya (Nagalingum et al., 2011; Condamine et al., 2015; Liu et al., 2022). Cycads are currently globally distributed in the tropics and subtropics, where they hold considerable cultural significance with a rich history of anthropological and ethnobotanical research (eg. Carrasco et al., 2022).

Cycads exhibit a striking amount of mutualism in their associations with insects. Entire lineages of both herbivores and pollinators are specialized on cycads, in what has been suggested to be classic examples of coevolution (Tang et al., 2020; Whitaker & Salzman, 2020). All of these insects must contend with a host of cycad secondary metabolites, some of which are rather rare in the known biological world, and many of the specialist insects are aposematically colored. (Whitaker & Salzman 2020). These dioecious gymnosperms appear to rely almost exclusively on insect vectors for pollination which they maneuver between cone sexes through a series of physiological events that includes cone thermogenesis (Terry et al., 2016). The brood-site pollination mutualists live their entire life cycles within the reproductive structures of their host cycad, feeding, breeding and laying eggs within the tissue (Terry et al. 2012).

Cycads also engage in symbiosis with microbiota. The plants produce morphologically distinct coralloid roots (Fig. 1h) that house fungi, nitrogen-fixing cyanobacteria and associated bacteria, the functions of which are only recently being elucidated. More recent research suggests that leaf microbial associates may also contribute to plant growth success and habitat diversity, including nitrogen fixation in the world’s only epiphytic gymnosperm species (Zamia pseudoparasitica : Bell-Doyon et al., 2020).

Position in land plant phylogeny

Cycads hold a pivotal position in understanding the origin and evolution of seed plants (Liu et al., 2022; Coiro et al., 2023), yet their relationship to other gymnosperms remains unresolved. Extant gymnosperms consist of five major lineages: cycads, Ginkgo , Pinaceae, cupressophytes, and gnetophytes. Although the monophyly of each group is well resolved, the relationships among the five lineages have long been controversial (Xi, Rest & Davis, 2013). As it currently stands, nuclear and plastid genomic data are congruent on a cycad and Ginkgo clade as sister to the rest of gymnosperms (Stull et al., 2021; Liu et al., 2022). However mitochondrial genomic data support cycads alone being sister to all other gymnosperms (Liu et al., 2022). Recent advances in cycad genomics (see section VII) places a clearer understanding of seed plant evolutionary history within our grasp.

Generic relationships within the order Cycadales are less controversial with universal consensus on branching order (Fig. 2, Nagalingum et al., 2011; Condamine et al., 2015; Liu et al., 2022). The two extant families,
Cycadaceae and Zamiaceae form sister clades. Most cycad diversity is found in *Cycas* (119 species) and *Zamia* (86 species) with rather disjunct non-overlapping distributions (Fig. 2). Divergence time estimates, using a total evidence approach including 60 fossil cycads, date the origin of Cycadales back to the Carboniferous (median 330 Mya, 296-358 Mya) (Coiro et al., 2023), a much older date than other landmark studies focused solely on extant taxa (Nagalingum et al., 2011; Condamine et al., 2015; Fig. 2).

While the lineage itself is ancient, most extant genera seem to have diversified in the last ~27 million years in a nearly synchronous fashion (Fig. 2; Calonje et al., 2019; Nagalingum et al., 2011). Long stem branches may represent either high extinction or low divergence during the origin of a clade; here the large number of cycad fossils seems to suggest high extinction (Nagalingum et al., 2011). Phylogenetic analysis using organellar loci has not been efficient in resolving such recent radiations at the species level (Calonje et al., 2019), initiating a shift to large molecular data sets generated from transcriptome data that provide a clearer understanding of divergence and diversification (Habib et al., 2022; 2023). For example, the most morphologically and ecologically diverse genus, *Zamia*, has been found to have diversified since the Miocene (~16-24 mya, Calonje et al., 2019) with several clades showing a rather surprising biogeographical pattern in the American continent (Calonje et al., 2019; Lindstrom et al., in review). In addition, diversifications have not been completely tracked by all loci with several cases of incomplete lineage sorting and phylogenomic incongruence in several key nodes in *Macrozamia*, *Ceratozamia* and *Zamia* (Habib et al., 2022; 2023; Lindstrom et al., in review). As in many plant studies, ancient and recent reticulation are widespread in all genera, ideal for further intraspecific studies using the wealth of genomic data available as recently done in species of the genus *Dioon* (Gutiérrez-Ortega et al., 2020), *Ceratozamia* (Gutiérrez-Ortega et al., 2023); *Cycas* (He et al., 2023) and *Zamia* using microsatellites (Salas-Leiva et al., 2017 and references therein).

**Pollination and herbivory**

Like other land plants, cycads' interactions with animals (mostly insects) include parasitic and mutualistic relationships that are largely mediated by the plants' chemical and morphological characteristics. But cycads' deep evolutionary history, thermogenesis, unique chemistries, and defensive traits make them somewhat unusual in the context of plant-insect interactions. These traits position cycads as especially promising for...
understanding a suite of ‘big questions’ in plant-insect interactions, relating to the evolutionary origins of insect pollination and phytophagy, the origins and maintenance of phytochemical diversity, insect counter-adaptations to plant defenses, and plant signaling and communication.

Ancient pollination and multi-modal partner-encounter signals

Cycads have maintained obligate insect pollination since before the rise of flowering plants. While cycads were historically thought to be exclusively wind-pollinated, brood-site pollination mutualists have now been found in most cycad species, mostly involving Coleoptera (beetles and weevils), but sometimes Thysanoptera (thrips) or Lepidoptera (moths) (Tang, 2004; Terry et al., 2012; Salzman et al., 2020). These brood-site pollinators feed, breed and develop on plant reproductive tissue - almost exclusively on pollen cones although one obligate ovule parasite may play a small role in pollination of at least one species (Donaldson, 1997). Pollinators must be induced to leave the pollen cone and visit an ovulate cone of a nearby conspecific to complete pollination. In both thrips and weevils, this occurs through a ‘push-pull’ pollination process whereby pollinators are attracted to lower quantities of cone VOCs and repelled by high VOC quantities (Terry et al., 2007; Salzman et al., 2020; Salzman et al., 2021; Fig. 1a). Pollen and ovulate cones undergo a daily process of increased respiration and thermogenesis that is followed closely by an increase in cone VOCs (Terry et al., 2016; reviewed in Salzman et al., 2020). Ovulate cones mimic pollen cone scent and emission patterns but have much lower emission rates, making them attractive at peak VOC emission times when pollen cones become repellent (Fig. 1a). This has been hypothesized to be a form of “pollination by mistake” whereby pollinators are tricked into visiting the ovulate cone where they would not otherwise aggregate in large numbers (Tang 2004; Salzman et al., 2020). This push-pull pollination appears to be ancestral in the lineage (Salzman et al., 2020) and fossil evidence dates Coleopteran-cycad brood site pollination to at least 185 Mya (Cai et al., 2018), placing this pollination mechanism before the rise of flowering plants and as the oldest pollination mechanism yet documented (Salzman et al., 2020). The likely antiquity of this pollination strategy makes cycad pollination a rich case-study on the mechanisms and maintenance of plant-insect mutualisms.

Cycads also provide an excellent system to investigate mechanisms of partner-encounter signaling with implications for understanding early insect pollination prior to the overt visual signaling of flowers. Thermogenesis produces many potential cues (temperature, CO2, humidity) that insects are innately capable of perceiving and likely preceded the evolution of visual or chemical signals. Indeed, cycad cone humidity has been shown to affect pollinator behavior (Terry et al., 2014; Salzman et al., 2023) and has recently been suggested to perform a signal-like function in pollination in general (Dahake et al., 2022; Salzman et al., 2023). As a byproduct of respiration, humidity cannot be under selective pressure, but morphological traits can and cycad cone morphology that amplifies or disseminates a humidity gradient matches pollinator humidity preferences in at least the two species that have been tested (Terry et al., 2014; Salzman et al., 2023). It is becoming increasingly clear that studies of plant signaling should be extended beyond visual and chemical signals to include humidity (Arx et al., 2012; Terry et al., 2014; Dahake et al., 2022; Salzman et al., 2023), temperature (Seymour & Matthews 2006; Terry et al., 2014) and CO2 (Goyret et al., 2008). Given the evolutionary history, pollination mutualisms, and variety of plant signals, research on cycad pollination mutualisms has the potential to provide new insight into the evolution of insect pollination through the lens of partner-encounter and multi-modal signaling.

Parasitic insect herbivores

Cycads’ unique chemistries are also important for understanding cycads’ associations with parasitic insects. Cycads produce potent lineage-specific phytotoxins such as methylazoxymethanol (MAM) glycosides, β-methylamino-L-alanine (BMAA), and lesser studied compounds such as steryl glycosides. Yet cycad-feeding has been documented in at least six insect orders, with most cycad herbivores belonging to the orders Lepidoptera (moths and butterflies), Coleoptera (beetles and weevils), and Hemiptera (true bugs). A previous review concluded that cycad-feeding has evolved independently among Lepidoptera multiple times, and that insect defensive traits may be especially important in determining whether cycad-feeding lineages diversify (Whitaker & Salzman, 2020). A systematic review of cycad-feeding Coleoptera is currently underway to
investigate the ecology and evolution of their incredible diversity of feeding ecologies, species interactions, and defensive traits.

Many cycad-associated insects are obligate cycad specialists and must possess adaptations to contend with cycads’ chemical defenses. In general, herbivorous insects can excrete, sequester, or detoxify phytotoxins, but it is not clear which strategies most cycad herbivores utilize. There is some evidence that pollinating weevils excrete BMAA in their frass and pupal casings, and some Lepidoptera have been shown to sequester MAGIC for protection from natural enemies (reviewed in Whitaker & Salzman, 2020). Some cycad-feeding Aulacoscelinae beetles reflexively bleed MAM—presumably sequestered from their foodplants—when disturbed or threatened (Prado et al., 2011), and at least one lepidopteran has been shown to be capable of detoxifying MAM (Teas, 1967). Recently, the obligate cycad herbivore Eumaeus atala butterfly was found to accumulate BMAA into their larval and adult tissues (Whitaker et al., 2023), although the defensive value of sequestering BMAA remains questionable given its latent toxicity (Whitaker et al., 2022). Genomic evidence does suggest, however, that toxin tolerance is a key adaptation in the radiation of Eumaeus butterflies, a wholly cycadivorous neotropical genus of six species (Robbins et al., 2021). It has also been suggested that insects’ gut microorganisms may contribute to degrading cycad toxins (Salzman et al., 2018), though this remains to be experimentally demonstrated.

Specialized cycad herbivores must also possess adaptations to locate and select feeding and oviposition sites, but very little is known about the chemical, thermal, and visual cues used in host selection. A more mechanistic view of host selection would help assess the potential for host switching and pest management as increasing cultivation of non-native cycads introduces opportunities for emerging pest dynamics (Rensburg et al., 2020). Invasive pests such as the cycad scale represent major threats to cycad conservation (Marler, Lindström & Watson, 2021), and even native herbivores may pose a threat in some circumstances: recent host expansions have been documented among cycad-feeding lepidoptera where native and non-native cycads are co-cultivated, with potentially dire effects for cycad conservation (Marler, Lindström & Terry, 2012; Normark et al., 2017; Whitaker et al., 2020). Cycad-herbivore interactions span from mutualistic to parasitic and provide the opportunity for investigating mechanisms of host selection, specialization, and toxin tolerance, a better understanding of which will improve species management and advance research on coevolution and plant-insect interactions.

**Novel insights into the evolution of coralloid roots symbiosis**

One of the first observations of cycad root symbionts is from the 19th century, in which various taxa were described by Schneider (1894), including fungi, bacteria and algae. Subsequent research however, focused mostly on Nostocales cyanobacteria as the key taxon in coralloid roots (Suárez-Moo et al., 2019; Bell-Doyon et al., 2020). These cyanobionts colonize a morphologically distinct cycad organ called the coralloid root (Fig. 1h). Cyanobacterial Nostocales in the coralloid root fix nitrogen in exchange for plant carbohydrates, but only a few species of the genera Nostoc, Desmonostoc, Calothrix and Aulosira have been found inside the coralloid root forming three distinct phylogenetic clades (Bustos-Díaz et al., in review; Fig. 3). While no “symbiotic genes” have been identified in cyanobiont genomes, experiments using cyanobionts isolated from cycads to infect other plants found multiple genes crucial for the establishment of the symbiosis (Alvarez et al., 2022; Wong & Meeks, 2002), suggesting that while specific traits are required for the symbiosis, gene specificity, if any, is not apparent with current data (Bustos-Díaz et al., in review).

The mechanism of initiating symbiosis and the necessary morphological changes in both plant and symbiont is not well described, although recent evidence shows that multiple molecular mechanisms and signal recognition systems in both the cyanobiont and the host underlie the symbiosis, including the hormogonium-inducing factor produced by precoralloid roots (Hashidoko et al., 2019; Figs. 1g & 3a). Based on Cycas panzhihuaensis coralloid roots transcriptomes (Lin et al., 2022), additionally, terpenoids have been implicated with establishment of the symbiosis. The bacterial community has also been shown to produce various metabolites which might influence the symbiosis (Fig. 3a; Dehm et al., 2019; Freitas et al., 2022). Among these metabolites, BMAA has been linked to impaired cyanobacterial nitrogen fixation (Berntzon et al., 2013) and may be produced by both the cyanobiont and host (Marler, Snyder & Shaw, 2010). The main challenge here is
analytically distinguishing amongst the vast number of BMAA isomers and deciphering their interactions with other toxic cycad metabolites.

We are beginning to explore the role of gene expression in symbiosis. The recently sequenced cycad genome (see below) recovered a two-fold upregulation of 10 genes from the common symbiotic pathway of legume-rhizobial and plant-mycorrhizal symbiosis (Delaux & Schornack, 2021) (RAD1, DHY, SymRK, EPP1, VAPYRIN, CASTOR/POLLUX, NFP, CYTB561, GRAS, HEP) in precoralloid roots compared to mature coralloid roots colonized by cyanobacteria. In contrast, five genes were highly expressed in colonized coralloid roots (CCaMK, CYCLOPS, LIN, SYN and TAU), while a large number of genes are preferentially up and down regulated in both root types (Fig. 3b). In addition, 24 conserved genes were found shared among plants with cyanobacterial symbiosis (Table 1). The specific role of all of these plant genes for symbiosis remains to be experimentally elucidated. Comparative genomics provides the necessary groundwork for an understanding of the evolutionary mechanisms of plant-cyanobacterial symbiosis.

![Figure 3.](image)

Figure 3. a. Chemical structures of cycad derived specialized metabolites found in leaves and coralloid roots: (1) BMAA, (2) Macrozamin and (3) cycacin. (4) hormogonium-inducing factor (HIF) 1-palmitoyl-2-linoleoyl-sn-glycerol produced in precoralloid roots. b. Chemical structures of microbiota derived metabolites found in coralloid roots: (5) Caulobacter- produced indigoidine-like metabolite, (6) cyanobiont produced desmamide A, (7) nostocyclopeptide A1/A3. Note that BMAA (1) is also produced by cyanobacteria in the coralloid roots. c. Host genome gene count differences between the coralloid roots (with Cyanobacteria) vs precoralloid roots (without Cyanobacteria) of Cycas panzhihuaensis. KEGG mapping of enriched genes in coralloid roots (>10000 counts) correspond to endopeptidase inhibitor activity (CYCAS_000803, CYCAS_008980, CYCAS_000805), manganese ion binding and nutrient reservoir activity (CYCAS_020515), response to stress, stimulus or wounding (CYCAS_000805), and Serine-type endopeptidase inhibitor activity (CYCAS_008985). Enriched genes in precoralloid roots (>10000 counts) correspond to endopeptidases inhibitor activity (CYCAS_018813, CYCAS_001064, CYCAS_008980). Data from Liu et al. (2022).

Regarding the community ecology of the coralloid, recent studies have confirmed the existence of multiple microorganisms including fungi and viruses inside the coralloid root and other nitrogen-fixers such as Hypomicrobiales (Rhizobiales) (Bustos-Díaz et al., in review and references therein.). While the cyanobiont appears to be the main nitrogen fixer, the role of the microbial community is not well characterized, although efforts are underway (Liu et al., 2023; Ndlovu et al., 2023). Along these lines, specialized metabolites produced under nitrogen starvation by associated bacteria in interaction with the cyanobiont in Dioon edule have been reported (Gutiérrez-García et al., 2019). It has also been suggested that the community itself might be recruited from the soil as a consortia because coralloid root microbial composition differs greatly from other tissues and the surrounding rhizosphere (Suárez-Moo et al., 2019; Zheng & Gong, 2019). Investigations into the recruitment of the community, the possibility of some taxa inherited via the seed, and the dynamics within coralloid roots are still in their infancy and will likely produce many important advances in symbiosis ecology and specialized metabolism in the coming decade.
The long standing hypothesis that coralloid roots are an ancestral trait, based on the near ubiquity of coralloid roots in all living cycad species, was recently challenged. In the absence of coralloid root fossils, a proxy method used to deduce the existence of symbiotic nitrogen fixation using nitrogen isotopic ratios of fossil cycad leaves found independent origins of the symbiosis in living Zamiaceae and Cycadaceae (Kipp et al., 2024). The authors concluded that a 35 million year old fossil *Zamia* entered in symbiosis with nitrogen-fixing bacteria, while inconclusive results were obtained from *Bowenia* leaf fossils from 50 million years ago. Interestingly, older fossils from extinct cycad genera showed no signs of being capable of symbiotic nitrogen fixation. The authors hypothesize that morphological similarities between coralloid roots from cycads in both families result from convergent evolution which, if true, would make cycads a rich system for investigations into plant morphology and evo-devo. Combined with the ongoing and recent advances in the chemical ecology and genetics of symbiosis described above, these data provide a full picture of the ecology, evolution, physiology, genetics, and development of root symbiosis.

Table 1. Twenty-four conserved cyanobacterial symbiotic associated genes shared among *C. panzhihuaensis*, the ferns *Azolla filiculoides* and *Azolla cf. caroliniana*, the liverwort *Blasia pusilla* and the hornwort *Anthoceros punctatus*. Gene sequences can be downloaded from (https://db.cngb.org/codeplot/datasets/PwRflGHfPs5qG3gE), data from Liu et al., (2022).

<table>
<thead>
<tr>
<th>Gene ID (Cycas panzhihuaensis)</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>CYCAS_027607</td>
<td>26S proteasome regulatory subunit 6B homolog</td>
</tr>
<tr>
<td>CYCAS_019641</td>
<td>Anion exchange protein 3 (AE 3) (Anion exchanger 3)</td>
</tr>
<tr>
<td>CYCAS_024481</td>
<td>Asparagine synthetase [glutamine-hydrolyzing]</td>
</tr>
<tr>
<td>CYCAS_014362</td>
<td>Band 3 anion transporter (Anion exchange protein 1)</td>
</tr>
<tr>
<td>CYCAS_011926</td>
<td>Biosynthetic arginine decarboxylase (ADC)</td>
</tr>
<tr>
<td>CYCAS_029123</td>
<td>Glycosyl family 1</td>
</tr>
<tr>
<td>CYCAS_021936</td>
<td>Pentatricopeptide repeat containing protein At3g09040</td>
</tr>
<tr>
<td>CYCAS_013170</td>
<td>Pentatricopeptide repeat containing protein At4g16470</td>
</tr>
<tr>
<td>CYCAS_023547</td>
<td>Pentatricopeptide repeat containing protein DOT4, chloroplastic</td>
</tr>
<tr>
<td>CYCAS_004639</td>
<td>Probable ubiquitin carboxyl-terminal hydrolase FAF-X</td>
</tr>
<tr>
<td>CYCAS_018734</td>
<td>Protein argonaute 10 (Protein PINHEAD)</td>
</tr>
<tr>
<td>CYCAS_003784</td>
<td>Protein EXECUTER 2, chloroplastic</td>
</tr>
<tr>
<td>CYCAS_001250, CYCAS_018283, CYCAS_018284, CYCAS_018285, CYCAS_021027, CYCAS_022196</td>
<td>Trihelix transcription factor ASIL2</td>
</tr>
<tr>
<td>CYCAS_031507</td>
<td>Zinc finger protein 606</td>
</tr>
<tr>
<td>CYCAS_013796, CYCAS_024257, CYCAS_024328, CYCAS_026696, CYCAS_024940</td>
<td>Unknown</td>
</tr>
</tbody>
</table>

Cycad phyllosphere microbiota and interaction with the plant metabolome

Above-ground plant tissues (i.e., the phyllosphere) microbiota may also play a role in plant ecology and fitness. Phyllosphere bacteria of cycads has only been published in *Cycas panzhihuaensis* (Zheng & Gong, 2019), but did not include leaves, the most prominent cycad organ. The authors found that reproductive structures (ovules, unfertilized and viable seeds) contain less bacterial diversity in comparison with roots (which are enriched in cyanobacteria). Bacterial communities among reproductive organs were highly similar and dominated by the family Enterobacteriaceae. Cycad pollen, however, was significantly enriched with Comamonadaceae, when compared to the other plant organs (Zheng & Gong, 2019). Here we present the first cycad leaf bacteriome and mycobiome (Fig. 4a, b). We find that Enterobacteriaceae are rarely represented in the phyllosphere community of six *Zamia* species, while Nostocaceae, Beijerinckiaceae and Acetobacteraceae are abundant. The rather surprising abundance of the nitrogen fixing bacteria *Desmonostoc* suggests functional redundancy with coralloid roots, possibly as a failsafe to ensure the plant acquires sufficient fixed nitrogen (Moyes et al., 2016) and represents the first evidence of potential additional symbionts in these long-lived plants. Research on leaf fungal endophytes is limited to a few studies, i.e., *Encephalartos* (Nesamani et al., 2017) and two *Zamia* species (Sierra et al., in review). The endophytic leaf fungi form a polyphyletic assemblage of mostly Ascomycota (Fig. 5b). These initial diversity studies of leaf endo-
and epiphytic microbes provide a veritable mine of data for further investigations into the ecological and evolutionary significance of the phyllosphere.

Cycads are known for harboring a number of anti-herbivore and toxic compounds (see Sect. V), whose origin and impact across trophic levels has been a source of great debate (Marler, Snyder & Shaw, 2010 and reference therein). Sierra et al. (in review) recently addressed the association of plant-microbiota (bacteria and fungi) on the intraspecific leaf metabolome composition of two species (Zamia nana and Z. pseudoparasitica). Among the 49 abundant metabolites, five compounds (Benzodioxoles, Biflavonoids and polyflavonoids, 1,2-diacylglycerols, N-acyl-alpha amino acids and glutamic acid and derivatives), were more expressed in Z. pseudoparasitica compared to Z. nana (Sierra et al., in review). The metabolite variation was correlated to bacterial and fungal leaf endophyte community composition (Fig. 4c), and significantly associated with the abundance of one bacterial order (Frankiales) and three fungal orders (Hypocreales, Sordariales and Chaetothyriales). We are just beginning to unravel the association of cycad metabolites with phyllosphere microbiota and the ecological consequences for associated insects. Defensive secondary metabolites in cycads seem to be induced by complex plant-microbe reciprocal interactions, which calls for experimental research to uncover the genomic and biochemical underpinnings derived by the host and endophytic microbes, or both.
Figure 4. Survey of Zamia phyllosphere microbiota and metabolites reveals potential ecological functions. 

a. Core members of the bacterial endophyte community (892 ASVs) of six species (Z. elegantissima, Z. nana, Z. pseudomonticola, Z. pseudoparasitica, Z. skinneri, Z. stevensonii) identified using 16S-RNA amplicon sequencing (n = 28). The nitrogen-fixing cyanobacteria Desmonostoc was observed in 30% of the samples with relative abundance up to 20%.

b. Leaf fungal endophyte variants (10343 ASVs) of two species (Z. pseudoparasitica and Z. nana) identified using ITS1 amplicon sequencing (n = 127). The genera Colletotrichum and...
is the most prevalent fungal endophyte in >50% of the samples studied. c. Foliar metabolome composition of Z. nana and Z. pseudoparasitica summarized into two dimensions with non-metric multidimensional scaling (NMDS) based on the Bray-Curtis dissimilarity index indicating the correlation between the metabolome and the microbiome. Solid lines correspond to regressions with p-value = <0.05 and dashed lines with p-value = 0.1. Modified from Sierra et al. (in review)

Cycas panzhihuaensis genome

Cycads offer a unique opportunity to uncover shared genomic traits exclusive to seed-producing plants. The first cycad genome was recently published and this high-quality 10.5 Gb chromosome-level assembly immediately provided a plethora of insights (Liu et al., 2022). The Cycas panzhihuaensis genome contains evidence of an ancient whole-genome duplication event in the common ancestor of living gymnosperms and key innovations enabling seed formation are evident in the genome, including expansions of gene families regulating embryo development, secondary growth, and pollen tube formation. The genome also shows remarkable expansions of genes involved in chemical defenses, stress tolerance, and disease resistance, which may help explain the higher extant species diversity of cycads compared to Gingko. We elaborate further on a few particularly exciting findings from the genome. Overall, the high-quality C. panzhihuaensis genome assembly and annotation sheds light on the origin and evolution of cycads and provides insights into genetic innovations underlying seed plant success (Liu et al., 2022).

Horizontal gene transfer

The C. panzhihuaensis genome contains four copies of an insecticidal toxin gene family named fit D that was acquired via ancient horizontal transfer from fungi (Lui et al., 2022). These fit D genes encode proteins similar to the fit toxin and MCF toxin of Photorhabdus bacteria, which are lethal to insects. Experimental injection of the synthesized C. panzhihuaensis FitD protein into moth larvae resulted in significantly higher mortality than controls, demonstrating that this protein provides resistance against at least some insect herbivores. Phylogenetic analyses indicate the fitD gene family originated from bacteria and the cycad sequences share a common ancestor with fungi fitD with subsequent gene family expansion in the ancestor of C. panzhihuaensis and Cycas debaoensis. The acquisition of this toxin gene family represents an important innovation that likely enhanced the defense capacities of Cycas against insect pests and illustrates how horizontal transfer of foreign genes can provide adaptive benefits.

Sex chromosomes

Genetic mapping and genome sequencing revealed a large non-recombining region spanning 124 Mb on chromosome-8 that functions as the sex-determining region in C. panzhihuaensis (Lui et al., 2022). Assembly of the male-specific Y chromosome region further revealed a candidate sex determination gene encoding a MADS-box transcription factor named MADS-Y. This gene and its autosomal paralog showed sex-specific expression patterns, with MADS-Y exclusively expressed in males. MADS-Y represents a potentially conserved mechanism for sex determination originating before the divergence of cycads and Gingko over 200 million years ago (Liao et al., 2020). The identification of this master regulator provides key insights into the genetic control of sex determination in cycads, the most ancient dioecious seed plant lineage.

Flagellum formation and the shift to non-motile sperm

Unlike most extant gymnosperms and all angiosperms, cycads retain motile sperm cells with flagella similar to more ancestral land plants (Fig. 1b; see Offer et al., 2023 for a comparative study of reproductive structures between cycads and Gingko), providing an excellent opportunity to understand the evolutionary transition to non-motile sperm as well as the intricate network of genes and signaling pathways involved in the development and functionality of flagella. Previous research has identified several genes associated with flagellum assembly and maintenance, such as dynein genes and intraflagellar transport (IFT) proteins (Hodges et al., 2011). Dyneins are molecular motors that play a vital role in flagellar movement and IFT proteins are crucial for cilia and flagella assembly and maintenance. The C. panzhihuaensis genome (Lui et al., 2022) contains intact genes encoding dynein proteins and IFTs as well as other flagellar components
including outer dense fibers and some of the regulatory mechanisms governing gene expression and signal transduction pathways associated with flagellum formation and function (Arnaiz et al., 2009). However, some genes encoding radial spoke proteins are missing (Fig. 5), suggesting partial flagellar defects compared to non-seed plants (Hodges et al., 2011). Cycads represent a transitional state in the evolutionary shift toward non-motile sperm in gymnosperms and angiosperms and the identification of core flagellar genes present in the cycad genome but not in other seed plant lineages strengthens our knowledge of the genetic machinery behind flagellum formation and the shift to non-motile sperm.

Figure 5. Distribution of flagellar genes across embryophyte genomes reveals key genes in the evolution of non-motile sperm. a. Illustration of Cycas sperm cell, scale = 40 μm (reproduced from The Biology of the Cycads, Norstog & Nicholls 1997 Cornell University Press). Each sperm cell possesses between 40,000-50,000 flagella arranged in 5-10 sinistral coils (Offer et al., 2023). b. Cladogram illustrating flagellum loss in green plants. c. Distribution pattern of outer dense fiber protein and other crucial flagellar proteins among representative embryophytes and green algae. Each lineage is color-coded. Protein nomenclature follows Liu et al., (2022)

Conclusion

Here we have reviewed the last decade’s worth of research on cycad phylogenetics, genomics, and symbioses. These exciting recent discoveries position cycads as integral to our understanding of seed plant evolution and ecology, with far-reaching implications for the biological sciences. Phylogenetic advances continue to shed light on the early diversification of seed plants, while investigations into cycads’ diverse associations with insects has directed new attention to a broader range of signaling methodologies and provided fertile ground for investigations into adaptation and coevolution. Studies on coralloid root symbioses have highlighted the importance of multi-species bacterial communities in cycad-cyanobacteria symbiosis and opened the door for studies on the ecology of specialized metabolites. Similarly, investigations of plant phyllosphere microbiota have identified previously overlooked putative symbionts. Finally, recent genomic advances have uncovered a wealth of information highlighting important aspects of land plant evolution from the shift to non-motile sperm to ancient horizontal gene transfer across biological kingdoms. Overall, cycads and their biotic associates provide a rich study system for investigating ancient co-evolution and microevolutionary and ecological interactions.

Much of the current research on cycads can be unified through the lens of chemical ecology, as we become increasingly aware of the myriad ways that cycads’ chemodiversity drives their interactions with other organisms. Investigations into chemical trait evolution and ecology will shed light on the origins and maintenance of plant chemodiversity over long timescales and the implications for mutualistic and parasitic insect associations. Further investigations into the metabolic diversity and ecology of microbionts will improve our understanding of the mechanisms underlying plant-microbe symbioses. Just as cycads provide a bridge between extant and early seed plants, they will also help scientists to connect diverse research disciplines and
continue to provide insights into the evolution and ecology of seed plants.

Data availability statement
Leaf microbiome data presented in Figure 4 from Sierra et al. (in review) are currently in review and are available at https://github.com/adrielmsierra/Zamia_Phyllosphere/. Raw data is found in the Genbank Bioproject PRJNA1061598.

Author contribution
JCV conceived of the idea for the review, SS spearheaded the manuscript, all authors contributed to writing.

Acknowledgments
We acknowledge the vast amount of historical and current research in cycad biology that we were unable to include in this review due to scope and citation limits, it has and continues to inspire us. JCVA acknowledges support from the Canada Research Chair (950-232698), the CRNSG- RGPIN 05967–2016, and the Canadian Foundation for Innovation (projects 36781, 39135). SKS acknowledges the 10KP project (https://db.cngb.org/10kp/). AMS acknowledges support from the doctoral scholarships of the NSERC CREATE program in Biodiversity, Ecosystem Services and Sustainability (BESS) and the Canada Research Chair (950-232698).

References


cads generates parallel accelerated evolution of toxin tolerance in two clades of *Eumaeus* caterpillars (Lepidoptera: Lycaenidae). *Proceedings of the National Academy of Science USA* **118**: e2018965118. https://doi.org/10.1073/pnas.2018965118


