The *Encephalartos natalensis*-cyanobacterial coralloid root partnership for nitrogen acquisition

by

CASSANDRA BIANCA SCHOEMAN

Submitted in partial fulfilment of the requirements for the degree

Magister Scientiae

In the Faculty of Natural and Agricultural Sciences

Department of Biochemistry, Genetics and Microbiology

University of Pretoria

Pretoria

24 June 2022

Supervisor: Prof. Eshchar Mizrachi

Co-supervisors: Prof. Thulani Makhalanyane and Dr Danielle Roodt

Declaration

I, Cassandra Bianca Schoeman, declare that the dissertation, which I hereby submit for the degree M.Sc. Genetics at the University of Pretoria, is my own work and has not previously been submitted by me for a degree at this or any other tertiary institution.

Cassandra Bianca Schoeman

24/06/2022

Date

To my mother,

for her love and support and

without whom this would not have been possible.

I love you lots and lots like jelly tots.

TABLE OF CONTENT

DISSERTATION SUMMARY

The *Encephalartos natalensis*-cyanobacterial coralloid root partnership for nitrogen acquisition

Cassandra Bianca Schoeman

Supervised by ***Prof. Eshchar Mizrachi***

Co-supervised by ***Prof. Thulani Makhalanyane*** *(University of Pretoria) and* ***Dr Danielle Roodt*** *(University of Pretoria)*

Submitted in partial fulfilment of the requirements for the degree ***Magister Scientiae***

Department of Biochemistry, Genetics and Microbiology

University of Pretoria

Plant-microbe symbioses have important environmental and ecological implications because of the role they played in plant diversification as well as their ongoing impact on nutrient acquisition in nutrient-poor environments. Land plants, the foundation of most terrestrial ecosystems today, are the descendants of ancestral aquatic algae that transitioned to land approximately 443 to 470 million years ago. The consensus is that symbiotic associations facilitated plant colonisation. Over 450 million years of coevolution has led to a tremendous diversity of mutualistic symbioses, which could have aided the eventual domination of angiosperms. Plant-microbe evolution studies have mainly focused on the angiosperms, especially focusing on the two well-known symbioses commonly established by them; the arbuscular mycorrhizal (fungal) and the root-nodule (bacterial) symbioses. Although they make up ~80% of extant land plant species, angiosperms represent but a single plant lineage, therefore, invaluable knowledge can be gained from studying the unique biology of symbioses in other land plant lineages. For example, cyanobacterial symbiosis independently evolved across unrelated land plant lineages (bryophytes, ferns, gymnosperms and angiosperms), and is amongst the major evolutionary innovations linked to the acquisition of nitrogen via partnerships with microorganisms. However, since the majority of published literature on cyanobioses are dated, information for these partnerships have remained inadequate – especially at a molecular and cellular resolution.

Using an integrated approach and utilising multiple techniques including microscopy, *de novo* transcriptome assembly and quantification, and comparative genomics, the objectives of this MSc project were to investigate gene expression in coralloid root and control root tissues from the cycad, *Encephalartos natalensis*, and analyse genes preferentially expressed within these tissues as well as genes involved in the more commonly studied symbiotic associations. For comparative purposes, the symbiotic and control tissues from three other cyanobiosis-forming plant lineages were also analysed. By integrating multiple techniques and scientific fields, this project is one of the first to show possible neofunctionalisation of common symbiotic pathway genes for cyanobacterial partnerships from pre-existing arbuscular mycorrhizal symbiosis in a similar way as that which occurred in nodule

symbioses. While adding integral knowledge to the field of plant-microbe evolution, major outputs of this study also include transcriptomic resources for two cyanobacterial hosts, a comprehensive candidate gene list as well as gene expression profiles.

PREFACE

This dissertation focuses on the evolution and biology of plant-microbe symbioses for nitrogen acquisition – in particular a cyanobacterial symbiosis with the indigenous cycad, *Encephalartos natalensis*. The research completed in this dissertation makes a unique contribution to our understanding of this multifaceted and under-explored symbiosis. This dissertation was completed in the Department of Biochemistry, Genetics and Microbiology with a focus on genetics, but utilised a variety of techniques, such as light and transmission electron microscopy, transcriptomics, comparative genomics as well as bioinformatics. The format of this dissertation has been divided into draft articles, and a concluding remarks chapter containing a summary of the findings, my thoughts on and future prospects of this research. As such, some information and references are duplicated and repeated between chapters.

Chapter 1 was written as a review article, which considers literature pertaining to nutrient acquisition and the origin and emergence of plant mutualistic symbioses as adaptations for nitrogen acquisition. The main focus of the review is on the evolution of mycorrhizae and root nodule symbioses following the terrestrialisation of land plants, concluding with what is currently known about the only other plant-microbe symbioses, namely those in partnership with cyanobacteria. This includes a discussion of the limited molecular information regarding plant-cyanobacterial partnerships and highlights specific and pertinent knowledge gaps. Known genetic mechanisms regulating plant mutualistic symbioses in general, with a particular emphasis on the "common symbiotic pathway" are discussed in detail. Given that molecular analyses and transcriptomic studies have been performed on two cyanobiosis-forming lineages (namely hornwort and *Azolla*), Chapter 2 focused on the major gaps, especially for cycad but also *Gunnera*. This study centred primarily around the cycad, *E. natalensis*-cyanobacterial partnership while representatives of other cyanobacterial hosts (*Anthoceros punctatus*, *Azolla filiculoides* and *Gunnera perpensa*) served as comparative analyses. The presence of nitrogen-fixing cyanobacteria was investigated using light and transmission electron microscopy in the cycad and two other Southern African cyanobacterial hosts, *Az. pinnatta* spp. *africana* and *G. perpensa*. Transcriptomic resources were generated and analysed for *E. natalensis* and *G. perpensa*, and the underlying genetic mechanisms responsible for this symbiosis investigated. Major outputs of this study, in addition to transcriptomic resources, include a comprehensive candidate gene list and gene expression profiles, which revealed that the cycad-cyanobacterial symbioses neofunctionalised arbuscular mycorrhizae (CSP) and sugar transport genes in a similar way as root nodule symbioses have.

The research conducted during the course of this dissertation was done by myself in the African Plant Systems biology for the Bioeconomy (APSB) programme at the Forestry and Agricultural Biotechnology Institute (FABI), Department of Biochemistry, Genetics and Microbiology, University of Pretoria, under the guidance of my supervisors, Prof. Eshchar Mizrachi, Dr Danielle Roodt and Prof. Thulani Makhalanyane, with financial support from South Africa's National Research Foundation (Grant UID 116239) and the Technology Innovation Agency

of South Africa through the Forest Molecular Genetics (FMG) Cluster programme and the University of Pretoria. Samples from *Az. pinnatta* spp. *africana*, *E. natalensis* and *G. perpensa* were provided by the University of Pretoria Manie van der Schijff Botanical Garden and the University of Pretoria Cycad and Indigenous Plant Nursery with kind assistance from the curators. At the time of submission, the following conference abstract emanated from this MSc dissertation:

NATIONAL CONFERENCE PROCEEDINGS

Schoeman, C. B., Roodt, D., Pinard, D., Hammerbacher, A., Makhalanyane, T., Mizrachi, E., The *Encephalartos natalensis*-cyanobacterial coralloid root partnership for nitrogen acquisition, Oral presentation. South African Genetics Society and South African Society for Bioinformatics BIO2022 joint conference, April 24-27 2022, Stellenbosch, South Africa.

ACKNOWLEDGEMENTS

To my supervisor, Prof. Eshchar Mizrachi, and co-supervisors, Dr Danielle Roodt and Prof. Thulani Makhalanyane, thank you for your guidance and support throughout this project. Esh, thank you for your motivation and encouragement, your passion for science and excellence, and for creating a research group in which we are free to learn and grow. Danielle, thank you for your mentorship, enduring motivation and encouragement, but most importantly your advice on any subject no matter the time. During the course of this project, I not only found a mentor with a passion for science but also a friend. Thulani, thank you for being a part of this project and for providing invaluable feedback.

To the co-authors of each study included in this dissertation, Desré Pinard and Oliver Bezuidt, thank you for your assistance, contributions and inputs. To Prof. Almuth Hammerbacher, thank you for your contributions and inputs in committee meetings and for your willingness to serve as a BGM representative. Thank you also to the staff at the H.G.W.J. Schweickerdt Herbarium and the Laboratory for Microscopy and Microanalysis for use of the facilities and for assistance with preparing and staining samples for the light and transmission electron microscopy work included in Chapter 2. In particular, I'd like to individually thank Dr Kenneth Oberlander, Dr Angelique Kritzinger and Ms Erna van Wilpe not only for their eager and friendly assistance but also for the knowledge they shared with me. For sampling of the species included in Chapter 2 of this project, thank you to the University of Pretoria Manie van der Schijff Botanical Gardens curator Jason Sampson and the University of Pretoria Cycad and Indigenous Plant Nursery curator Arnold Frisby for eager and friendly assistance.

I am grateful to the members of my research group, the African Plant Systems biology for the Bioeconomy (APSB) programme for their support, motivation, friendship and good humour. A special thank you to the directors, leaders and the members of the Forest Molecular Genetics (FMG) programme, Forestry and Agricultural Biotechnology Institute (FABI) and the Department of Biochemistry, Genetics and Microbiology (BGM) for creating a stimulating environment in which young researchers can grow and learn.

I am grateful for the financial support I received throughout this project. I acknowledge South Africa's National Research Foundation (Grant UID 116239) and the Technology Innovation Agency of South Africa through the Forest Molecular Genetics (FMG) Cluster programme for their annual support and the University of Pretoria for support through the UP Postgraduate Bursary.

Thank you to my friends, both those in and outside FMG, for neverending support and encouragement. A special thanks to Mandy for being there since the start of my undergraduate studies. Thank you for being such a wonderful friend.

To my family, thank you for your support and encouragement. To my father, Louis, thank you for your love and support throughout my academic career. To my mother, Bronwyn, thank you for raising me to be independent, for

guiding me and shaping the person I am today. Thanks to you and oom Clifton for being there not only in the best of times but through the hardest of times as well, your enduring support and love made completing this research project possible.

To my second family, tannie Gerda and oom Mike, I am grateful for your love and support throughout not only my academic career but my personal life as well. Thank you for this and so much more.

To my significant other, Ruwé Kapp, words cannot fully describe the enormous contribution you've made to my life. You have made my life an adventure and I am better for knowing you. Thank you for your unwavering love, support, motivation, encouragement and so much more.

CHAPTER 1

LITERATURE REVIEW

LAND PLANT SYMBIOSES FOR NITROGEN ACQUISITION

Cassandra Schoeman[1], Danielle Roodt[1], Thulani Makhalanyane[2], Eshchar Mizrachi[1*]

[1] *Department of Biochemistry, Genetics and Microbiology, Forestry and Agricultural Biotechnology Institute, University of Pretoria, Private bag X28, Pretoria 0028, South Africa*

[2] *Department of Biochemistry, Genetics and Microbiology, University of Pretoria, Pretoria 0028, South Africa*

*Corresponding author email: eshchar.mizrachi@fabi.up.ac.za

This chapter has been prepared in the format of a manuscript for a peer-reviewed research journal. I drafted the manuscript and prepared all of the figures. Danielle Roodt, Thulani Makhalanyane and Eshchar Mizrachi helped edit the manuscript.

1.1. Abstract

Nutrients like nitrogen (N) are essential for plant survival, supporting both their growth and development. Nitrogen, although abundant in the atmosphere, is the most common growth-limiting nutrient in terrestrial ecosystems. Plants, however, lack the capacity to fix atmospheric nitrogen (N_2) instead relying on reactive forms, ammonia (NH_3) and/or nitrate (NO_3^-), for their survival. Therefore, most land plants associate with microorganisms to overcome this inability, particularly when available nitrogen is scarce. Mycorrhizae and diazotrophic bacteria (rhizobia and *Frankia*) are well-known examples of symbionts that play crucial roles in nitrogen acquisition. The former of which coevolved with plants since plant terrestrialisation approximately 443 to 470 million years ago. Despite the evolutionary time separating the origin of mycorrhizal and nodule symbioses, these symbioses share numerous conserved features, including a plant-specific symbiotic signalling pathway. Today, symbioses are found in most of the plant kingdom, with plant-cyanobacterial symbiosis being but one of only four major innovations in plant evolution for acquiring nitrogen through partnership with microbes. Unlike plant symbioses established by arbuscular mycorrhizae, or symbioses involving rhizobia or *Frankia*, cyanobacterial symbiosis evolved independently and in parallel across unrelated lineages of land plants. Although work has been done in understanding the establishment and functioning of cyanobacterial symbioses, much of the underlying genes and pathways remain unknown.

1.2. Introduction

Plant terrestrialisation was a major evolutionary event not only for the future of land plants but for all terrestrial life (Pirozynski and Malloch, 1975; Kenrick and Crane, 1997). As primary producers, plants depend on abiotic factors to support their growth and development, such as carbon dioxide, sunlight, water and soil nutrients (reviewed in Oldroyd and Leyser, 2020). For some time, the availability of soil nutrients have been known to constrain plant growth. Nitrogen (N) being the most common "growth-limiting nutrient" in terrestrial ecosystems (particularly on the African continent; Hengl et al., 2017). Plants rely on inorganic nitrate (NO_3^-) or ammonium (NH_4^+), released into the soil via the nitrogen cycle as decomposed nitrogen-containing matter (for a review on the nitrogen cycle, see Thamdrup (2012). Despite the availability of various sources in soil, the majority of nitrogen (95-99%) is either inorganic or in the form of inert atmospheric nitrogen (N_2), which are biologically inaccessible to plants (Thamdrup, 2012).

Nitrogen fixation – the process of converting atmospheric nitrogen into inorganic or organic nitrogen forms – occurs naturally either through abiotic processes or enzymatically in free-living fungi and diazotrophic bacteria or by those involved in symbiotic associations (reviewed in Reed et al., 2011). Plants, however, are unable to perform nitrogen fixation (Merrick and Dixon, 1984; Cheng, 2008). Instead, plants optimise their nitrogen uptake through modifications to their root systems. For instance, by promoting lateral root proliferation (Drew, 1975; Zhang and Forde, 1998), or interacting with microorganisms (reviewed in Hodge, 2004; Oldroyd and Leyser, 2020). The

plant's ability to associate with fungi and diazotrophic bacteria, termed symbiosis, represents an important natural system for nitrogen capture enabling them to adapt to nutrient-poor environments (reviewed in Mahmud et al., 2020; Oldroyd and Leyser, 2020). Symbioses, as such, are widespread among plant species today where they have not only shaped plant diversity but have also had an unmistakable impact on plant mineral nutrition (reviewed in MacLean et al., 2017; Roy et al., 2020).

Plant diversification and evolution was dependent on plant-microbe symbioses, which not only arose during but likely facilitated land plant terrestrialisation around ~450 million years ago (Pirozynski and Malloch, 1975; Kenrick and Crane, 1997; Selosse and Le Tacon, 1998). Since then, plants and symbionts have coevolved leading to the development of specialised nutrient strategies and the eventual domination of angiosperms (reviewed in Brundrett, 2002; Brundrett and Tedersoo, 2018). The main plant symbioses are those established by mycorrhizal fungi, arbuscular mycorrhizal and ectomycorrhizal, and those with nitrogen-fixing bacteria, rhizobium and *Frankia* (reviewed in Oldroyd et al., 2011; Santi et al., 2013; Genre et al., 2020). Mycorrhizal-plant symbioses are widespread, representing 70-90% of all terrestrial symbioses, and even facilitated major transitions, such as the colonisation of plants into terrestrial habitats (Brundrett, 2002; Brundrett and Tedersoo, 2018). Nitrogen-fixing bacterial symbioses, despite being a crucial system for capturing and processing atmospheric nitrogen (N_2), are rare across the plant kingdom (Santi et al., 2013). Hosts mainly include a few families of angiosperms (Soltis et al., 1995; Swensen, 1996) and diverse individuals involved in cyanobacterial symbioses, respectively (reviewed in Rai et al., 2000; Rai et al., 2002; Rai et al., 2019). In this review, we focus on the evolution of symbiosis, the known genes of a common symbiosis signalling pathway, as well as those responsible for symbiosis establishment and maintenance in specific plant species. We place special emphasis on the less-understood and diverse cyanobacterial-plant symbioses, focusing on establishment, maintenance and nutrient exchange. We also address key differences between this lesser-known nitrogen-fixing symbiosis and those established by well-known symbiosis in angiosperm lineages and highlight important gaps and questions in plant cyanobacterial symbioses.

1.3. Plant acquisition and assimilation of nitrogen

Accessibility of essential nutrients in the soil, such as nitrogen, is limited which causes variable spatiotemporal landscapes both locally and on a global scale (reviewed in Stark, 1994; Hinsinger et al., 2005). A plant's growth and developmental responses hinge on the ability to sense nitrogen in the local environment as well as perceive and regulate its overall nitrogen status in source and sink tissues (reviewed in Tegeder and Masclaux-Daubresse, 2018; Oldroyd and Leyser, 2020). To this end, plants rely on local and systemic foraging responses to sense and acquire accessible nitrogen, and subsequently optimise uptake through the promotion of lateral root growth and proliferation thereof into local nitrogen patches (reviewed in Bellegarde et al., 2017; Ueda et al., 2017; Oldroyd and Leyser, 2020). Plant growth mainly depends on inorganic nitrogen, ammonium (NH_4^+) and nitrate (NO_3^-) uptake from the rhizosphere (Näsholm et al., 2009) but root uptake of organic nitrogen (amino acids and urea) also contributes (Tegeder and Rentsch, 2010). Plants have, therefore, developed highly diverse nitrogen foraging

responses and uptake systems, with a variety of inorganic and organic plasma membrane-localised transporters having been identified in plant roots (reviewed in Tegeder and Masclaux-Daubresse, 2018).

Plants use a variety of nitrogen sources to support their growth, particularly in fluctuating soil nitrogen environments where prevalent nitrogen sources such as NO_3^- are not always available (Miller and Cramer, 2005). Nitrate and Ammonium Transporters (NRT1s and NRT2s, and AMTs, respectively) facilitate the uptake of inorganic nitrogen from the soil in various plant species (Figure 1.1; reviewed in Tegeder and Masclaux-Daubresse, 2018; Baslam et al., 2021). In *Arabidopsis thaliana*, key nitrate (NRT2.4 and NRT2.5) and ammonium transporters (AMT1;1, AMT1;3 and AMT1;5) facilitate direct uptake from the soil, while others (NRT2.1, NRT2.2 and AMT1;2) are responsible for the import of nitrate or ammonium into endodermal and cortical cells (Loqué et al., 2006; Yuan et al., 2007; Kiba et al., 2012; Lezhneva et al., 2014). Amino acid uptake by roots is the predominant source of organic nitrogen to plants even though other organic nitrogen compounds (peptides and/or proteins) are also absorbed (Rentsch et al., 2007; Paungfoo-Lonhienne et al., 2008). Amino acid importers characterised to date include Amino Acid Permeases (AAPs), Lysine/Histidine-like Transporters (LHTs), Proline and Glycine Betaine Transporters (ProTs), and Usually Multiple Acids Move In and Out Transporters (UmamiTs) all of which transport numerous amino acids (Rentsch et al., 2007; Pratelli and Pilot, 2014). Arabidopsis AAP1, for example, is responsible for the uptake of glutamate and neutral amino acids into roots (Lee et al., 2007; Perchlik et al., 2014), while AAP5 acquires arginine (Arg), histidine (His) and/or lysine (Lys; Svennerstam et al., 2008; Svennerstam et al., 2011). These amino acids are subsequently transported to different sink tissues for growth and development (Tegeder and Masclaux-Daubresse, 2018; Baslam et al., 2021). However, since amino acids represent but a small proportion of nitrogen directly taken up from the soil, other sources need to be assimilated into amino acids prior to transport and subsequent use (Tegeder and Masclaux-Daubresse, 2018; Baslam et al., 2021).

Nitrate, once imported by NRT transporters into plant roots, is reduced to ammonium (Figure 1.1) – initially to nitrite (NO_2^-) by nitrate reductase in the cytosol and then by nitrite reductase in plastids to produce ammonium (Miflin, 1974; Srivastava, 1980). NO_2^- transport into plastids (for further reduction) is accomplished by the Nitrite Transporter 2;1 (NITR2;10 in Arabidopsis; Maeda et al., 2014). Glutamine Synthetases (GSs) and glutamine-2-oxoglutarate aminotransferases (GOGATs) are the principal route for assimilating ammonium into amino acids through nitrate reduction or ammonium assimilation (Lam et al., 1996; Kojima et al., 2014). It is well established that nitrite and nitrate reductase activity is regulated by both metabolic and environmental conditions, such as glutamine, sucrose, cytokinin and light, causing nitrate reduction to mainly occur during the day (Galangau et al., 1988; Reisdorf-Cren and Hirel, 1999). *De novo* biosynthesis of amides and amino acids occur both in plastids and the cytosol for subsequent transport through the xylem and phloem (Tegeder and Weber, 2007). Plasma membrane-localised transporters are essential to nitrogen transfer and regulates not only uptake but shoot and sink transport as well (Tegeder, 2014).

Transport of numerous amino acids throughout the plant occurs at the cellular level within the xylem and phloem but often the most abundant of those transported are glutamine, glutamate, asparagine and aspartate (Delrot et al., 2001). Xylem-to-phloem transfer of nitrogen to rapidly growing sinks in Arabidopsis and most crop plants (Pate et al., 1975; van Bel, 1984) requires nitrogen retrieval into xylem parenchyma cells (van Bel, 1990), symplastic movement of nitrogen towards and subsequent apoplastic phloem loading thereof (Okumoto and Pilot, 2011; Tegeder, 2012). As such, xylem not only functions in the direct supply of shoot tissue (typically photosynthetically active leaves) but also facilitates the retrieval of nitrogen for metabolism or storage within roots, stem and leaf major veins (Bailey and Leegood, 2016). Xylem retrieval and xylem-to-phloem exchange of nitrogen is accomplished by NPF7.2/NRT1.8 and AAP6, respectively (Hunt et al., 2009; Li et al., 2010). While AAP2 together with AAP6, expressed in the phloem throughout the plant, are essential to the xylem-to-phloem transfer of amino acids (Hunt et al., 2009; Zhang et al., 2010a). Once transported throughout the plant, nitrate and/or amino acids act as proxies for overall nitrogen status by contributing to local nitrogen sensing (Imsande and Touraine, 1994; Lejay et al., 1999; Gifford et al., 2008).

Plants assess their overall nitrogen status at the tissue or organ level and use this information to control their growth and metabolism (Gansel et al., 2001). For example, an effective strategy employed by plants to control local nitrogen acquisition, transport and assimilation is to accumulate nitrogen stores either within the cytosol or vacuole as short- or long-term storage pools (Staswick, 1994; Miller and Cramer, 2005). Integration of local and systemic nitrogen responses is, however, needed to block nitrogen foraging when shoot nitrogen is deemed sufficient (Zhang et al., 1999). Long-distance signalling mechanisms (i.e. systemic nitrogen responses) have evolved in plants which allow them to senses nitrogen stress, causing compensatory growth of roots into nitrogen-rich environments for improved nitrogen uptake (Tabata et al., 2014; Ohkubo et al., 2017). C-terminally encoded peptides (CEPs) along with cytokinin have been identified as root-to-shoot nitrogen signalling regulators (Sakakibara et al., 2006; Landrein et al., 2018; Taleski et al., 2018). These peptides (and cytokinin) are produced in roots experiencing nitrogen starvation and following synthesis travels to the shoot, where they are recognised by CEP receptors ultimately leading to CEP DOWNSTREAM 1 (CEPD) production (Ruffel et al., 2011; Tabata et al., 2014; Ohkubo et al., 2017). CEPDs act as long-distance signals since they are produced in leaves but function in roots where they up-regulate nitrate transporter expression, thereby stimulating nitrate uptake in locations where nitrate is ample (Ohkubo et al., 2017). Oldroyd and Leyser (2020), therefore, reviews that this root-to-shoot signalling allows plants to respond to their local nutrient landscape. It is clear that plants are able to adapt and survive in nutrient-poor environments. Nevertheless, to help facilitate nutrient capture and guarantee their growth and survival in nutrient-poor environments many plants turn to microorganisms (Oldroyd and Leyser, 2020).

1.4. Nitrogen acquisition in plants relies on symbiotic relationships with fungi and bacteria in the absence of available nitrogen

Plant symbioses, defined as mutualistic interactions which benefit the host as well as the microsymbionts (Bronstein, 2015), are widespread among extant plant species, where a major driver for its emergence was and remains to be a need to access scarce nutrients (reviewed in Martin et al., 2017). During the emergence of embryophytes, a single evolutionary event with lineage-specific adjustments along the way, for example, the evolution of ligand-recognising receptors in certain angiosperms (Bozsoki et al., 2020), gave rise to the diverse intracellular types of symbiotic associations we see today (reviewed in Delaux and Schornack, 2021). Since all extant plants are equipped with both general cell processes and defence strategies, it has been proposed that such genes served as a starting point from which a set of genetic regulators, known as the "common symbiotic pathway" (CSP), evolved that not only gave plants the capability to benefit from but also to support symbionts (reviewed in (Delaux and Schornack, 2021). Aside from a few families of angiosperms which harbour nitrogen-fixing bacteria (Soltis et al., 1995; Swensen, 1996), intracellular symbioses in plants are mainly established by endomycorrhiza, and in particular arbuscular mycorrhizal (AM) fungi (reviewed in Brundrett, 2002; Brundrett and Tedersoo, 2018). By contrast, intercellular symbiosis for example the well-characterised ectomycorrhiza seem to have repeatedly evolved through convergent mechanisms (Tedersoo and Brundrett, 2017). Here we focus on the characteristics and evolution of well-established mycorrhizal and nitrogen-fixing bacterial symbioses, while some of the key genes and signalling pathways underlying these associations are expanded on in section 1.4.3.

1.4.1. Mycorrhizal plant-symbiosis

Mycorrhizas (meaning 'fungus' and 'root') are ubiquitous associations between specialised soil fungi (myco) and plants roots (rhiza; Genre et al., 2020). Mycorrhizal fungi are present in most ecosystems (from tropical forests to arable land to deserts) where they develop extensive hyphal networks in the soil and specialised symbiotic interfaces (arbuscules, Hartig nets and mantels, fungal coils or pelotons) which connect the vast majority of plants with scarce and essential nutrients (reviewed in van der Heijden et al., 2015; Genre et al., 2020). As such, the establishment of mycorrhizal-plant symbioses was a crucial event in the evolutionary history of land plants; sustaining their growth in natural ecosystems with limited soil nutrient availability (Pirozynski and Malloch, 1975; Smith and Read, 2008). Mycorrhizal symbioses are broadly classified into ectomycorrhizas (EMs) and endomycorrhizas depending on whether the fungi colonises the root surface, inter- or intracellular spaces – the latter of which has several forms, namely orchid, ericoid and arbuscular mycorrhizas (AMs; reviewed in Martin et al., 2016; Genre et al., 2020).

The origin of arbuscular mycorrhizal fungi approximately 393 to 419 million years ago (mya; Remy et al., 1994) initiated the evolution of mycorrhizal associations in terrestrial plants (reviewed in Brundrett, 2002; Brundrett and Tedersoo, 2018). Primitive plant-microbe associations reminiscent of earlier-diverging AM-like fungi developed

when the earliest plants (simple rhizoid-based, leafless plants) made the transition to land 443 to 470 mya (Remy et al., 1994; Taylor et al., 1995). The consensus is that this colonisation drove fungi and early land plants alliances toward symbiotic relationships (Figure 1.2) because these plants faced major limitations to nutrients and water in the absence of existing soil (Pirozynski and Malloch, 1975; Kenrick and Crane, 1997; Kenrick and Strullu-Derrien, 2014). Mycorrhizal associations have since played key roles in land plant evolution and diversification and is said to have a single origin (Brundrett, 2002; Bravo et al., 2016) supported both by fossil evidence (Remy et al., 1994; Taylor et al., 1995) and the recent discovery that land plant ancestors were preadapted for fungal interaction (Delaux et al., 2015). Mycorrhizal evolution continued with a period of major root diversification, such as new types of mycorrhizas (independently evolved ectomycorrhizal, ericoid, and orchid associations; Figure 1.2) and non-mycorrhizal roots (parasitic and carnivorous plants, as well as nitrogen-fixing symbioses; reviewed in Brundrett, 2002; Brundrett and Tedersoo, 2018). This not only coincided with the origin of many plant families but might have played a role in the eventual domination of angiosperms in many habitats (reviewed in Brundrett, 2002; Brundrett and Tedersoo, 2018). Mycorrhizal evolution today is strongly linked to increasing habitat and soil complexity such that specialised nutrition strategies and multifunctional roots have emerged in a few plant families (reviewed in Brundrett, 2002; Brundrett and Tedersoo, 2018). Despite the success of mycorrhizal diversification, such that plant families developed new root traits, the majority of land plants (72-80%) have retained arbuscular mycorrhizal fungi throughout evolution (Figure 1.2; Brundrett and Tedersoo, 2018).

Mycorrhizal symbionts are distinguished based on their characteristic symbiotic interfaces, all of which have a plant-derived membrane-cell wall matrix which separates them from the plant cytoplasm (reviewed in Genre et al., 2020). Of the common mycorrhizal symbioses, the ancient arbuscular mycorrhiza, which has established symbioses with approximately 200,000 plant species, and the more recent ectomycorrhiza symbioses are the most widespread (reviewed in van der Heijden et al., 2015; Brundrett and Tedersoo, 2018). In the former, glomeromycete hyphae form tree-like arbuscules after growing in between plant root cells and penetrating cortical cells (reviewed in Peterson and Massicotte, 2004). While, ectomycorrhizal basidiomycetes and ascomycetes preferentially colonise lateral roots of nearly 6,000 woody plants, including beeches, dipterocarps, eucalypts, oaks, pines and poplars, at a site known as the “mycorrhizal infection zone” (Peterson and Massicotte, 2004). Colonisation precedes with sheathing mycorrhizas, in which a fungal mantle covers the root tip, and the Hartig net of intercellular hyphae surround epidermal cells in angiosperms and both epidermal and outer cortical cells in gymnosperms (Peterson and Massicotte, 2004). Contrary to this, ericoid and orchid mycorrhiza are confined to single plant families, namely the Ericaceae (commonly known as heath plants) and Orchidaceae family, respectively (Peterson and Massicotte, 2004). In ericoid mycorrhiza, fungal hyphae penetrate thick epidermal cell walls of root hairs to form intracellular hyphal complexes, termed fungal coils, within these cells (Peterson and Massicotte, 2004). Whereas orchid mycorrhizas penetrate roots of orchids and parenchyma cells of protocorms forming fungal pelotons (Peterson and Massicotte, 2004). The symbiotic structures of orchid, ericoid, arbuscular

mycorrhizas and ectomycorrhizas are efficient interfaces for the bi-directional transport of nutrients to and from plant cells (Peterson and Massicotte, 2004).

Arbuscular mycorrhizal fungi colonise pre-existing, intercellular spaces within roots which the plant cell actively prepares prior to infection (Genre et al., 2005; Genre et al., 2008). Within epidermal cells, the nucleus migrates towards the emerging fungal hyphopodium flagging the point of fungal entry (reviewed in Parniske, 2008). The nucleus then guides the growing prepenetration apparatus (a thick cytoplasmic bridge) through the epidermal cell where it forms a 'transcellular tunnel' through which the fungal hypha penetrates these cells (Parniske, 2008). Fungal hyphae subsequently grow into cortical cells where they repeatedly branch to produce arbuscules (Peterson and Massicotte, 2004). Once formed, a cortical cell, plant-derived membrane, namely the periarbuscular membrane, separates the arbuscules from the host cytoplasm (Balestrini and Bonfante, 2005, 2014). Together the periarbuscular membrane, the periarbuscular space and the fungal plasma membrane make up the symbiotic interface, and the site of nutrient exchange (Harrison, 2005). A fungal extraradical mycelium grows into the surrounding soil to capture mineral nutrients, such as nitrogen, to ultimately be transferred to the plant via arbuscules (reviewed in MacLean et al., 2017). The fungal extraradical mycelium absorbs ammonium (NH_4^+) from the soil via Ammonium Transporters (AMT; López-Pedrosa et al., 2006; Pérez-Tienda et al., 2011; Calabrese et al., 2016) which is then assimilated into arginine (Figure 1.3; Johansen et al., 1996; Jin et al., 2005). Arginine following transport to the intraradical mycelium is broken down into ammonia to be taken up by plant AMTs (Figure 1.3; Govindarajulu et al., 2005; Jin et al., 2005; Kobae et al., 2010).

While the bulk of carbon supplied to arbuscular mycorrhizal fungi was believed to be transported across the symbiotic interface in the form of free hexoses (reviewed in Parniske, 2008). Recently, in addition to sugars, host-synthesised lipids or fatty acids were found to be a key source of carbon directly transported to arbuscular mycorrhizal fungi to produce their own lipids (Figure 1.3; Luginbuehl et al., 2017). *De novo* fatty acid biosynthesis occurs in mycorrhizal colonised roots but not in spores or extraradical hyphae (Bago et al., 1999; Pfeffer et al., 1999). In plants, *de novo* fatty acid biosynthesis takes place inside plastids where malonyl-ACP, together with acetyl-CoA, is used to produce long-chain fatty acids in the form of 16- or 18-carbon molecules (Figure 1.3; Ohlrogge et al., 1993). *De novo* fatty acid biosynthesis ends when the ACP and acyl groups are released from fatty acid chains which in mycorrhizal plants involves an additional acyl-ACP thioesterase (FatM; Jones et al., 1995; Bravo et al., 2017). Free fatty acids are exported from the plastid, bound to coenzyme A (CoA) and transported to the endoplasmic reticulum where sn-2 monoacylglycerol (β-MAG), most likely with C16:0 acyl groups, is produced (Figure 1.3; Yang et al., 2010). The sn-2 monoacylglycerols (β-MAG), or derivatives thereof, are exported by the periarbuscular membrane-localised ABC transporters (named STR and STR2) into the periarbuscular apoplast and subsequently accessed by the fungus (Figure 1.3; Zhang et al., 2010b; Gutjahr et al., 2012). The *de novo* fatty acid biosynthesis of β-MAG and its subsequent export to the apoplast is unique to mycorrhizal plants (reviewed in Bravo et al., 2017).

1.4.2. Plant-symbioses in the nitrogen-fixing clade

Some angiosperms and nitrogen-fixing bacteria have coevolved to establish endosymbiotic associations (Soltis et al., 1995). These symbioses gave rise to specialised root organs, called nodules, that host prokaryotes and fix atmospheric nitrogen (N_2; Roy et al., 2020). All nodulating plant species occur within a monophyletic group, the nitrogen-fixing clade (NFC), which comprises the orders Fabales, Fagales, Cucurbitales, and Rosales (Soltis et al., 1995). Nodules, occurring in ten plant lineages that span across these four distinct taxonomic orders, are the heart of symbiotic interaction established by rhizobia and *Frankia* (Figure 1.2; Swensen, 1996). They carry within them bacteroids (the nitrogen-fixing form of rhizobia; reviewed in Oldroyd et al., 2011) and nitrogen-fixing vesicles formed at the ends of *Frankia* hyphae (Huss-Danell, 1997) thereby improving the nitrogen acquisition of host plants.

Nodulation represents a major event in angiosperm evolution and is hypothesised to have originated either as (1) massive parallel losses in most nitrogen-fixing clade descendants following a single gain of nodulation in the common ancestor, or (2) parallel evolution of nodulation in some descendants (reviewed in Battenberg et al., 2018; van Velzen et al., 2019). Since nodulating lineages differ in colonisation strategies, type of microsymbionts, nodule ontogeny, and development (Pawlowski and Demchenko, 2012; Svistoonoff et al., 2014), it is often regarded as evidence that nodulation arose independently in different lineages (reviewed in van Velzen et al., 2019). Previous studies suggest that a predisposition event (possibly the differentiation of nitrogen-fixing nodules) evolved around 110 mya at the base of the nitrogen-fixing clade (NFC), the predecessor of nodulation which ultimately facilitated the independent evolution of nodulation in each host plant lineage (Soltis et al., 1995; Werner et al., 2014). However, despite divergence between the main nodulating lineages more than 100 mya, these lineages share similarities in both development and structure (Doyle, 1994; Sprent et al., 2013; Svistoonoff et al., 2014). van Velzen et al. (2019) argues that these similarities suggest a stronger phylogenetic link than what can be expected from independent evolution. van Velzen et al. (2019) also interpret these similarities as a result of a single recruitment of all of the components necessary for a homologous nodulation symbiosis (Doyle, 1994; Swensen, 1996). Since a single gain hypothesis suggests a massive parallel loss of nodulation, non-nodulating taxa would have to harbour “fossil” evidence of nodulation similar to hosts that have lost arbuscular mycorrhizal symbiosis (Delaux et al., 2014; Doyle, 2016). This evidence may include pseudogenization of nodulation-specific genes, as is the case for non-nodulating *Prunus* and *Trema* species (Griesmann et al., 2018; van Velzen et al., 2018). Assuming parallel evolution of nodulation, gain-of-function events are proposed to result in two types of symbioses: rhizobial symbiosis occurring twice – once in legumes (order Fabaceae) and in the non-legume *Parasponia* (Cannabaceae, order Rosales) – and actinorhizal symbioses, an evolutionary diverse group of plants in the Fagales, Rosales and Cucurbitales orders, occurring multiple times with *Frankia* bacterial species (Doyle, 1994; Swensen, 1996). By contrast, a single gain of nodulation implies that at least two switches between *Frankia*

and rhizobia bacterial partners must have occurred (van Velzen et al., 2019). van Velzen et al. (2019) proposes that the single evolutionary origin of nodulation was driven by nitrogen-fixing *Frankia*.

Rhizobial root nodule symbioses (RNS) define associations between α- or β-proteobacteria (reviewed in van Velzen et al., 2019). These associations are collectively known as rhizobia and occur in plants from the Fabaceae and Cannabaceae family (Doyle, 2011; Pawlowski and Demchenko, 2012). Legumes (Fabaceae) is the third-largest family of angiosperms consisting of more than 18,000 species (Group et al., 2013), most of which associate with rhizobia (Roy et al., 2020). Well-known legume hosts include species within the genera *Lotus* and *Medicago*, including model hosts *Lotus japonicus* (Szczyglowski et al., 1998) and *Medicago truncatula* (Penmetsa and Cook, 2000). Surprisingly, *Parasponia* (Cannabaceae), a genus of non-leguminous plants, are the only other plants known to enter into symbiosis with rhizobia (Becking, 1992). Legume nodules are divided into two types – determinate and indeterminate – with each arising from different positions in the root and on different legume hosts (reviewed in Oldroyd et al., 2011). Determinate nodules, with a transient meristem, arise from the central cortex while indeterminate nodules, with a tip-growing meristem and cylindrical appearance, originate from inner cortical cells bordering the endodermis (reviewed in Oldroyd et al., 2011). Despite rhizobial hosts being phylogenetically diverse, they share many cellular, molecular, and genetic characteristics responsible for nodule organogenesis and rhizobial infection (reviewed in Vessey et al., 2005; van Velzen et al., 2019).

Colonising nitrogen-fixing bacterial symbionts are housed within *de novo* constructed root organs, known as nodules (Crespi and Frugier, 2008). The cellular changes induced within legume and non-legume (*Parasponia* and actinorhizal plants) hosts are, therefore, more complex such that root cells divide, nodule meristems form, and eventually, nitrogen-fixing bacteroids fill thousands of unique organelle-like compartments called “symbiosomes” (reviewed in Vessey et al., 2005; Oldroyd et al., 2011). Legume and actinorhizal plant roots share commonalities in rhizobial or *Frankia* entry, in both cases, bacteria enter via tubular infection threads which originate from epidermal root hair cells (Miller and Baker, 1986; Sprent and James, 2007) whereas in *Parasponia* hosts the root epidermis below rhizobia colonisation is simply eroded (Becking, 1992). Leguminous rhizobia or *Frankia* attach to root hair cells causing them to curl, “trapping” bacteria within developing infection threads which not only contains but will allow the eventual entry of dividing rhizobia colonies or branching *Frankia* hypha (Vessey et al., 2005; Oldroyd et al., 2011). Nodule organogenesis is triggered by cell divisions in the root cortex and pericycle following successful infection (Guan et al., 2013; Xiao et al., 2014). Nodule tissue encases a central, infected region which in indeterminate legume nodules is organised into developmental zones: the meristem (zone I), infection zone (zone II), nitrogen fixation zone (zone III), and senescence zone (zone IV) while only consisting of a nitrogen fixation zone in determinate nodules (Xiao et al., 2014). Bacterial colonisation in *Parasponia* and actinorhizal plants initially result in prenodules, while prenodules develop the nodule primordium is induced in the pericycle and continued colonisation results in nodule development similar to intermediate legume nodules (Callaham and Torrey, 1977; Bender et al., 1987). When the infection threads reach emerging nodule cortices,

they branch and release bacteria into large, polyploid plant cells, via unwalled infection "droplets" in rhizobia or intense branching in *Frankia*, resulting in symbiosomes (Schwintzer et al., 1982; Price et al., 1984; Burgess and Peterson, 1987). Bacteria within symbiosomes grow and differentiate into nitrogen-fixing bacteroids separated from the plant cytoplasm by a host-produced symbiosome membrane, where nutrient exchange occurs (Figure 1.3; Roth and Stacey, 1989; Huss-Danell, 1990).

In most plants, including legumes, sucrose is the primary carbon resource supplied to roots from the shoot (Giaquinta, 1983). Sucrose is unloaded via several Sucrose Transporters (SUTs), Sugars Will Eventually be Exported Transporters (SWEETs) and Monosaccharide Transporters (MSTs) where it is cleaved into fructose and glucose to supply arbuscular mycorrhizal fungi and nitrogen-fixing bacteroids (Figure 1.3; Doidy et al., 2012). The degradation of shoot-originated sucrose into UDP-glucose (UDP-Glc), glucose and fructose occur in legume nodules and a range of arbuscular mycorrhizal plant species (reviewed in Liu et al., 2018). In nodules, legumes provide bacteroids with dicarboxylates, particularly malate and succinate (Figure 1.3; Day and Copeland, 1991). Nitrogen fixed in bacteroids are exported as ammonia (NH_3) from the symbiosome into infected host cells, where it is assimilated into glutamine (Gln) and glutamate (Glu; Liu et al., 2018). Amino acids, aspartate (Asp) and asparagine (Asn), and/or ureides are then exported from legume, *Parasponia* and actinorhizal plant nodules to the rest of the plant (Figure 1.3; Liu et al., 2018).

1.4.3. Symbiotic establishment in arbuscular mycorrhizal, rhizobial and actinorhizal symbioses

Diverse types of intracellular symbiotic associations (arbuscular mycorrhizal, rhizobial, and actinorhizal associations) mostly rely on a conserved specialised pathway, now referred to as the "common symbiotic pathway" (CSP), supplemented with lineage- or species-specific changes (Oldroyd, 2013; Radhakrishnan et al., 2020; Delaux and Schornack, 2021). Adjustments to this conserved symbiotic signalling pathway refined an ancestral intracellular symbiotic state to recognise and accommodate other mycorrhizal and bacterial partners (Delaux and Schornack, 2021). For recent reviews on arbuscular mycorrhizal symbiosis see MacLean et al. (2017) and on nitrogen-fixing root-nodule symbiosis see Oldroyd et al. (2011) and Roy et al. (2020). Here we compare known intracellular symbioses, focusing on symbiotic establishment and functioning.

The development of symbiotic interactions is triggered in response to phosphate or nitrogen deprivation in plants and subsequent perception of symbionts in the rhizosphere, such as arbuscular mycorrhizal fungi, rhizobia or *Frankia* nitrogen-fixing bacteria (reviewed in MacLean et al., 2017; Roy et al., 2020). Consequently, microbes and their host plants have evolved a perception system that informs the plant of a potential symbiont in its vicinity and the microbe of a receptive host (reviewed in MacLean et al., 2017; Roy et al., 2020). Plant roots release strigolactones and flavonoids or isoflavonoids as a signal to arbuscular mycorrhizal fungi (Akiyama et al., 2005) and nitrogen-fixing bacteria (Peters et al., 1986; Redmond et al., 1986; Liu and Murray, 2016), respectively (Figure 1.4). Strigolactones induce a so-called presymbiotic stage in arbuscular mycorrhizal fungi characterised by profuse

branching and continued hyphal growth (Akiyama et al., 2005; Besserer et al., 2006). In turn, fungi produce lipo-chitooligosaccharides (Maillet et al., 2011) and short-chain chitin oligomers (Genre et al., 2013) collectively known as mycorrhizal factors (Myc factors) while rhizobia produce structurally similar nodulation factors (Nod factors; Dénarié et al., 1996). Although *Frankia* species produce symbiotic signals, they are structurally different to those produced by fungi and rhizobia (Granqvist et al., 2015; Chabaud et al., 2016). Despite differences in symbiotic signals, these compounds are recognised by plant lysine motif (LysM) receptor-like kinases, ultimately activating the CSP (Figure 1.4; MacLean et al., 2017; Roy et al., 2020).

Non-leguminous plant hosts, such as rice, tomato (arbuscular mycorrhizal hosts) and *Parasponia* (dual arbuscular mycorrhizal and rhizobia host), rely on a single LysM receptor-like kinase, encoded by *OsCERK1* (Miyata et al., 2014; Carotenuto et al., 2017), *SlLYK10* (Buendia et al., 2016) and *PanLYK3* (Op den Camp et al., 2011) respectively, to recognise symbiosis-associated microbial signatures. By contrast, in legume hosts, symbiotic signals perception is accomplished by LysM receptor-like kinases in a heterodimeric complex localised to the plasma membrane (Moling et al., 2014), namely Nod Factor Receptor (NFR1) and NFR5 in *L. japonicus* (Madsen et al., 2003; Radutoiu et al., 2003) and LysM Receptor Kinase (LYK3) and Nod Factor Perception (NFP) in *M. truncatula* (Limpens et al., 2003; Arrighi et al., 2006). An interacting co-receptor, known as the symbiotic receptor-like kinase (SYMRK; known as DMI2 in *M. truncatula*), a protein essential to the perception of intracellular symbioses signals is as such said to be "the entry point into the common symbiotic pathway" (Endre et al., 2002; Stracke et al., 2002). Perception of symbiotic signals, whether Myc factors, Nod factors or *Frankia* signals, results in transmission of this signal from the receptive plant host's plasma membrane to the nucleus where calcium spiking drives changes in gene expression associated with root invasion, as well as arbuscule or nodule development (Figure 1.4; Charpentier and Oldroyd, 2013).

The recognition of symbiotic signals either by one (*OsCERK1, SlLYK10* and *PanLYK3*) or two (LjNFR1/NFR5 and MtLYK3/NFP) LysM receptor-like kinases and its associated co-receptor SYMRK/DMI2 triggers nuclear calcium (Ca^{2+}) oscillations (Figure 1.4; Kevei et al., 2007; Venkateshwaran et al., 2015). Mevalonate, produced by an isoprenoid biosynthesis reductase (3-Hydroxy-3-Methylglutaryl CoA Reductase 1, HMGR1), was linked to the CSP when its interaction with the common symbiotic receptor-like kinase (SYMRK/DMI2) was discovered in *M. truncatula* (Kevei et al., 2007; Venkateshwaran et al., 2015). Mevalonate, or its phosphorylated metabolites, act as secondary messengers that link the perception of symbiotic factors from the plasma membrane with the generation of nuclear Ca^{2+} oscillations (Venkateshwaran et al., 2015). Various nuclear envelope proteins help generate calcium spiking signals including voltage-activated Cyclic Nucleotide Gated Channels (CNGC a/b/c; (Charpentier et al., 2016), calcium symporter(s) and counter-ion channels (Ané et al., 2004; Imaizumi-Anraku et al., 2005; Charpentier et al., 2008). Calcium channels (in this case CNGC15 a/b/c channels) mediate the influx of calcium from the endoplasmic reticulum – the symbiotic calcium store – into the nucleoplasm while calcium symporters (possibly MCA8) pump calcium back into the nuclear envelope (reviewed in Oldroyd, 2013). Nuclear

localised potassium (K^+)-permeable cation channels, CASTOR and POLLUX discovered in *L. japonicus* (Charpentier et al., 2008), counters the influx of calcium into the nucleoplasm via potassium (K^+) efflux thereby repolarising the nuclear membrane to enable sustained Ca^{2+} oscillations (Charpentier et al., 2016). Interestingly, a single nuclear localised potassium-permeable cation channel, DMI1, serves this same purpose in *M. truncatula* (Ané et al., 2004). In *L. japonicus* it was demonstrated that genes encoding proteins like nucleoporin subunits, Nucleoporin 85 (NUP85) and NUP133, act upstream of this Ca^{2+} spiking and might be facilitating the transport of CASTOR and POLLUX/DMI1 to the inner nuclear envelope (Kanamori et al., 2006; Saito et al., 2007).

The nuclear calcium and calmodulin-dependent kinase (CCaMK; known as DMI3 in *M. truncatula*) acts as a "master decoder and regulatory kinase", deciphering nuclear Ca^{2+} oscillations and marking the start of the symbiotic transcriptional response (Figure 1.4; Miller et al., 2013). Nuclear Ca^{2+} spiking induces binding of Ca^{2+}-calmodulin to CCaMK/DMI3, causing conformational changes that stimulate its association to and phosphorylation of CYCLOPS (known as IPD3 in *M. truncatula*; Yano et al., 2008; Horváth et al., 2011). CYCLOPS/IPD3 presumably interacts with CCaMK/DMI3 together with Gibberellic-Acid Insensitive, Repressor of Gai and Scarecrow (GRAS) domain regulatory proteins, DELLA, to initiate the transcriptional responses required to promote symbioses (Jin et al., 2016; Pimprikar et al., 2016). In *M. truncatula* and *Pisum sativum* (pea), DELLAs (MtDELLA1, MtDELLA2, MtDELLA3, PsLA and PsCRY) function redundantly to integrate symbiosis signalling with plant growth and development (Floss et al., 2013; Foo et al., 2013; Fonouni-Farde et al., 2016), while but one DELLA protein in rice (OsSLR1) does the same (Yu et al., 2014). Direct regulation of GRAS transcription factors discovered in *M. truncatula* – Required for Arbuscular Mycorrhization 1 (RAM1; Gobbato et al., 2012; Gobbato et al., 2013; Park et al., 2015; Rich et al., 2017), Nodulation Signalling Pathway (NSP1; Smit et al., 2005), NSP2 (Kaló et al., 2005), Required for Arbuscule Development 1 (RAD1; Xue et al., 2015), Mycorrhiza-Induced GRAS 1 (MIG1; Heck et al., 2016), ERF Required for Nodulation 1 (ERN1; Cerri et al., 2017; Kawaharada et al., 2017) and Nodule Inception (NIN; Schauser et al., 1999) – by CCaMK, CYCLOPS and DELLAs reflects their central role in arbuscule or nodule development. GRAS transcription factors, therefore, likely have interconnected roles in the symbiotic signalling pathway to either promote mycorrhiza- or rhizobia-specific responses (reviewed in Oldroyd, 2013).

1.5. Free-living cyanobacteria and their role in nitrogen-fixing symbioses

Amidst the biodiversity within the bacterial world, the photosynthetic cyanobacteria represent a unique and widely distributed monophyletic group (reviewed in Flores and Herrero, 2010). Since their origin between 3.7 and 2.4 billion years ago (bya), cyanobacteria have diversified into unicellular, multicellular, and filamentous forms with some cells being able to differentiate and carry out specialised functions (Flores and Herrero, 2010). They bear gram-negative cell walls comprised of a thickened peptidoglycan layer sandwiched between a plasma and outer membrane (Drews and Weckesser, 1982; Hoiczyk and Hansel, 2000). The outer membrane of filamentous

cyanobacteria is continuous so that it never enters adjacent cells' septum instead the plasma membrane and peptidoglycan layer surround individual cells (Flores et al., 2006; Schneider et al., 2007; Wilk et al., 2011). This means that the periplasmic space, which contains the peptidoglycan layer and can be found between the plasma and outer membranes, is also continuous (Flores et al., 2006). Cyanobacteria are defined by their extensive internal thylakoid membranes, which enable them to perform photosynthesis and in the absence of combined nitrogen are even capable of performing nitrogen fixation (reviewed in Flores and Herrero, 2010; Sciuto and Moro, 2015). Filamentous, nitrogen-fixing cyanobacteria confine photosynthesis to vegetative cells and N_2 fixation to metabolically specialised cells, known as the heterocyst (Gallon, 1981; Fay, 1992; Wolk et al., 1994).

Heterocysts devoted to N_2 fixation arise from semiregular vegetative cells along cyanobacterial filaments and create within them the microoxic environment required for N_2 fixation (Wolk et al., 1994; Wolk, 1999; Meeks et al., 2002). As heterocysts differentiate, their morphology changes to limit oxygen entry, among which the deposition of additional envelope layers external to the already existing cell wall is key (Wolk et al., 1994; Wolk, 1999; Meeks et al., 2002). This newly formed envelope consists of an external layer of polysaccharides (heterocyst envelope polysaccharides, Hep) and inner layers of glycolipids (heterocyst-specific glycolipids, Hgl), which together functions as a barrier to oxygen (Cardemil and Wolk, 1979, 1981; Nicolaisen et al., 2009). Mature heterocysts, in addition, have respiratory oxidases located at heterocyst poles and lack the O_2-evolving photosystem II and CO_2-fixing enzyme, all of which favours the establishment of the intracellular microoxic environment (Zhao and Wolk, 2008). Since heterocysts are unable to photosynthetically fix CO_2, a reciprocal source-sink relationship is established where vegetative cells are provided with fixed nitrogen, and in return, heterocysts receive carbon and reductants (Wolk, 1968; Cumino et al., 2007). Heterocyst development relies on the combined effects of external and internal signals, as well as temporal and spatial regulation of cellular processes and genes; a basic mechanism that is conserved in filamentous, heterocyst-forming cyanobacteria (reviewed in Flores and Herrero, 2010; Flores et al., 2019).

Heterocyst differentiation is triggered as a suite of nitrogen stress responses in clusters of cells (proheterocysts) and culminates in a single cell committed to nitrogen fixation (mature heterocysts; Flores and Herrero, 2010; Flores et al., 2019). In the absence of a fixed nitrogen source in the environment a Krebs cycle intermediate, 2-oxoglutarate (2-OG), a link between nitrogen and carbon balance, accumulates (Muro-Pastor et al., 2001; Laurent et al., 2005). The carbon-nitrogen imbalance caused by 2-oxoglutarate (2-OG) accumulation is perceived by NtcA, a global transcriptional regulator in the cyclic AMP receptor protein family, responsible for orchestrating nitrogen stress responses in all cyanobacteria (Frías et al., 1994; Wei et al., 1994; Herrero et al., 2004; Zhao et al., 2010). The DNA binding activity of NtcA to its target promoters is enhanced by 2-oxoglutarate and is therefore required for transcription of several nitrogen assimilation genes (Vázquez-Bermúdez et al., 2002; Olmedo-Verd et al., 2008; Valladares et al., 2008). Heterocyst differentiation, if nitrogen stress persists, is induced by the combined action of two regulators: NtcA and heterocyst differentiation control protein (HetR) linked by NrrA, a NtcA-regulated

response regulator (Buikema and Haselkorn, 1991; Ehira and Ohmori, 2006a, b). The two regulatory genes, induced in a mutually dependent manner, are also dependent on positive autoregulation to increase their own expression and the expression of each other in spatially distributed proheterocysts (Black et al., 1993; Muro-Pastor et al., 2002; Rajagopalan and Callahan, 2010). HetR (a serine-type protease with autoprotease and DNA binding activity) activates several heterocyst-related genes which play key roles in differentiation and pattern formation and is therefore called the "master regulator of heterocyst development" (Huang et al., 2004; Khudyakov and Golden, 2004; Higa and Callahan, 2010; Feldmann et al., 2011; Du et al., 2012). Heterocyst distribution along the filament relies on the inhibitory roles of two gene products, PatS (the heterocyst inhibition signalling pentapeptide, RGSGR) and HetN (a ketoacyl reductase) which prevent mature heterocysts from occupying adjacent positions (Callahan and Buikema, 2001; Yoon and Golden, 2001; Huang et al., 2004). Whereas PatS products act laterally along the filament to suppress HetR activity in *de novo* pattern formation, HetN products participate in heterocyst pattern maintenance once it has been established (Black et al., 1993; Yoon and Golden, 2001).

Heterocyst development ends when dinitrogen (N_2) is reduced to ammonia (during the process of nitrogen fixation) via the enzyme nitrogenase (Flores and Herrero, 1994; Herrero et al., 2001). The well-conserved enzyme, nitrogenase, consists of two components: dinitrogenase and dinitrogenase reductase (Rubio and Ludden, 2008). Dinitrogenase reductase transfers electrons to dinitrogenase from the heterocyst-specific ferredoxin (encoded by *FdxH*; (Masepohl et al., 1997; Bothe et al., 2010). In cyanobacteria, the nitrogen-fixation gene cluster (*nif*HDK) is responsible for encoding a molybdenum-containing nitrogenase enzyme complex (Herrero et al., 2001). During heterocyst differentiation, the *nif* gene cluster (*nif*HDK) undergoes major rearrangements to excise specific regions (11 kb in *nifD* gene and 55 kb in *fdxN* gene) thus producing the functional *nif*HDK operon and ultimately the nitrogenase enzyme (Golden et al., 1985; Brusca et al., 1989). Ammonium fixed from atmospheric nitrogen (N_2) is incorporated by glutamine synthetase to produce glutamine (Gln) which is, at least in part, transferred to vegetative cells (Wolk et al., 1976; Thomas et al., 1977; Martín-Figueroa et al., 2000). In addition to glutamine, cyanophycin degradation product, β-aspartyl arginine, is transported back into vegetative cells where it is hydrolysed into aspartate (Asp) and arginine (Arg; Burnat et al., 2014). Heterocyst metabolism is supported by sucrose, the reduced carbon form provided by the vegetative cells (López-Igual et al., 2010; Vargas et al., 2011).

1.5.1. Cyanobacteria in plant symbioses

Some cyanobacterial genera, while still able to thrive in a free-living form, also establish symbiosis with a wide range of eukaryotic hosts (Table 1.1; Rai et al., 2000; Rai et al., 2002; Adams et al., 2006; Rai et al., 2019). In embryophytes, instances of cyanobacterial symbioses occur in all extant lineages barring lycophytes; in bryophytes (liverworts and hornworts; Meeks, 1990a), ferns in the aquatic genus *Azolla* (Braun-Howland and Nierzwicki-Bauer, 1990), gymnosperms (exclusively in and in all species of the cycadopsida; Lindblad and Bergman, 1990) and the angiosperm genus *Gunnera* (Bonnett, 1990). With the exception of the water fern genus *Azolla* (dated 90 million years ago;(Metzgar et al., 2007) and the angiosperm genus *Gunnera* (dated 115 mya;

Vekemans et al., 2012), cycads (265-290 mya; Brenner et al., 2003; Condamine et al., 2015) and bryophytes, in particular, represent ancient lineages (470 mya; Edwards et al., 2014; Laenen et al., 2014; Brown et al., 2015) suggesting that these partnerships are ancient (~500 million years old) as well. Despite evidence of nitrogen-fixing cyanobacteria in symbiosis with lichens existing by the Neoproterozoic (Raven, 2002; Yuan et al., 2005), the origin and evolution of land plant-cyanobacterial partnerships remains an enigma.

The "earliest direct fossil evidence of an intracellular cyanobacterial symbiosis" in land plants was discovered when filamentous cyanobacteria were observed in *Aglaophyton major* prostrate axes (Krings et al., 2009), an early non-vascular land plant (dated to ~400 mya) commonly found in the Rhynie chert flora (Remy and Hass, 1996). The *A. major* cyanobacterial endophyte – non-heterocystous cyanobacteria in the order Oscillatoriales – enters prostrate axes via stomata, colonising chambers and intercellular spaces of the outer cortex, and ultimately extends through the well-defined mycorrhizal arbuscule-zone penetrating cells harbouring arbuscules (for details on the *A. major* mycorrhiza, see Taylor et al. (1995). Krings et al. (2009) interpreted this as an "on-again–off-again association" linked only by habitat flooding which brought them into close contact with cyanobacteria. While this on-again–off-again association is unlikely to be an ancestor to present-day mutualistic land plant-cyanobacterial partnerships, the physical association of *A. major* and Oscillatoriales cyanobacteria highlights how such relationships might have developed. We know now that all plant-cyanobacterial symbioses are established by filamentous heterocystous cyanobacteria mainly found in the genus *Nostoc* (reviewed in Rai et al., 2000; Meeks et al., 2002; Meeks and Elhai, 2002; Rai et al., 2002; Adams et al., 2006; Rai et al., 2019). Recently, Warshan et al. (2018) dated the ancestors of today's symbiotic *Nostoc* spp. to the Mesoproterozoic-Neoproterozoic era, which predates not only land plants but their hosts as well. Their results together with a metagenomic study by Knack et al. (2015) hypothesises an origin for symbiotic *Nostoc* spp. with streptophyte algae predating embryophytes. The *N. punctiforme* and glomeromycotinan fungus, *Geosiphon pyriformis*, intracellular symbiosis (once believed to be a widespread fungal-cyanobacterial symbiosis) also predates present-day land plant-cyanobacterial partnerships (Schüßler et al., 1994; Mollenhauer et al., 1996; Schüßler, 2012). Thus, Warshan et al. (2018) proposes that the last common ancestor of symbiotic *Nostoc* spp. entered into streptophyte algae-intracellular symbioses, and subsequently with early land plants and fungi as well, but that this ability was lost barring *Geosiphon* and *Gunnera Nostoc* symbioses. The prevailing hypothesis is that land plant-cyanobacterial association evolved independently in the plant lineages as they emerged following the origin of symbiotic competent *Nostoc* species (Raven, 2002).

Filamentous heterocystous cyanobacteria – mainly members of the order Nostocales (*Anabaena*, *Calothrix*, *Nostoc*, and *Scytone*) – are in themselves versatile (Rai et al., 2000; Meeks et al., 2002; Meeks and Elhai, 2002; Rai et al., 2002; Adams et al., 2006; Rai et al., 2019). These microbes and their host plants represent diverse symbiotic associations that lead to highly successful, long-lived, and persistent symbioses (Rai et al., 2000; Rai et al., 2002; Rai et al., 2019), compartmentalised inter- or intracellularly in specialised host organs such as cavities in thalli of liverworts and hornworts, root nodules of cycads, and stem glands of *Gunnera* spp. (Table 1.1).

Irrespective of the plant species or mode of transmission (i.e. horizontal or vertical), the plant structures colonised develop independent of cyanobacterial infection (Rai et al., 2000; Rai et al., 2002; Rai et al., 2019). Plant's attraction, internalisation of symbiotic cyanobacteria and subsequent regulation of heterocyst differentiation is key to establishing a successful interaction (Gorelova, 2006). The plant-controlled conversion of free-living vegetative filaments into motile, short-lived hormogonia (transient, non-growing, short gliding filaments) pre-colonisation, and their subsequent differentiation into nitrogen-fixing heterocysts once inside the host plant cell/tissue are essential steps in the symbiotic process (reviewed in Santi et al., 2013). Plants exude several hormogonia-inducing factors (HIF), contained within the mucus, during nitrogen stress to stimulate hormogonia induction in nearby *Nostoc* species (most likely *Nostoc punctiforme*; Rai et al., 2000; Rai et al., 2002; Rai et al., 2019). Hormogonia covered with pili break off at vegetative cell-heterocyst connections and move towards the symbiotic plant by a gliding mechanism (Duggan et al., 2007). Hormogonia-inducing factors (thought to be phenolic compounds) are produced by the hornwort *Anthoceros punctatus* (Campbell and Meeks, 1989; Meeks, 2003), the liverwort *Blasia pusilla* (Meeks, 2006), cycad coralloid root (Ow et al., 1999) and the stem glands of *Gunnera* species (Johansson and Bergman, 1992; Rasmussen et al., 1994). Although the identity of hormogonia-inducing factors remains mostly unknown, a recent study by Hashidoko et al. (2019) named diacylglycerols, in particular 1-palmitoyl-2-linoleoyl-*sn*-glycerol, as a potential hormogonia-inducing factor (HIF) in cycad *Cycas revoluta*. Although hormogonia formation is a prerequisite for symbiosis, it is not sufficient for its establishment, because, unsurprisingly, hormogonia-inducing factors are also produced in the mucus of non-host plants such as wheat and rice in artificial cyanobacterial associations (Gantar et al., 1993; Gusev et al., 2002; Nilsson et al., 2005; Álvarez et al., 2021). Migration of hormogonia into preexisting symbiotic structures is crucial for the establishment of these nitrogen-fixing symbioses, and chemoattraction (such as by soluble sugars though to be present in this mucus as well) is an important step in this process (Knight and Adams, 1996; Nilsson et al., 2006). Once colonised, hormogonia turn back into non-motile vegetative filaments and the plant cavities close off from the external environment (Rai et al., 2000; Rai et al., 2002; Rai et al., 2019).

Available literature has established that *Nostoc* cyanobionts are housed within preexisting plant cavities all of which, even if their development is not completely understood, contains a secondary metabolite (most likely HIFs) and polysaccharide-rich mucilage (Rai et al., 2000; Rai et al., 2002; Rai et al., 2019). In cycads, hormogonia enter precoralloid roots (adventitious, apogeotropic roots with a swollen appearance) at any stage of maturation (Milindasuta, 1975). Entry occurs via openings on the root surface either through local injuries, lenticles or breaks in the dermal layer (hallmark signs of precoralloid root maturation; see Nathanielsz and Staff (1975). A switch to geotrophic growth occurs following cyanobacterial colonisation and repeated dichotomy gives rise to the branched appearance characteristic of coralloid roots (reviewed in Vessey et al., 2005). Cyanobacteria spreads from the root surface via intercellular spaces, and ultimately becomes confined to a well-defined region of the cortex, called the cyanobacterial zone (Milindasuta, 1975; Ahern and Staff, 1994). Instead in *Gunnera*, cyanobacteria invade intercellular channels which extend from a gland (a hemispherical protuberance) at the base of each petiole into

stem tissue becoming intracellular within them (reviewed in Bergman et al., 1992a; Bergman, 2002; Bergman and Osborne, 2002). *Nostoc* enters between cells (possibly those lining the channels) after which cyanobacteria become engulfed by stem cells (Silvester and McNamara, 1976; Bonnett and Silvester, 1981). The stem cells resume meristematic activity after colonisation and give rise to clusters of stem gland cells (Silvester and McNamara, 1976; Bonnett and Silvester, 1981). Once inside these preexisting cavities, entry of which occurs through stomata-like (i.e. mucilage clefts) or cavity pores in bryophytes and *Azolla* respectively (Adams, 2002), hosts produced chemical signals promote the heterocyst differentiation in cyanobacteria, the specialised bacterial cells devoted to fixing dinitrogen (Adams et al., 2006).

The repression of hormogonia and subsequent differentiation of heterocysts is associated with the release of an as yet unknown hormogonia-repressing factor (HRF) by the host (Rai et al., 2000; Rai et al., 2002; Rai et al., 2019). In free-living cyanobacteria, flavonoids (naringenin and naringin) significantly affect cyanobacterial growth (Żyszka et al., 2017) and the expression of a heterocyst-related gene highlighting their role as hormogonia-repressing factors (Cohen and Yamasaki, 2000). Recently, the genomes of three hornwort species (*Anthoceros agrestis*, *An. punctatus*, and *An. angustus*; (Li et al., 2020) and the fern *Azolla filiculoides* (Li et al., 2018) were sequenced and, in all species, flavonoid biosynthesis genes were found but their involvement in heterocyst development from hormogonia still needs to be elucidated. (Delaux and Schornack, 2021) hypothesises that flavonoids could have been independently recruited for bacterial symbioses, independently in cyanobacterial and nodule symbioses. Heterocyst differentiation is enhanced in cyanobionts when compared to free-living cyanobacteria (5-10%), with a frequency of 35% in plant associations, resulting in elevated rates of nitrogen fixation and ammonium release in all known associations barring in cycads (Table 1.2; Rai et al., 2000; Rai et al., 2002; Rai et al., 2019). Cyanobacteria assimilate dinitrogen, nitrate, nitrite, urea or organic nitrogen into ammonium via glutamine synthetase and glutamate synthase (Flores and Herrero, 1994). Thus, ammonium fixed from dinitrogen is only released into symbiotic cavities after it has been transferred to vegetative cells (Meeks, 2009). In the cycad, nitrogen is hypothesised to instead be released as organic nitrogen (glutamine, glutamate and citrulline; Pate et al., 1988; Lindblad and Bergman, 1990). A *Gunnera-Nostoc* association study revealed that ammonia is assimilated, outside the heterocyst cells, into glutamine and glutamate and finally exported as asparagine (Silvester et al., 1996) indicating that organic nitrogen may be produced by plant cells in this and other cyanobacterial associations. In contrast to free-living cyanobacteria, nitrogen stress signals do not regulate heterocyst differentiation in cyanobionts, since vegetative cells do not show signs of nitrogen stress, and differentiation into additional heterocysts continues despite cyanobionts being immersed in a pool of N_2-derived ammonium (Meeks et al., 1983; Meeks, 2006). Symbiotic plants, therefore, produce unknown signals that replace nitrogen stress signals to increase heterocyst frequency beyond free-living levels, resulting in multiple singular heterocysts with shortened vegetative cell intervals or multiple contiguous heterocysts (Meeks, 2009).

1.6. Conclusion

The origin and evolution of plant-microbe symbioses was a major evolutionary innovation linked to the acquisition of nutrients in nutrient-poor environments. Since then plant life and symbionts have diversified together which led to the development of specialised nutrient strategies and the eventual domination of angiosperms (reviewed in Brundrett, 2002; Brundrett and Tedersoo, 2018). While much work has been done in studying aspects of symbiosis establishment and development and the plant's role therein, the underlying processes and components elucidated thus far has largely focused on angiosperms lineages (reviewed in MacLean et al., 2017; Radhakrishnan et al., 2020; Roy et al., 2020). However, ~200 million years ago, the terrestrial landscape was dominated by early gymnosperms, ferns and lycophytes (Chamberlain, 1957; Kramer and Green, 1990; Taylor et al., 2009). Symbiosis with nitrogen-fixing cyanobacteria, a nitrogen-fixing symbiosis that has received far less attention, established symbiosis with various terrestrial plants such as hornworts, liverworts, ferns, and cycads (reviewed in Rai et al., 2000; Rai et al., 2002; Rai et al., 2019). To date, the majority of published literature on cyanobacteria has focused on these prokaryotes as free-living entities, while insightful, far less progress has been made on the identification of factors and genes regulating plant-cyanobacterial symbioses which still relies on dated literature.

While the identification of the common symbiotic pathway (CSP) in legumes has given us valuable insight into the molecular mechanisms governing arbuscular mycorrhizal and root nodule symbioses (reviewed in Oldroyd, 2013; MacLean et al., 2017; Roy et al., 2020), the genetic mechanism underlying cyanobacterial symbioses have remained largely undetermined. An interesting consideration with regards to the establishment and regulation of plant-cyanobacterial symbioses is, therefore, the role common symbiotic pathway genes have (if any) in facilitating these symbioses. Since early land plants had already evolved complex mutualistic associations (arbuscular mycorrhizal and/or cyanobacterial) by the Early Devonian which, in the case of mycorrhiza have remained largely conserved for the last 450 million years, suggests that cyanobacterial partnerships could have recruited genes and pathways utilised by the CSP, mycorrhizal or bacterial interactions or combinations thereof for their establishment and subsequent functioning. Investigations into the interactions of non-model species across the diversity of plants (diverse algae, bryophytes, ferns, and the so far under-sampled diversity of seed plants) will help reconstruct the "common toolkit" applied across terrestrial plants for symbiosis, and enable us to decode the molecular mechanisms that have led to the emergence of new interactions, signalling pathways and enzymes that regulate the scope of these types of interactions occurring in plants. The availability of more and high quality genomic and transcriptomic data from ever-increasing plant lineages, particularly the underrepresented lineages, will undoubtedly aid in a better and more holistic understanding of the numerous aspects of plant-cyanobacterial symbioses and their evolution.

1.7. References

Adams DG (2002) Cyanobacteria in symbiosis with hornworts and liverworts. In Cyanobacteria in Symbiosis, AN Rai, B Bergman, and U Rasmussen, eds (Dordrecht: Springer Netherlands), pp. 117-135.

Adams DG, Duggan PS, Jackson O (2012) Cyanobacterial symbioses. In Ecology of cyanobacteria II: their diversity in space and time, B.A. Whitton, ed (Dordrecht: Springer Netherlands), pp. 593-647.

Adams DG, Bergman B, Nierzwicki-Bauer S, Rai, AN, Schüßler A (2006) Cyanobacterial-plant symbioses. In The prokaryotes: a handbook on the biology of bacteria, M Dworkin, S Falkow, E Rosenberg, K Schleifer, and E Stackebrandt, eds (New York: Springer), pp. 331-363.

Ahern CP, Staff IA (1994) Symbiosis in cycads: the origin and development of coralloid roots in *Macrozamia communis* (Cycadaceae). Am J Bot **81**: 1559-1570.

Akiyama K, Matsuzaki KI, Hayashi H (2005) Plant sesquiterpenes induce hyphal branching in arbuscular mycorrhizal fungi. Nature **435**: 824-827.

Álvarez C, Brenes-Álvarez M, Molina-Heredia FP, Mariscal V (2021) Early events of the endophytic symbiotic between *Oryza sativa* and *Nostoc punctiforme* involve the SYM pathway. bioRxiv: 2021.2004.2009.436957.

Ané JM, Kiss GB, Riely BK, Penmetsa RV, Oldroyd GED et al. (2004) *Medicago truncatula DMI1* required for bacterial and fungal symbioses in legumes. Science **303**: 1364-1367.

Arrighi JF, Barre A, Ben Amor B, Bersoult A, Soriano LC et al. (2006) The *Medicago truncatula* lysine motif-receptor-like kinase gene family includes *NFP* and new nodule-expressed genes. Plant Physiol **142**: 265-279.

Bago B, Pfeffer PE, Douds DD, Brouillette J Jr., Bécard G et al. (1999) Carbon metabolism in spores of the arbuscular mycorrhizal fungus *Glomus intraradices* as revealed by nuclear magnetic resonance spectroscopy. Plant Physiol **121**: 263-272.

Bailey KJ, Leegood RC (2016) Nitrogen recycling from the xylem in rice leaves: dependence upon metabolism and associated changes in xylem hydraulics. J Exp Bot **67**: 2901-2911.

Balestrini R, Bonfante P (2005) The interface compartment in arbuscular mycorrhizae: a special type of plant cell wall? Plant Biosyst **139**: 8-15.

Balestrini R, Bonfante P (2014) Cell wall remodeling in mycorrhizal symbiosis: a way towards biotrophism. Front Plant Sci **5**: 237-237.

Baslam M, Mitsui T, Sueyoshi K, Ohyama T (2021) Recent advances in carbon and nitrogen metabolism in C3 plants. Int J Mol Sci **22**: 318.

Battenberg K, Potter D, Tabuloc CA, Chiu JC, Berry AM (2018) Comparative transcriptomic analysis of two actinorhizal plants and the legume *Medicago truncatula* supports the homology of root nodule symbioses and is congruent with a two-step process of evolution in the nitrogen-fixing clade of angiosperms. Front Plant Sci **9**: 1256.

Becking J (1992) The *Rhizobium* symbiosis of the nonlegume *Parasponia*. In Biological nitrogen fixation: achievements and objectives (New York: Chapman and Hall), pp. 497-559.

Bellegarde F, Gojon A, Martin A (2017) Signals and players in the transcriptional regulation of root responses by local and systemic N signaling in *Arabidopsis thaliana*. J Exp Bot **68**: 2553-2565.

Bender GL, Nayudu M, Goydych W, Rolfe BG (1987) Early infection events in the nodulation of the non-legume *Parasponia andersonii* by *Bradyrhizobium*. Plant Sci **51**: 285-293.

Bergman B (2002) The *Nostoc-Gunnera* symbiosis. In Cyanobacteria in Symbiosis, AN Rai, B Bergman, and U. Rasmussen, eds (Dordrecht: Springer Netherlands), pp. 207-232.

Bergman B, Osborne B (2002) The Gunnera: Nostoc symbiosis. Biol Environ Proc R Ir Acad **102**: 35-39.

Bergman B, Johansson C, Söderbäck E (1992a) The *Nostoc–Gunnera* symbiosis. New Phytol **122**: 379-400.

Bergman B, Rai AN, Johansson C, Söderbäck E (1992b) Cyanobacterial-plant symbioses. Symbiosis **14**: 61-81.

Besserer A, Puech-Pagès V, Kiefer P, Gomez-Roldan V, Jauneau A et al. (2006) Strigolactones stimulate arbuscular mycorrhizal fungi by activating mitochondria. PLoS Biol **4**: e226.

Black TA, Cai Y, Wolk CP (1993) Spatial expression and autoregulation of hetR, a gene involved in the control of heterocyst development in *Anabaena*. Mol Microbiol **9**: 77-84.

Bonnett HT (1990) The *Nostoc-Gunnera* association. In CRC handbook of symbiotic cyanobacteria, AN Rai, ed (Boca Raton, Florida: CRC Press), pp. 137-159.

Bonnett HT, Silvester WB (1981) Specificity in the *Gunnera-Nostoc* endosymbiosis. New Phytol **89**: 121-128.

Bothe H, Schmitz O, Yates MG, Newton WE (2010) Nitrogen fixation and hydrogen metabolism in cyanobacteria. Microbiol Mol Biol Rev **74**: 529-551.

Bozsoki Z, Gysel K, Hansen SB, Lironi D, Krönauer C et al. (2020). Ligand-recognizing motifs in plant LysM receptors are major determinants of specificity. Science **369**: 663-670.

Braun-Howland EB, Nierzwicki-Bauer SA (1990) *Azolla-Anabaena* symbiosis: biochemistry, physiology, ultrastructure, and molecular biology. In CRC Handbook of Symbiotic Cyanobacteria, AN Rai, ed (Boca Raton, Florida: CRC Press), pp. 65-117.

Bravo A, York T, Pumplin N, Mueller LA, Harrison MJ (2016) Genes conserved for arbuscular mycorrhizal symbiosis identified through phylogenomics. Nat Plants **2**: 15208.

Bravo A, Brands M, Wewer V, Dörmann P, Harrison MJ (2017) Arbuscular mycorrhiza-specific enzymes FatM and RAM2 fine-tune lipid biosynthesis to promote development of arbuscular mycorrhiza. New Phytol **214**: 1631-1645.

Brenner ED, Stevenson DW, Twigg RW (2003) Cycads: evolutionary innovations and the role of plant-derived neurotoxins. Trends Plant Sci **8**: 446-452.

Bronstein JL (2015) The study of mutualism. In Mutualism, JL Bronstein, ed (Oxford University Press, USA).

Brown RC, Lemmon BE, Shimamura M, Villarreal JC, Renzaglia KS (2015) Spores of relictual bryophytes: diverse adaptations to life on land. Rev Palaeobot Palynol **216**: 1-17.

Brundrett MC (2002) Coevolution of roots and mycorrhizas of land plants. New Phytol **154**: 275-304.

Brundrett MC, Tedersoo L (2018) Evolutionary history of mycorrhizal symbioses and global host plant diversity. New Phytol **220**: 1108-1115.

Brusca JS, Hale MA, Carrasco CD, Golden JW (1989) Excision of an 11-kilobase-pair DNA element from within the *nifD* gene in *Anabaena variabilis* heterocysts. J Bacteriol **171**: 4138-4145.

Buendia L, Wang T, Girardin A, Lefebvre B (2016) The LysM receptor-like kinase SlLYK10 regulates the arbuscular mycorrhizal symbiosis in tomato. New Phytol **210**: 184-195.

Buikema WJ, Haselkorn R (1991) Characterization of a gene controlling heterocyst differentiation in the cyanobacterium *Anabaena* 7120. Genes Dev **5**: 321-330.

Burgess D, Peterson R (1987) Development of *Alnus japonica* root nodules after inoculation with *Frankia* strain HFPArI3. Can J Bot **65**: 1647-1657.

Burja AM, Hill RT (2001) Microbial symbionts of the Australian Great Barrier Reef sponge, *Candidaspongia flabellata*. Hydrobiologia **461**: 41-47.

Burnat M, Herrero A, Flores E (2014) Compartmentalized cyanophycin metabolism in the diazotrophic filaments of a heterocyst-forming cyanobacterium. Proc Natl Acad Sci USA **111**: 3823-3828.

Calabrese S, Pérez-Tienda J, Ellerbeck M, Arnould C, Chatagnier O et al. (2016). GintAMT3 – a low-affinity ammonium transporter of the arbuscular mycorrhizal *Rhizophagus irregularis*. Front Plant Sci **7**: 679.

Callaham D, Torrey JG (1977) Prenodule formation and primary nodule development in roots of *Comptonia* (Myricaceae). Can J Bot **55**: 2306-2318.

Callahan SM, Buikema WJ (2001) The role of HetN in maintenance of the heterocyst pattern in *Anabaena* sp. PCC 7120. Mol Microbiol **40**: 941-950.

Campbell EL, Meeks JC (1989) Characteristics of hormogonia formation by symbiotic *Nostoc* spp. in response to the presence of *Anthoceros punctatus* or its extracellular products. Appl Environ Microbiol **55**: 125-131.

Cardemil L, Wolk CP (1979) The polysaccharides from heterocyst and spore envelopes of a blue-green alga. Structure of the basic repeating unit. J Biol Chem **254**: 736-741.

Cardemil L, Wolk CP (1981) Isolated heterocysts of *Anabaena variabilis* synthesize envelope polysaccharide. Biochim Biophys Acta **674**: 265-276.

Carotenuto G, Chabaud M, Miyata K, Capozzi M, Takeda N et al. (2017) The rice LysM receptor-like kinase OsCERK1 is required for the perception of short-chain chitin oligomers in arbuscular mycorrhizal signaling. New Phytol **214**: 1440-1446.

Carpenter EJ, Foster RA (2002) Marine cyanobacterial symbioses. In Cyanobacteria in Symbiosis, AN Rai, B Bergman, and U Rasmussen, eds (Dordrecht: Springer Netherlands), pp. 11-17.

Cerri MR, Wang Q, Stolz P, Folgmann J, Frances L et al. (2017) The *ERN1* transcription factor gene is a target of the CCaMK/CYCLOPS complex and controls rhizobial infection in *Lotus japonicus*. New Phytol **215,** 323-337.

Chabaud M, Gherbi H, Pirolles E, Vaissayre V, Fournier J et al. (2016). Chitinase-resistant hydrophilic symbiotic factors secreted by *Frankia* activate both Ca^{2+} spiking and *NIN* gene expression in the actinorhizal plant *Casuarina glauca*. New Phytol **209**: 86-93.

Chamberlain CJ (1957) Gymnosperms: structure and evolution. (Johnson Reprint).

Charpentier M, Oldroyd GE (2013) Nuclear calcium signaling in plants. Plant Physiol **163**: 496-503.

Charpentier M, Bredemeier R, Wanner G, Takeda N, Schleiff E et al. (2008) *Lotus japonicus* CASTOR and POLLUX are ion channels essential for perinuclear calcium spiking in legume root endosymbiosis. Plant Cell **20**: 3467-3479.

Charpentier M, Sun J, Martins TV, Radhakrishnan GV, Findlay K et al. (2016). Nuclear-localized cyclic nucleotide–gated channels mediate symbiotic calcium oscillations. Science **352**: 1102-1105.

Cheng Q (2008) Perspectives in biological nitrogen fixation research. J Integr Plant Biol **50**: 786-798.

Cohen MF, Yamasaki H (2000) Flavonoid-induced expression of a symbiosis-related gene in the cyanobacterium *Nostoc punctiforme*. J Bacteriol **182**: 4644-4646.

Condamine FL, Nagalingum NS, Marshall CR, Morlon H (2015) Origin and diversification of living cycads: a cautionary tale on the impact of the branching process prior in Bayesian molecular dating. BMC Evol Biol **15**: 65.

Crespi M, Frugier F (2008) *De novo* organ formation from differentiated cells: root nodule organogenesis. Sci Signal **2**: re11.

Cumino A, Marcozzi C, Barreiro R, Salerno G (2007) Carbon cycling in *Anabaena* sp. PCC 7120. sucrose synthesis in the heterocysts and possible role in nitrogen fixation. Plant Physiol **143**: 1385-1397.

Day DA, Copeland L (1991) Carbon metabolism and compartmentation in nitrogen-fixing legume nodules. Plant Physiol Biochem **29**: 185-201.

Delaux PM, Schornack S (2021) Plant evolution driven by interactions with symbiotic and pathogenic microbes. Science **371**: eaba6605.

Delaux PM, Varala K, Edger PP, Coruzzi GM, Pires JC et al. (2014). Comparative phylogenomics uncovers the impact of symbiotic associations on host genome evolution. PLoS Genet **10**: e1004487.

Delaux PM, Radhakrishnan GV, Jayaraman D, Cheema J, Malbreil M et al. (2015). Algal ancestor of land plants was preadapted for symbiosis. Proc Natl Acad Sci USA **112**: 13390.

Delrot S, Rochat C, Tegeder M, Frommer W (2001) Amino acid transport. In Plant nitrogen (Springer), pp. 213-235.

Dénarié J, Debellé F, Promé JC (1996) Rhizobium lipo-chitooligosaccharide nodulation factors: signaling molecules mediating recognition and morphogenesis. Annu Rev Biochem **65**: 503-535.

Doidy J, van Tuinen D, Lamotte O, Corneillat M, Alcaraz G, Wipf D (2012) The *Medicago truncatula* sucrose transporter family: characterization and implication of key members in carbon partitioning towards arbuscular mycorrhizal fungi. Mol Plant **5**: 1346-1358.

Doyle JJ (1994) Phylogeny of the legume family: an approach to understanding the origins of nodulation. Annu Rev Ecol Syst **25**: 325-349.

Doyle JJ (2011) Phylogenetic perspectives on the origins of nodulation. Mol Plant-Microbe Interact **24**: 1289-1295.

Doyle JJ (2016) Chasing unicorns: nodulation origins and the paradox of novelty. Am J Bot **103**: 1865-1868.

Drew MC (1975) Comparison of the effects of a localised supply of phosphate, nitrate, ammonium and potassium on the growth of the seminal root system, and the shoot, in barley. New Phytol **75**: 479-490.

Drews G, Weckesser J (1982) Function, structure and composition of cell walls and external layers. In The Biology of Cyanobacteria, NG Carr and BA Whitton, eds (Oxford: Blackwell Scientific Publications), pp. 333-357.

Du Y, Cai Y, Hou S, Xu X (2012) Identification of the HetR recognition sequence upstream of *hetZ* in *Anabaena* sp. strain PCC 7120. J Bacteriol **194**: 2297-2306.

Duggan PS, Gottardello P, Adams DG (2007) Molecular analysis of genes in *Nostoc punctiforme* involved in pilus biogenesis and plant infection. J Bacteriol **189**: 4547-4551.

Edwards D, Morris JL, Richardson JB, Kenrick P (2014) Cryptospores and cryptophytes reveal hidden diversity in early land floras. New Phytol **202**: 50-78.

Ehira S, Ohmori M (2006a) NrrA, a nitrogen-responsive response regulator facilitates heterocyst development in the cyanobacterium *Anabaena* sp. strain PCC 7120. Mol Microbiol **59**: 1692-1703.

Ehira S, Ohmori M (2006b) NrrA directly regulates expression of *hetR* during heterocyst differentiation in the cyanobacterium *Anabaena* sp. strain PCC 7120. J Bacteriol **188**: 8520-8525.

Endre G, Kereszt A, Kevei Z, Mihacea S, Kaló P et al. (2002) A receptor kinase gene regulating symbiotic nodule development. Nature **417**: 962-966.

Erwin PM, Thacker RW (2008) Phototrophic nutrition and symbiont diversity of two Caribbean sponge–cyanobacteria symbioses. Mar Ecol Prog Ser **362**: 139-147.

Fay P (1992) Oxygen relations of nitrogen fixation in cyanobacteria. Microbiol Rev **56**: 340-373.

Feldmann EA, Ni S, Sahu ID, Mishler CH, Risser DD et al. (2011) Evidence for direct binding between HetR from *Anabaena* sp. PCC 7120 and PatS-5. Biochemistry **50**: 9212-9224.

Fine M, Loya Y (2002) Endolithic algae: an alternative source of photoassimilates during coral bleaching. Proc R Soc B: Biol Sci **269**: 1205-1210.

Flores E, Herrero A (1994) Assimilatory nitrogen metabolism and its regulation. In The Molecular Biology of Cyanobacteria, DA Bryant, ed (Dordrecht: Springer Netherlands), pp. 487-517.

Flores E, Herrero A (2010) Compartmentalized function through cell differentiation in filamentous cyanobacteria. Nat Rev Microbiol **8**: 39-50.

Flores E, Herrero A, Wolk CP, Maldener I (2006) Is the periplasm continuous in filamentous multicellular cyanobacteria? Trends Microbiol **14**: 439-443.

Flores E, Picossi S, Valladares A, Herrero A (2019) Transcriptional regulation of development in heterocyst-forming cyanobacteria. Biochim Biophys Acta **1862**: 673-684.

Floss DS, Levy JG, Lévesque-Tremblay V, Pumplin N, Harrison MJ (2013) DELLA proteins regulate arbuscule formation in arbuscular mycorrhizal symbiosis. Proc Natl Acad Sci USA **110**: 5025-5034.

Fonouni-Farde C, Tan S, Baudin M, Brault M, Wen J et al. (2016) DELLA-mediated gibberellin signalling regulates Nod factor signalling and rhizobial infection. Nat Commun **7**: 12636.

Foo E, Ross JJ, Jones WT, Reid JB (2013) Plant hormones in arbuscular mycorrhizal symbioses: an emerging role for gibberellins. Ann Bot **111**: 769-779.

Försterra G, Häussermann V (2008) Unusual symbiotic relationships between microendolithic phototrophic organisms and azooxanthellate cold-water corals from Chilean fjords. Mar Ecol Prog Ser **370**: 121-125.

Foster RA, Collier JL, Carpenter EJ (2006a) Reverse transcription PCR amplification of cyanobacterial symbiont 16S rRNA sequences from single non-photosynthetic eukaryotic marine planktonic host cells. J Phycol **42**: 243-250.

Foster RA, Carpenter EJ, Bergman B (2006b) Unicellular cyanobionts in open ocean dinoflagellates, radiolarians, and tintinnids: ultrastructural characterization and immuno-localization of phycoerythrin and nitrogenase. J Phycol **42**: 453-463.

Foster RA, Subramaniam A, Zehr JP (2009) Distribution and activity of diazotrophs in the Eastern Equatorial Atlantic. Environ Microbiol **11**: 741-750.

Frías JE, Flores E, Herrero A (1994) Requirement of the regulatory protein NtcA for the expression of nitrogen assimilation and heterocyst development genes in the cyanobacterium *Anabaena* sp. PCC7120. Mol Microbiol **14**: 823-832.

Galangau F, Daniel-Vedele FO, Moureaux TRS, Dorbe MF, Leydecker MTRS et al. (1988) Expression of leaf nitrate reductase genes from tomato and tobacco in relation to light-dark regimes and nitrate supply. Plant Physiol **88**: 383-388.

Gallon JR (1981) The oxygen sensitivity of nitrogenase: a problem for biochemists and micro-organisms. Trends Biochem Sci **6**: 19-23.

Gansel X, Muños S, Tillard P, Gojon A (2001) Differential regulation of the NO_3^- and NH_4^+ transporter genes *AtNrt2.1* and *AtAmt1.1* in *Arabidopsis*: relation with long-distance and local controls by N status of the plant. Plant J **26**: 143-155.

Gantar M, Kerby NW, Rowell P (1993) Colonization of wheat (*Triticum vulgare* L.) by N_2-fixing cyanobacteria: III. The role of a hormogonia-promoting factor. New Phytol **124**: 505-513.

Genre A, Lanfranco L, Perotto S, Bonfante P (2020) Unique and common traits in mycorrhizal symbioses. Nat Rev Microbiol **18**: 649-660.

Genre A, Chabaud M, Timmers T, Bonfante P, Barker DG (2005) Arbuscular mycorrhizal fungi elicit a novel intracellular apparatus in *Medicago truncatula* root epidermal cells before infection. Plant Cell **17**: 3489-3499.

Genre A, Chabaud M, Faccio A, Barker DG, Bonfante P (2008) Prepenetration apparatus assembly precedes and predicts the colonization patterns of arbuscular mycorrhizal fungi within the root cortex of both *Medicago truncatula* and *Daucus carota*. Plant Cell **20**: 1407-1420.

Genre A, Chabaud M, Balzergue C, Puech-Pagès, V, Novero, M et al. (2013) Short-chain chitin oligomers from arbuscular mycorrhizal fungi trigger nuclear Ca^{2+} spiking in *Medicago truncatula* roots and their production is enhanced by strigolactone. New Phytol **198**: 190-202.

Giaquinta RT (1983) Phloem loading of sucrose. Annu Rev Plant Physiol **34**: 347-387.

Gifford ML, Dean A, Gutierrez RA, Coruzzi GM, Birnbaum KD (2008) Cell-specific nitrogen responses mediate developmental plasticity. Proc Natl Acad Sci USA **105**: 803-808.

Gobbato E, Wang E, Higgins G, Bano SA, Henry C et al. (2013) *RAM1* and *RAM2* function and expression during arbuscular mycorrhizal symbiosis and *Aphanomyces euteiches* colonization. Plant Signal Behav **8**: e26049.

Gobbato E, Marsh JF, Vernié T, Wang E, Maillet F et al. (2012) A GRAS-type transcription factor with a specific function in mycorrhizal signaling. Curr Biol **22**: 2236-2241.

Golden JW, Robinson SJ, Haselkorn R (1985) Rearrangement of nitrogen fixation genes during heterocyst differentiation in the cyanobacterium *Anabaena*. Nature **314**: 419-423.

Gómez F, Furuya K, Takeda S (2005) Distribution of the cyanobacterium *Richelia intracellularis* as an epiphyte of the diatom *Chaetoceros compressus* in the western Pacific Ocean. J Plankton Res **27**: 323-330.

Gorelova OA (2006) Communication of cyanobacteria with plant partners during association formation. Microbiology **75**: 465-469.

Govindarajulu M, Pfeffer PE, Jin H, Abubaker J, Douds DD et al. (2005) Nitrogen transfer in the arbuscular mycorrhizal symbiosis. Nature **435**: 819-823.

Granqvist E, Sun J, Op den Camp R, Pujic P, Hill L et al. (2015) Bacterial-induced calcium oscillations are common to nitrogen-fixing associations of nodulating legumes and nonlegumes. New Phytol **207**: 551-558.

Griesmann M, Chang Y, Liu X, Song Y, Haberer G et al. (2018) Phylogenomics reveals multiple losses of nitrogen-fixing root nodule symbiosis. Science **361**: eaat1743.

Group TLPW, Bruneau A, Doyle JJ, Herendeen P, Hughes C et al. (2013) Legume phylogeny and classification in the 21st century: Progress, prospects and lessons for other species–rich clades. Taxon **62**: 217-248.

Guan D, Stacey N, Liu C, Wen J, Mysore KS et al. (2013) Rhizobial infection is associated with the development of peripheral vasculature in nodules of *Medicago truncatula*. Plant Physiol **162**: 107-115.

Gusev MV, Baulina OI, Gorelova OA, Lobakova ES, Korzhenevskaya TG (2002) Artificial cyanobacterium-plant symbioses. In Cyanobacteria in Symbiosis, AN Rai, B Bergman, and U Rasmussen, eds (Dordrecht: Springer Netherlands), pp. 253-312.

Gutjahr C, Radovanovic D, Geoffroy J, Zhang Q, Siegler H et al. (2012) The half-size ABC transporters STR1 and STR2 are indispensable for mycorrhizal arbuscule formation in rice. Plant J **69**: 906-920.

Harrison MJ (2005) Signaling in the arbuscular mycorrhizal symbiosis. Annu Rev Microbiol **59**: 19-42.

Hashidoko Y, Nishizuka H, Tanaka M, Murata K, Murai Y et al. (2019) Isolation and characterization of 1-palmitoyl-2-linoleoyl-*sn*-glycerol as a hormogonium-inducing factor (HIF) from the coralloid roots of *Cycas revoluta* (Cycadaceae). Sci Rep **9**: 4751.

Heck C, Kuhn H, Heidt S, Walter S, Rieger N et al. (2016) Symbiotic fungi control plant root cortex development through the novel GRAS transcription factor MIG1. Curr Biol **26**: 2770-2778.

Hengl T, Leenaars JGB, Shepherd KD, Walsh MG, Heuvelink GBM et al. (2017) Soil nutrient maps of Sub-Saharan Africa: assessment of soil nutrient content at 250 m spatial resolution using machine learning. Nutr Cycling Agroecosyst **109**: 77-102.

Herrero A, Muro-Pastor AM, Flores E (2001) Nitrogen control in cyanobacteria. J Bacteriol **183**: 411-425.

Herrero A, Muro-Pastor AM, Valladares A, Flores E (2004) Cellular differentiation and the NtcA transcription factor in filamentous cyanobacteria. FEMS Microbiol Rev **28**: 469-487.

Higa KC, Callahan SM (2010) Ectopic expression of *hetP* can partially bypass the need for *hetR* in heterocyst differentiation by *Anabaena* sp. strain PCC 7120. Mol Microbiol **77**: 562-574.

Hinsinger P, Gobran GR, Gregory PJ, Wenzel WW (2005) Rhizosphere geometry and heterogeneity arising from root-mediated physical and chemical processes. New Phytol **168**: 293-303.

Hirose E, Hirose M (2007) Morphological process of vertical transmission of photosymbionts in the colonial ascidian *Tridemnum miniatum* Kott, 1977. Mar Biol **150**: 359-367.

Hodge A (2004) The plastic plant: root responses to heterogeneous supplies of nutrients. New Phytol **162**: 9-24.

Hoiczyk E, Hansel A (2000) Cyanobacterial cell walls: news from an unusual prokaryotic envelope. J Bacteriol **182**: 1191-1199.

Horváth B, Yeun LH, Domonkos Á, Halász G, Gobbato E et al. (2011) *Medicago truncatula IPD3* is a member of the common symbiotic signaling pathway required for rhizobial and mycorrhizal symbioses. Mol Plant-Microbe Interact **24**: 1345-1358.

Huang X, Dong Y, Zhao J (2004) HetR homodimer is a DNA-binding protein required for heterocyst differentiation, and the DNA-binding activity is inhibited by PatS. Proc Natl Acad Sci USA **101**: 4848-4853.

Hunt E, Gattolin S, Newbury HJ, Bale JS, Tseng HM et al. (2009) A mutation in amino acid permease AAP6 reduces the amino acid content of the *Arabidopsis* sieve elements but leaves aphid herbivores unaffected. J Exp Bot **61**: 55-64.

Huss-Danell K (1990) The physiology of actinorhizal nodules. In The biology of *Frankia* and actinorhizal plants, CR Schwintzer and JD Tjepkema, eds (San Diego, USA: Academic Press), pp. 129-156.

Huss-Danell K (1997) Tansley Review No. 93. Actinorhizal symbioses and their N_2 fixation. New Phytol **136**: 375-405.

Imaizumi-Anraku H, Takeda N, Charpentier M, Perry J, Miwa H et al. (2005) Plastid proteins crucial for symbiotic fungal and bacterial entry into plant roots. Nature **433,** 527-531.

Imsande J, Touraine B (1994) N demand and the regulation of nitrate uptake. Plant Physiol **105**: 3-7.

Jin H, Pfeffer PE, Douds DD, Piotrowski E, Lammers PJ et al. (2005) The uptake, metabolism, transport and transfer of nitrogen in an arbuscular mycorrhizal symbiosis. New Phytol **168**: 687-696.

Jin Y, Liu H, Luo D, Yu N, Dong W et al. (2016) DELLA proteins are common components of symbiotic rhizobial and mycorrhizal signalling pathways. Nat Commun **7**: 12433.

Johansen A, Finlay RD, Olsson PA (1996) Nitrogen metabolism of external hyphae of the arbuscular mycorrhizal fungus *Glomus intraradices*. New Phytol **133**: 705-712.

Johansson C, Bergman B (1992) Early events during the establishment of the *Gunnera/Nostoc* symbiosis. Planta **188**: 403-413.

Jones A, Davies HM, Voelker TA (1995) Palmitoyl-acyl carrier protein (ACP) thioesterase and the evolutionary origin of plant acyl-ACP thioesterases. Plant Cell **7**: 359-371.

Jyothibabu R, Madhu N, Maheswaran P, Devi CA, Balasubramanian T et al. (2006) Environmentally-related seasonal variation in symbiotic associations of heterotrophic dinoflagellates with cyanobacteria in the western Bay of Bengal. Symbiosis **42**: 51-58.

Kaló P, Gleason C, Edwards A, Marsh J, Mitra RM et al. (2005) Nodulation signaling in legumes requires NSP2, a member of the GRAS family of transcriptional regulators. Science **308**: 1786-1789.

Kanamori N, Madsen LH, Radutoiu S, Frantescu M, Quistgaard EMH et al. (2006) A nucleoporin is required for induction of Ca^{2+} spiking in legume nodule development and essential for rhizobial and fungal symbiosis. Proc Natl Acad Sci USA **103**: 359-364.

Kawaharada Y, James EK, Kelly S, Sandal N, Stougaard J (2017) The ethylene responsive factor required for nodulation 1 (ERN1) transcription factor is required for infection-thread formation in *Lotus japonicus*. Mol Plant Microbe Interact **30**: 194-204.

Kenrick P, Crane PR (1997) The origin and early evolution of plants on land. Nature **389**: 33-39.

Kenrick P, Strullu-Derrien C (2014) The origin and early evolution of roots. Plant Physiol **166**: 570-580.

Kevei Z, Lougnon G, Mergaert P, Horváth GV, Kereszt A et al. (2007) 3-Hydroxy-3-methylglutaryl coenzyme a reductase 1 interacts with NORK and is crucial for nodulation in *Medicago truncatula*. Plant Cell **19**: 3974-3989.

Khudyakov IY, Golden JW (2004) Different functions of HetR, a master regulator of heterocyst differentiation in *Anabaena* sp. PCC 7120, can be separated by mutation. Proc Natl Acad Sci USA **101**: 16040-16045.

Kiba T, Feria-Bourrellier AB, Lafouge F, Lezhneva L, Boutet-Mercey S et al. (2012) The *Arabidopsis* nitrate transporter NRT2.4 plays a double role in roots and shoots of nitrogen-starved plants Plant Cell **24**: 245-258.

Knack JJ, Wilcox LW, Delaux PM, Ané JM, Piotrowski MJ et al. (2015) Microbiomes of streptophyte algae and bryophytes suggest that a functional suite of microbiota fostered plant colonization of land. Int J Plant Sci **176**: 405-420.

Knight CD, Adams DG (1996) A method for studying chemotaxis in nitrogen fixing cyanobacterium-plant symbioses. Physiol Mol Plant Pathol **49**: 73-77.

Kobae Y, Tamura Y, Takai S, Banba M, Hata S (2010) Localized expression of arbuscular mycorrhiza-inducible ammonium transporters in soybean. Plant Cell Physiol **51**: 1411-1415.

Kojima S, Konishi N, Beier MP, Ishiyama K, Maru I et al. (2014) NADH-dependent glutamate synthase participated in ammonium assimilation in *Arabidopsis* root. Plant Signal Behav **9**: e29402.

Kramer KU, Green P (1990) Introduction to pteridophytes and gymnosperms. In Pteridophytes and Gymnosperms. The families and generaof vascular plants, KU Kramer and P Green, eds (Berlin: Springer), pp. 1.

Krings M, Hass H, Kerp H, Taylor TN, Agerer R et al. (2009) Endophytic cyanobacteria in a 400-million-yr-old land plant: A scenario for the origin of a symbiosis? Rev Palaeobot Palynol **153**: 62-69.

Laenen B, Shaw B, Schneider H, Goffinet B, Paradis E et al. (2014) Extant diversity of bryophytes emerged from successive post-Mesozoic diversification bursts. Nat Commun **5**: 5134.

Lam HM, Coschigano KT, Oliveira IC, Melo-Oliveira R, Coruzzi GM (1996) The molecular-genetics of nitrogen assimilation into amino acids in higher plants. Annu Rev Plant Physiol Plant Mol Biol **47**: 569-593.

Landrein B, Formosa-Jordan P, Malivert A, Schuster C, Melnyk CW et al. (2018) Nitrate modulates stem cell dynamics in *Arabidopsis* shoot meristems through cytokinins. Proc Natl Acad Sci USA **115**: 1382-1387.

Laurent S, Chen H, Bédu S, Ziarelli F, Peng L et al. (2005) Nonmetabolizable analogue of 2-oxoglutarate elicits heterocyst differentiation under repressive conditions in *Anabaena* sp. PCC 7120. Proc Natl Acad Sci USA **102**: 9907-9912.

Lee JJ (2006) Algal symbiosis in larger foraminifera. Symbiosis **42**: 63-75.

Lee YH, Foster J, Chen J, Voll LM, Weber AP et al. (2007) AAP1 transports uncharged amino acids into roots of *Arabidopsis*. Plant J **50**: 305-319.

Lejay L, Tillard P, Lepetit M, Olive FD, Filleur S et al. (1999) Molecular and functional regulation of two NO_3^- uptake systems by N- and C-status of *Arabidopsis* plants. Plant J **18**: 509-519.

Lesser MP, Mazel CH, Gorbunov MY, Falkowski PG (2004) Discovery of symbiotic nitrogen-fixing cyanobacteria in corals. Science **305**: 997-1000.

Lesser MP, Falcon LI, Rodriguez-Roman A, Enriquez S, Hoegh-Guldberg O et al. (2007) Nitrogen fixation by symbiotic cyanobacteria provides a source of nitrogen for the scleractinian coral *Montastraea cavernosa*. Mar Ecol Prog Ser **346**: 143-152.

Lezhneva L, Kiba T, Feria-Bourrellier AB, Lafouge F, Boutet-Mercey S et al. (2014) The *Arabidopsis* nitrate transporter NRT2.5 plays a role in nitrate acquisition and remobilization in nitrogen-starved plants. Plant J **80**: 230-241.

Li FW, Nishiyama T, Waller M, Frangedakis E, Keller J et al. (2020) *Anthoceros* genomes illuminate the origin of land plants and the unique biology of hornworts. Nat Plants **6**: 259-272.

Li FW, Brouwer P, Carretero-Paulet L, Cheng S, de Vries J et al. (2018) Fern genomes elucidate land plant evolution and cyanobacterial symbioses. Nat Plants **4**: 460-472.

Li JY, Fu YL, Pike SM, Bao J, Tian W et al. (2010) The *Arabidopsis* nitrate transporter NRT1.8 functions in nitrate removal from the xylem sap and mediates cadmium tolerance Plant Cell **22**: 1633-1646.

Limpens E, Franken C, Smit P, Willemse J, Bisseling T et al. (2003) LysM domain receptor kinases regulating rhizobial Nod factor-induced infection. Science **302**: 630-633.

Lindblad P, Bergman B (1990) The cycad-cyanobacterial symbiosis. In CRC handbook of symbiotic cyanobacteria, AN Rai, ed (Boca Raton, Florida: CRC Press), pp. 137-159.

Lindquist N, Barber PH, Weisz JB (2005) Episymbiotic microbes as food and defence for marine isopods: unique symbioses in a hostile environment. Proc R Soc B: Biol Sci **272**: 1209-1216.

Liu A, Contador CA, Fan K, Lam HM (2018) Interaction and regulation of carbon, nitrogen, and phosphorus metabolisms in root nodules of legumes. Front Plant Sci **9**: 1860.

Liu CW, Murray JD (2016) The role of flavonoids in nodulation host-range specificity: an update. Plants **5**: 33.

López-Igual R, Flores E, Herrero A (2010) Inactivation of a heterocyst-specific invertase indicates a principal role of sucrose catabolism in heterocysts of *Anabaena* sp. J Bacteriol **192**: 5526-5533.

López-Pedrosa A, González-Guerrero M, Valderas A, Azcón-Aguilar C, Ferrol N (2006) *GintAMT1* encodes a functional high-affinity ammonium transporter that is expressed in the extraradical mycelium of *Glomus intraradices*. Fungal Genet Biol **43**: 102-110.

Loqué D, Yuan L, Kojima S, Gojon A, Wirth J et al. (2006) Additive contribution of AMT1;1 and AMT1;3 to high-affinity ammonium uptake across the plasma membrane of nitrogen-deficient *Arabidopsis* roots. Plant J **48**: 522-534.

Luginbuehl LH, Menard GN, Kurup S, Van Erp H, Radhakrishnan GV et al. (2017) Fatty acids in arbuscular mycorrhizal fungi are synthesized by the host plant. Science **356**: 1175-1178.

MacLean AM, Bravo A, Harrison MJ (2017) Plant signaling and metabolic pathways enabling arbuscular mycorrhizal symbiosis. Plant Cell **29**: 2319-2335.

Madsen EB, Madsen LH, Radutoiu S, Olbryt M, Rakwalska M et al. (2003) A receptor kinase gene of the LysM type is involved in legume perception of rhizobial signals. Nature **425**: 637-640.

Maeda SI, Konishi M, Yanagisawa S, Omata T (2014) Nitrite transport activity of a novel HPP family protein conserved in cyanobacteria and chloroplasts. Plant Cell Physiol **55**: 1311-1324.

Mahmud K, Makaju S, Ibrahim R, Missaoui A (2020) Current progress in nitrogen fixing plants and microbiome research. Plants **9**: 97

Maillet F, Poinsot V, André O, Puech-Pagès V, Haouy A et al. (2011) Fungal lipochitooligosaccharide symbiotic signals in arbuscular mycorrhiza. Nature **469**: 58-63.

Martín-Figueroa E, Navarro F, Florencio FJ (2000) The GS-GOGAT pathway is not operative in the heterocysts. Cloning and expression of *glsF* gene from the cyanobacterium *Anabaena* sp. PCC 7120. FEBS Lett **476**: 282-286.

Martin F, Kohler A, Murat C, Veneault-Fourrey C, Hibbett DS (2016) Unearthing the roots of ectomycorrhizal symbioses. Nat Rev Microbiol **14**: 760-773.

Martin FM, Uroz S, Barker DG (2017) Ancestral alliances: plant mutualistic symbioses with fungi and bacteria. Science **356**: eaad4501.

Masepohl B, Schölisch K, Görlitz K, Kutzki C, Böhme H (1997) The heterocyst-specific *fdxH* gene product of the cyanobacterium *Anabaena* sp. PCC 7120 is important but not essential for nitrogen fixation. Mol Gen Genet **253**: 770-776.

Meeks JC (1990a) Cyanobacterial-bryophyte associations. In CRC Handbook of Symbiotic Cyanobacteria, AN Rai, ed (Boca Raton, Florida: CRC Press), pp. 21.

Meeks JC (1990b) Cyanobacterial-bryophyte associations. In Handbook of Symbiotic Cyanobacteria, AN Rai, ed (Boca Raton, Florida: CRC Press), pp. 43-63.

Meeks JC (2003) Symbiotic interactions between *Nostoc punctiforme*, a multicellular cyanobacterium, and the hornwort *Anthoceros punctatus*. Symbiosis **35**: 55-71.

Meeks JC (2006) Molecular mechanisms in the nitrogen-fixing *Nostoc*-bryophyte symbiosis. Prog Mol Subcell Biol **41**: 165-196.

Meeks JC (2009) Physiological adaptations in nitrogen-fixing *Nostoc*–plant symbiotic associations. In Prokaryotic Symbionts in Plants, K Pawlowski, ed (Berlin, Heidelberg: Springer Berlin Heidelberg), pp. 181-205.

Meeks JC, Elhai J (2002) Regulation of cellular differentiation in filamentous cyanobacteria in free-living and plant-associated symbiotic growth states. Microbiol Mol Biol Rev **66**: 94-121.

Meeks JC, Wycoff KL, Chapman JS, Enderlin CS (1983) Regulation of expression of nitrate and dinitrogen assimilation by *Anabaena* species. Appl Environ Microbiol **45**: 1351-1359.

Meeks JC, Campbell EL, Summers ML, Wong FC (2002) Cellular differentiation in the cyanobacterium *Nostoc punctiforme*. Arch Microbiol **178**: 395-403.

Meeks JC, Enderlin CS, Joseph CM, Chapman JS, Lollar MWL (1985) Fixation of [^{13}N]N_2 and transfer of fixed nitrogen in the *Anthoceros-Nostoc* symbiotic association. Planta **164**: 406-414.

Meeks JC, Steinberg NA, Enderlin CS, Joseph CM, Peters GA (1987) *Azolla-Anabaena* relationship: XIII. Fixation of [^{13}N]N_2. Plant Physiol **84**: 883-886.

Merrick M, Dixon R (1984) Why don't plants fix nitrogen? Trends Biotechnol **2**: 162-166.

Metzgar JS, Schneider H, Pryer KM (2007) Phylogeny and divergence time estimates for the fern genus *Azolla* (Salviniaceae). Int J Plant Sci **168**: 1045-1053.

Miflin BJ (1974) The location of nitrite reductase and other enzymes related to amino acid biosynthesis in the plastids of root and leaves. Plant Physiol **54**: 550-555.

Milindasuta BE (1975) Developmental anatomy of coralloid roots in cycads. Am J Bot **62**: 468-472.

Miller AJ, Cramer MD (2005) Root nitrogen acquisition and assimilation. Plant Soil **274**: 1-36.

Miller IM, Baker DD (1986) Nodulation of actinorhizal plants by *Frankia* strains capable of both root hair infection and intercellular penetration. Protoplasma **131**: 82-91.

Miller JB, Pratap A, Miyahara A, Zhou L, Bornemann S et al. (2013) Calcium/calmodulin-dependent protein kinase is negatively and positively regulated by calcium, providing a mechanism for decoding calcium responses during symbiosis signaling. Plant Cell **25**: 5053-5066.

Miyata K, Kozaki T, Kouzai Y, Ozawa K, Ishii K et al. (2014) The bifunctional plant receptor, OsCERK1, regulates both chitin-triggered immunity and arbuscular mycorrhizal symbiosis in rice. Plant Cell Physiol **55**: 1864-1872.

Moling S, Pietraszewska-Bogiel A, Postma M, Fedorova E, Hink MA et al. (2014) Nod factor receptors form heteromeric complexes and are essential for intracellular infection in *Medicago* nodules. Plant Cell **26**: 4188-4199.

Mollenhauer D, Mollenhauer R, Kluge M (1996) Studies on initiation and development of the partner association in *Geosiphon pyriforme* (Kütz.) v. Wettstein, a unique endocytobiotic system of a fungus (Glomales) and the cyanobacterium *Nostoc punctiforme* (Kütz.) Hariot. Protoplasma **193**: 3-9.

Muro-Pastor AM, Valladares A, Flores E, Herrero A (2002) Mutual dependence of the expression of the cell differentiation regulatory protein HetR and the global nitrogen regulator NtcA during heterocyst development. Mol Microbiol **44**: 1377-1385.

Muro-Pastor MI, Reyes JC, Florencio FJ (2001) Cyanobacteria perceive nitrogen status by sensing intracellular 2-oxoglutarate levels. J Biol Chem **276**: 38320-38328.

Näsholm T, Kielland K, Ganeteg U (2009) Uptake of organic nitrogen by plants. New Phytol **182**: 31-48.

Nathanielsz CP, Staff IA (1975) A mode of entry of blue-green algae into the apogeotrophic roots of *Macrozamia communis*. Am J Bot **62**: 232-235.

Nicolaisen K, Hahn A, Schleiff E (2009) The cell wall in heterocyst formation by *Anabaena* sp. PCC 7120. J Basic Microbiol **49**: 5-24.

Nilsson M, Rasmussen U, Bergman B (2005) Competition among symbiotic cyanobacterial *Nostoc* strains forming artificial associations with rice (*Oryza sativa*). FEMS Microbiol Lett **245**: 139-144.

Nilsson M, Rasmussen U, Bergman B (2006) Cyanobacterial chemotaxis to extracts of host and non-host plants. FEMS Microbiol Ecol **55**: 382-390.

Ohkubo Y, Tanaka M, Tabata R, Ogawa-Ohnishi M, Matsubayashi Y (2017) Shoot-to-root mobile polypeptides involved in systemic regulation of nitrogen acquisition. Nat Plants **3**: 17029.

Ohlrogge JB, Jaworski JG, Post-Beittenmiller D (1993) *De novo* fatty acid biosynthesis. In Lipid Metabolism in Plants, TS Moore, ed (Boca Raton, Florida: CRC Press), pp. 3-32.

Okumoto S, Pilot G (2011) Amino acid export in plants: a missing link in nitrogen cycling. Mol Plant **4**: 453-463.

Oldroyd GED (2013) Speak, friend, and enter: signalling systems that promote beneficial symbiotic associations in plants. Nat Rev Microbiol **11**: 252-263.

Oldroyd GED, Leyser O (2020) A plant's diet, surviving in a variable nutrient environment. Science **368**: eaba0196.

Oldroyd GED, Murray JD, Poole PS, Downie JA (2011) The rules of engagement in the legume-rhizobial symbiosis. Annu Rev Genet **45**: 119-144.

Olmedo-Verd E, Valladares A, Flores E, Herrero A, Muro-Pastor AM (2008) Role of two NtcA-binding sites in the complex *ntcA* gene promoter of the heterocyst-forming cyanobacterium *Anabaena* sp. strain PCC 7120. J Bacteriol **190**: 7584-7590.

Olson ND, Ainsworth TD, Gates RD, Takabayashi M (2009) Diazotrophic bacteria associated with Hawaiian *Montipora* corals: diversity and abundance in correlation with symbiotic dinoflagellates. J Exp Mar Biol Ecol **371**: 140-146.

Op den Camp R, Streng A, De Mita S, Cao Q, Polone E et al. (2011) LysM-type mycorrhizal receptor recruited for rhizobium symbiosis in nonlegume *Parasponia*. Science **331**: 909-912.

Ow MC, Gantar M, Elhai J (1999) Reconstitution of a cycad-cyanobacterial association. Symbiosis **27**: 125-134.

Park HJ, Floss DS, Levesque-Tremblay V, Bravo A, Harrison MJ (2015) Hyphal branching during arbuscule development requires *Reduced Arbuscular Mycorrhiza 1*. Plant Physiol **169**: 2774-2788.

Parniske M (2008) Arbuscular mycorrhiza: the mother of plant root endosymbioses. Nat Rev Microbiol **6**: 763-775.

Pate JS, Sharkey PJ, Lewis OAM (1975) Xylem to phloem transfer of solutes in fruiting shoots of legumes, studied by a phloem bleeding technique. Planta **122**: 11-26.

Pate JS, Lindblad P, Atkins CA (1988) Pathways of assimilation and transfer of fixed nitrogen in coralloid roots of cycad-*Nostoc* symbioses. Planta **176**: 461-471.

Paungfoo-Lonhienne C, Lonhienne TG, Rentsch D, Robinson N, Christie M et al. (2008) Plants can use protein as a nitrogen source without assistance from other organisms. Proc Natl Acad Sci USA **105**: 4524-4529.

Pawlowski K, Demchenko KN (2012) The diversity of actinorhizal symbiosis. Protoplasma **249**: 967-979.

Penmetsa RV, Cook DR (2000) Production and characterization of diverse developmental mutants of *Medicago truncatula*. Plant Physiol **123**: 1387-1398.

Perchlik M, Foster J, Tegeder M (2014) Different and overlapping functions of Arabidopsis LHT6 and AAP1 transporters in root amino acid uptake. J Exp Bot **65**: 5193-5204.

Pérez-Tienda J, Testillano PS, Balestrini R, Fiorilli V, Azcón-Aguilar C et al. (2011) GintAMT2, a new member of the ammonium transporter family in the arbuscular mycorrhizal fungus *Glomus intraradices*. Fungal Genet Biol **48**: 1044-1055.

Peters NK, Frost JW, Long SR (1986) A plant flavone, luteolin, induces expression of *Rhizobium meliloti* nodulation genes. Science **233**: 977-980.

Peterson R, Massicotte H (2004) Exploring structural definitions of mycorrhizas, with emphasis on nutrient-exchange interfaces. Can J Bot **82**: 1074-1088.

Pfeffer PE, Douds DD, Bécard G Jr., Shachar-Hill Y (1999) Carbon uptake and the metabolism and transport of lipids in an arbuscular mycorrhiza. Plant Physiol **120**: 587-598.

Pimprikar P, Carbonnel S, Paries M, Katzer K, Klingl V et al. (2016) A CCaMK-CYCLOPS-DELLA complex activates transcription of RAM1 to regulate arbuscule branching. Curr Biol **26**: 1126.

Pirozynski KA, Malloch DW (1975) The origin of land plants: a matter of mycotrophism. BioSyst **6**: 153-164.

Pratelli R, Pilot G (2014) Regulation of amino acid metabolic enzymes and transporters in plants. J Exp Bot **65**: 5535-5556.

Price G, Mohapatra S, Gresshoff P (1984) Structure of nodules formed by *Rhizobium* strain ANU289 in the nonlegume *Parasponia* and the legume *Siratro* (Macroptilium atropurpureum). Bot Gaz **145**: 444-451.

Radhakrishnan GV, Keller J, Rich MK, Vernié T, Mbadinga Mbadinga DL et al. (2020). An ancestral signalling pathway is conserved in intracellular symbioses-forming plant lineages. Nat Plants **6**: 280-289.

Radutoiu S, Madsen LH, Madsen EB, Felle HH, Umehara Y et al. (2003) Plant recognition of symbiotic bacteria requires two LysM receptor-like kinases. Nature **425**: 585-592.

Rai AN, Söderbäck E, Bergman B (2000) Tansley Review No. 116 Cyanobacterium–plant symbioses. New Phytol **147**: 449-481.

Rai AN, Bergman B, Rasmussen U (2002). Cyanobacteria in symbiosis. (Berlin: Springer).

Rai AN, Singh AK, Syiem MB (2019) Plant growth-promoting abilities in cyanobacteria. In Cyanobacteria from Basic Science to Applications, AK Mishra, DN Tiwari, and AN Rai, eds (Academic Press), pp. 459-476.

Rajagopalan R, Callahan SM (2010) Temporal and spatial regulation of the four transcription start sites of *hetR* from *Anabaena* sp. strain PCC 7120. J Bacteriol **192**: 1088-1096.

Rasmussen U, Johansson C, Bergman B (1994) Early communication in the *Gunnera-Nostoc* symbiosis: plant-induced cell differentiation and protein synthesis in the cyanobacterium. Mol Plant Microbe Interact **7**: 696-702.

Raven JA (2002) Evolution of cyanobacterial symbioses. In Cyanobacteria in Symbiosis, AN Rai, B Bergman, and U Rasmussen, eds (Dordrecht: Springer Netherlands), pp. 329-346.

Redmond JW, Batley M, Djordjevic MA, Innes RW, Kuempel PL et al. (1986) Flavones induce expression of nodulation genes in *Rhizobium*. Nature **323**: 632-635.

Reed SC, Cleveland CC, Townsend AR (2011) Functional ecology of free-living nitrogen fixation: a contemporary perspective. Annu Rev Ecol Evol Syst **42**: 489-512.

Reisdorf-Cren M, Hirel B (1999) Glutamine synthetase in higher plants regulation of gene and protein expression from the organ to the cell. Plant Cell Physiol **40**: 1187-1193.

Remy W, Hass H (1996) New information on gametophytes and sporophytes of *Aglaophyton major* and inferences about possible environmental adaptations. Rev Palaeobot Palynol **90**: 175-193.

Remy W, Taylor TN, Hass H, Kerp H (1994) Four hundred-million-year-old vesicular arbuscular mycorrhizae. Proc Natl Acad Sci USA **91**: 11841-11843.

Rentsch D, Schmidt S, Tegeder M (2007) Transporters for uptake and allocation of organic nitrogen compounds in plants. FEBS Lett **581**: 2281-2289.

Rich MK, Courty PE, Roux C, Reinhardt D (2017) Role of the GRAS transcription factor ATA/RAM1 in the transcriptional reprogramming of arbuscular mycorrhiza in *Petunia hybrida*. BMC Genom **18**: 589.

Romagnoli T, Bavestrello G, Cucchiari EM, De Stefano M, Di Camillo CG et al. (2007) Microalgal communities epibiontic on the marine hydroid *Eudendrium racemosum* in the Ligurian Sea during an annual cycle. Mar Biol **151**: 537-552.

Roth LE, Stacey G (1989) Bacterium release into host cells of nitrogen-fixing soybean nodules: the symbiosome membrane comes from three sources. Eur J Cell Biol **49**: 13-23.

Roy S, Liu W, Nandety RS, Crook A, Mysore KS et al. (2020) Celebrating 20 years of genetic discoveries in legume nodulation and symbiotic nitrogen fixation. Plant Cell **32**: 15-41.

Rubio LM, Ludden PW (2008) Biosynthesis of the iron-molybdenum cofactor of nitrogenase. Annu Rev Microbiol **62**: 93-111.

Ruffel S, Krouk G, Ristova D, Shasha D, Birnbaum KD et al. (2011) Nitrogen economics of root foraging: transitive closure of the nitrate–cytokinin relay and distinct systemic signaling for N supply vs. demand. Proc Natl Acad Sci USA **108**: 18524-18529.

Saito K, Yoshikawa M, Yano K, Miwa H, Uchida H et al. (2007). Nucleoporin 85 is required for calcium spiking, fungal and bacterial symbioses, and seed production in *Lotus japonicus*. Plant Cell **19**: 610-624.

Sakakibara H, Takei K, Hirose N (2006) Interactions between nitrogen and cytokinin in the regulation of metabolism and development. Trends Plant Sci **11**: 440-448.

Santi C, Bogusz D, Franche C (2013) Biological nitrogen fixation in non-legume plants. Ann Bot **111**: 743-767.

Schauser L, Roussis A, Stiller J, Stougaard J (1999) A plant regulator controlling development of symbiotic root nodules. Nature **402**: 191-195.

Schneider D, Fuhrmann E, Scholz I, Hess WR, Graumann PL (2007) Fluorescence staining of live cyanobacterial cells suggest non-stringent chromosome segregation and absence of a connection between cytoplasmic and thylakoid membranes. BMC Cell Biol **8**: 39.

Schnepf E, Schlegel I, Hepperle D (2002) *Petalomonas sphagnophila* (Euglenophyta) and its endocytobiotic cyanobacteria: a unique form of symbiosis. Phycologia **41**: 153-157.

Schüßler A (2012) The *Geosiphon–Nostoc* endosymbiosis and its role as a model for arbuscular mycorrhiza research. In Fungal Associations, B Hock, ed (Berlin, Heidelberg: Springer Berlin Heidelberg), pp. 77-91.

Schüßler A, Mollenhauer D, Schnepf E, Kluge M (1994) *Geosiphon pyriforme*, an endosymbiotic association of fungus and cyanobacteria: the spore structure resembles that of arbuscular mycorrhizal (AM) fungi. Botanica Acta **107**: 36-45.

Schwintzer CR, Berry AM, Disney LD (1982) Seasonal patterns of root nodule growth, endophyte morphology, nitrogenase activity, and shoot development in *Myrica gale*. Can J Bot **60**: 746-757.

Sciuto K, Moro I (2015) Cyanobacteria: the bright and dark sides of a charming group. Biodivers Conserv **24**: 711-738.

Selosse MA, Le Tacon F (1998) The land flora: a phototroph-fungus partnership? Trends Ecol Evol **13**: 15-20.

Silvester WB, McNamara PJ (1976) The infection process and ultrastructure of the *Gunnera-Nostoc* symbiosis. New Phytol **77**: 135-141.

Silvester WB, Parsons R, Watt PW (1996) Direct measurement of release and assimilation of ammonia in the *Gunnera–Nostoc* symbiosis. New Phytol **132**: 617-625.

Smit P, Raedts J, Portyanko V, Debellé F, Gough C et al. (2005) NSP1 of the GRAS protein family is essential for rhizobial Nod factor-induced transcription. Science **308**: 1789-1791.

Smith SE, Read DJ (2008). Mycorrhizal symbiosis. (London: Academic press).

Soltis DE, Soltis PS, Morgan DR, Swensen SM, Mullin BC et al. (1995) Chloroplast gene sequence data suggest a single origin of the predisposition for symbiotic nitrogen fixation in angiosperms. Proc Natl Acad Sci USA **92**: 2647-2651.

Sprent JI, James EK (2007) Legume evolution: where do nodules and mycorrhizas fit in? Plant Physiol **144**: 575-581.

Sprent JI, Ardley JK, James EK (2013) From North to South: A latitudinal look at legume nodulation processes. S Afr J Bot **89**: 31-41.

Srivastava HS (1980) Regulation of nitrate reductase activity in higher plants. Phytochemistry **19**: 725-733.

Stark, JM (1994) Causes of soil nutrient heterogeneity at different scales. In Exploitation of Environmental Heterogeneity by Plants, MM Caldwell and RW Pearcy, eds (Boston: Academic Press), pp. 255-284.

Staswick PE (1994) Storage proteins of vegetative plant tissues. Annu Rev Plant Physiol Plant Mol Biol **45**: 303-322.

Stracke S, Kistner C, Yoshida S, Mulder L, Sato S et al. (2002). A plant receptor-like kinase required for both bacterial and fungal symbiosis. Nature **417**: 959-962.

Svennerstam H, Ganeteg U, Näsholm T (2008) Root uptake of cationic amino acids by *Arabidopsis* depends on functional expression of amino acid permease 5. New Phytol **180**: 620-630.

Svennerstam H, Jämtgård S, Ahmad I, Huss-Danell K, Näsholm T et al. (2011) Transporters in *Arabidopsis* roots mediating uptake of amino acids at naturally occurring concentrations. New Phytol **191**: 459-467.

Svistoonoff S, Hocher V, Gherbi H (2014) Actinorhizal root nodule symbioses: what is signalling telling on the origins of nodulation? Curr Opin Plant Biol **20**: 11-18.

Swensen SM (1996) The evolution of actinorhizal symbioses: evidence for multiple origins of the symbiotic association. Am J Bot **83**: 1503-1512.

Szczyglowski K, Shaw RS, Wopereis J, Copeland S, Hamburger D et al. (1998) Nodule organogenesis and symbiotic mutants of the model legume *Lotus japonicus*. Mol Plant-Microbe Interact **11**: 684-697.

Tabata R, Sumida K, Yoshii T, Ohyama K, Shinohara H et al. (2014). Perception of root-derived peptides by shoot LRR-RKs mediates systemic N-demand signaling. Science **346**: 343-346.

Taleski M, Imin N, Djordjevic MA (2018) CEP peptide hormones: key players in orchestrating nitrogen-demand signalling, root nodulation, and lateral root development. J Exp Bot **69**: 1829-1836.

Tarangkoon W, Hansen G, Hansen PJ (2010) Spatial distribution of symbiont-bearing dinoflagellates in the Indian Ocean in relation to oceanographic regimes. Aquat Microb Ecol **58**: 197-213.

Taylor EL, Taylor TN, Krings M (2009) Paleobotany: the biology and evolution of fossil plants. (Oxford: Academic Press).

Taylor TN, Remy W, Hass H, Kerp H (1995) Fossil arbuscular mycorrhizae from the early Devonian. Mycologia **87**: 560-573.

Tedersoo L, Brundrett MC (2017) Evolution of ectomycorrhizal symbiosis in plants. In Biogeography of Mycorrhizal Symbiosis, L. Tedersoo, ed (Cham: Springer International Publishing), pp. 407-467.

Tegeder M (2012) Transporters for amino acids in plant cells: some functions and many unknowns. Curr Opin Plant Biol **15**: 315-321.

Tegeder M (2014) Transporters involved in source to sink partitioning of amino acids and ureides: opportunities for crop improvement. J Exp Bot **65**: 1865-1878.

Tegeder M, Weber A (2007) Metabolite transporters in the control of plant primary metabolism. In Control of Primary Metabolism in Plants, WC Plaxton and MT McManus, eds (Oxford, USA: Blackwell Publishing), pp. 85-120.

Tegeder M, Rentsch D (2010) Uptake and partitioning of amino acids and peptides. Mol Plant **3**: 997-1011.

Tegeder M, Masclaux-Daubresse C (2018) Source and sink mechanisms of nitrogen transport and use. New Phytol **217**: 35-53.

Thamdrup B (2012) New pathways and processes in the global nitrogen cycle. Annu Rev Ecol Evol Syst **43**: 407-428.

Thomas J, Meeks JC, Wolk CP, Shaffer PW, Austin SM (1977) Formation of glutamine from [^{13}N]ammonia, [^{13}N]dinitrogen, and [^{14}C]glutamate by heterocysts isolated from *Anabaena cylindrica*. J Bacteriol **129**: 1545-1555.

Thurber RV, Willner-Hall D, Rodriguez-Mueller B, Desnues C, Edwards RA et al. (2009). Metagenomic analysis of stressed coral holobionts. Environ. Microbiol. **11**: 2148-2163.

Ueda Y, Konishi M, Yanagisawa S (2017) Molecular basis of the nitrogen response in plants. Soil Sci Plant Nutr **63**: 329-341.

Valladares A, Flores E, Herrero A (2008) Transcription activation by NtcA and 2-oxoglutarate of three genes involved in heterocyst differentiation in the cyanobacterium *Anabaena* sp. strain PCC 7120. J Bacteriol **190**: 6126-6133.

van Bel AJE (1984) Quantification of the xylem-to-phloem transfer of amino acids by use of inulin [^{14}C]carboxylic acid as xylem transport marker. Plant Sci Lett **35**: 81-85.

van Bel AJE (1990) Xylem-phloem exchange via the rays: the undervalued route of transport. J Exp Bot **41**: 631-644.

van der Heijden MG, Martin FM, Selosse MA, Sanders IR (2015) Mycorrhizal ecology and evolution: the past, the present, and the future. New Phytol **205**: 1406-1423.

van Velzen R, Doyle JJ, Geurts R (2019) A resurrected scenario: single gain and massive loss of nitrogen-fixing nodulation. Trends Plant Sci **24**: 49-57.

van Velzen R, Holmer R, Bu F, Rutten L, van Zeijl A et al. (2018) Comparative genomics of the nonlegume *Parasponia* reveals insights into evolution of nitrogen-fixing rhizobium symbioses. Proc Natl Acad Sci USA **115**: 4700-4709.

Vargas WA, Nishi CN, Giarrocco LE, Salerno GL (2011) Differential roles of alkaline/neutral invertases in *Nostoc* sp. PCC 7120: Inv-B isoform is essential for diazotrophic growth. Planta **233**: 153-162.

Vázquez-Bermúdez MAF, Herrero A, Flores E (2002) 2-Oxoglutarate increases the binding affinity of the NtcA (nitrogen control) transcription factor for the *Synechococcus glnA* promoter. FEBS Lett **512**: 71-74.

Vekemans D, Proost S, Vanneste K, Coenen H, Viaene T et al. (2012) Gamma paleohexaploidy in the stem lineage of core eudicots: significance for MADS-box gene and species diversification. Mol Biol Evol **29**: 3793-3806.

Venkateshwaran M, Jayaraman D, Chabaud M, Genre A, Balloon AJ et al. (2015) A role for the mevalonate pathway in early plant symbiotic signaling. Proc Natl Acad Sci USA **112**: 9781-9786.

Vessey JK, Pawlowski K, Bergman B (2005) Root-based N_2-fixing symbioses: Legumes, actinorhizal plants, *Parasponia* sp. and cycads. In Root Physiology: from Gene to Function, H Lambers and TD Colmer, eds (Dordrecht: Springer Netherlands), pp. 51-78.

Warshan D, Liaimer A, Pederson E, Kim SY, Shapiro N et al. (2018) Genomic changes associated with the evolutionary transitions of *Nostoc* to a plant symbiont. Mol Biol Evol **35**: 1160-1175.

Watanabe I, Van Hove C (1996) Phylogenetic, molecular and breeding aspects of *Azolla-Anabaena* symbiosis. In Pteridology in perspective, J Camus, M Gibby and R. Johns, eds (Royal Botanic Gardens: Kew, UK.), pp. 611-619.

Wei TF, Ramasubramanian TS, Golden JW (1994) *Anabaena* sp. strain PCC 7120 *ntcA* gene required for growth on nitrate and heterocyst development. J Bacteriol **176**: 4473-4482.

Werner GDA, Cornwell WK, Sprent JI, Kattge J, Kiers ET (2014) A single evolutionary innovation drives the deep evolution of symbiotic N_2-fixation in angiosperms. Nat Commun **5**: 4087.

Wilk L, Strauss M, Rudolf M, Nicolaisen K, Flores E et al. (2011) Outer membrane continuity and septosome formation between vegetative cells in the filaments of Anabaena sp. PCC 7120. Cell Microbiol **13**: 1744-1754.

Wolk CP (1968) Movement of carbon from vegetative cells to heterocysts in *Anabaena cylindrica*. J Bacteriol **96**: 2138-2143.

Wolk CP (1999) Heterocyst formation in Anabaena. In Prokaryotic Development, YV Brun and L Shimkets, eds (Washington, DC: American Society for Microbiology), pp. 83-104.

Wolk CP, Ernst A, Elhai J (1994) Heterocyst metabolism and development. In The Molecular Biology of Cyanobacteria, DA Bryant, ed (Dordrecht: Springer Netherlands), pp. 769-823.

Wolk CP, Thomas J, Shaffer PW, Austin SM, Galonsky A (1976) Pathway of nitrogen metabolism after fixation of ^{13}N-labeled nitrogen gas by the cyanobacterium, *Anabaena cylindrica*. J Biol Chem **251**: 5027-5034.

Xiao TT, Schilderink S, Moling S, Deinum EE, Kondorosi E et al. (2014) Fate map of *Medicago truncatula* root nodules. Development **141**: 3517-3528.

Xue L, Cui H, Buer B, Vijayakumar V, Delaux PM et al. (2015) Network of GRAS transcription factors involved in the control of arbuscule development in *Lotus japonicus*. Plant Physiol **167**: 854-871.

Yang W, Pollard M, Li-Beisson Y, Beisson F, Feig M et al. (2010). A distinct type of glycerol-3-phosphate acyltransferase with sn-2 preference and phosphatase activity producing 2-monoacylglycerol. Proc Natl Acad Sci USA **107**: 12040-12045.

Yano K, Yoshida S, Müller J, Singh S, Banba M et al. (2008). CYCLOPS, a mediator of symbiotic intracellular accommodation. Proc Natl Acad Sci USA **105**: 20540-20545.

Yoon HS, Golden JW (2001) PatS and products of nitrogen fixation control heterocyst pattern. J Bacteriol **183**: 2605-2613.

Yu N, Luo D, Zhang X, Liu J, Wang W et al. (2014). A DELLA protein complex controls the arbuscular mycorrhizal symbiosis in plants. Cell Res **24**: 130-133.

Yuan L, Loqué D, Kojima S, Rauch S, Ishiyama K et al. (2007) The organization of high-affinity ammonium uptake in *Arabidopsis* roots depends on the spatial arrangement and biochemical properties of AMT1-type transporters. Plant Cell **19**: 2636-2652.

Yuan X, Xiao S, Taylor TN (2005) Lichen-like symbiosis 600 million years ago. Science **308**: 1017-1020.

Zhang H, Forde BG (1998) An *Arabidopsis* MADS box gene that controls nutrient-induced changes in root architecture. Science **279**: 407-409.

Zhang H, Jennings A, Barlow PW, Forde BG (1999) Dual pathways for regulation of root branching by nitrate. Proc Natl Acad Sci USA **96**: 6529-6534.

Zhang L, Tan Q, Lee R, Trethewy A, Lee YH, Tegeder M (2010a) Altered xylem-phloem transfer of amino acids affects metabolism and leads to increased seed yield and oil content in *Arabidopsis*. Plant Cell **22**: 3603-3620.

Zhang Q, Blaylock LA, Harrison MJ (2010b) Two *Medicago truncatula* half-ABC transporters are essential for arbuscule development in arbuscular mycorrhizal symbiosis. Plant Cell **22**: 1483-1497.

Zhao J, Wolk CP (2008) Developmental biology of heterocysts. In Myxobacteria: Multicellularity and differentiation, DE Whitworth, ed (Washington, DC: ASM Press), pp. 397–418.

Zhao MX, Jiang YL, He YX, Chen YF, Teng YB et al. (2010) Structural basis for the allosteric control of the global transcription factor NtcA by the nitrogen starvation signal 2-oxoglutarate. Proc Natl Acad Sci USA **107**: 12487-12492.

Żyszka B, Anioł M, Lipok J (2017) Modulation of the growth and metabolic response of cyanobacteria by the multifaceted activity of naringenin. PLoS One **12**: e0177631.

1.8. Figures

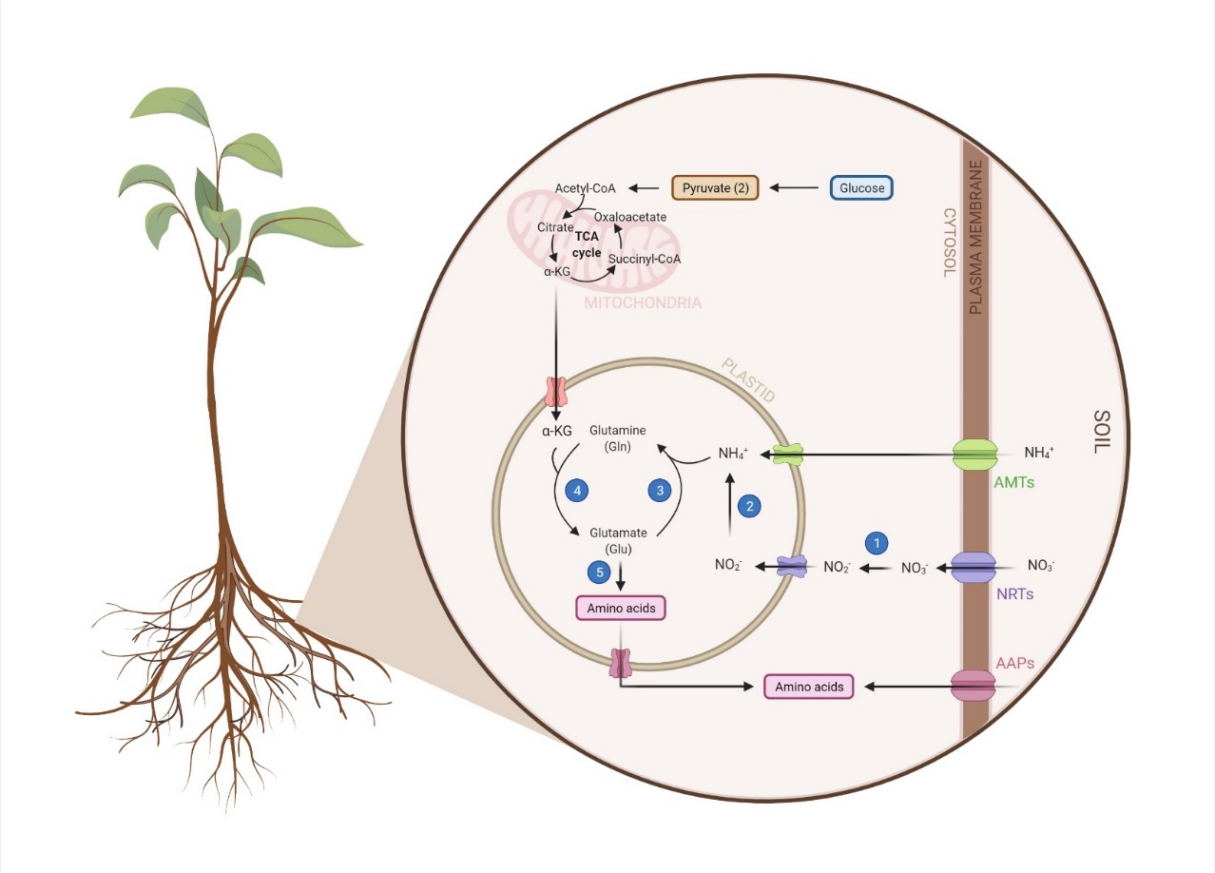


Figure 1.1. Nitrogen uptake and assimilation in plant roots. Inorganic nitrogen, nitrate (NO_3^-) and ammonium (NH_4^+) are absorbed through transporters at the root surface. Nitrate is reduced to nitrite (NO_2^-) by nitrate reductase (1), incorporated into a plastid and further reduced to ammonium via nitrite reductase (2). Ammonium, either absorbed directly or reduced from nitrate, is combined with glutamate and assimilated into glutamine by glutamine synthetase (3). An intermediate of the tricarboxylic acid cycle, 2-oxoglutarate (also known as α-ketoglutarate) receives an amide group to make glutamate via glutamate synthase (4) which in turn transfers an amino group to an organic acid to produce other amino acids via transaminases (5). These amino acids are subsequently transported throughout the plant. NRTs, Nitrate Transporters; AMTs, Ammonium Transporters; AAPs, Amino Acid Permeases. Adapted from Baslam et al. (2021) and created with BioRender.com.

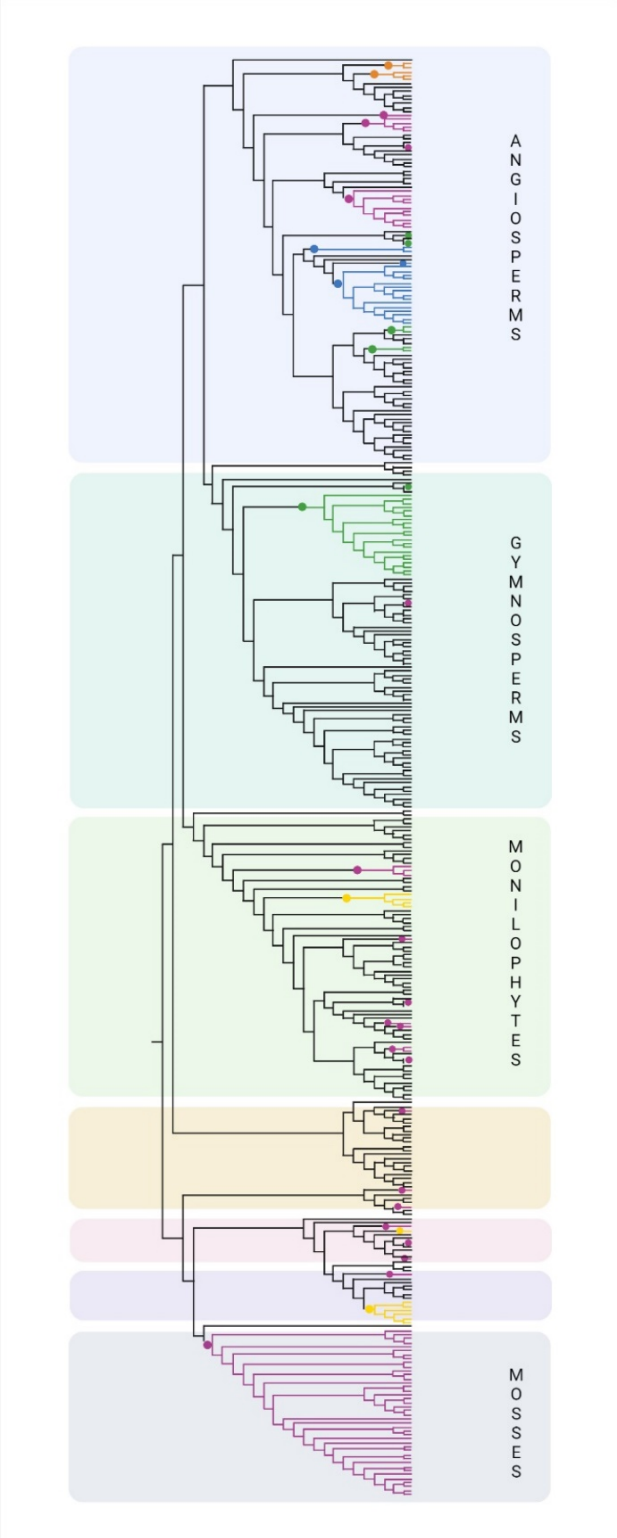


Figure 1.2. Distribution of symbioses across land plants. The phylogenetic tree depicts the theoretical plant phylogeny with the types of symbiosis formed by each species indicated. Loss of arbuscular mycorrhizae (pink), cyanobioses (yellow), ectomycorrhizal symbiosis (green), root-nodule symbiosis (blue) and orchid mycorrhiza and ericoid symbioses (orange). Adapted from Radhakrishnan et al. (2020) and created with BioRender.com.

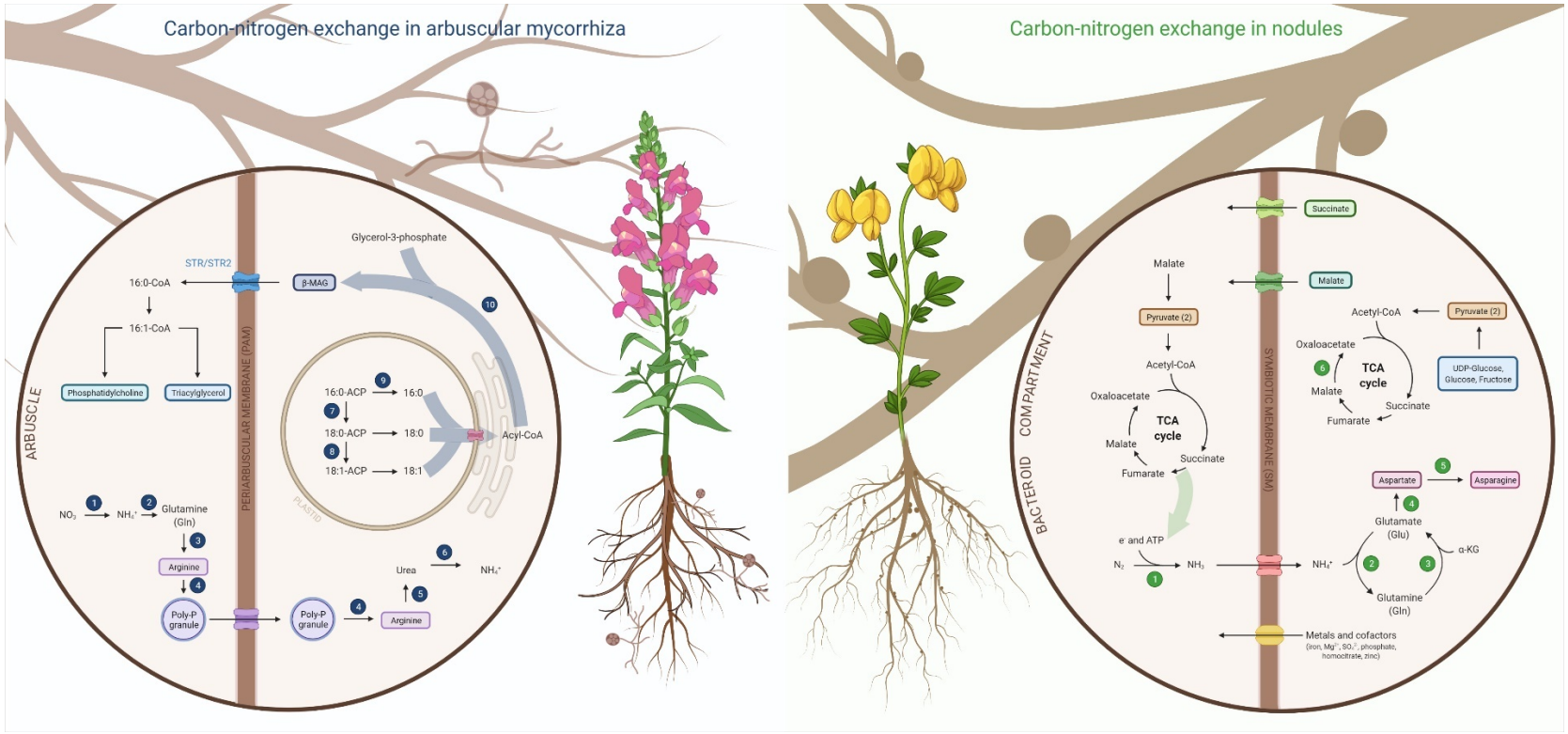


Figure 1.3. Carbon and nitrogen metabolism in arbuscular mycorrhiza (left) and nodules (right). *(Left)* Nitrate (NO_3^-) is absorbed by the extraradical hyphae and reduced to ammonium (NH_4^+) via nitrite reductase (1). Ammonium is combined with glutamate and assimilated into glutamine by glutamine synthetase (2). Glutamine transfers an amino group to produce arginine via transaminases (3). Arginine is incorporated into polyphosphate granules and likely transported across the periarbuscular membrane within these granules (4). Nitrogen is released from arginine as urea (5) and cleavage into ammonium (6) for plant use. Plants *de novo* synthesise fatty acids (palmitic acid, stearic acid and oleic acid) inside plastids (6-9) where they are exported to the endoplasmic reticulum, bound to coenzyme A (CoA) and sn-2 monoacylglycerol (β-MAG) is produced (10). The sn-2 monoacylglycerols (β-MAG) are transported across the periarbuscular membrane via the STR/STR2 transporter for use by arbuscular mycorrhiza. *(Right)* The ammonia (NH_4^+) produced by nitrogen fixation (1) is transported from the bacteriod compartment to the plant and assimilated by glutamine synthetase (2) and glutamate synthase (3) into glutamine and glutamate which are further converted to aspartate (4) and asparagine (5) by aspartate aminotransferase (AAT) and asparagine synthetase (AS), respectively. UDP-glucose, glucose and fructose are catabolised into pyruvate which enters the TCA cycle to produce intermediates such as malate, succinate or fumarate (6). These carbon sources, particularly malate and succinate, are transported across the symbiotic membrane to enter the TCA cycle in the bacteroid to be metabolised. Adapted from Parniske (2008), Bravo et al. (2017) and Liu et al. (2018) and created with BioRender.com.

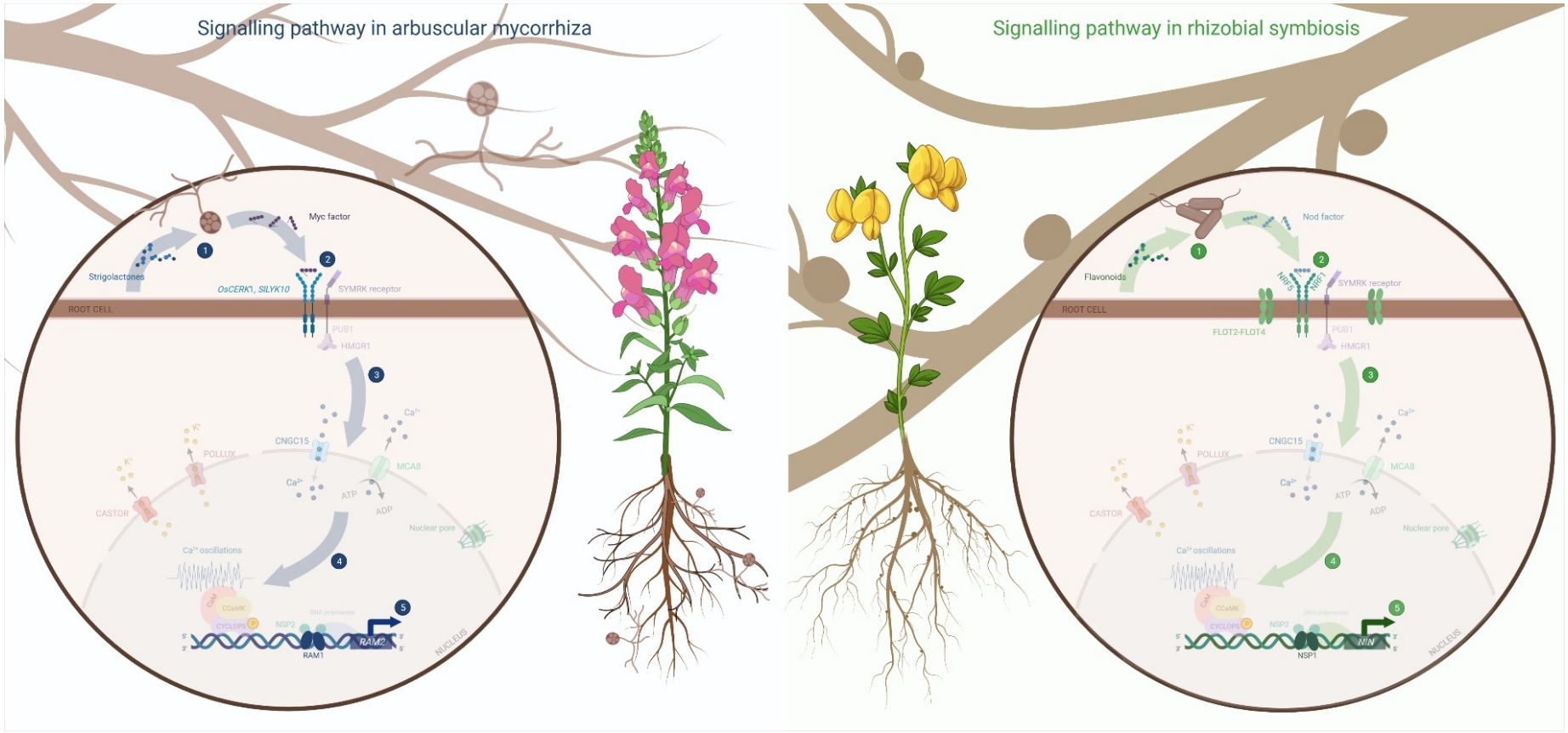


Figure 1.4. Common symbiosis signalling pathway for arbuscular mycorrhizal (left) and rhizobia symbioses (right). (1) Signal production: Strigolactones and flavonoids exuded by plant roots induces signal production by arbuscular mycorrhizae and rhizobia respectively. (2) Signal perception: Mycorrhizal (Myc) and nodulation (Nod) factors are perceived by LysM receptor-like kinases, *OsCERK1* (left), *LjNFR1* and *LjNFR5* (right), which together with SYMRK and HMGR induces nuclear calcium oscillations. SYMREM1 and FLOT2 and FLOT4 additionally associate with this receptor complex to aid in Nod factors recognition (right). (2) Calcium oscillation production: downstream of the symbiotic receptors, nuclear calcium oscillation is caused by the movement of calcium into the nucleoplasm. Potassium-permeable channels, CASTOR and POLLUX, compensate for the resulting depolarisation while MCA8 (a calcium symporter) pumps calcium back into the nuclear envelope. (3) Calcium oscillation perception: CCaMK forms a complex with phosphorylated CYCLOPS to decode calcium oscillations. (4) Transcription factor complex formation: the GRAS transcription factors, NSP1 and RAM1 (left) and NSP2 (right) forms a complex and controls expression of mycorrhiza-and rhizobium-specific genes. Adapted from Parniske (2008) and Oldroyd (2013) and created with BioRender.com.

1.9. Tables

Table 1.1. Cyanobacterial symbioses in eukaryotic hosts. Adapted from Rai et al. (2000), Adams et al. (2012) and Rai et al. (2019)

Eukaryotic hosts		Cyanobiont	Symbioses		Symbiotic tissue
Lineage	**Partner**		Location	Type	
Autotrophs Algae: Diatoms	Freshwater: *Rhopalodia, Epithemia, Denticulata, Climacodium, Leptocylindrus*[1]	Coccoid cyanobacteria	Extracellular	Facultative	Periplasmic space
	Marine: *Bacteriastrum, Chaetoceros*[2, 3]*, Hemiaulus, Rhizosolenia, Steptotheca*[1]*, Neostreptotheca*[1]*, Roperia*	*Richelia, Cyanothece, Calothrix*			
Dinoflagellates	*Ornithocercus, Histioneis, Parahistioneis, Citharistes, Dinophysis, Amphisolenia*[1, 4, 5]	*Ornithocercus, Histioneis, Parahistioneis, Citharistes, Dinophysis, Amphisolenia*	Intracellular/ Extracellular	Facultative	
Radiolarians	*Spongostaurus, Dictyocoryne*[6, 7]	*Prochlorococcus*			
Tintinnids	*Codonella*[1, 6]	*Synechococcus*			
Euglenoids	*Petalomonas sphagnophila*[8]	Unicellular cyanobacteria			
Foramenifera	*Marginopora vertebralis, Amphisorus hemprichii*[9]				
Bryophytes	Hornworts: *Anthoceros, Notothylas, Dendroceros, Phaeoceros*				
	Liverworts: *Blasia, Cavicularia*	*Nostoc* spp.	Extracellular	Facultative	Slime cavities[10]
	Mosses: *Sphagnum*				
Pteridophytes	*Azolla*	*Nostoc* spp.	Extracellular	Obligate	Epiphytic Leaf cavities[11, 12]
Gymnosperm	Cycadaceae: *Cycas*				
	Stangeriaceae: *Stangeria*	*Nostoc* spp.	Extracellular	Facultative	Coralloid roots
	Zamiaceae: *Bowenia, Ceratozamia, Dioon, Encephalartos, Lepidozamia, Macrozamia, Microcycas, Zamia*				
Angiosperm	*Gunnera*[13, 14]	*Nostoc* spp.	Intracellular	Facultative	Stem glands
Heterotrophs Fungi	Bipartite cyanolichens: *Pseudocyphellaria intricate, Stricta* spp.*, Stricta weigelii, Peltigera* spp.*, Collema* spp.	*Nostoc* spp., *Scytonema*, *Fishcerella*, *Calothrix*, *Dichothrix*, *Strigonema*, *Tolypothrix*	Extracellular	Facultative	Fungal thallus
	Tripartite cyanolichens: *Lobaria* spp.*, Pseudocyphellaria aurata, Nephroma arcticum*				
	Geosiphon pyriformis	*Nostoc punctiforme*	Intracellular	Obligate	Bladders
Sponges	*Candidaspongia flabellate*[15]*, Aplysina* spp.[16]*, Neopetrosia subtriangularis*[16]*, Lamellodysidea chlorea, Xestospongia* spp.*, Ircinia fasciculata*	*Aphanocapsa* spp., *Synechococcus* spp., *Prochloron* spp., *Oscillatoria spongeliae*	Intracellular	Facultative	Cyanocytes
Corals	*Montastreae* spp.[17, 18]*, Diploria strigose, Porites* sp.[19]*, Montipora* spp.[20]*, Oculina patagonica*[21]*, Desmophyllum dianthus, Caryophyllia huinayensis*[22]	*Synechococcus* spp., *Prochlorococcus* spp., *Prochloron* spp., *Synechocystis trididemni*	Intracellular/ Extracellular	Facultative/ Obligate	Surface, cloacal cavity, tunic
Ascidians	Didemnidae: *Didemnum, Trididemnum, Lissoclinum, Diplosoma*[23]				
Echiuroid worms	*Ikedosoma gogoshimense, Bonellia fuliginosa*[1]	Unicellular cyanobacteria	Extracellular	Facultative	Surface
Isopods	*Santia*[24]	Unicellular cyanobacteria			

Hydroid	*Eudendrium racemosum*[25]	*Oscillatoria lutea, Spirulina subsalsa*	Extracellular	Facultative
Midge larvae	*Cricotopus nostocicola*[26]	*Nostoc parmeloides*	Extracellular	Facultative

[1] Carpenter et al. (2002); [2] Gómez et al. (2005); [3] Foster et al. (2009); [4] Jyothibabu et al. (2006); [5] Tarangkoon et al. (2010); [6] Foster et al. (2006a); [7] Foster et al. (2006b); [8] Schnepf et al. (2002); [9] Lee (2006); [10] Meeks (1990b); [11] Braun-Howland et al. (1990); [12] Watanabe et al. (1996); [13] Bergman et al. (1992a); [14] Bergman et al. (1992b); [15] Burja et al. (2001); [16] Erwin et al. (2008); [17] Lesser et al. (2004); [18] Lesser et al. (2007); [19] Thurber et al. (2009); [20] Olson et al. (2009); [21] Fine et al. (2002); [22] Försterra et al. (2008); [23] Hirose et al. (2007); [24] Lindquist et al. (2005); [25] Romagnoli et al. (2007); [26] Adams et al. (2012)

Table 1.2. Features of cyanobacterial symbioses. Adapted from Rai et al. (2000) and Rai et al. (2019)

Eukaryotic hosts	**Cyanobiont**			**Nutrient exchange**	
	Photosynthesis activity	**Nitrogen-fixation**	**Heterocyst frequency**	**Fixed carbon (C) transfer**	**Fixed nitrogen (N) transfer**
Autotrophs Algae	Active	Active GS repressed	Unchanged	Indeterminate	As ammonia (NH_3) to host
Bryophytes	Inactive Rubisco inhibited	Active GS decreased	40-45%	As sucrose to cyanobiont	As ammonia (NH_3) to host[1, 2]
Azolla	Inactive Rubisco repressed	Active GS decreased	25-30%		
Cycads	Inactive Rubisco normal	Active GS as in free-living	~45%	As fixed carbon to cyanobiont	As amino acids (glutamate and/or citrulline) to host[3]
Gunnera	Inactive Rubisco normal	Active GS decreased	60-80%		As ammonia (NH_3) to host[4]
Heterotrophs Fungi	Active	Active GS repressed	Bipartite lichen: Unchanged Tripartite lichen: 15-35%	As glucose to host No transfer of fixed carbon occurs	As ammonia (NH_3) to host

[1] Meeks et al. (1985); [2] Meeks et al. (1987); [3] Pate et al. (1988); [4] Silvester et al. (1996)

CHAPTER 2

ANATOMICAL AND GENE EXPRESSION ANALYSES OF CORALLOID ROOTS IN THE CYCAD *ENCEPHALARTOS NATALENSIS* WITH NITROGEN-FIXING CYANOBIONTS

Cassandra Schoeman[1], Danielle Roodt[1], Oliver Bezuidt[2], Desré Pinard[1], Thulani Makhalanyane[3], Eshchar Mizrachi[1*]

[1] *Department of Biochemistry, Genetics and Microbiology, Forestry and Agricultural Biotechnology Institute, University of Pretoria, Private bag X28, Pretoria 0028, South Africa*

[2] *Department of Biochemistry, Genetics and Microbiology, Centre for Microbial Ecology and Genomics, University of Pretoria, Private bag X28, Pretoria 0028, South Africa*

[3] *Department of Biochemistry, Genetics and Microbiology, University of Pretoria, Pretoria 0028, South Africa*

*Corresponding author email: eshchar.mizrachi@fabi.up.ac.za

This chapter has been prepared in the format of a manuscript and parts of it will be submitted to *New Phytologist*. At the time of submission, the manuscript was in preparation. I drafted the manuscript, sampled the material, prepared the RNA and DNA samples for sequencing, performed the microscopy, constructed the comprehensive candidate gene list and analysed the data. Danielle Roodt participated in material sampling and RNA preparation, assembled and annotated the transcriptomic data, helped analyse the data, helped co-supervise, plan and design the research as well as helped edit the manuscript. Desré Pinard developed a custom R app which was used for preferential gene analysis while Oliver Bezuidt provided the preliminary metagenomic data. Thulani Makhalanyane helped supervise, plan, and design the research and edit the manuscript. Eshchar Mizrachi conceived of and supervised the study, helped plan and design the research, analyse the data and draft and edit the manuscript.

2.1. Abstract

Cyanobacterial symbiosis evolved independently across unrelated lineages of land plants, unlike those involving symbioses with arbuscular mycorrhizae, rhizobia, *Frankia* diazotrophic bacteria and flowering plants. However, information regarding the precise genes and processes linked to cyanobacterial partnerships are dated and lacking. We addressed this knowledge deficit by investigating the metabolic function, transcriptional regulation and morphology of a suspected active nitrogen-fixing cyanobiosis in the cycad (*Encephalartos natalensis*). We investigated gene expression in symbiotic and control tissues from four cyanobacterial partnership plant lineages, including a comprehensive tissue-specific transcriptomic dataset from cycad. We analysed genes preferentially expressed in symbiotic tissue from *Anthoceros punctatus*, *Azolla filiculoides*, *E. natalensis* and *Gunnera perpensa*, along with known genes involved in other symbiotic partnerships. These data were combined with light and transmission electron microscopy on selected cyanobioses as supporting evidence towards an active nitrogen-fixing symbiosis. We provide evidence of an active and ongoing nitrogen-fixing symbiosis in the cycad, *E. natalensis*, and find tissue-specific regulation of citrulline and ornithine metabolising enzymes in coralloid and control root tissues. We also demonstrate evidence for the high conservation of common symbiotic pathway genes in the cycad-cyanobacterial partnership, despite their absence or loss in other cyanobiosis plant lineages. Together these results suggest that cycads may have neofunctionalised the same genes and pathways for its cyanobacterial partnership from an existing arbuscular mycorrhizal symbiosis, apart from those involved in carbon-nitrogen source biosynthesis. Addressing these open questions has provided new insight regarding cycad-cyanobacterial symbioses and their evolution.

2.2. Introduction

Plants have been foundational to terrestrial ecosystems since their colonisation of land more than 450 million years ago (mya; Gensel, 2008). Such a successful colonisation occurred only once in the Viridiplantae, and it has been proposed that this transition from exclusively aquatic environments to land was facilitated by symbiotic associations (Pirozynski and Malloch, 1975; Heckman et al., 2001). Plant diversification has since led to the emergence of diverse symbiotic relationships, ranging from loosely associated extracellular to fully mutualistic intracellular associations (reviewed in Vessey et al., 2005; Santi et al., 2013). Symbioses are widespread among extant plant species, where it has not only shaped plant diversity but has also had an unmistakable impact on nutrient acquisition by plants (reviewed in MacLean et al., 2017; Roy et al., 2020). As primary producers, plants rely on large above- and belowground structures to capture the nutrients needed to support their growth and development (reviewed in Oldroyd and Leyser, 2020). Although plants are dependent on several mineral nutrients from the soil, the key nutrient that limits their overall growth and productivity is nitrogen (N; Vitousek and Howarth, 1991; Ågren et al., 2012). As such, many land plants have evolved associations with microorganisms, such as mycorrhizal fungi

and diazotrophic (i.e., nitrogen-fixing) bacteria, to improve nutrient acquisition to overcome nitrogen shortages in particular (Soltis et al., 1995; Swensen, 1996; Brundrett, 2002).

Our understanding of the molecular mechanisms governing the establishment and function of symbiosis is limited. In a recent synthesis, Martin et al. (2017) suggest that the data are largely restricted to intracellularly housed arbuscular mycorrhizal (AM) and root-nodule symbiosis (RNS) in model legume species and a few other angiosperms. Arbuscular mycorrhizal symbioses are the most common; known to occur in most extant plants (72-80%), except for moss (Bryophyta; Brundrett, 2002; Brundrett and Tedersoo, 2018). Mycorrhizal associations, having been retained throughout the evolutionary history of land plants with the exception of some agricultural crops and weeds including the model plant *Arabidopsis thaliana* (Tester et al., 1987; Brundrett, 2002; Brundrett and Tedersoo, 2018), point towards an ancestral origin (Brundrett, 2002; Bravo et al., 2016). Whereas root-nodule symbioses (symbiotic interaction established by rhizobia and *Frankia* nitrogen-fixing bacteria) occur in flowering plant families such as legumes, *Parasponia*, and many actinorhizal plants (Doyle, 1994; Swensen, 1996). Even though nodulation represents a major event in angiosperm evolution, there is much debate over its origin (reviewed in van Velzen et al., 2019). Previous studies suggest that a precursor state evolved in a common ancestor of nodulating plants around 110 mya (Soltis et al., 1995; Werner et al., 2014). Evolution continued either as massive parallel losses in most descendants or as independent evolution in others (van Velzen et al., 2019). Despite more than 100 million years of divergence, the main nodulating lineages share similarities in both development and structure (Doyle, 1994; Sprent et al., 2013; Svistoonoff et al., 2014). This suggests a stronger phylogenetic link than one would expect from independent evolution (van Velzen et al., 2019). However, differences in colonisation strategies, type of microsymbionts, and nodule ontology are often regarded as evidence for the independent evolution of nodulation (Pawlowski and Demchenko, 2012; Svistoonoff et al., 2014). Despite differences in these intracellular symbiotic associations, recent genetic studies have identified a suite of conserved genes (the common symbiosis pathway, CSP), which regulate their establishment and maintenance complemented by lineage-specific adjustments (reviewed in Oldroyd, 2013; Radhakrishnan et al., 2020; Delaux and Schornack, 2021).

Comparative phylogenomics, and various forward- and reverse-genetic approaches in angiosperms, have uncovered nearly 200 genes required for symbiosis establishment and maintenance (reviewed in MacLean et al., 2017; Roy et al., 2020). Most of the research has focused on the conserved symbiosis signalling pathway, the so-called common symbiosis pathway (CSP). This pathway is required for symbioses with arbuscular mycorrhizal fungi and rhizobia (reviewed in Oldroyd, 2013). Several studies have accelerated our understanding of this pathway and, more recently, of the symbiosis-specific downstream responses that it controls, for example, the bidirectional exchange of nutrients across symbiotic interfaces (reviewed in MacLean et al., 2017; Roy et al., 2020). In most plants, sucrose is the primary carbon resource supplied to roots from the shoot (Giaquinta, 1983). Sucrose is unloaded via several Sucrose Transporters (SUTs), Sugars Will Eventually be Exported Transporters (SWEETs) and Monosaccharide Transporters (MSTs) where it is cleaved into uridine diphosphate (UDP) glucose, glucose,

and fructose to supply arbuscular mycorrhizal fungi and nitrogen-fixing bacteroids (Doidy et al., 2012). The *de novo* fatty acid biosynthesis of β-monoacylglycerol (β-MAG), by acyl-ACP thioesterases (FatM), and its subsequent export to the apoplast via periarbuscular membrane-localised ABC transporters, named Stunted arbuscule (STR) and STR2, are unique to mycorrhizal plants (Zhang et al., 2010; Gutjahr et al., 2012; Bravo et al., 2017). Driven by carbon supply from the root, the fungus develops an extraradical mycelium into the surrounding soil to capture ammonium (NH_4^+; López-Pedrosa et al., 2006; Pérez-Tienda et al., 2011; Calabrese et al., 2016) whereas bacteroids within symbiosomes fix nitrogen (reviewed in Liu et al., 2018). Ammonia (NH_3), either fixed or captured from the soil, is exported across the symbiotic interface into the cytosol of infected host cells, where it is assimilated into glutamine (Gln) and glutamate (Glu; Govindarajulu et al., 2005; Jin et al., 2005; Kobae et al., 2010). Amino acids aspartate (Asp) and asparagine (Asn), and/or ureides are then exported to the rest of the plant (Liu et al., 2018).

A third nitrogen-fixing symbiosis exists, in addition to the well-studied mycorrhizal and nitrogen-fixing root nodule symbioses, that have received far less attention: symbiosis with nitrogen-fixing cyanobacteria. This association has independently evolved in plant lineages as diverse as liverworts and hornworts (reviewed in Meeks, 1990), species in the *Azolla* genus of the pteridophytes (reviewed in Braun-Howland and Nierzwicki-Bauer, 1990), the cycad lineage of the gymnosperms (reviewed in Lindblad and Bergman, 1990), and species in the angiosperm genus *Gunnera* (reviewed in Bonnett, 1990). Although these plant species represent diverse lineages, the cyanobacteria are, in all cases, accommodated in pre-existing cavities that are intercellular structures (with the exception of *Gunnera* wherein cyanobionts are housed intracellularly) filled with a secondary metabolite and polysaccharide-rich mucilage (reviewed in Rai et al., 2000; Rai et al., 2002; Rai et al., 2019). Pre-existing structures are present prior to cyanobacterial invasion and develop into the symbiotic structures upon invasion. These structures include slime cavities in bryophytes (Meeks, 1990), leaf cavities in ferns (Braun-Howland and Nierzwicki-Bauer, 1990), specialised roots in gymnosperms (Lindblad and Bergman, 1990), and intracellular stem glands in angiosperms (Bonnett, 1990). This contrasts with legume or actinorhizal nodules which only differentiate upon symbiont invasion. Once inside these cavities, chemical signals produced by the hosts (hormogonia-inducing and -repressing factors) promote the differentiation of the cyanobacteria into heterocysts – specialised cyanobacterial cells that fix dinitrogen (reviewed in Rai et al., 2000; Rai et al., 2002; Rai et al., 2019). While the identity of hormogonia-inducing and -repressing factors remain mostly unknown, 1-palmitoyl-2-linoleoyl-*sn*-glycerol was recently identified and characterised as a potential hormogonia-inducing factor (HIF) in the cycad, *Cycas revoluta* (Hashidoko et al., 2019). In cycads, this relationship is established through specialised roots, called coralloid roots. Several studies suggest that this relationship is absent from all other gymnosperms and of particular interest to this study (Lindblad and Bergman, 1990; Norstog and Nicholls, 1997; Lindblad, 2009). Although it is well-known that precoralloid roots represent an early stage of coralloid root development (Milindasuta, 1975; Ahern and Staff, 1994), the host genetic and molecular mechanisms determining symbiosis with cyanobacteria remain largely

unknown. Knowledge about the genetic mechanisms involved in carbon-nitrogen exchange is lacking, apart from speculation on putative carbon or nitrogen sources (reviewed in Rai et al., 2000; Rai et al., 2002; Rai et al., 2019).

In this study, we performed light and transmission electron microscopy on cyanobacteria containing tissue from the cycad *Encephalartos natalensis*, the angiosperm *Gunnera perpensa*, and the aquatic fern *Azolla pinnatta* spp. *africana*. We also constructed a comprehensive gene catalogue from ten *E. natalensis* tissues and organs to investigate the molecular regulation of plant-cyanobacterial interactions in coralloid roots. In addition, a comparative transcriptomics analysis was performed across land plants hosting cyanobacteria as well as rhizobia and *Frankia*, which necessitated the construction of a gene catalogue from cyanobiont containing and control tissues of the angiosperm *G. perpensa*. We theorised that we would observe a high frequency of heterocyst in the cyanobacteria containing tissues of naturally growing plant species investigated, suggesting active nitrogen supply by the cyanobiont to the plant. We further hypothesised that specific pre-existing genes and pathways, including the CSP ancestrally evolved for mycorrhizae associations, would be co-opted to interact with cyanobacteria, although to what extent was unknown. We also interrogated evidence for molecular mechanisms governing carbon and nitrogen exchange, which has been a long-standing open question. Addressing these questions provides new insight into how plant-cyanobacterial symbioses are generally and specifically governed, how they evolved, and how to view these considering other plant-microbe symbioses in an evolutionary context.

2.3. Results

2.3.1. Heterocysts are dominant and active in all cyanobacteria containing tissues from naturally growing *Encephalartos natalensis*, *Gunnera perpensa* and *Azolla pinnatta* spp. *africana*

Prior to the construction of gene catalogues from naturally growing *E. natalensis* and *G. perpensa*. We profiled tissues and organs from these plants, using light and transmission electron microscopy. The presence of cyanobacteria and the relative abundances of heterocysts may be a reliable indicator that nitrogen is actively being supplied to the plant host. We, therefore, visualised plastic embedded symbiotic tissue and their cyanobacterial morphology, using toluidine blue O and 4% aqueous uranyl acetate and Reynold's lead citrate staining. The morphology was compared with those of non-symbiotic *E. natalensis* and *G. perpensa* samples and the cells were quantified. Given that recent 16S rRNA gene amplicon-based studies have identified other bacterial phyla in the coralloid roots of some cycad species (Zheng et al., 2018; Suárez-Moo et al., 2019) while morphological studies indicated that these species also harbour fungi (Muthukumar and Udaiyan, 2002; Fisher and Vovides, 2004). This also gave us the opportunity to determine whether *E. natalensis* coralloid roots harbour any other microbes. For comparison, naturally growing *Az. pinnatta* spp. *africana* was also profiled. We found evidence of cyanobacterial cells in the cyanobacterial zone of *E. natalensis* coralloid roots (Figure 2.1 A and B), the stem glands of *G. perpensa* (Supplementary Figure S2.1 A and B) and the leaf cavity of *Az. pinnatta* spp. *africana* (Supplementary

Figure S2.1 F and G). In both *E. natalensis* and *G. perpensa*, the presence of mucilage was observed although this was absent from *Az. pinnatta* spp. africana (Figure 2.1 and Supplementary Figure S2.1).

Precoralloid and coralloid roots are thought to be highly specialised types of lateral roots emerging as "transformed normal, lateral root apices" (Milindasuta, 1975; Ahern and Staff, 1994). We, therefore, included *E. natalensis* primary and lateral roots as control tissue for comparative purposes (Supplementary Figure S2.2 A and B). Primary and lateral *E. natalensis* roots differed morphologically and differed between precoralloid and coralloid roots. Primary roots contained large parenchymal cells, which harboured starch granules (Supplementary Figure S2.2 A), while lateral roots consisted of oblong parenchymal cells with numerous interspersed tannin channels (Supplementary Figure S2.2 B). Our observations confirm distinct differences in the morphology of lateral, and precoralloid or coralloid roots, such as in vascular tissue organisation (Supplementary Figure S2.3).

E. natalensis coralloid roots harbour the so-called cyanobacterial zone. This is the site of cyanobacterial colonisation located between the inner and outer root cortex (Figure 2.1 A). In all coralloid roots, cyanobionts occupied mucilage-filled intercellular spaces within the cyanobacterial zone which are separated by elongated, columnar host cells (Figure 2.1 A and B). The same mucilage-filled intercellular spaces separated by columnar cells were seen in the precoralloid roots of *E. natalensis* (Figure 2.1 C and D). Therefore, the cyanobacterial zone of both coralloid and precoralloid roots revealed the presence of these interspersed mucilage-filled intercellular spaces within columnar cells irrespective of cyanobacterial presence (Figure 2.1 A-D). The cyanobacterial zone is initiated as a new meristematic layer that de-differentiated in the cortex as observed in precoralloid roots (Figure 2.1 E). Cyanobionts within the cyanobacterial zone of *E. natalensis* coralloid roots consisted of vegetative cells (Figure 2.2 C) and heterocysts (Figure 2.2 A and B). The vegetative cells contained thylakoids with numerous phycobilisomes, carboxysomes and cyanophycin granules (Figure 2.2 C) with narrow heterocyst poles also present between vegetative cells and heterocysts (Figure 2.2 B). Interestingly, we also saw a similar narrow "plug" between two heterocysts (Figure 2.2 A), which has also been reported for other cycad species with double heterocysts (Caiola, 1974; Cheng, 1982; Lindblad et al., 1985).

G. perpensa house their cyanobionts intracellularly within stem glands (Supplementary Figure S2.1 A and B). Stem glands consisted of clusters of cyanobacterial containing cells with interspersed non-symbiotic cells surrounding these clusters (Supplementary Figure S2.1 A). A sheath composed of a few layers of oblong, non-symbiotic cells encircled those that contained cyanobacteria creating a barrier around the cyanobacterial containing tissue. These barriers were clearly distinguished from the surrounding stem parenchymal tissue. These comprised of large isodiametric cells containing starch granules and served as control tissue (Supplementary Figure S2.2 C). The *Az. pinnatta* spp. *africana* symbiotic tissue (i.e. the dorsal leaf lobe), as is reported for all *Azolla* species, contained a leaf cavity (Supplementary Figure S2.1 F and G) – an extracellular leaf compartment surrounded by two epidermal cell layers (Veys et al., 1999; Lechno-Yossef and Nierzwicki-Bauer, 2002). The leaf cavity, as shown in Supplementary Figure S2.1 F and G, contained symbionts including cyanobionts and/or other

bacteria. These were localised to the periphery of the cavity wall, consistent with previous reports (Lechno-Yossef and Nierzwicki-Bauer, 2002; van Hove and Lejeune, 2002). Cyanobionts within the leaf cavity of *Az. pinnatta* spp. *africana* and stem glands of *G. perpensa*, as with those observed in *E. natalensis*, consisted of vegetative cells (Supplementary Figure S2.1 D and I) and heterocysts (Supplementary Figure S2.1 C and H). Again, in both species, the vegetative cells contained thylakoids with numerous phycobilisomes, carboxysomes and cyanophycin granules (Supplementary Figure S2.1 D and I). In their free-living state, heterocysts differentiate at semiregular intervals between stretches of vegetative cells. Therefore, heterocyst poles with narrow "plugs" exist to limit the contact between these cell types (Kumar et al., 2010; Flores et al., 2019), which could be seen for *G. perpensa* (Supplementary Figure S2.1 C and E). This is contrary to those between vegetative cells as observed in *Az. pinnatta* spp. *africana* (Supplementary Figure S2.1 J).

To provide supporting evidence that active nitrogen fixation by cyanobionts is likely occurring in the samples analysed, we determined the proportion of heterocysts in the symbiotic tissue of *E. natalensis*, *G. perpensa* and *Az. pinnatta* spp. *africana*. In plant partnerships, heterocyst frequencies as high as 30 - 80% are indicative of active nitrogen fixation (reviewed in Rai et al., 2000; Rai et al., 2002; Rai et al., 2019). Cell count analysis suggests that the ratio of vegetative cells to heterocysts was 1:2 in *E. natalensis*, 2:1 in *G. perpensa* and the highest in *Az. pinnatta* spp. *africana*, with a ratio of 6:1 (Figure 2.3 and Supplementary Table S2.1). Taken together, an average heterocyst percentage of 57.6%, 33.7% and 25.6% was observed in *E. natalensis*, *G. perpensa* and *Az. pinnatta* spp. *africana*, respectively (Figure 2.3 and Supplementary Table S2.1). Together, these results provide supporting evidence towards active nitrogen fixation occurring in the cyanobionts in *E. natalensis* and *G. perpensa* intended for transcriptomic analysis, as evidenced by a remarkably large proportion of heterocysts (i.e., heterocysts dominated these plant tissues and organs) in each host despite their morphological differences in other ways.

2.3.2. The tissue-specific transcriptome assembly of *Encephalartos natalensis* highlighted the uniqueness of coralloid root biology and its carbon and nitrogen exchange genes

To investigate plant host genes preferentially expressed in the *E. natalensis*-cyanobacterial symbiosis and to identify nutrient exchange related genes, we performed mRNA-seq in replicated samples of cyanobacteria-containing (coralloid root) and control (upper primary root) tissues. mRNA-seq of the cyanobacteria containing and control tissue samples obtained from *E. natalensis* plants yielded ~140 million and ~200 million clean reads for *E. natalensis* coralloid root and upper primary root, respectively, with average GC contents of 44.3% and 44.5% (Supplementary Table S2.2). Additional mRNA-seq data obtained for mature and immature leaflet, mature and immature rachis, leaf base, stem, lower primary root, and germinating seed were used to assemble the transcriptome and are summarised in Supplementary Table S2.2. Following trimming, *de novo* assembly, annotation and clustering, a reference *E. natalensis* transcriptome was generated (Supplementary Table S2.3). The resultant *E. natalensis* assembly contained 34,713 contigs (of which 24,996 contigs were functionally annotated) with a mean length of 1,371 bp (N50 = 2,222 bp and N75 = 3,237 bp; see Materials and Methods for details), as

expected since no recent whole-genome duplication events occurred (Roodt et al., 2017). Furthermore, a BUSCO analysis against the plant dataset (Simão et al., 2015) revealed that the gene catalogue contained 1,238 (85.9%) complete genes, low duplication (8.2%), low fragmentation (3.3%) and some missing genes (10.7%). For downstream comparative analyses, a total of 5,246 transcripts were identified by DESeq2 (Love et al., 2014) as differentially expressed ($p \geq 0.05$) when comparing *E. natalensis* coralloid root tissue and upper primary root tissue.

Principal components analysis (PCA) was applied to the *E. natalensis* transcriptome to identify structure within the dataset (Supplementary Figure S2.4 A). The PCA plot indicated significant ($p \geq 0.05$) variation between photosynthetic and sink tissues and organs with photosynthetic tissues themselves showing distinct separation of leaflet and rachis samples (Supplementary Figure S2.4 A). Coralloid root samples, which clustered quite close together, were distinct and separate from all other tissues and crucially, from all other root types (Supplementary Figure S2.4 A). Despite sequencing fails for immature leaflet, mature rachis, leaf base, stem and lower primary root tissues, the tissues clustered primarily by biological replicates (from which there were at least two biological replicates per tissue) creating distinct and separate groups for each tissue as expected. To further investigate coralloid root distinctness from all other root types, a transcriptional profile was generated from the *E. natalensis* transcriptome to show distinct tissue expression patterns (Supplementary Figure S2.4 B). We would expect to see root tissues cluster together because these structures are expected to be similar based on developmental signals retained in these tissues. However, the coralloid root tissue clustered separately from other root tissue, instead they were most similar to germinating seed (Supplementary Figure S2.4 B). The PCA plot matches the expression profile patterns predicted with hierarchical clustering indicating significant variation between coralloid roots and primary roots, further emphasising coralloid roots as a unique root tissue (Supplementary Figure S2.4 A and B).

To gain insights into the carbon and nitrogen source investments during this active cyanobiosis, we performed a metabolic pathway enrichment analysis of differentially expressed genes in coralloid and primary root. We specifically queried differentially expressed genes with tissue-specific expression or preferential gene expression (Figure 2.4 A). To ensure that genes with low expression levels were filtered out, we only included relatively highly expressed genes (50th percentile and over) with tissue-specific expression when compared to all other tissues sequenced. The heat map analysis presented in Figure 2.4 A showed a clear profile of highly expressed genes (cluster 3) in coralloid root, while cluster 1 showed high expression of genes in the control tissue, upper and lower primary root. Analysis of the coralloid and primary root differentially expressed *Arabidopsis* homologs using KEGG (Kanehisa and Goto, 2000) indicated that these clusters were highly enriched for processes including amino acid biosynthesis and carbohydrate metabolism. Several pathways showed a high proportion of preferentially expressed genes (≥ 50% of expression) in coralloid root tissue (Supplementary Figures S2.5 and S2.6). As expected, genes preferentially expressed in coralloid root, but not primary root (control) tissue were related to overall nitrogen metabolism (particularly ammonia assimilation) as well as citrulline and ornithine biosynthesis (intermediates of the arginine biosynthesis pathway; Figure 2.5).

The functional importance of these genes in coralloid roots is highlighted by the fact that most enzymatic steps for ammonia (NH_3) assimilation into ornithine and citrulline contained genes expressed above the 75th, 95th (5 steps) and 99th percentile (2 steps; Figure 2.5). Many steps except for the alternative pathway to ammonia production from nitrate (NO_3^-) and nitrite (NO_2^-) via NRTs → NR → NiR, contained members showing highly coralloid root-specific expression (>50% of expression in coralloid root compared to other tissues). Although it is well-established that the transported nitrogen is in the organic form (glutamine, citrulline and possibly glutamate; (Rai et al., 2019), it still remains unclear whether (1) the cyanobiont produces these amino acids and subsequently provides it to the cycad (Pate et al., 1988; Rai et al., 2000; Rai et al., 2002; Rai et al., 2019), (2) the cyanobiont provides ammonia (NH_3) and the cycad uses this precursor to synthesise the amino acids, or (3) the cyanobiont produces NH_3 which is subsequently converted to NO_3^- by nitrifying microbes (theorized in Chang et al. 2019). The canonical thought is that in all other species their cyanobionts provide NH_3 (Rai et al., 2000; Rai et al., 2002; Rai et al., 2019). The complete absence of any detectable expression of NO_3^- and NO_2^- transporters or metabolising enzymes in coralloid root likely rules out scenario 3 (Figure 2.5). Our metabolic pathway enrichment analysis indicated that NH_3 conversion into glutamine and glutamate by glutamine synthetase [EC: 6.3.1.2] and glutamate dehydrogenase [EC: 1.4.1.3] were highly expressed, 100th and 83rd percentile respectively, in coralloid root which strongly supports scenario 2. Although heterocysts are capable of producing glutamine, aspartate and arginine (Wolk et al., 1976; Flores et al., 2019), we saw no evidence of highly coralloid root-specific expression (>50% of expression in coralloid root compared to other tissues) of aspartate aminotransferase (GOT2; Figure 2.5). Therefore, while scenario 1 cannot conclusively be ruled out, the relatively high expression of glutamine synthetase [EC: 6.3.1.2] and glutamate dehydrogenase [EC: 1.4.1.3] along with the absence of a detectable level of aspartate aminotransferase [EC: 2.6.1.1] provides conclusive evidence that *E. natalensis* is at least partially (if not entirely) responsible for producing glutamine, glutamate and citrulline from NH_3.

Perhaps unexpected was the root-specific investment in starch and sucrose metabolism such that enzymes involved in catabolism to simple sugars (fructose, mannose, or glucose) are coralloid root-specific, whereas those involved in anabolism to more complex sugars are not (Supplementary Figure S2.6). Fructose and mannose metabolism enrichments supports the breakdown of starch to simple sugars in coralloid root (Supplementary Figure S2.6). For example, the breakdown of D-glucose and D-fructose into D-glucose-6P and D-fructose-6P by hexokinase [EC: 2.7.1.1] or fructokinase [EC: 2.7.1.4] showed highly coralloid root-specific expression (>50% of expression in coralloid root compared to other tissues; Supplementary Figure S2.6). While legumes provide malate and/or succinate to their symbionts and fatty acids, most likely β-MAG, is provided to arbuscular mycorrhizal fungi (Day and Copeland, 1991; Bravo et al., 2017), the complete absence of any detectable expression of such metabolising enzymes in coralloid root likely rules out these carbon sources as likely candidates (Supplementary Figure S2.6). In addition, sucrose and/or fixed carbon has been proposed as putative carbon sources to the cyanobionts in all other species (Rai et al., 2000; Rai et al., 2002; Rai et al., 2019). Together, these findings point to a possible source of carbon to the cyanobiont in the form of simple sugars such as fructose, mannose, and glucose, while indicating

that coralloid roots have an active and ongoing nitrogen assimilation process that can be linked to the presence of a nitrogen-fixing cyanobiont.

2.3.3. Common symbiotic pathway genes seem to have been neofunctionalised for cyanobiosis in lineages that also harbour arbuscular mycorrhizae

To investigate the molecular basis of plant-cyanobacterial symbioses in cycads and its parallel evolution in land plants, we set up an experiment interrogating the data from *E. natalensis* in the context of other plant-cyanobacterial partnerships. To this end, we obtained mRNA-seq data and reference transcriptomes for the hornwort, *An. punctatus* (Li et al., 2020), and the water fern, *Az. filiculoides* (Li et al., 2018a), which were quantified with the same parameters as the *E. natalensis* data. Given that no studies have interrogated *Gunnera* species and their cyanobacterial interactions yet, we also performed mRNA-seq to *de novo* assemble a comprehensive gene catalogue for the southern African species *G. perpensa* by sampling (three biological replicates each) stem gland tissue which harbours cyanobacteria and surrounding stem tissue as control (Supplementary Figure S2.7). mRNA-seq of the cyanobacteria containing and control tissue samples obtained from *G. perpensa* plants yielded ~140 million and ~150 million clean reads for *G. perpensa* colonised stem glands and stem tissue respectively, with an average GC content of 45.6% (Supplementary Table S2.2). Following trimming, *de novo* assembly, annotation and clustering, a reference *G. perpensa* transcriptome was generated (Supplementary Table S2.3). The resultant *G. perpensa* assembly contained 33,346 contigs (of which 23,101 were functionally annotated) with a mean length of 1,120 bp (N50 = 1,714 bp and N75 = 2,572 bp), as expected (see Materials and Methods for details). Furthermore, a BUSCO analysis against the plant dataset (Simão et al., 2015) revealed that the gene catalogue contained 1,301 (90.4%) complete genes, low duplication (1.9%), low fragmentation (2.8%) and some missing genes (6.8%). A total of 11,526, 5,493 and 1,059 transcripts were identified by DESeq2 (Love et al., 2014) as differentially expressed ($p \geq 0.05$) in *An. punctatus, Az. filiculoides* and *G. perpensa* cyanobacterial containing and control tissues, respectively. Like *E. natalensis* (Supplementary Figure S2.4 A), all species showed separate clustering of cyanobacteria containing (symbiotic) tissues compared to control tissues (Supplementary Figure S2.8). For additional comparative analyses and to provide context of genes expressed in plant-bacterial symbioses in general, we also obtained mRNA-seq data and reference transcriptomes for the rhizobial host *M. truncatula* (Pecrix et al., 2018) and the *Frankia* host *D. glomerata* (Battenberg et al., 2018).

In plants, two mutualistic associations – the fungal arbuscular mycorrhizal symbioses and the bacterial (rhizobia and *Frankia*) nitrogen-fixing root nodule symbioses – have been well characterised. Evidence to date strongly supports the claim that the CSP, co-opted from fungal associations, is highly conserved in all land plants that retained either fungal or bacterial nodule symbioses (Delaux et al., 2015). Therefore, to explore whether there might have been co-option during the evolution of the *E. natalensis*-cyanobacterial symbiosis, we constructed a comprehensive query gene list to interrogate against the *E. natalensis* transcriptome by mining literature on well-

characterised symbioses, including rhizobia, *Frankia* and arbuscular mycorrhiza (refer to Materials and Methods and Supplementary Data Set 2.1 for details). The same gene list was interrogated against the transcriptomic data from other cyanobacterial (*An. punctatus, Az. filiculoides* and *G. perpensa*) and nitrogen-fixing root nodule (*M. truncatula* and *D. glomerata*) hosts for comparison. We then identified and prioritised homologs of genes from this list that are preferentially upregulated in plant host symbiont-containing tissue when compared to control tissue (Figure 2.6 A and B). The heat map analysis of *E. natalensis* presented in Figure 2.6 A showed a clear profile of highly and specifically expressed genes (cluster 4) in coralloid root, while other species investigated also demonstrated symbiont-specific clusters for some of these homologs (Figure 2.6 B). Overlap of common homologs in these expression clusters between species was determined and filtered for comparative analyses to identify whether there is evidence for commonly (and potentially independently) co-opted genes involved in interactions with cyanobacteria.

Previous studies suggest that orthologs for the six essential CSP genes (SYMRK, CCaMK, CYCLOPS, CASTOR, POLLUX and Vapyrin) are absent in the *Az. filiculoides* genomes, but that these were present in the genomes of other ferns (Figure 2.7 A; Li et al., 2018a). In contrast, orthologs for all the key angiosperm arbuscular mycorrhizal symbiosis genes are present in the hornwort *An. punctatus* genome (Figure 2.7 A; Li et al., 2020). Our comparative analyses indicate that many of the essential CSP genes, such as SYMRK, CCaMK, CYCLOPS, POLLUX and Vapyrin, were preferentially up-regulated (>50% of expression compared to other tissues) in symbiont-containing tissue in the cycad *E. natalensis* and angiosperm *G. perpensa* (Figure 2.7 A). In addition, some nutrient-specific homologs either related to arbuscular mycorrhizal (STR, FatM and RAD1), rhizobia or *Frankia* nitrogen-fixing symbioses (MOT1;3 and SWEET11) were also preferentially up-regulated in symbiont-containing tissue in *E. natalensis* and *G. perpensa*, respectively (Figure 2.7 A). Interestingly, homologs of the molybdate transporter (MOT1) – a transporter that most likely supplies the co-factor molybdenum to the symbiotic nitrogenase enzyme (Tejada-Jiménez et al., 2017) – were preferentially up-regulated in all species of interest.

Upon closer inspection, we confirmed not only the preferential up-regulation of certain CSP genes but also found them to be significantly higher expressed in *E. natalensis* coralloid root compared to control root tissue (Figure 2.7 B). Among these were CCaMK (Enat_0024269), CYCLOPS (Enat_0010760), POLLUX (Enat_0013225) and Vapyrin (Enat_0001797; Figure 2.7 B). Of course, the possibility that *E. natalensis* coralloid roots also host arbuscular mycorrhizal fungi, as other cycad species do, exists (Muthukumar and Udaiyan, 2002; Fisher and Vovides, 2004). Therefore, to support our anatomical investigation of *E. natalensis* coralloid roots, where no evidence of fungi was observed, we performed a metagenomic analysis of coralloid roots to rule out this possibility. Preliminary metagenomics and associated analyses indicated an overwhelming abundance of cyanobacteria (Nostocales representatives; 98.9% ± 0.55%) in coralloid roots of *E. natalensis* (Supplementary Figure S2.9). Relatively no other bacteria (1.07% ± 0.05%), fungi (0.04% ± 0.01%) or viruses were present (see Supplementary Data Set 2.3).

This proves definitively (and potentially for the first time) that calcium oscillation and decoding via POLLUX, CCaMK, and CYCLOPS, which has been ancestrally present in land plants for mycorrhizal associations, has been independently co-opted across all land plants for rhizobia, *Frankia* and (likely multiple times independently) for cyanobacterial symbioses in *An. punctatus* and *E. natalensis* (Figure 2.7 A). Similarly, Vapyrin has been demonstrated to be essential in both arbuscule and nodule formation, through (potentially) the suppression of defence and cellular rearrangement (Pumplin et al., 2010; Bapaume et al., 2019; Liu et al., 2019), and here we saw that homologs of Vapyrin are preferentially up-regulated in the same two extracellular cyanbiont hosts (Figure 2.7 A). The conservation of certain CSP genes presented here and found in other studies indicates that some of these pre-existing genes and pathways were neofunctionalised at least by *An. punctatus* and *E. natalensis* specifically for cyanobacterial symbioses.

2.3.4. The phylogenetic relationships of sugar transport genes in cyanobiosis lineage representatives

The availability of comparative transcriptomics from a variety of plant-cyanobacterial symbioses also enabled a detailed investigation into similarities and differences in carbon and nitrogen transport in these lineages. Plant symbioses are finely regulated throughout development, because plant membrane transporters are responsible for the exchange of many metabolites in these interactions, including nutrients (Bapaume and Reinhardt, 2012). As such, key transporters have been identified as essential for nutrient exchange to and from symbionts. SWEETs, among others, have been implicated in carbon source transport to both arbuscular mycorrhizae and rhizobia across diverse plant families (Kryvoruchko et al., 2016; Manck-Götzenberger and Requena, 2016; Sugiyama et al., 2016; An et al., 2019). Notably, different SWEET family members are induced in these plant families for sugar supply and have in cases neofunctionalised for providing carbon to bacteria or fungal symbionts, where both are hosted by the plant. For example, in *M. truncatula* SWEET1b and SWEET11 are induced in arbuscule-containing root cells and nodules, respectively (Kryvoruchko et al., 2016; An et al., 2019). Given that simple sugars were identified as a carbon source in *E. natalensis* coralloid roots (Supplementary Figure S2.6) and that other symbioses have neofunctionalised SWEETs for carbon transport to symbionts, we wanted to investigate if cyanobacterial species independently neofunctionalised SWEETs for this purpose. Using a phylogenetic approach, candidate homologs of these genes were identified in the four representative cyanobiosis lineages. The protein sequences included were based on the best BLASTp result using the *A. thaliana* protein sequences as query, filtered based on differential expression, and their relationship to the respective differential genes in the rhizobia (*M. truncatula*) and *Frankia* (*D. glomerata*) hosts observed.

The Arabidopsis SWEETs were classified into four clades, namely AtSWEET1-3 (Clade I), AtSWEET4-8 (Clade II), AtSWEET9-15 (Clade III), and AtSWEET16–17 (Clade IV) as expected (Li et al., 2018b). The phylogenetic analysis of the SWEET protein sequences indicated cyanobiosis lineage sequences scattered among the *A. thaliana* SWEETs (Figure 2.8 and Supplementary Figure S2.10). The candidate cyanobioses SWEET protein sequences showed no distinct grouping according to the phylogenetic analysis performed here (Figure 2.8). Despite this,

different SWEET family members were identified across cyanobiosis lineages highlighting their potential role in carbon supply. Interestingly, only a single transcript from each cyanobiosis lineage gene catalogue generated here, as well as in *An. punctatus*, had SWEET proteins with significantly higher expression in symbiotic tissue; *An. punctatus* Apun_evm.model.utg000039l.841.1 (88th percentile), *E. natalensis* Enat_0023088 (99th percentile) and Gper_0003345 in *G. perpensa* (99th percentile; Figure 2.8). Given the expression level of the candidate SWEETs, we infer that these cyanobiosis lineages independently recruited SWEET paralogs for carbon supply (glucose, sucrose and fructose) to cyanobionts in a similar way as mycorrhizae and rhizobia hosts have (Kryvoruchko et al., 2016; Manck-Götzenberger and Requena, 2016; Sugiyama et al., 2016; An et al., 2019).

2.4. Discussion

In the 450 million years of land plant evolution, independent symbiont switches in many plant lineages have led to a tremendous diversity of mutualistic symbioses (reviewed in Werner et al., 2018). The discovery of a shared signal transduction pathway, the common symbiotic pathway (CSP), in legumes has given us valuable insight into the molecular mechanisms governing arbuscular mycorrhizal and root nodule symbioses (reviewed in Oldroyd, 2013). Since then, nearly 200 genes have been implicated in the establishment and functioning of mutualistic plant symbioses, with more likely to be discovered (reviewed in MacLean et al., 2017; Roy et al., 2020). Advances in comparative genomics and transcriptomics across the diversity of land plants have revealed a conserved ancestral signalling pathway utilised by intracellular symbioses, irrespective of the symbiont or plant lineage involved (Radhakrishnan et al., 2020). Despite these major advances, our understanding of the molecular mechanisms underlying extracellular mutualistic symbioses, such as cyanobacterial partnerships, are poorly understood (reviewed in Martin et al., 2017). Symbiosis with nitrogen-fixing cyanobacteria is a rare trait that may well represent the oldest nitrogen-fixing symbiosis (~500 million years old; Raven, 2002; Yuan et al., 2005). Due to the majority of published literature on cyanobacteria focusing on these prokaryotes as free-living entities and the scarcity of available genomic information for many non-flowering plants, the genetic mechanism underlying these persistent symbioses have remained largely undetermined. In this study, we investigated the uniqueness of cyanobacterial symbioses between plant lineage representatives, and used transcriptome assemblies from a representative species from each of the remaining cyanobiosis lineages; a bryophyte (*An. punctatus*; Li et al., 2020), fern (*Az. filiculoides*; Li et al., 2018a), gymnosperm (*E. natalensis*) and angiosperm (*G. perpensa*), to identify and study candidate genes underlying this complex trait.

In this work, we show that, at least in two cyanobacterial symbioses, the hornwort *An. punctatus* (Li et al., 2020) and cycad *E. natalensis* (Figure 2.7 A), have independently neofunctionalised arbuscular mycorrhizal association genes for cyanobacterial interactions. The conservation of common symbiotic pathway homologs (SYMRK, POLLUX, CCaMK, CYCLOPS and Vapyrin) in these evolutionary diverse cyanobiont hosts could instead be linked to them also hosting arbuscular mycorrhizae (Schüßler, 2000; Muthukumar and Udaiyan, 2002; Fisher and Vovides, 2004). This hypothesis is further supported by the presence of genes evolutionarily linked to arbuscular

mycorrhizal symbiosis (STR, STR2 and RAD1) in *An. punctatus* (Li et al., 2020), other bryophytes (Wang et al., 2010; Delaux et al., 2015) and *E. natalensis* (Figure 2.7 A). A recent study by Radhakrishnan et al. (2020) also implicated SYMRK, CCaMK and CYCLOPS in various cyanobacterial hosts (*An. angustus*, *Cycas micholitzii*, *Dioon edule*, *E. barteris* and *Stangeria eriopus*; Leebens-Mack et al., 2019) as essential for their intracellular arbuscular mycorrhizal symbioses. Despite independent lines of evidence suggesting the presence of mycorrhizae fungi in hornworts and cycads, no morphological or metagenomic evidence has been found to support this at least for the cycad, *E. natalensis* (Figure 2.1, 2.2 and Supplementary Figure S2.9). To what extent other cyanobacterial species have neofunctionalised common symbiotic pathway homologs is still an open question. Given the complete lack of orthologs for these genes in *Az. filiculoides* (Li et al., 2018a) and *Az. cf. caroliniana* (Radhakrishnan et al., 2020), Li et al. (2018a) proposes that Azollaceae species rely on an alternative pathway attributed to a secondary loss of mycorrhiza symbiosis in these species (Harley and Harley, 1987b, a; Gemma et al., 1992). On the contrary, *Gunnera* species might harbour arbuscular mycorrhizae as shown for *G. petaloidea* (Koske et al., 1992) and evident by the presence of SYMRK in *G. perpensa* investigated here (Figure 2.7 A). These genes, individually or taken together, might, therefore, not solely be responsible for arbuscular mycorrhizal symbioses, but might have been neofunctionalised for cyanobioses in plant lineages that have retained both arbuscular mycorrhizal and cyanobacterial symbioses. The co-option of common symbiotic pathway genes (CASTOR, POLLUX, CCaMK and CYCLOPS) has recently been shown for an artificial symbiosis between *Oryza sativa* (rice, an arbuscular mycorrhizal host) and *Nostoc* cyanobacteria (Álvarez et al., 2021). This seems to suggest that cyanobacterial symbioses recruited arbuscular mycorrhizae genes (CSP genes) in a similar way as root nodule symbioses have. The results of this study set up good hypotheses and resources (Supplementary Data Set 2.1 and 2.2) to perform new experiments reconstructing transcriptional networks in these various species, to see how and to what extent pre-existing genes and pathways were recruited and repurposed for cyanobacterial interactions.

The involvement of SWEET sugar transporters in mycorrhizal symbiosis has been established (An et al., 2019), as has their role in rhizobial root nodule symbioses (Kryvoruchko et al., 2016; Sugiyama et al., 2016) but much is still to be resolved about our understanding of carbon and nitrogen exchange in cyanobioses. We could link genes involved in starch and sucrose catabolism to simple sugars (fructose and mannose) with *E. natalensis* coralloid roots by mining metabolic pathway enrichments (Supplementary Figure S2.6; differential genes preferentially expressed in coralloid root tissue from *E. natalensis*, Figure 2.4). Anatomical investigation of precoralloid roots revealed high starch content in elongated, columnar host cells surrounding the cyanobacterial zone (Figure 2.1 D), supporting starch as a source of carbon for the cyanobionts that will reside here. Although we did not observe the same for coralloid roots (Figure 2.1 B), the mucilage matrix that surrounds these host columnar cells and wherein cyanobacteria can be found (Figure 2.1 A and B) is rich in polysaccharides (reviewed in Rai et al., 2000; Rai et al., 2002; Rai et al., 2019). It is, therefore, possible that the high starch content present in precoralloid roots supply sugars to this mucilage matrix so that by the time cyanobionts are present the roots no longer need to catabolise

starch. The discovery of mucilage-secreting aerial roots in a landrace of maize (Sierra Mixe maize; van Deynze et al., 2018) said to promotes nitrogen-fixing symbiosis with diazotrophic bacteria supports this hypothesis. As with cycad coralloid roots, these modified aerial roots produce large amounts of a complex polysaccharide-rich mucilage which, upon catabolism into simple sugars (arabinose, fucose and galactose), supports nitrogen fixation (van Deynze et al., 2018). While van Deynze et al. (2018) prove that aerial root mucilage is sufficient to support nitrogen fixation, similar investigations are required to confirm this in cycad coralloid roots. Despite the presence of genes involved in simple sugar catabolism from starch in *E. natalensis*, the symbiotic gene analysis presented in Figure. 2.7 A did not reveal the presence of a SWEET 11 homolog in coralloid roots. However, its presence in *G. perpensa* (Figure. 2.7 A) whose anatomical investigation also revealed high starch content (Supplementary Figure S2.1 A and B) supports a potential role for SWEET in cyanobioses. According to our phylogenetic analysis, different SWEET family members have convergently been recruited by cyanobiosis-forming plant lineages (Figure 2.8 and Supplementary Figure S2.10), a member of which has independently from our analysis been implicated in *An. punctatus*-cyanobiosis for carbon supply (Li et al., 2020). Taken together, our results suggest an independent acquisition of sugar transporters in cyanobiosis-forming plant lineages, utilising, like the root nodule angiosperms, genes from an existing and ancestral family.

Published literature emphasises that cyanobacteria, whether free-living or as symbionts, fix nitrogen in specialised cells, termed heterocysts (reviewed in Flores et al., 2019). Since their discovery, heterocysts have been characterised as large, round cells with thickened cell envelopes and “honeycomb” structures that surround prominent cyanophycin granules at their poles (Kumar et al., 2010; Maldener et al., 2014). This appearance of a thickened cell envelope is due to the deposition of additional envelope layers, an inner glycolipid and an outer polysaccharide layer (Cardemil and Wolk, 1979, 1981; Nicolaisen et al., 2009). Anatomical investigation of symbiotic tissue from our cyanobiont plant lineage representatives, coralloid roots of *E. natalensis* (Figure 2.1 and 2.2), stem glands of *G. perpensa* (Supplementary Figure S2.1 A-E), and leaf cavities of *Az. pinnatta* spp. *africana* (Supplementary Figure S2.1 F-J), not only revealed the presence of cyanobionts but also confirmed that heterocysts dominated these symbiotic tissues (Figure 2.3 and Supplementary Table S2.1). Plant hosts are known to induce the differentiation of heterocysts in cyanobionts causing a remarkably higher heterocyst frequency in symbioses compared to the 5-10% of heterocysts observed in free-living cyanobacteria (reviewed in Rai et al., 2000; Rai et al., 2002; Rai et al., 2019). Previous reports indicate heterocyst frequencies of 25-30% for *Azolla* species, ~45% in cycad species and as much as 60-80% in *Gunnera* species (reviewed in Rai et al., 2000; Rai et al., 2002; Rai et al., 2019). Heterocyst frequencies alone seem insufficient to conclude the presence of active nitrogen fixation. Therefore, studies providing direct measures of nitrogen fixation, through acetylene reduction assays, N difference and ^{15}N-labelling (Carlsson and Huss-Danell, 2008) are needed.

The enrichment and presence of genes involved in citrulline and ornithine biosynthesis (Figure 2.5) supports a previous study done by Pate et al. (1988) which provided evidence that coralloid root xylem from various cycad species, including *E. natalensis*, is abundant in amino acids such as glutamine, glutamate and citrulline. Most

enriched genes were specifically associated with NH_3 assimilation into glutamine and glutamate in coralloid roots with no detectable expression of NO_3^- and/or NO_2^- transporters or metabolising enzymes (Figure 2.5). Previous studies have directly measured the assimilation and release of NH_3 in *Anthoceros* (Meeks et al., 1985), *Azolla* (Meeks et al., 1987) and *Gunnera* (Silvester et al., 1996) cyanobacterial interactions, and in all cases found that NH_3 is provided to the plant as a precursor for subsequent assimilation into glutamate and asparagine (Asn), respectively, prior to transport to the rest of the plant. The relatively high expression (>50% expression compared to other tissues) of glutamine synthetase and glutamate dehydrogenase (step 1 and 2 in Figure 2.5) in coralloid roots together with genes up-regulated in all other roots (step 5, 11 and 12 in Figure 2.5) strongly supports a scenario where NH_3 is provided as a precursor by the cyanobiont and these amino acids are subsequently transported as a nitrogen source to other tissues. Together with heterocyst frequencies, we provide strong gene expression evidence supporting the presence of active nitrogen fixation. Surprisingly, a molybdate transporter (MOT1) shows complete conservation in not only cyanobioses but also root nodule symbioses lineages investigated (Figure 2.7 A), this transporter is likely specialised for molybdenum (a co-factor) supply to the cyanobiont nitrogenase as was found for root nodule nitrogenase activity (Tejada-Jiménez et al., 2017) indicating an active and ongoing nitrogen process. This evidence suggests that the same molybdate transporter is essential in all interactions and has been independently co-opted multiple times in plant evolution for symbiosis with bacteria.

An important finding in this study is the presence of an already well-delineated so-called cyanobacterial zone in precoralloid and coralloid roots, and their apparent anatomical similarity to primary roots (Figure 2.1). Since the discovery of coralloid roots (Reinke, 1872), studies by Milindasuta (1975) and Ahern and Staff (1994) on various cycad species have shaped our understanding of the development and anatomy of coralloid roots. In this work, we show that *E. natalensis*, does indeed harbour a cyanobacterial zone, complete with mucilage-filled intercellular spaces separated by radially elongated, columnar host cells (Figure 2.1 A and B). This discernable cell layer emerges due to the partial separation and radial elongation of cells while those in adjacent cortices enlarge longitudinally (reviewed in Milindasuta, 1975), giving them their columnar appearance as shown in Figure 2.1 B. Previous studies suggest that coralloid roots irreversibly develop from precoralloid roots (Caiola, 1974; Ahern and Staff, 1994; Lindblad, 2009), but their morphological similarities to one another have not been investigated. From our observations, the cyanobacterial zone and its associated mucilage is present prior to cyanobacterial invasion (Figure 2.1 C-E). This suggests that cycads actively develop defined and contained extracellular zones in anticipation which is consistent with observations made by Webb (1983) and Nathanielsz and Staff (1975) regarding the deposition of mucilage prior to invasion. It is commonly believed that precoralloid roots and, by extension, coralloid roots are highly specialised types of lateral roots emerging as “transformed normal, lateral root apices” (Ahern and Staff, 1994; Norstog and Nicholls, 1997). However, from our observations, it seems more likely that precoralloid roots arise from primary roots as either adventitious or specialised roots due to their morphological similarities rather than from lateral roots (Supplementary Figure S2.2 and S2.3). Our observations are to some extent supported by Ahern and Staff (1994) who reported that *Macrozamia communis* precoralloid

roots mainly emerged as secondary laterals along the primary (tap) root and that distinct differences in terms of vascular tissue organisation exist between lateral and precoralloid roots, the latter of which we also observed (Supplementary Figure S2.3). Further anatomical analysis to determine if precoralloid or coralloid root ontogenically resemble primary roots more than lateral roots will need to be done.

2.5. Conclusion

We employed a multifaceted approach to investigate the characteristics of the *E. natalensis*-cyanobacterial partnership. This was done by comparing representative species harbouring cyanobionts anatomically, as well as using genomic and transcriptomic datasets to elucidate genes potentially underlying this symbiosis. This study confirmed cyanobacterial presence in three naturally growing Southern African cyanobacterial host species. Taken together, the data showed that active nitrogen fixation in all three is likely occurring as visualised by high heterocyst presence. Although the *E. natalensis*-cyanobacterial symbiosis is unique in many ways, it bears a remarkable resemblance to angiosperms with nitrogen-fixing root nodule and arbuscular mycorrhizal symbioses. We highlighted unique carbon (simple sugars in the form of fructose and mannose) and nitrogen (amino acids particularly citrulline and ornithine) sources in the *E. natalensis*-cyanobacterial partnership with transporters being co-opted multiple times for nutrient exchange. Our study definitively (and potentially for the first time) identifies metabolising enzymes involved in arginine biosynthesis, as highly coralloid root-specific. We also hypothesise that the cyanobiont provides ammonia as a precursor, as is the case for all other cyanobacterial symbioses, which the cycad uses to assimilate these amino acids. The comparative transcriptomic analysis revealed that common symbiotic pathway homologs were neofunctionalised for cyanobiosis in the cycad, *E. natalensis*, from a pre-existing ancestral mycorrhizae symbiosis. Further investigation of the identified homologs involved in the *E. natalensis*-cyanobacterial partnership through functional and reverse genetics methods will be invaluable to better our understanding of interactions between host plants and various symbioses with microorganisms and may enable new approaches for research in bioengineering symbioses into commercially important crops.

2.6. Materials and methods

2.6.1. Sample collection and preparation

Tissue from *Az. pinnatta* spp. *africana, E. natalensis* and *G. perpensa* were obtained from naturally grown plants at the Manie van der Schijff Botanical Garden and the Cycad and Indigenous Plant Nursery at the University of Pretoria, South Africa. Cyanobacteria containing stem gland, the surrounding stem, leaf and root tissues were sampled from three *G. perpensa* plants. In addition, mature and immature leaflet, and rachis, leaf base, stem, upper primary root, lower primary root, precoralloid root and coralloid root tissues supplementing previous data were collected from three *E. natalensis* plants (Supplementary Table S2.4). Entire *Az. pinnatta* spp. *africana* plants were used for light and transmission electron microscopy.

RNA was extracted from *E. natalensis* and *G. perpensa* samples using a standard CTAB RNA extraction (Chang et al., 1993) and the Plant/Fungi RNA Purification Kit (Norgen Biotek Corp., Canada) methods. This was followed by a DNase treatment, using the DNase Max and RNase-Free DNase Kits (Qiagen, Germany) to remove genomic DNA. The quality and quantity of RNA extracts were determined by electrophoresis on a 1% (w/v) agarose gel and spectrophotometry using a Nanodrop ND_1000 (Nanodrop, United States) instrument, respectively. RNA was sent for poly(A) and total RNA sequencing using the Illumina HiSeq 2500 system at BGI, China, and Macrogen, the Netherlands, using the 2 x 150-bp chemistry.

For microbiome analysis, total genomic DNA was extracted from *E. natalensis* coralloid root samples using the DNeasy® PowerSoil® Kit (Qiagen, Germany) following numerous rounds of surface sterilisation according to the protocol reported by Gutiérrez-García et al. (2018). Briefly, coralloid roots were submerged in 15 ml of each solution and continuously stirred for 3 minutes in hydrogen peroxide (H_2O_2), 7 minutes in 70% ethanol, 30 seconds in Sabax water, 4 minutes in 6% sodium hypochlorite (NaClO), and three 1 minute washes in Sabax water. As a control, the last Sabax wash was plated onto agar bacteriological media (Merck, Germany) and incubated overnight at 37°C to ensure the absence of exogenous bacteria. The total genomic DNA was RNase treated (Omega Bio-tek, United States) and ethanol precipitated. The quality and quantity of the resultant DNA was determined by electrophoresis on a 1% (w/v) agarose gel and spectrophotometry using a Qubit 4 Fluorometer (Thermo Fisher Scientific, United States) instrument, respectively. DNA was sent for sequencing using the Illumina HiSeq platform at the DNA Sequencing Center, Brigham Young University, using the 2 x 250-bp chemistry.

2.6.2. Microscopy sample preparation and imaging for morphological comparison

2.6.2.1. *Light microscopy*

To determine the presence of symbiosis, cyanobacteria containing tissue (dorsal lobes, stem glands, precoralloid and coralloid roots), as well as control tissue (stem tissue, upper primary and lateral roots) were sampled from *Az. pinnatta* spp. *africana, G. perpensa* and *E. natalensis*, respectively. These samples were then placed in formalin for fixation (Sigma-Aldrich, United States) for at least 24 hours at room temperature. The plant material was dehydrated with 2-methoxyethanol, 100% ethanol, 100% propan-1-ol and 100% 1-butanol for 4 to 6 hours. Each dehydration solution was performed twice. Glycol methacrylate (GMA) infiltration was done through the addition of a used and new solution for 24 hours each at room temperature, followed by a final GMA infiltration step of 5 days. The material was transferred to gelatine embedding capsules and left to polymerise in a 52°C oven for 24 hours (Feder and O'Brien, 1968; Bruhl and Ashford, 1986).

Sections of 5 µm were made with a Leica-Jung 2045 Multicut microtome (Leica Biosystems, Germany) and stained both with 0.05% toluidine blue and following a periodic acid/Schiff (Pas) reaction with 0.5% toluidine blue O for 3 minutes as described previously (Feder and O'Brien, 1968). The sections were imaged with a Nikon Eclipse E2000 light microscope and Olympus EP50 camera.

2.6.2.2. *Transmission electron microscopy (TEM)*

To confirm the presence of cyanobacterial heterocysts, tissues were sampled and placed in a fixative solution (2.5% glutaraldehyde/formaldehyde in 1.5M sodium phosphate buffer), for at least 24 hours at room temperature. The plant material was washed three times with 1.5M sodium phosphate buffer for 15 minutes and dehydrated with a graded series of ethanol (30%, 50%, 70%, 90% and three times with 100%) for 15 minutes each except for the last 100% dehydration which was left for 30 minutes. Resin infiltration was done through the addition of increasing concentrations (20%, 40%, 60% and 80%) of LR White resin (SPI® Supplies, United States) in 100% ethanol for 1 hour each at room temperature, followed by the addition of 100% LR White left overnight at 4°C. Additional 100% LR White resin exchanges were performed prior to transferring the material to gelatine embedding capsules, after which they were left to polymerise in a 60°C oven for 96 hours.

Ultrathin (0.05-0.09 µm) and semi-thin (0.5 µm) sections were made with a Reichert Ultracut E microtome. Ultrathin sections were stained in 4% aqueous uranyl acetate and Reynold's lead citrate (Reynolds, 1963) for 3 minutes each, while semi-thin sections were stained with 1% toluidine blue O in 1% sodium borate buffer for 2 minutes. Ultrathin sections were imaged with a JEOL JEM-2100F transmission electron microscope and semi-thin sections with a Zeiss Axio Imager.M2 light microscope. Cyanobacterial heterocyst and vegetative cell ratios were determined by counting these cell types on their respective *Az. pinnatta* spp. *africana*, *E. natalensis* and *G. perpensa* grinds (Supplementary Table S2.1). Three biological repeats for each plant tissue were used to determine the ratio with at least five grid blocks counted for each and considered as representative of the samples.

2.6.3. Metagenomic analysis of *Encephalartos natalensis* cyanobacterial containing tissue

Raw paired-end (2 x 150-bp) reads from the three biological repeats (CR002, CR003 and CR005) were collected, and quality control was performed using FastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Low-quality reads and adapter sequences were removed with Trimmomatic (v0.36; Bolger et al., 2014) using the default settings (Supplementary Table S2.5). High-quality filtered reads were classified with Kraken (v2.0.8-beta; Wood et al., 2019) using a custom non-redundant (NR) protein database acquired from NCBI (Sayers et al., 2020) and confidence scores of 0.1 (Supplementary Table S2.5). Following classification, output files were saved in a MetaPhlan (Segata et al., 2012) tab-delimited format which indicated a small fraction of reads (with ranges 14-15.69%) with LCA taxonomic classifications (generated based on exact k-mer matches to the database). These were further parsed using a custom python script to represent taxonomic assignments in family level sub-classification, resulting in ~500K (CR002), ~76K (CR003) and ~880K (CR005) reads, respectively, for microbe abundance analysis (see Supplementary Data Set 2.3).

2.6.4. Transcriptome assemblies, annotations, differential, and preferential gene expression

2.6.4.1. *Transcriptome assemblies and annotations*

FastQC was used to trim low-quality reads and to remove sequencing adaptors from raw mRNA-Seq data. The resultant high-quality data were subsequently used for *de novo* transcriptome assemblies. The transcriptomes for *E. natalensis* and *G. perpensa* were constructed using a combinatorial approach of SOAPdenovo-Trans (v1.03; Xie et al., 2014) and Trinity (v-2.6.5; Haas et al., 2013) using default parameters and k-mer lengths of 21 to 111 and 25 and 31, respectively. After constructing individual k-mer assemblies, these were merged into multi k-mer assemblies, and redundancy removed using the EvidentialGene pipeline following author guidelines (Gilbert, 2013). BUSCO (v3.0.2) was used to assess assembly completeness. A custom script was used to compute the number of contigs, the total number of bases, N50, N75, maximum and minimum contig length, first and third quartile contig length, and median and mean contig length. Finally, the Trinity Ex90N50 Statistic and Ex90 Gene count scripts were used to further filter out potentially fragmented contigs (Haas et al., 2013).

Assemblies for *E. natalensis* and *G. perpensa* generated here, and for *An. punctatus* (Li et al., 2020), *Az. filiculoides* (Li et al., 2018a), *Medicago truncatula* (Pecrix et al., 2018), *Dactylis glomerata* (Battenberg et al., 2018) and *Ceanothus thyrsiflorus* (Battenberg et al., 2018), were annotated with the following: BLASTx against the TAIR10 database (Berardini et al., 2015) and BLASTx against each abovementioned transcriptome retaining only significant BLASTx hits, MapMan bin ID and descriptions (Thimm et al., 2004), KEGG ID and descriptions (Kanehisa and Goto, 2000), InterProScan (Blum et al., 2020), Pfam (Mistry et al., 2020) and Gene Ontology (Ashburner et al., 2000; Consortium, 2020), and orthology assignment using OrthoFinder (Emms and Kelly, 2015, 2019).

2.6.4.2. *Differential and preferential gene expression*

mRNA-seq libraries were mapped to assembled transcriptomes using Salmon (Patro et al., 2017) and transcripts per million (TPM) values were added to the annotation files of each species. Differential gene expression was performed on cyanobacteria containing and control tissues for all species with DESeq2, using default parameters (Love et al., 2014), and significant ($p \geq 0.05$) differential transcripts indicated in the annotation files.

We analysed these genes for preferential expression in cyanobacteria containing tissues of *An. punctatus*, *Az. filiculoides*, *E. natalensis* and *G. perpensa* using a custom-designed R app (Pinard et al. in progress). This custom script constructs hierarchical clustering diagrams, and comparisons between homologous or orthologous gene sets, from annotated transcriptome files. The gene profiles were further divided into clusters, based on dendrograms of hierarchical clustering and a heat map constructed using percentile scores. These percentile scores are normalised scores based on expression values (≥50% relative expression was used), using a list of differential gene IDs as reference. The gene IDs were linked between the species by OrthoFinder and significant reciprocal BLASTp analyses, correlating annotations and percentile scores for clusters specific to cyanobacteria containing tissue within each species. Gene sets of interest were extracted into separate gene lists for further analyses. The resulting sets of preferentially expressed differential genes were used to determine KEGG pathway enrichment and Venn

diagrams drawn to compare presence. For the relative expression levels of differential genes (Additional Tables A2-A7 in Supplementary Data Set 2.2).

2.6.5. Identifying genes of interest

2.6.5.1. *Candidate gene list analysis*

Known genes involved in the establishment and maintenance of arbuscular mycorrhizal and rhizobial symbiosis were identified from literature. These data were then used to construct a candidate gene list for further analysis (MacLean et al., 2017; Roy et al., 2020). Additional genes, such as those involved in flavonoid and coumarin biosynthesis, were also included in this candidate gene list, since little is known about their involvement in cyanobioses (Winkel-Shirley, 2001, 2002; Falcone-Ferreyra et al., 2012; Stassen et al., 2021). The respective transcriptomes were each annotated with those genes that indicated a significant ($p \leq 0.05$) BLASTp result using the gene list protein sequences as query. We further analysed the preferential expression of these candidate genes in *An. punctatus*, *Az. filiculoides*, *E. natalensis*, *G. perpensa*, *M. truncatula* and *D. glomerata* using the same custom-designed R app (Pinard *et al.* in progress). The resulting sets of preferentially expressed genes were used to determine enrichment and Venn diagrams drawn to compare presence. For the comprehensive candidate gene list and relative expression levels see Additional Table A1 and Tables A8-A13 in Supplementary Data Set 2.1 and 2.2, respectively.

2.6.5.2. *Phylogenetic analysis*

Genes for carbon exchange, from the candidate gene list as well as literature, were investigated in the *E. natalensis, G. perpensa, Az. filiculoides, An. punctatus, M. truncatula* and *D. glomerata* gene catalogues, and their phylogenetic relationships and expression explored. To this end, protein sequences for the SWEET *A. thaliana* genes were obtained (Berardini et al., 2015) and used for BLASTx analyses retaining only significant BLASTx hits in all six species. Multiple alignments of the differential candidate sequences were completed using MAFFT, alignments curated using trimAI (both on the https://ngphylogeny.fr/; Dereeper et al., 2008; Lemoine et al., 2019), and maximum likelihood phylogenetic trees were constructed using 1,000 bootstrap iterations. For the relative expression levels of candidate SWEET genes see Additional Table A14 in Supplementary Data Set 2.2.

2.7. Acknowledgements

The authors would like to acknowledge Jason Sampson, Arnold Frisby, Kenneth Oberlander, and Erna van Wilpe for their technical assistance with plant sampling, light and transmission electron microscopy, respectively. This study was funded by South Africa's National Research Foundation (Grant UID 116239) and the Technology Innovation Agency of South Africa through the Forest Molecular Genetics (FMG) Cluster programme at the University of Pretoria. Further support was received from the University of Pretoria.

2.8. References

Ågren GI, Wetterstedt JÅM, Billberger MFK (2012) Nutrient limitation on terrestrial plant growth – modeling the interaction between nitrogen and phosphorus. New Phytol **194**: 953-960.

Ahern CP, Staff IA (1994) Symbiosis in cycads: the origin and development of coralloid roots in *Macrozamia communis* (Cycadaceae). Am J Bot **81**: 1559-1570.

Álvarez C, Brenes-Álvarez M, Molina-Heredia FP, Mariscal V (2021) Early events of the endophytic symbiotic between *Oryza sativa* and *Nostoc punctiforme* involve the SYM pathway. bioRxiv: 2021.2004.2009.436957.

An J, Zeng T, Ji C, de Graaf S, Zheng Z, Xiao TT, Deng X, Xiao S, Bisseling T, Limpens E, Pan Z (2019) A *Medicago truncatula* SWEET transporter implicated in arbuscule maintenance during arbuscular mycorrhizal symbiosis. New Phytol **224**: 396-408.

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H et al. (2000) Gene Ontology: tool for the unification of biology. Nat Genet **25**: 25-29.

Bapaume L, Reinhardt D (2012) How membranes shape plant symbioses: signaling and transport in nodulation and arbuscular mycorrhiza. Front Plant Sci **3**: 223.

Bapaume L, Laukamm S, Darbon G, Monney C, Meyenhofer F et al. (2019) Vapyrin marks an endosomal trafficking compartment involved in arbuscular mycorrhizal symbiosis. Front Plant Sci **10**: 666.

Battenberg K, Potter D, Tabuloc CA, Chiu, JC, Berry, AM (2018) Comparative transcriptomic analysis of two actinorhizal plants and the legume *Medicago truncatula* supports the homology of root nodule symbioses and is congruent with a two-step process of evolution in the nitrogen-fixing clade of angiosperms. Front Plant Sci **9**: 1256.

Berardini TZ, Reiser, L, Li, D, Mezheritsky, Y, Muller, R et al. (2015) The Arabidopsis information resource: Making and mining the "gold standard" annotated reference plant genome. Genesis **53,** 474-485.

Blum M, Chang, HY, Chuguransky, S, Grego, T, Kandasaamy, S et al. (2020) The InterPro protein families and domains database: 20 years on. Nucleic Acids Res **49**: 344-354.

Bolger AM, Lohse, M, Usadel, B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics **30**: 2114-2120.

Bonnett HT (1990) The *Nostoc-Gunnera* association. In CRC handbook of symbiotic cyanobacteria, AN Rai, ed (Boca Raton, Florida: CRC Press), pp. 137-159.

Braun-Howland EB, Nierzwicki-Bauer, SA (1990) *Azolla-Anabaena* symbiosis: biochemistry, physiology, ultrastructure, and molecular biology. In CRC Handbook of Symbiotic Cyanobacteria, AN Rai, ed (Boca Raton, Florida: CRC Press), pp. 65-117.

Bravo A, York, T, Pumplin, N, Mueller, LA, Harrison, MJ (2016) Genes conserved for arbuscular mycorrhizal symbiosis identified through phylogenomics. Nat Plants **2**: 15208.

Bravo A, Brands, M, Wewer, V, Dörmann, P, Harrison, MJ (2017) Arbuscular mycorrhiza-specific enzymes FatM and RAM2 fine-tune lipid biosynthesis to promote development of arbuscular mycorrhiza. New Phytol **214**: 1631-1645.

Bruhl J, Ashford, A (1986) Application of glycol methacrylate embedding to herbarium specimens of fruits. Stain Technol **61**: 163-167.

Brundrett MC (2002) Coevolution of roots and mycorrhizas of land plants. New Phytol **154**: 275-304.

Brundrett MC, Tedersoo, L (2018) Evolutionary history of mycorrhizal symbioses and global host plant diversity. New Phytol **220**: 1108-1115.

Caiola MG (1974) A light and electron microscopic study of the blue-green algae living either in the coralloid roots of *Macrozamia communis* or isolated in culture. G Bot Ital **108**: 161-173.

Calabrese S, Pérez-Tienda J, Ellerbeck M, Arnould C, Chatagnier O et al. (2016) GintAMT3 – a low-affinity ammonium transporter of the arbuscular mycorrhizal *Rhizophagus irregularis*. Front Plant Sci **7**: 679.

Cardemil L, Wolk CP (1979) The polysaccharides from heterocyst and spore envelopes of a blue-green alga. Structure of the basic repeating unit. J Biol Chem **254**: 736-741.

Cardemil L, Wolk CP (1981) Isolated heterocysts of *Anabaena variabilis* synthesize envelope polysaccharide. Biochim Biophys Acta **674**: 265-276.

Carlsson G, Huss-Danell K (2008) How to quantify biological nitrogen fixation in forage legumes in the field. In Biological Nitrogen Fixation: Towards Poverty Alleviation through Sustainable Agriculture, FD Dakora, SBM Chimphango, AJ Valentine, C Elmerich, and WE Newton, eds (Dordrecht: Springer Netherlands), pp. 47-48.

Chang ACG, Chen T, Li N, Duan J (2019) Perspectives on endosymbiosis in coralloid roots: association of cycads and cyanobacteria. Front Microbiol **10**: 1888.

Chang S, Puryear J, Cairney J (1993) A simple and efficient method for isolating RNA from pine trees. Plant Mol Biol Rep **11**: 113-116.

Cheng Z (1982) Fine structure of blue-green algae and the cells lined along the endophyte cavity in the coralloid root of *Cycas*. J Integr Plant Biol **24**: 109-114.

The Gene Ontology Consortium (2020) The Gene Ontology resource: enriching a GOld mine. Nucleic Acids Res **49**: 325-334.

Day DA, Copeland L (1991) Carbon metabolism and compartmentation in nitrogen-fixing legume nodules. Plant Physiol Biochem **29**: 185-201.

Delaux PM, Schornack S (2021) Plant evolution driven by interactions with symbiotic and pathogenic microbes. Science **371**: eaba6605.

Delaux PM, Radhakrishnan GV, Jayaraman D, Cheema J, Malbreil M et al. (2015) Algal ancestor of land plants was preadapted for symbiosis. Proc Natl Acad Sci USA **112**: 13390.

Dereeper A, Guignon V, Blanc G, Audic S, Buffet S et al. (2008) Phylogeny.fr: robust phylogenetic analysis for the non-specialist. Nucleic Acids Res **36**: 465-469.

Doidy J, van Tuinen D, Lamotte O, Corneillat M, Alcaraz G. et al. (2012) The *Medicago truncatula* sucrose transporter family: characterization and implication of key members in carbon partitioning towards arbuscular mycorrhizal fungi. Mol Plant **5**: 1346-1358.

Doyle JJ (1994) Phylogeny of the legume family: an approach to understanding the origins of nodulation. Annu Rev Ecol Syst. **25**: 325-349.

Emms DM, Kelly S (2015) OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. Genome Biol **16**: 157.

Emms DM, Kelly S (2019) OrthoFinder: phylogenetic orthology inference for comparative genomics. Genome Biol **20**: 238.

Falcone-Ferreyra ML, Rius SP, Casati P (2012) Flavonoids: biosynthesis, biological functions, and biotechnological applications. Front Plant Sci **3**: 1-15.

Feder N, O'Brien TP (1968) Plant microtechnique: some principles and new methods. Am J Bot **55**: 123-142.

Fisher JB, Vovides AP (2004) Mycorrhizae are present in cycad roots. Bot Rev **70**: 16-23.

Flores E, Picossi S, Valladares A, Herrero A (2019) Transcriptional regulation of development in heterocyst-forming cyanobacteria. Biochim Biophys Acta **1862**: 673-684.

Gemma JN, Koske RE, Flynn T (1992) Mycorrhizae in Hawaiian pteridophytes: occurrence and evolutionary significance. Am J Bot **79**: 843-852.

Gensel PG (2008) The emerald planet: how plants changed earth's history (Oxford, USA: University Press).

Giaquinta RT (1983) Phloem loading of sucrose. Annu Rev Plant Physiol **34**: 347-387.

Gilbert D (2013) Gene-omes built from mRNA seq not genome DNA. In 7th annual arthropod genomics symposium (Notre Dame).

Govindarajulu M, Pfeffer PE, Jin H, Abubaker J, Douds DD et al. (2005) Nitrogen transfer in the arbuscular mycorrhizal symbiosis. Nature **435**: 819-823.

Gutiérrez-García K, Bustos-Díaz ED, Corona-Gómez JA, Ramos-Aboites HE, Sélem-Mojica N et al. (2018) Cycad coralloid roots contain bacterial communities including cyanobacteria and *Caulobacter* spp. that encode niche-specific biosynthetic gene clusters. Genome Biol. Evol. **11**: 319-334.

Gutjahr C, Radovanovic D, Geoffroy J, Zhang Q, Siegler H et al. (2012) The half-size ABC transporters STR1 and STR2 are indispensable for mycorrhizal arbuscule formation in rice. Plant J **69**: 906-920.

Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD et al. (2013) *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat Protoc **8**: 1494-1512.

Harley JL, Harley EL (1987a) A check-list of mycorrhiza in the British flora-addenda, errata and index. New Phytol **107**: 741-749.

Harley JL, Harley EL (1987b) A check-list of mycorrhiza in the British flora. New Phytol **105**: 1-102.

Hashidoko Y, Nishizuka H, Tanaka M, Murata K, Murai Y et al. (2019) Isolation and characterization of 1-palmitoyl-2-linoleoyl-*sn*-glycerol as a hormogonium-inducing factor (HIF) from the coralloid roots of *Cycas revoluta* (Cycadaceae). Sci Rep **9**: 4751.

Heckman DS, Geiser DM, Eidell BR, Stauffer RL, Kardos NL et al. (2001) Molecular evidence for the early colonization of land by fungi and plants. Science **293**: 1129-1133.

Jin H, Pfeffer PE, Douds DD, Piotrowski E, Lammers PJ et al. (2005) The uptake, metabolism, transport and transfer of nitrogen in an arbuscular mycorrhizal symbiosis. New Phytol **168**: 687-696.

Kanehisa M, Goto S (2000) KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res **28**: 27-30.

Kobae Y, Tamura Y, Takai S, Banba M, Hata S (2010) Localized expression of arbuscular mycorrhiza-inducible ammonium transporters in soybean. Plant Cell Physiol **51**: 1411-1415.

Koske RE, Gemma JN, Flynn T (1992) Mycorrhizae in Hawaiian angiosperms: a survey with implications for the origin of the native flora. Am J Bot **79**: 853-862.

Kryvoruchko IS, Sinharoy S, Torres-Jerez I, Sosso D, Pislariu CI et al. (2016) MtSWEET11, a nodule-specific sucrose transporter of *Medicago truncatula* root nodules. Plant Physiol **171**: 554-565.

Kumar K, Mella-Herrera RA, Golden JW (2010) Cyanobacterial heterocysts. Cold Spring Harb. Perspect Biol **2**: a000315.

Lechno-Yossef S, Nierzwicki-Bauer SA (2002) *Azolla-Anabaena* symbiosis. In Cyanobacteria in Symbiosis, AN Rai, B Bergman, and U Rasmussen, eds (Dordrecht: Springer Netherlands), pp. 153-178.

Leebens-Mack JH, Barker MS, Carpenter EJ, Deyholos MK, Gitzendanner MA et al. (2019) One thousand plant transcriptomes and the phylogenomics of green plants. Nature **574**: 679-685.

Lemoine F, Correia D, Lefort V, Doppelt-Azeroual O, Mareuil F et al. (2019) NGPhylogeny.fr: new generation phylogenetic services for non-specialists. Nucleic Acids Res **47**: 260-265.

Li FW, Nishiyama T, Waller M, Frangedakis E, Keller J et al. (2020) *Anthoceros* genomes illuminate the origin of land plants and the unique biology of hornworts. Nat Plants **6**: 259-272.

Li FW, Brouwer P, Carretero-Paulet L, Cheng S, de Vries J et al. (2018a) Fern genomes elucidate land plant evolution and cyanobacterial symbioses. Nat Plants **4**: 460-472.

Li X, Si W, Qin Q, Wu H, Jiang H (2018b) Deciphering evolutionary dynamics of SWEET genes in diverse plant lineages. Sci Rep **8**: 13440.

Lindblad P (2009) Cyanobacteria in symbiosis with cycads. Microbiol Monogr **8**: 225-233.

Lindblad P, Bergman B. (1990) The cycad-cyanobacterial symbiosis. In CRC handbook of symbiotic cyanobacteria, AN Rai, ed (Boca Raton: CRC Press), pp. 137-159.

Lindblad P, Bergman B, Hofsten AV, Hällbom L, Nylund JE (1985) The cyanobacterium-*Zamia* symbiosis: an ultrastructural study. New Phytol **101**: 707-716.

Liu A, Contador CA, Fan K, Lam HM (2018) Interaction and regulation of carbon, nitrogen, and phosphorus metabolisms in root nodules of legumes. Front Plant Sci **9**: 1860.

Liu CW, Breakspear A, Stacey N, Findlay K, Nakashima J et al. (2019) A protein complex required for polar growth of rhizobial infection threads. Nat Commun **10**: 2848.

López-Pedrosa A, González-Guerrero M, Valderas A, Azcón-Aguilar C, Ferrol N (2006) *GintAMT1* encodes a functional high-affinity ammonium transporter that is expressed in the extraradical mycelium of *Glomus intraradices*. Fungal Genet Biol **43**: 102-110.

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

MacLean AM, Bravo A, Harrison MJ (2017) Plant signaling and metabolic pathways enabling arbuscular mycorrhizal symbiosis. Plant Cell **29**: 2319-2335.

Maldener I, Summers ML, Sukenik A (2014) Cellular differentiation in filamentous cyanobacteria. In The cell biology of cyanobacteria, E Flores and A Herrero, eds (Norfolk, UK: Caister Academic Press), pp. 263-291.

Manck-Götzenberger J, Requena N (2016) Arbuscular mycorrhiza symbiosis induces a major transcriptional reprogramming of the potato SWEET sugar transporter family. Front Plant Sci **7**: 487.

Martin FM, Uroz S, Barker DG (2017) Ancestral alliances: plant mutualistic symbioses with fungi and bacteria. Science **356**: eaad4501.

Meeks JC (1990) Cyanobacterial-bryophyte associations. In CRC Handbook of Symbiotic Cyanobacteria, AN Rai, ed (Boca Raton, Florida: CRC Press), pp. 21.

Meeks JC, Enderlin CS, Joseph CM, Chapman JS, Lollar MWL (1985) Fixation of [^{13}N]N_2 and transfer of fixed nitrogen in the *Anthoceros-Nostoc* symbiotic association. Planta **164**: 406-414.

Meeks JC, Steinberg NA, Enderlin CS, Joseph CM, Peters GA (1987) *Azolla-Anabaena* relationship: XIII. Fixation of [^{13}N]N_2. Plant Physiol **84**: 883-886.

Milindasuta BE (1975) Developmental anatomy of coralloid roots in cycads. Am J Bot **62**: 468-472.

Mistry J, Chuguransky S, Williams L, Qureshi M, Salazar GA et al. (2020) Pfam: The protein families database in 2021. Nucleic Acids Res **49**: 412-419.

Muthukumar T, Udaiyan K (2002) Arbuscular mycorrhizas in cycads of southern India. Mycorrhiza **12**: 213-217.

Nathanielsz CP, Staff IA (1975) A mode of entry of blue-green algae into the apogeotropic roots of *Macrozamia communis*. Am J Bot **62**: 232-235.

Nicolaisen K, Hahn A, Schleiff E (2009) The cell wall in heterocyst formation by *Anabaena* sp. PCC 7120. J Basic Microbiol **49**: 5-24.

Norstog K, Nicholls TJ (1997) The biology of the cycads. (Ithaca, N.Y.: Comstock Pub. Associates).

Oldroyd GED (2013) Speak, friend, and enter: signalling systems that promote beneficial symbiotic associations in plants. Nat Rev Microbiol **11**: 252-263.

Oldroyd GED, Leyser O (2020) A plant's diet, surviving in a variable nutrient environment. Science **368**: eaba0196.

Pate JS, Lindblad P, Atkins CA (1988) Pathways of assimilation and transfer of fixed nitrogen in coralloid roots of cycad-*Nostoc* symbioses. Planta **176**: 461-471.

Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C (2017) Salmon provides fast and bias-aware quantification of transcript expression. Nat Methods **14**: 417-419.

Pawlowski K, Demchenko KN (2012) The diversity of actinorhizal symbiosis. Protoplasma **249**: 967-979.

Pecrix Y, Staton SE, Sallet E, Lelandais-Brière C, Moreau S et al. (2018) Whole-genome landscape of *Medicago truncatula* symbiotic genes. Nat Plants **4**: 1017-1025.

Pérez-Tienda J, Testillano PS, Balestrini R, Fiorilli V, Azcón-Aguilar C et al. (2011) GintAMT2, a new member of the ammonium transporter family in the arbuscular mycorrhizal fungus *Glomus intraradices*. Fungal Genet Biol **48**: 1044-1055.

Pirozynski KA, Malloch DW (1975) The origin of land plants: a matter of mycotrophism. BioSyst **6**: 153-164.

Pumplin N, Mondo SJ, Topp S, Starker CG, Gantt JS et al. (2010) *Medicago truncatula* Vapyrin is a novel protein required for arbuscular mycorrhizal symbiosis. Plant J **61**: 482-494.

Radhakrishnan GV, Keller J, Rich MK, Vernié T, Mbadinga Mbadinga DL et al. (2020) An ancestral signalling pathway is conserved in intracellular symbioses-forming plant lineages. Nat Plants **6**: 280-289.

Rai AN, Söderbäck E, Bergman B (2000) Tansley Review No. 116 Cyanobacterium–plant symbioses. New Phytol **147**: 449-481.

Rai AN, Bergman B, Rasmussen U (2002) Cyanobacteria in symbiosis. (Springer).

Rai AN, Singh AK, Syiem MB (2019) Plant growth-promoting abilities in cyanobacteria. In Cyanobacteria from Basic Science to Applications, AK Mishra, DN Tiwari, and AN Rai, eds (Academic Press), pp. 459-476.

Raven JA (2002) Evolution of cyanobacterial symbioses. In Cyanobacteria in Symbiosis, AN Rai, B Bergman, and U Rasmussen, eds (Dordrecht: Springer Netherlands), pp. 329-346.

Reinke J (1872) Parasitische Anabaena in Wurzelen der Cycadeen. Gottingen Nachr **57**: 107.

Reynolds ES (1963) The use of lead citrate at high pH as an electron-opaque stain in electron microscopy. J Cell Biol **17**: 208-212.

Roodt D, Lohaus R, Sterck L, Swanepoel RL, van de Peer Y, Mizrachi E (2017) Evidence for an ancient whole genome duplication in the cycad lineage. PLoS One **12**: e0184454.

Roy S, Liu W, Nandety RS, Crook A, Mysore KS et al. (2020) Celebrating 20 years of genetic discoveries in legume nodulation and symbiotic nitrogen fixation. Plant Cell **32**: 15-41.

Santi C, Bogusz D, Franche C (2013) Biological nitrogen fixation in non-legume plants. Ann Bot **111**: 743-767.

Sayers EW, Beck J, Brister JR, Bolton EE, Canese K et al. (2020) Database resources of the national center for biotechnology information. Nucleic Acids Res **48**: 9-16.

Schüßler A (2000) *Glomus claroideum* forms an arbuscular mycorrhiza-like symbiosis with the hornwort *Anthoceros punctatus*. Mycorrhiza **10**: 15-21.

Segata N, Waldron L, Ballarini A, Narasimhan V, Jousson O et al. (2012) Metagenomic microbial community profiling using unique clade-specific marker genes. Nat Methods **9**: 811-814.

Silvester WB, Parsons R, Watt PW (1996) Direct measurement of release and assimilation of ammonia in the *Gunnera–Nostoc* symbiosis. New Phytol **132**: 617-625.

Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM (2015) BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics **31**: 3210-3212.

Soltis DE, Soltis PS, Morgan DR, Swensen SM, Mullin BC et al. (1995) Chloroplast gene sequence data suggest a single origin of the predisposition for symbiotic nitrogen fixation in angiosperms. Proc. Natl Acad. Sci. USA **92**: 2647-2651.

Sprent JI, Ardley JK, James EK (2013) From North to South: A latitudinal look at legume nodulation processes. S Afr J Bot **89**: 31-41.

Stassen MJJ, Hsu SH, Pieterse CMJ, Stringlis IA (2021) Coumarin communication along the microbiome–root–shoot axis. Trends Plant Sci **26**: 169-183.

Suárez-Moo PDJ, Vovides AP, Griffith MP, Barona-Gómez F, Cibrián-Jaramillo A (2019) Unlocking a high bacterial diversity in the coralloid root microbiome from the cycad genus *Dioon*. PLoS One **14**: e0211271.

Sugiyama A, Saida Y, Yoshimizu M, Takanashi K, Sosso D et al. (2016) Molecular characterization of LjSWEET3, a sugar transporter in nodules of *Lotus japonicus*. Plant Cell Physiol **58**: 298-306.

Svistoonoff S, Hocher V, Gherbi H (2014) Actinorhizal root nodule symbioses: what is signalling telling on the origins of nodulation? Curr Opin Plant Biol **20**: 11-18.

Swensen SM (1996) The evolution of actinorhizal symbioses: evidence for multiple origins of the symbiotic association. Am J Bot **83**: 1503-1512.

Tejada-Jiménez M, Gil-Díez P, León-Mediavilla J, Wen J, Mysore KS et al. (2017) *Medicago truncatula* Molybdate Transporter type 1 (MtMOT1.3) is a plasma membrane molybdenum transporter required for nitrogenase activity in root nodules under molybdenum deficiency. New Phytol **216**: 1223-1235.

Tester M, Smith SE, Smith FA (1987) The phenomenon of "nonmycorrhizal" plants. Can J Bot **65**: 419-431.

Thimm O, Bläsing O, Gibon Y, Nagel A, Meyer S et al. (2004) MapMan: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. Plant J **37**: 914-939.

van Deynze A, Zamora P, Delaux PM, Heitmann C, Jayaraman D et al. (2018) Nitrogen fixation in a landrace of maize is supported by a mucilage-associated diazotrophic microbiota. PLoS Biol. **16**: e2006352.

van Hove C, Lejeune A (2002) Applied aspects of *Azolla-Anabaena* symbiosis. In Cyanobacteria in Symbiosis, AN Rai, B Bergman, and U Rasmussen, eds (Dordrecht: Springer Netherlands), pp. 179-193.

van Velzen R, Doyle JJ, Geurts R (2019) A resurrected scenario: single gain and massive loss of nitrogen-fixing nodulation. Trends Plant Sci **24**: 49-57.

Vessey JK, Pawlowski K, Bergman B (2005) Root-based N_2-fixing symbioses: Legumes, actinorhizal plants, *Parasponia* sp. and cycads. In Root Physiology: from Gene to Function, H Lambers and TD Colmer, eds (Dordrecht: Springer Netherlands), pp. 51-78.

Veys P, Waterkeyn L, Lejeune A, Hove C (1999) The pore of the leaf cavity of *Azolla*: Morphology, cytochemistry and possible functions. Symbiosis **27**: 33-57.

Vitousek PM, Howarth RW (1991) Nitrogen limitation on land and in the sea: how can it occur? Biogeochemistry **13**: 87-115.

Wang B, Yeun LH, Xue JY, Liu Y, Ané JM et al. (2010) Presence of three mycorrhizal genes in the common ancestor of land plants suggests a key role of mycorrhizas in the colonization of land by plants. New Phytol **186**: 514-525.

Webb DT (1983) Nodulation in light- and dark-grown *Macrozamia communis* L. Johnson seedlings in sterile culture. Ann Bot **52**: 543-547.

Werner GDA, Cornwell WK, Sprent JI, Kattge J, Kiers ET (2014) A single evolutionary innovation drives the deep evolution of symbiotic N_2-fixation in angiosperms. Nat Commun **5**: 4087.

Werner GDA, Cornelissen JHC, Cornwell WK, Soudzilovskaia NA, Kattge J et al. (2018) Symbiont switching and alternative resource acquisition strategies drive mutualism breakdown. Proc Natl Acad Sci USA **115**: 5229-5234.

Winkel-Shirley B (2001) Flavonoid biosynthesis. A colorful model for genetics, biochemistry, cell biology, and biotechnology. Plant Physiol **126**: 485-493.

Winkel-Shirley B (2002) Biosynthesis of flavonoids and effects of stress. Curr Opin Plant Biol **5**: 218-223.

Wolk CP, Thomas J, Shaffer PW, Austin SM, Galonsky A (1976) Pathway of nitrogen metabolism after fixation of ^{13}N-labeled nitrogen gas by the cyanobacterium, *Anabaena cylindrica*. J Biol Chem **251**: 5027-5034.

Wood DE, Lu J, Langmead B (2019) Improved metagenomic analysis with Kraken 2. Genome Biol **20**: 257.

Xie Y, Wu G, Tang J, Luo R, Patterson J et al. (2014) SOAPdenovo-Trans: *de novo* transcriptome assembly with short RNA-Seq reads. Bioinformatics **30**: 1660-1666.

Yuan X, Xiao S, Taylor TN (2005) Lichen-like symbiosis 600 million years ago. Science **308**: 1017-1020.

Zhang Q, Blaylock LA, Harrison MJ (2010) Two *Medicago truncatula* half-ABC transporters are essential for arbuscule development in arbuscular mycorrhizal symbiosis. Plant Cell **22**: 1483-1497.

Zheng Y, Chiang TY, Huang CL, Gong X (2018) Highly diverse endophytes in roots of *Cycas bifida* (Cycadaceae), an ancient but endangered gymnosperm. J Microbiol **56**: 337-345.

2.9. Figures

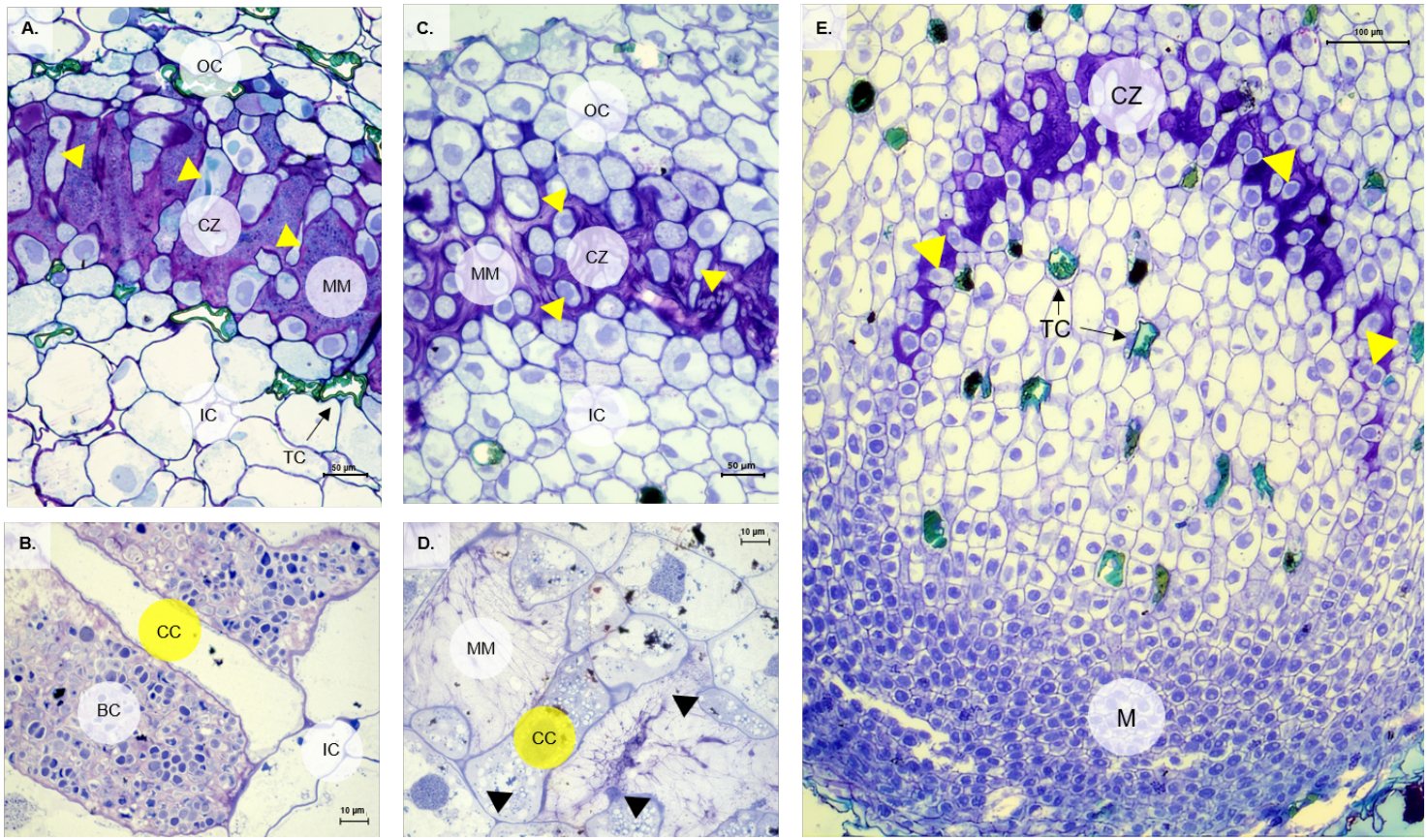


Figure 2.1. Light microscopy images of the cyanobacterial zone in *Encephalartos natalensis* coralloid and precoralloid root. *(A)* Transverse section of a coralloid root showing the cyanobacterial zone (CZ) between the outer (OC) and inner cortex (IC). Cyanobacteria are restricted to the mucilaginous matrix (MM) between columnar cells (yellow arrowheads) which radially transverse the zone. *(B)* Magnified view of the CZ in coralloid root showing columnar cells (CC) and cyanobacteria (BC). *(C)* Transverse section of a precoralloid root showing the (as yet not colonised) CZ between the OC and IC. The MM and CC (yellow arrowheads) are visible. *(D)* Magnified view of the CZ in precoralloid root showing CC containing starch granules (black arrowheads). *(E)* Transverse section of a precoralloid root showing the CZ with interspersed CC (yellow arrowheads) and the meristem clearly visible (M). TC, tannin channels.

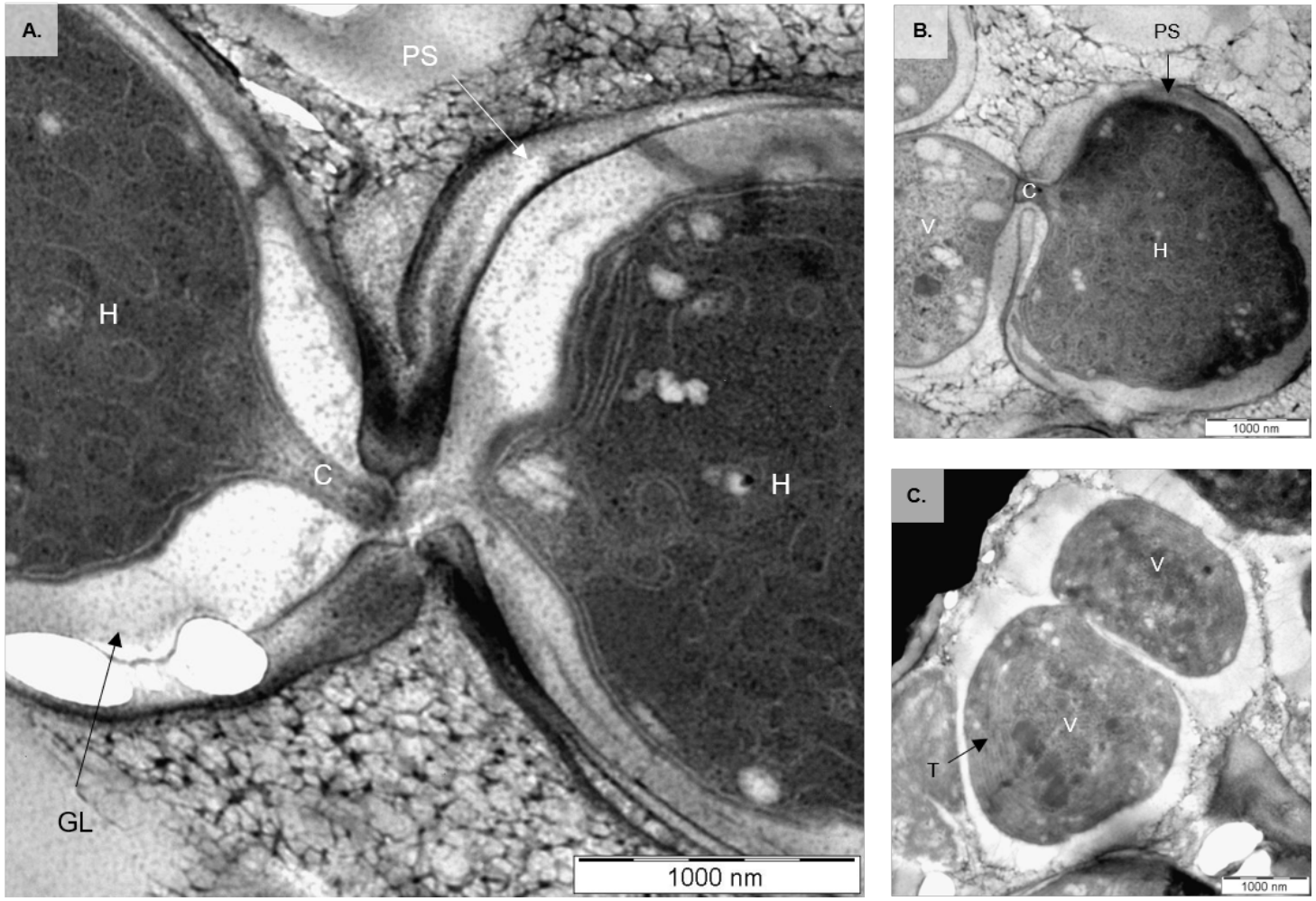


Figure 2.2. Transmission electron micrographs of *Encephalartos natalensis* symbiotic tissue. Heterocysts (H) and vegetative cells (V) in *E. natalensis (A-C)* are indicated. Numerous carboxysomes can be seen in both heterocysts and vegetative cells. C, polar cyanophycin granule; GL, glycolipid layer; PL, plasmalemma; PS, polysaccharide layer; T, thylakoid membranes.

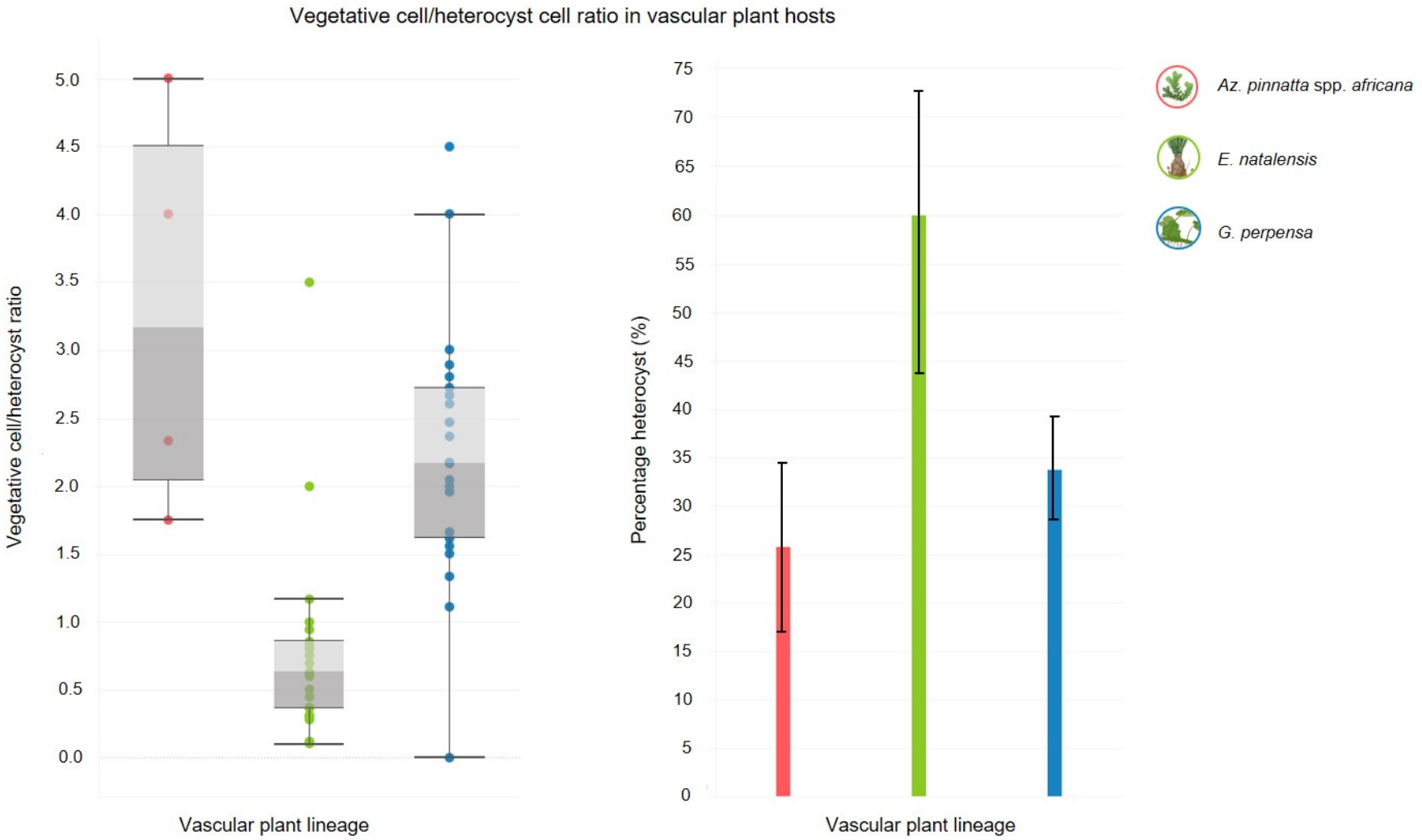


Figure 2.3. Box-and-whisker plots of vegetative cell/heterocyst ratio (left) and bar charts of percentage heterocysts (right) for the vascular plant hosts.

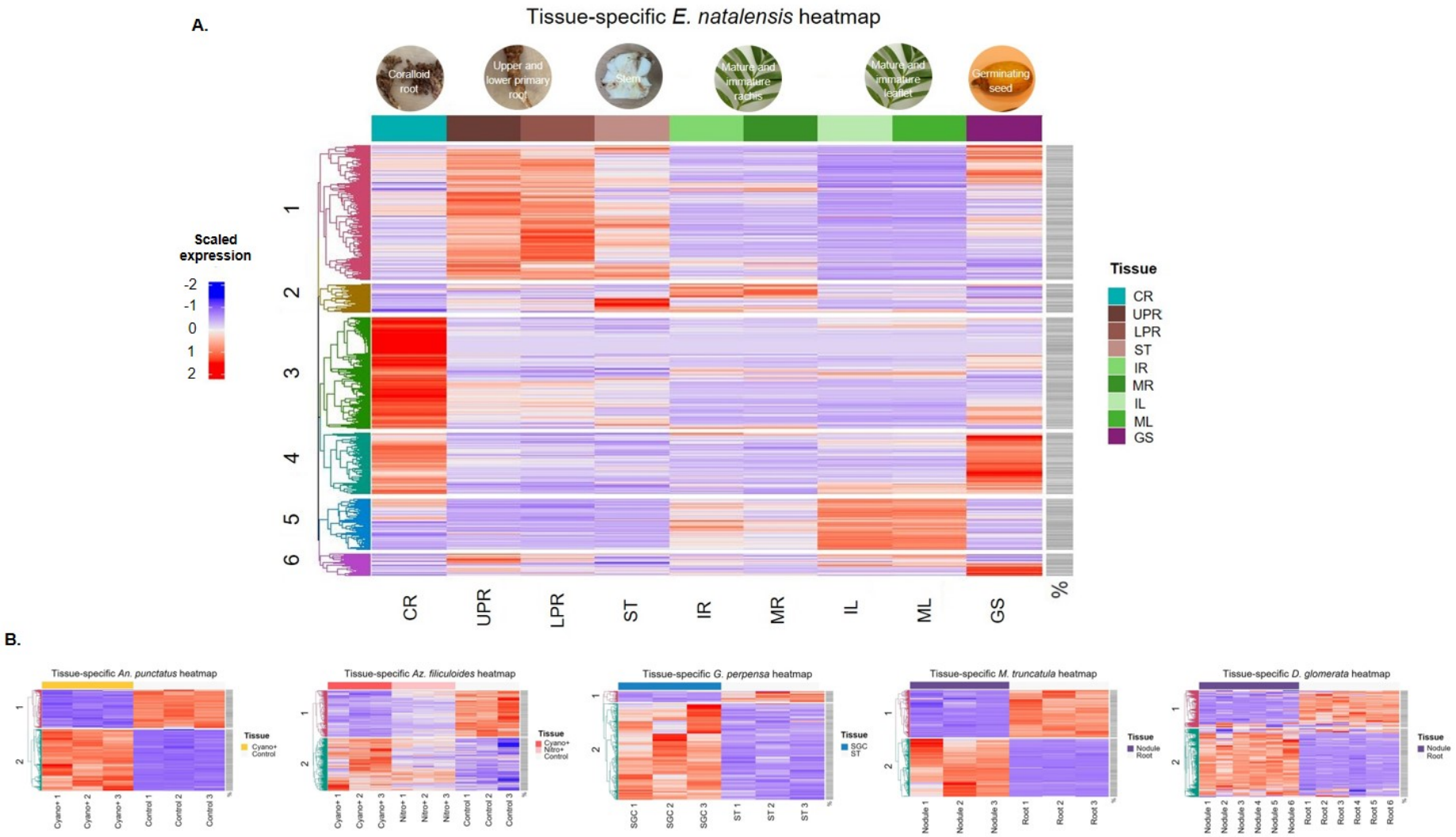


Figure 2.4. Expression of differential genes related to symbiotic tissues in cyanobiont, rhizobia and *Frankia* symbioses. *(A, B)* Heatmaps showing those differentially expressed genes with tissue specific expression or preferential gene expression (above median). *(A)* In the *E. natalensis* heatmap, cluster 3 (1,011 genes) shows a profile of highly expressed genes in coralloid root, while cluster 1 (1,217 genes) shows those highly expressed in primary root. *(B)* In the heatmaps, clusters 2 (2,083 *An. punctatus*, 1,697 *Az. filiculoides*, 508 *G. perpensa*, 3,947 *M. truncatula* and 4,158 *D. glomerata* genes) show profiles of highly expressed genes in symbiotic tissues.

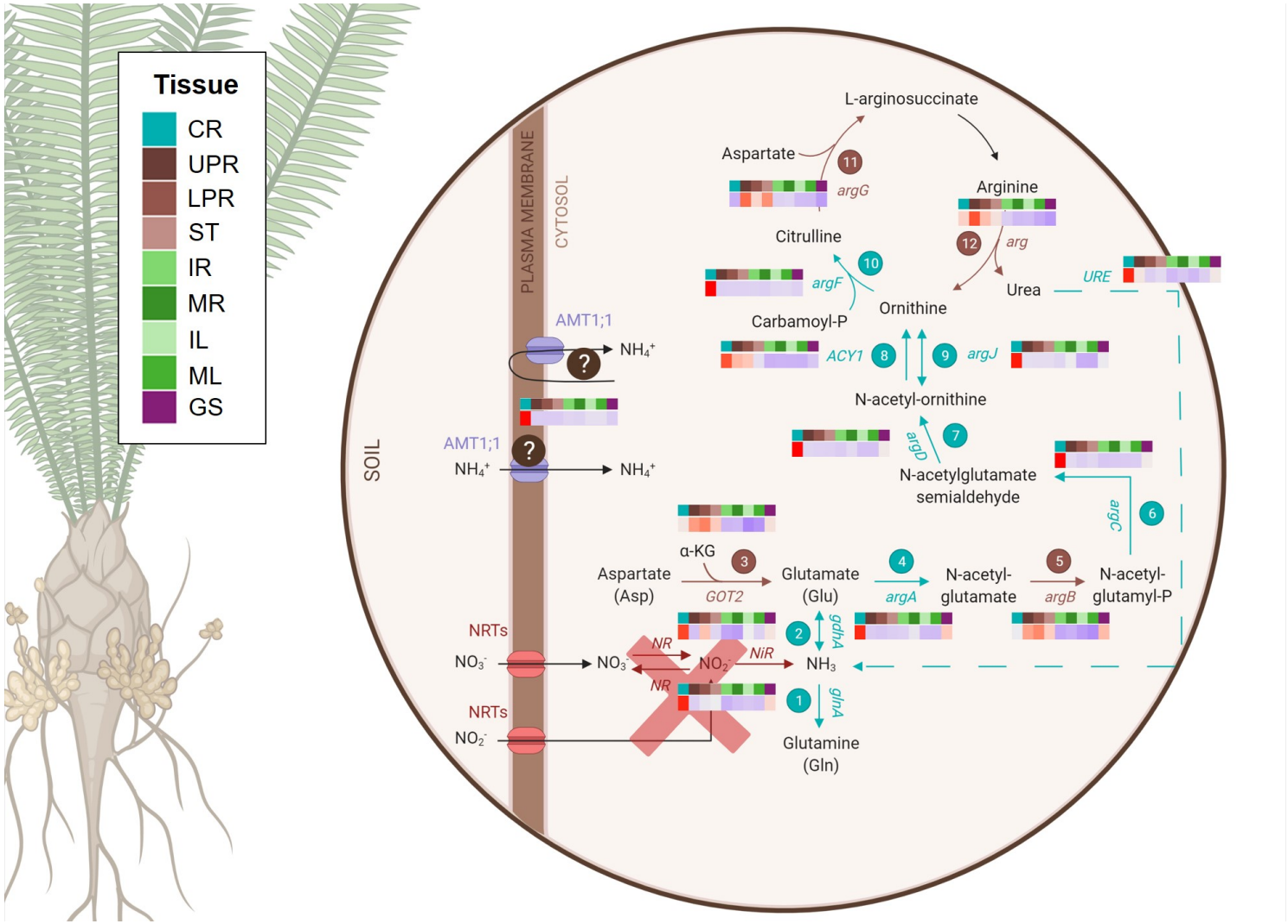


Figure 2.5. Nitrogen metabolism and arginine biosynthesis enrichments. Genes differentially expressed in cycad, *E. natalensis*, coralloid root (teal) and primary root (brown) for ammonium (NH_4^+) assimilation into arginine intermediates, ornithine and citrulline. As indicated here, nitrate (NO_3^-) and/or nitrite (NO_2^-) transporters and reduction genes responsible for reducing ammonium are not significantly expressed in coralloid or primary root. The expression of an ammonium transporter, nitrogen assimilation, arginine biosynthesis and urea breakdown genes in all the *E. natalensis* tissues are indicated by heatmaps. Created with BioRender.com.

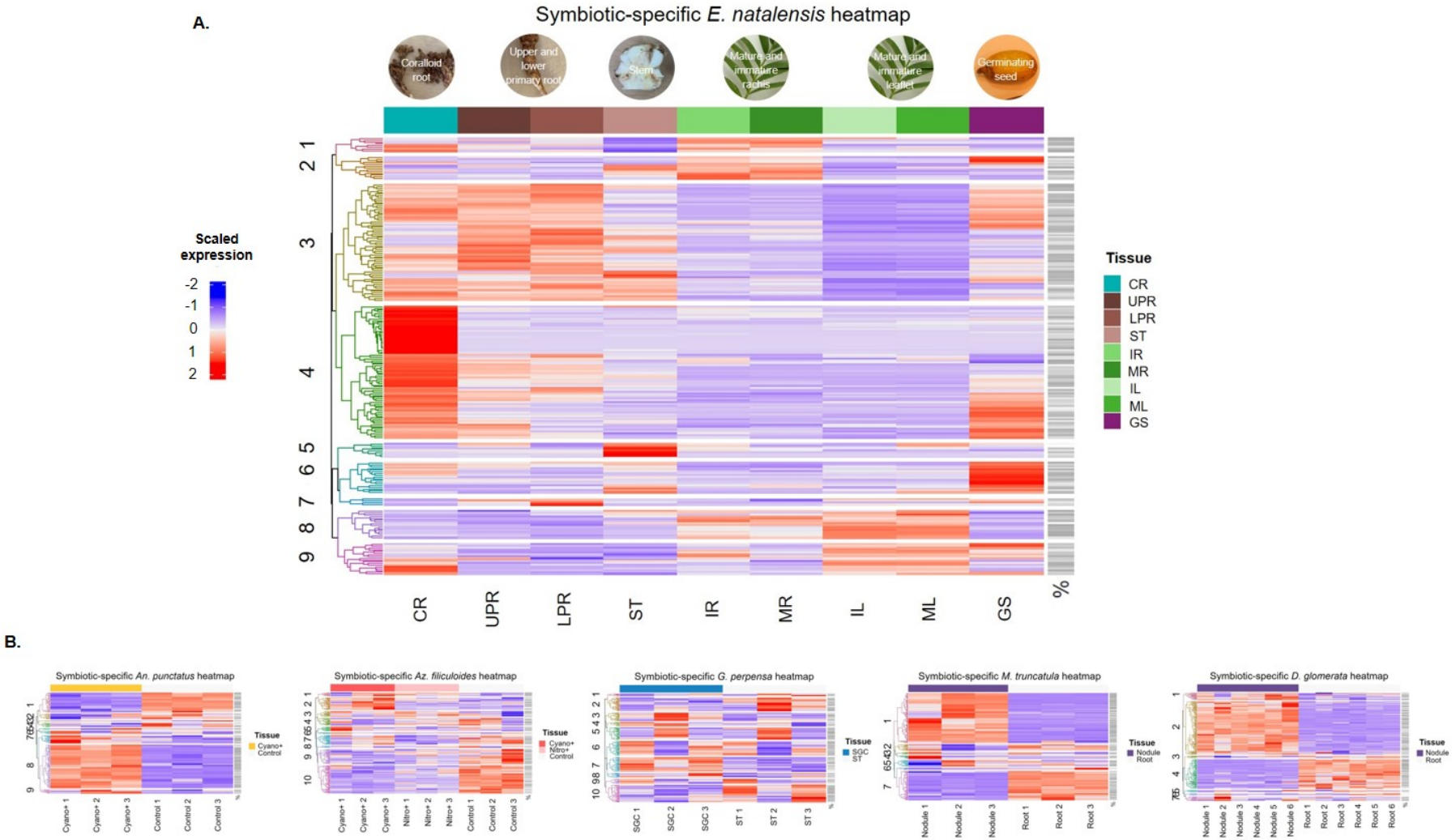


Figure 2.6. Prioritisation of candidate plant host genes functionally involved in cyanobiont, rhizobia and *Frankia* symbioses. *(A, B)* Heatmaps showing those genes with tissue specific expression or preferential gene expression (above median). *(A)* In the *E. natalensis* heatmap, cluster 4 (76 genes) shows a profile of highly and specifically expressed genes in coralloid root. *(B)* In the heatmaps, clusters 8 (88 *An. punctatus* genes), cluster 1-2 (37 *Az. filiculoides* genes), cluster 7 (35 *G. perpensa* genes), cluster 1 (122 *M. truncatula* genes) and cluster 2 (125 *D. glomerata* genes) show profiles of highly and specifically expressed genes in symbiotic tissues.

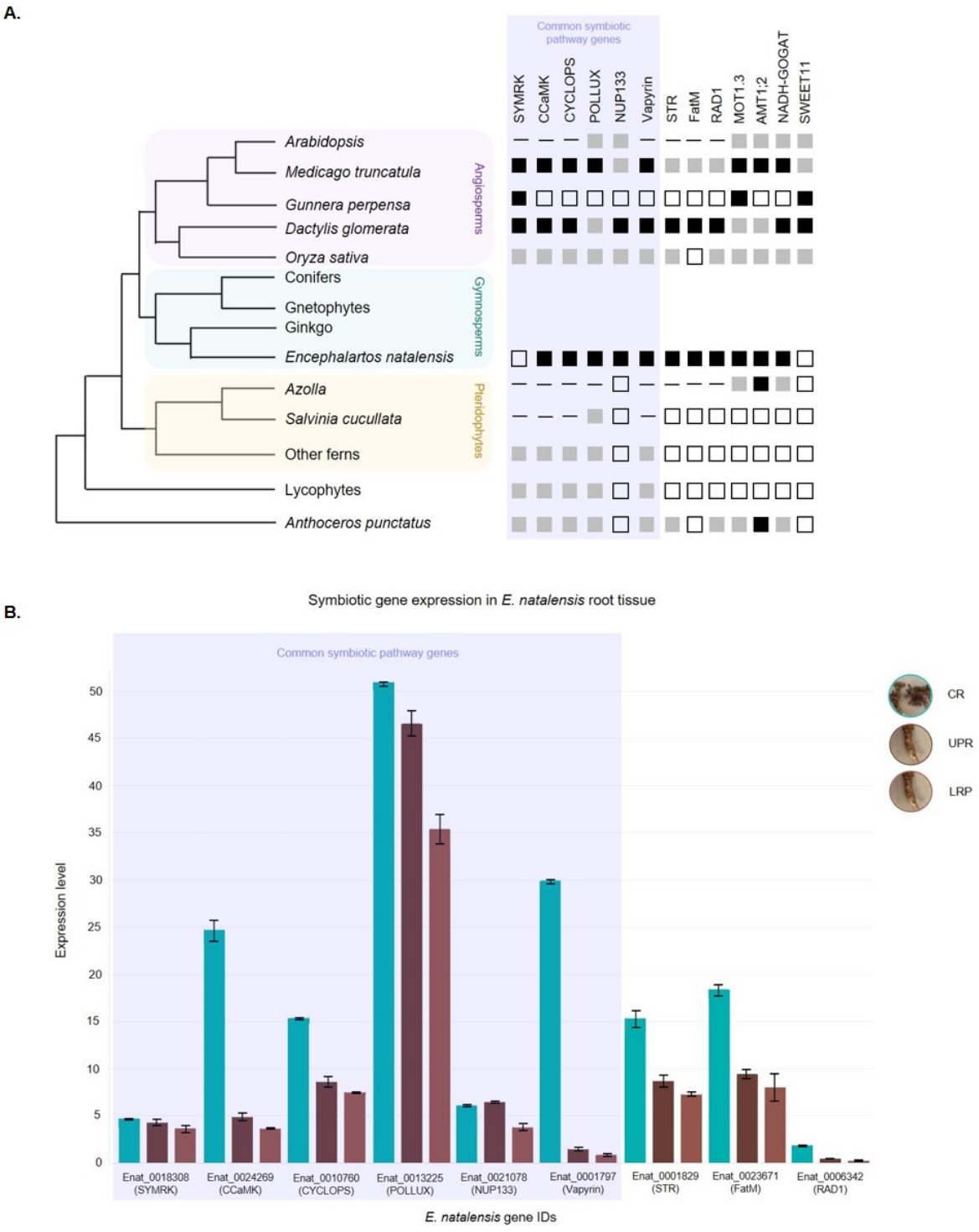


Figure 2.7. Candidate symbiosis-related genes. *(A)* Homologs of candidate symbiosis-related genes can be found in the cycad *E. natalensis* and the angiosperm *G. perpensa* (black boxes), those without information are indicated by empty boxes. Previous literature (Li *et al.* 2018a; Li *et al.* 2020) indicated common symbiotic pathway (CSP) genes were present in *An. punctatus*, lycophytes and other ferns (grey boxes) but were lost in *Az. filiculoides* and *Salvinia* (lines). *(B)* Expression of candidate symbiosis-related genes in *E. natalensis* coralloid, upper and lower primary root with CSP genes indicated by a blue box.

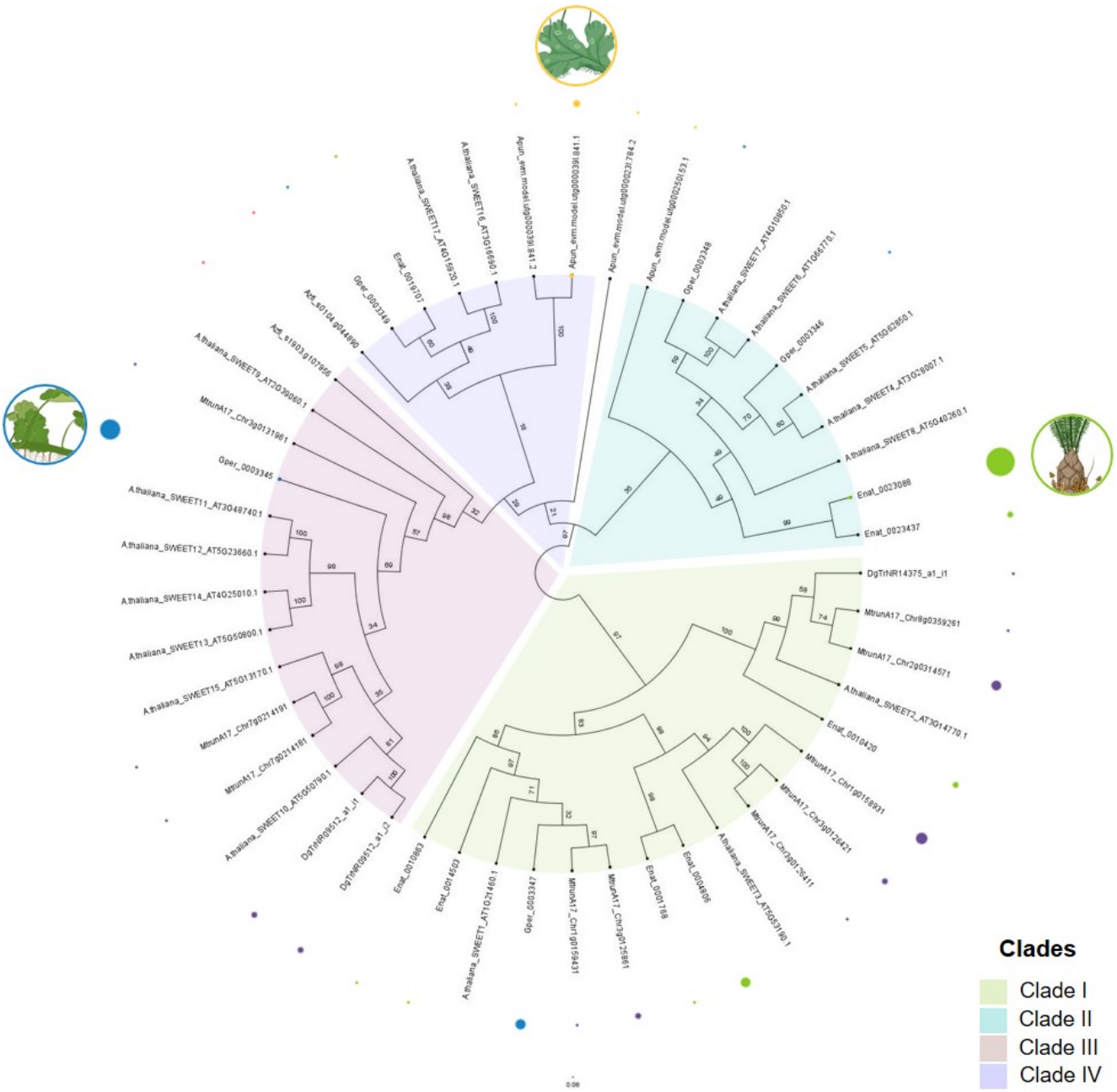


Figure 2.8. Trimmed phylogenetic analyses of protein sequences from the Sugars Will Eventually be Exported Transporters (SWEETs) proteins. The *Arabidopsis thaliana* SWEET sequences formed four clades; SWEET1-3 (Clade I, green shading), SWEET4-8 (Clade II, blue shading), SWEET9-15 (Clade III, pink shading) and SWEET16-17 (Clade IV, purple shading) indicated by the colour legend. The cyanobiosis lineage representative differential homologous sequences from *An. punctatus*, *Az. filiculoides*, *E. natalensis* and *G. perpensa* are scattered among the Arabidopsis SWEET sequences in all clades. Homologous sequences are included for *A. thaliana* and best differential BLASTp hits from *An. punctatus*, *Az. filiculoides*, *E. natalensis*, *G. perpensa*, *M. truncatula* and *D. glomerata*. Bootstrap values are indicated on branches. Species and protein IDs are included in each sequence label. Expression in symbiotic tissues are indicated for *An. punctatus* (orange), *Az. filiculoides* (red), *E. natalensis* (green), *G. perpensa* (blue), *M. truncatula* and *D. glomerata* (purple) next to each sequence label.

2.10. Supplementary data

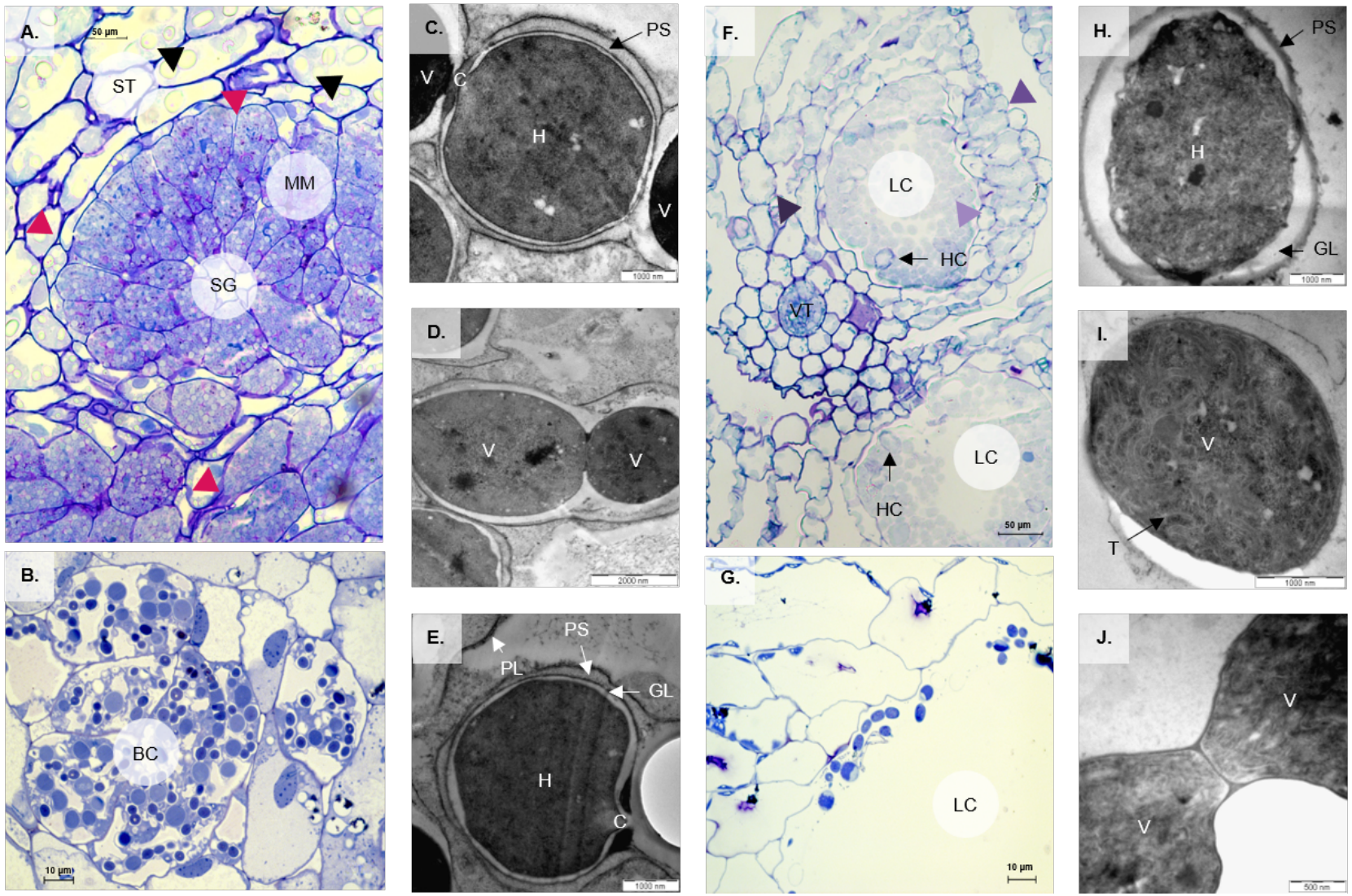


Figure S2.1. Light microscopy (A, B, F, G) and transmission electron microscopy (C-E, H-J) images of the symbiotic tissue in *Gunnera perpensa* (A-E) and *Azolla pinnatta* spp. *africana* (F-J). *(A)* Transverse section of stem tissue showing stem glands (SG) visible as cell clusters filled with cyanobacteria and a mucilaginous matrix (MM) interspersed by non-symbiotic cells. The SG are surrounded by a sheath (red arrowheads) consisting of oblong, non-symbiotic cells containing starch granules (black arrowheads). *(B)* Magnified view of the SG in *(A)* containing cyanobacteria. *(F)* Transverse section of a dorsal leaf lobe showing cyanobacteria within the leaf

cavity (LC). Hair cells (HC) and epidermis layers (purple arrowheads), the abaxial (medium purple arrowhead), inner (light purple arrowhead) and adaxial (dark purple arrowhead), are indicated. *(G)* Magnified view of the leaf cavity containing cyanobacteria. Heterocysts (H) and vegetative cells (V) in *Az. pinnatta* spp. *africana* (C-E) and *G. perpensa* (H-J) are indicated. Numerous carboxysomes can be seen in both heterocysts and vegetative cells. VT, vascular trace; C, polar cyanophycin granule; GL, glycolipid layer; PL, plasmalemma; PS, polysaccharide layer; T, thylakoid membranes.

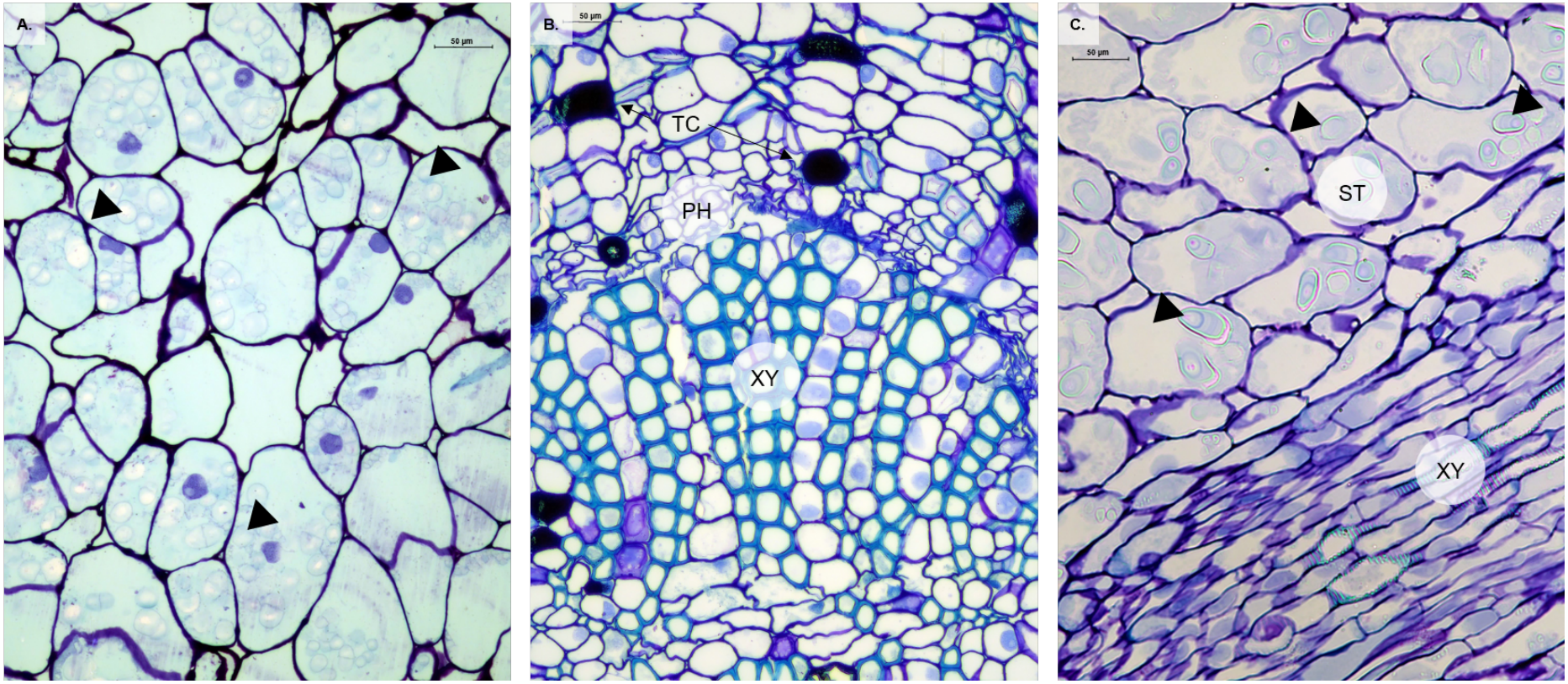


Figure S2.2 Light microscopy images of control tissue in *Encephalartos natalensis* and *Gunnera perpensa*. *(A, B) E. natalensis. (A)* Transverse section of primary root with cells containing starch granules (black arrowheads). *(B)* Transverse section of lateral root showing xylem (XY), phloem (PH) and tannin channels (TC). *(C)* Transverse section of *G. perpensa* stem tissue (ST) containing starch granules (black arrowheads).

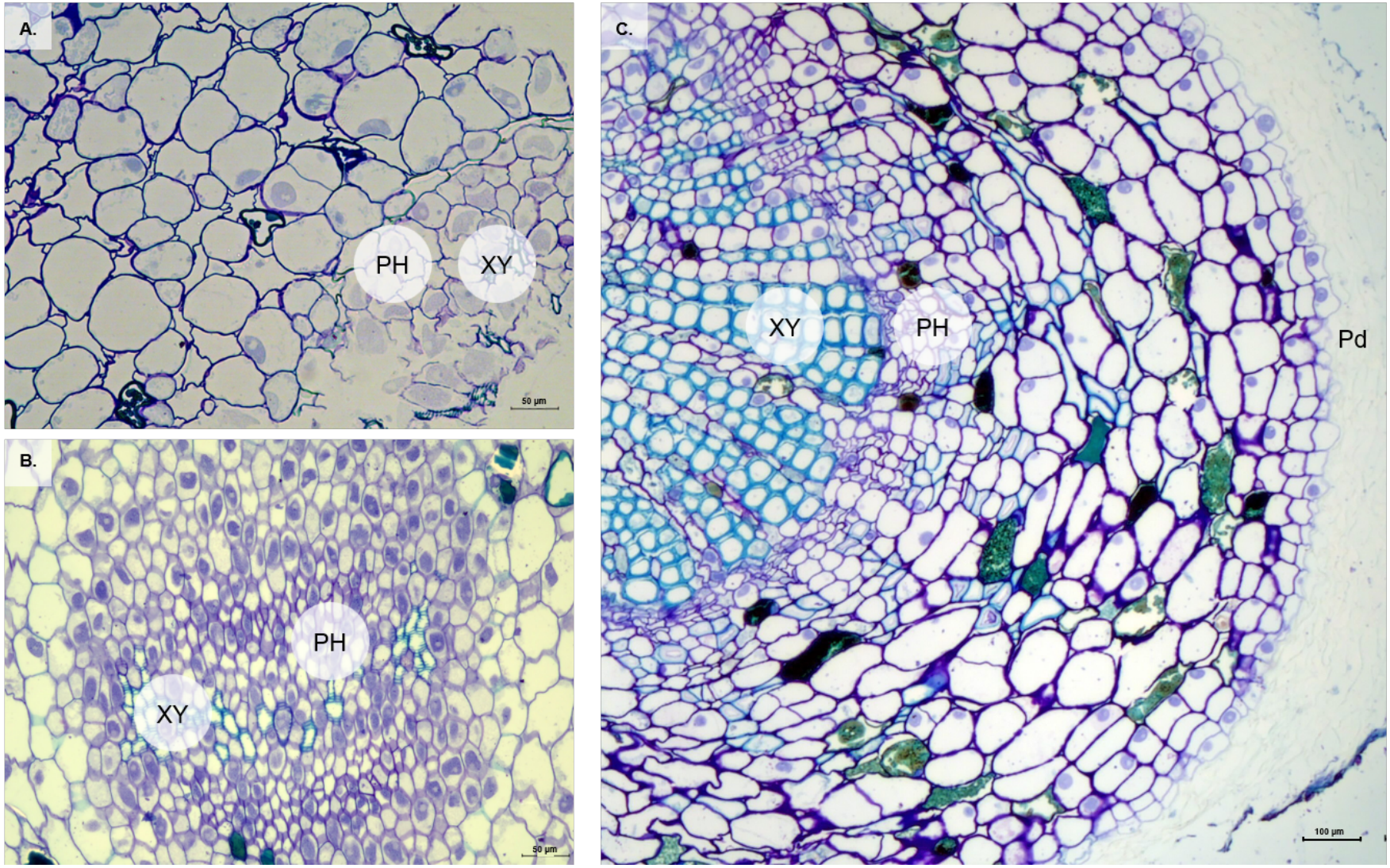


Figure S2.3. Light microscopy images of vascular tissue in *Encephalartos natalensis* roots. Xylem (XY) and phloem (PH) in *E. natalensis* coralloid root (A), precoralloid root (B) and lateral root (C) are indicated**.**

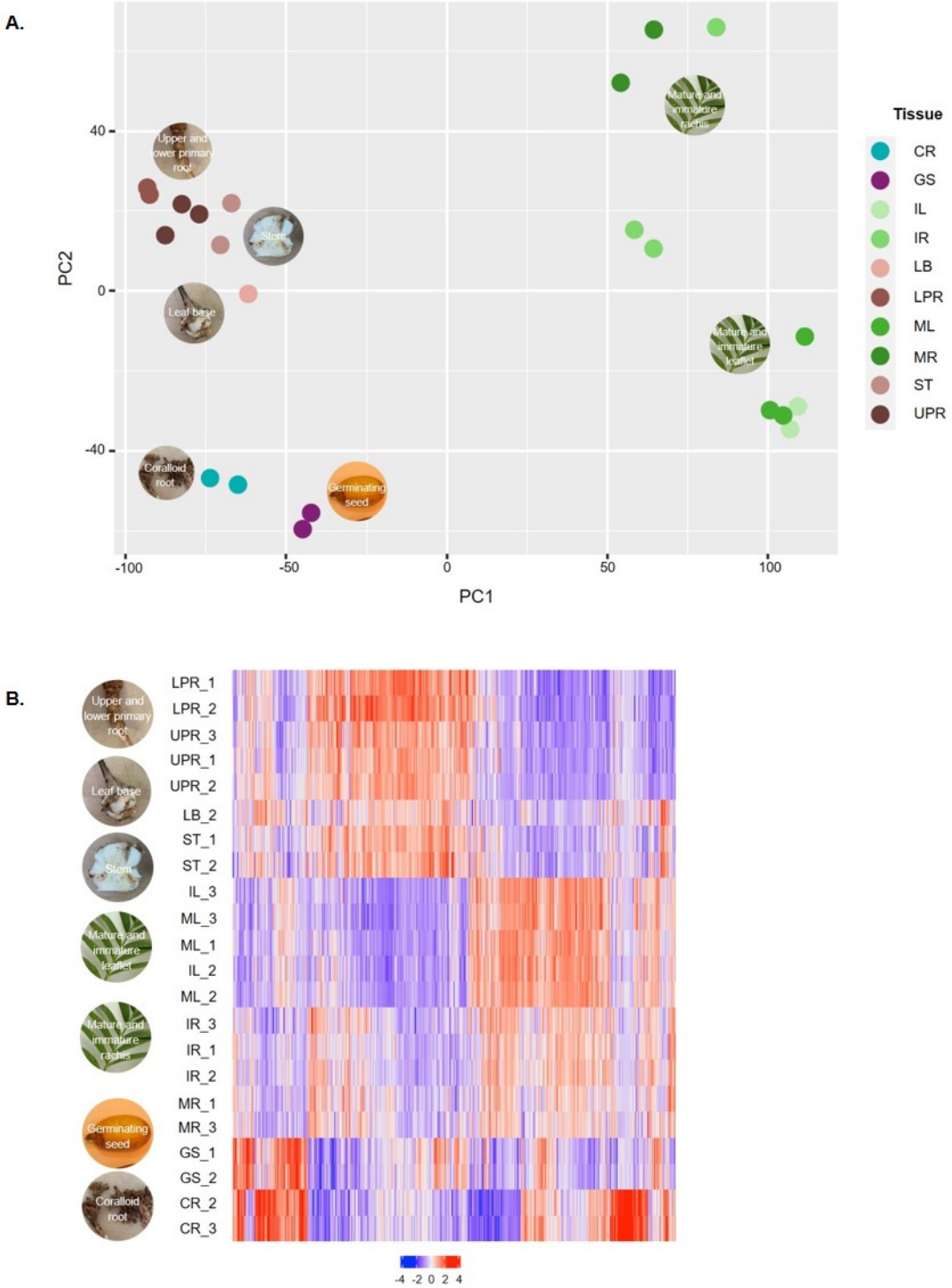


Figure S2.4. Transcriptomic data for the cycad *(Encephalartos natalensis).* *(A)* PCA plot for *E. natalensis* tissues. Coralloid roots group with germinating seed and are distinct from primary (upper and lower) root, stem and leaf base samples. *(B)* Hierarchical clustering of tissues and expression heatmap of genes (24,996 genes) in *E. natalensis*. Coralloid roots and germinating seeds cluster.

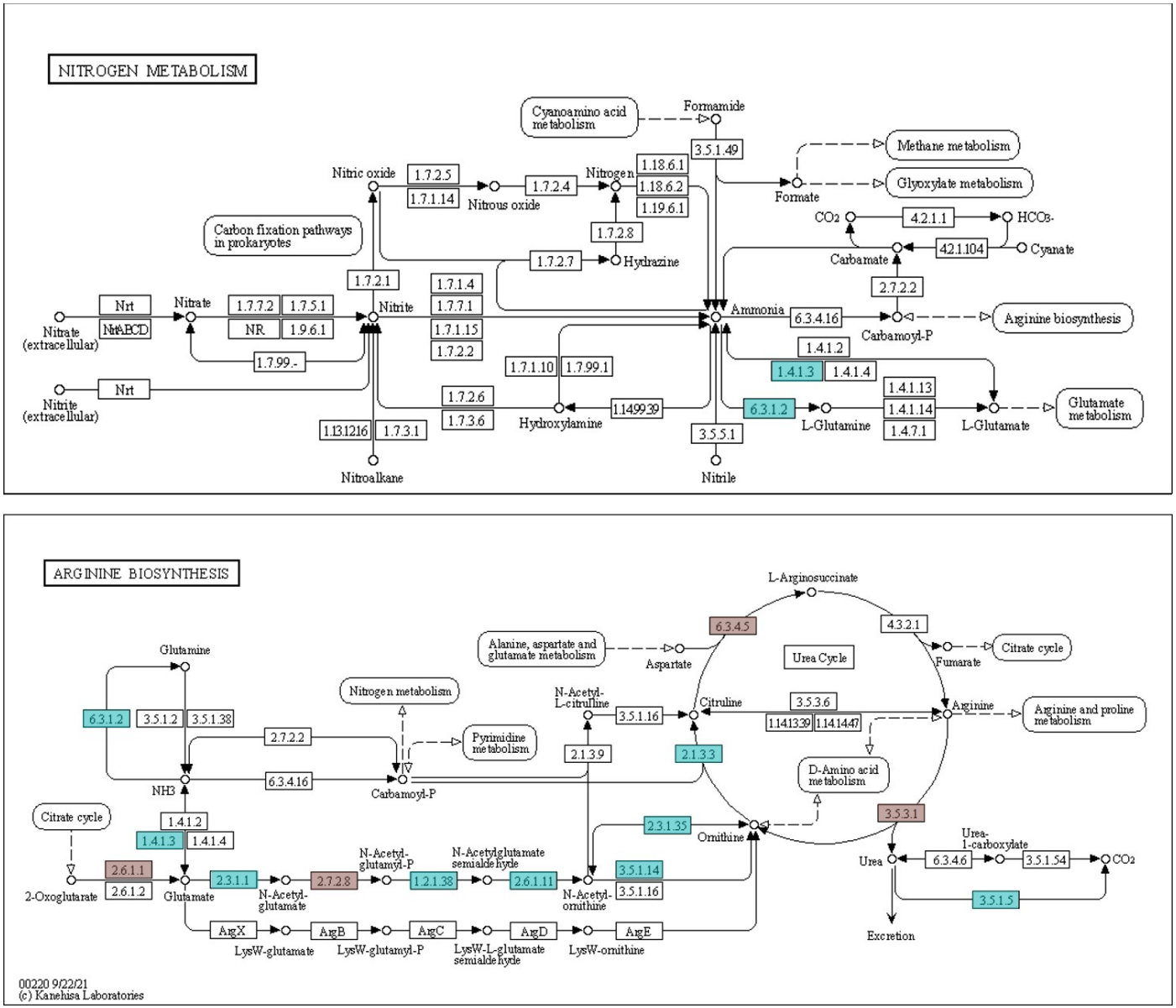
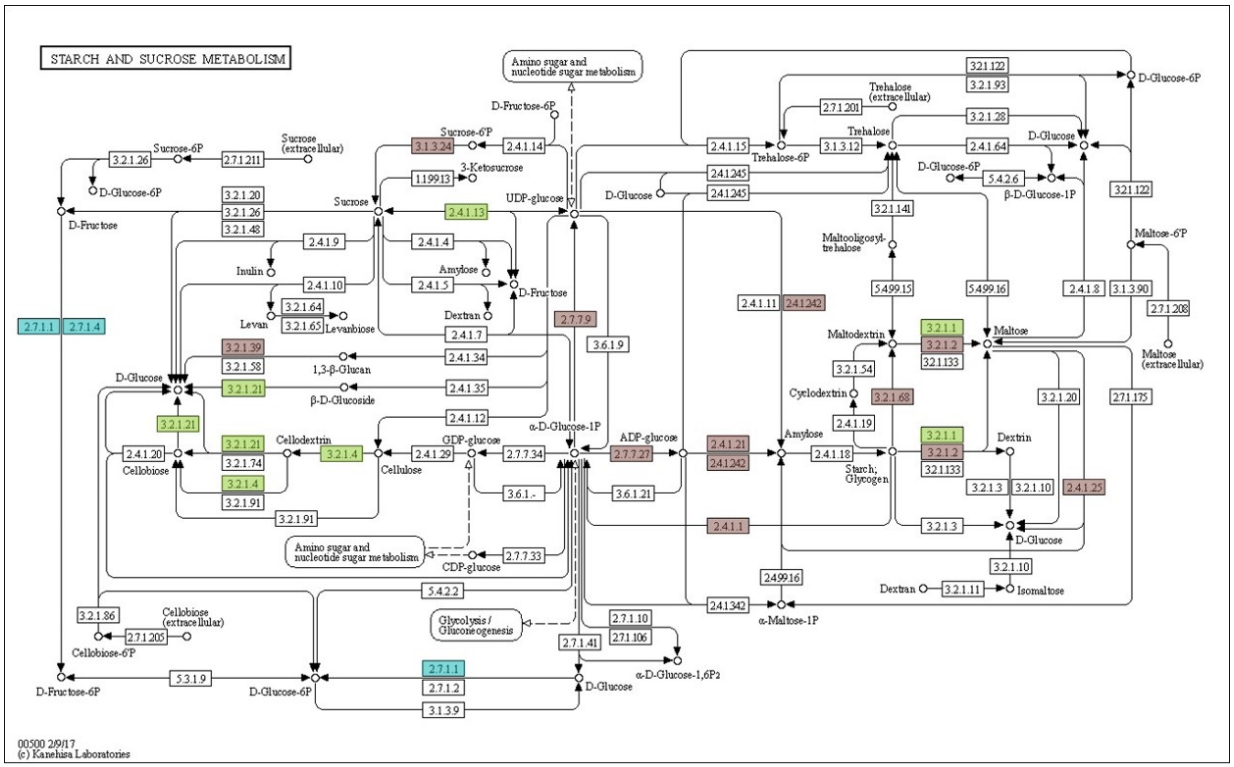


Figure S2.5. Nitrogen metabolism and arginine biosynthesis enrichments. Genes differentially expressed in cycad, E. natalensis, coralloid root (teal) or primary root (brown).



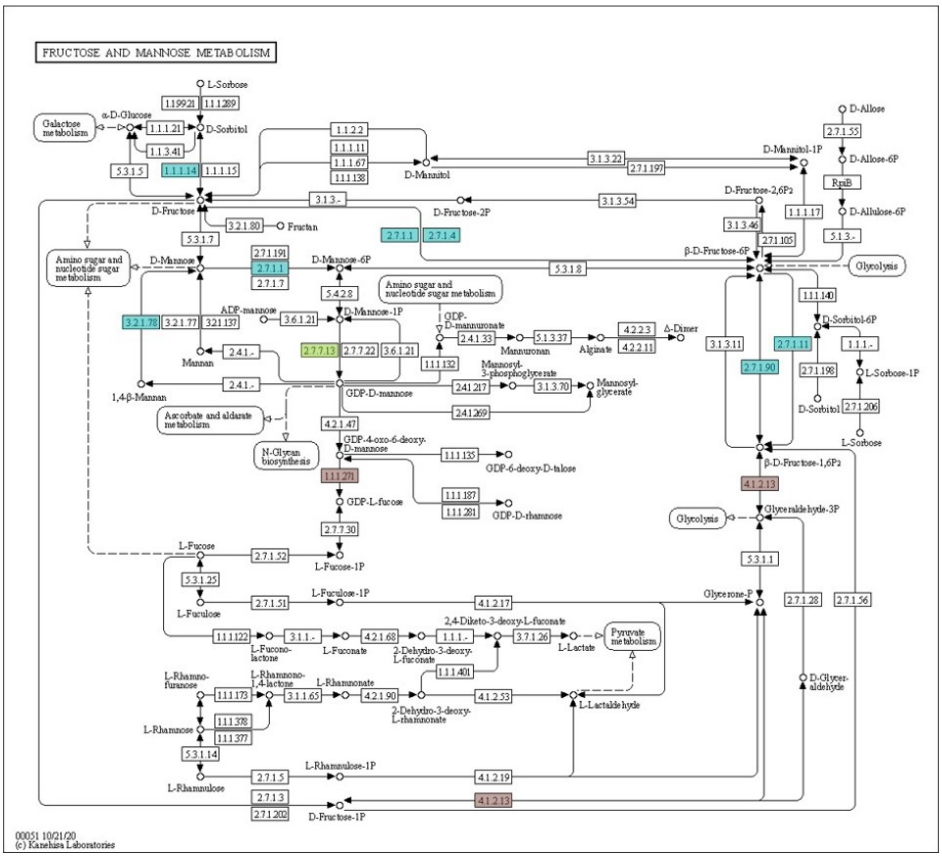


Figure S2.6. Starch, sucrose, fructose and mannose metabolism enrichments. Genes differentially expressed in cycad, *E. natalensis*, coralloid root (teal), primary root (brown) or having members expressed in coralloid and primary root (green).

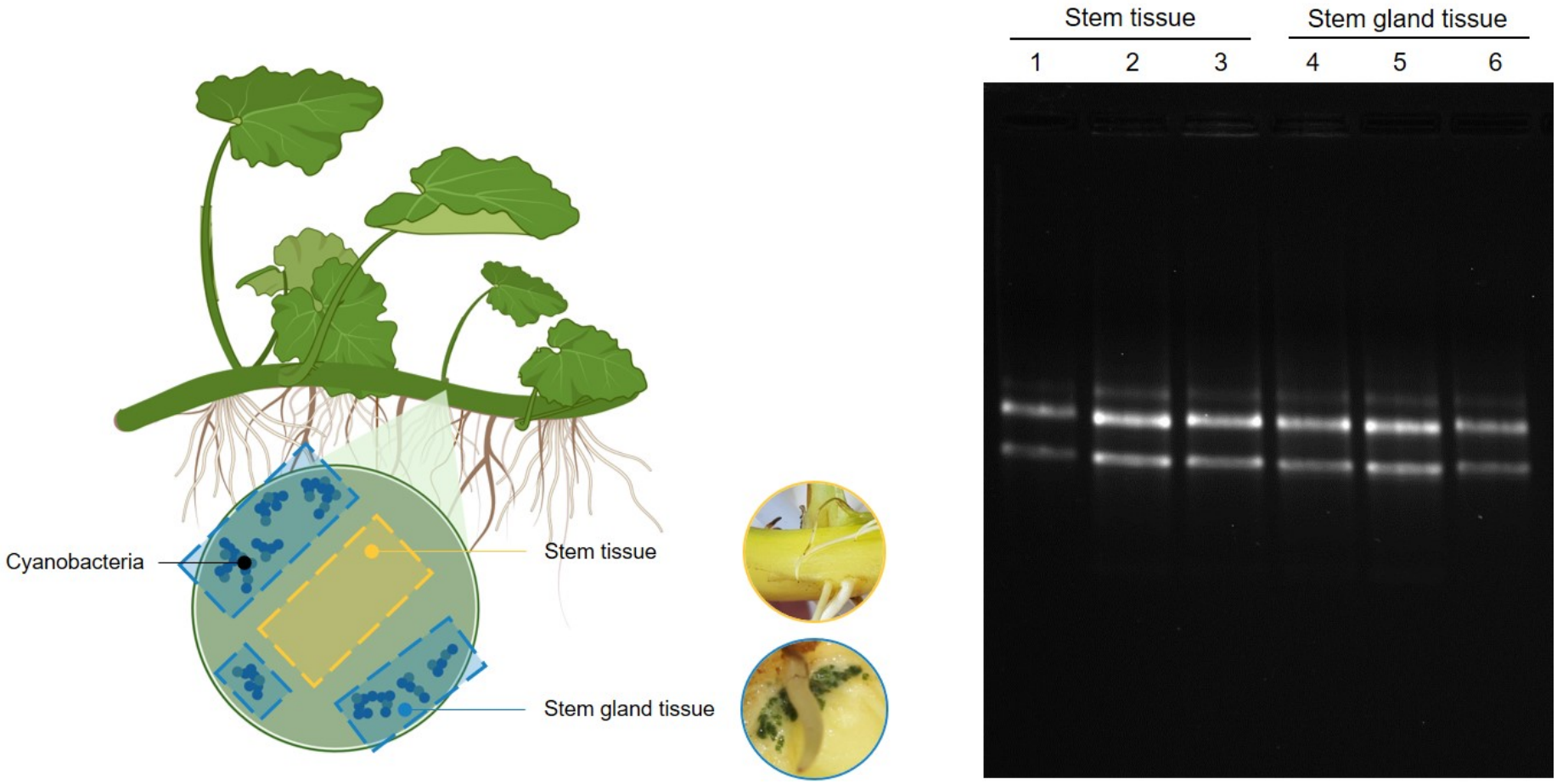


Figure S2.7. Samples from *Gunnera perpensa* used for RNA extraction. Illustration of stem and stem gland tissue sampling done for *G. perpensa* plants (left). A 1% (w/v) agarose gel indicating RNA from stem and stem gland tissue from *G. perpensa* (right). Extracted *G. perpensa* stem tissue RNA from plant 1 (lane 1), plant 2 (lane 2), and plant 3 (lane 3) along with stem gland RNA from plant 1 (lane 4), plant 2 (lane 5), and plant 3 (lane 6) are indicated.

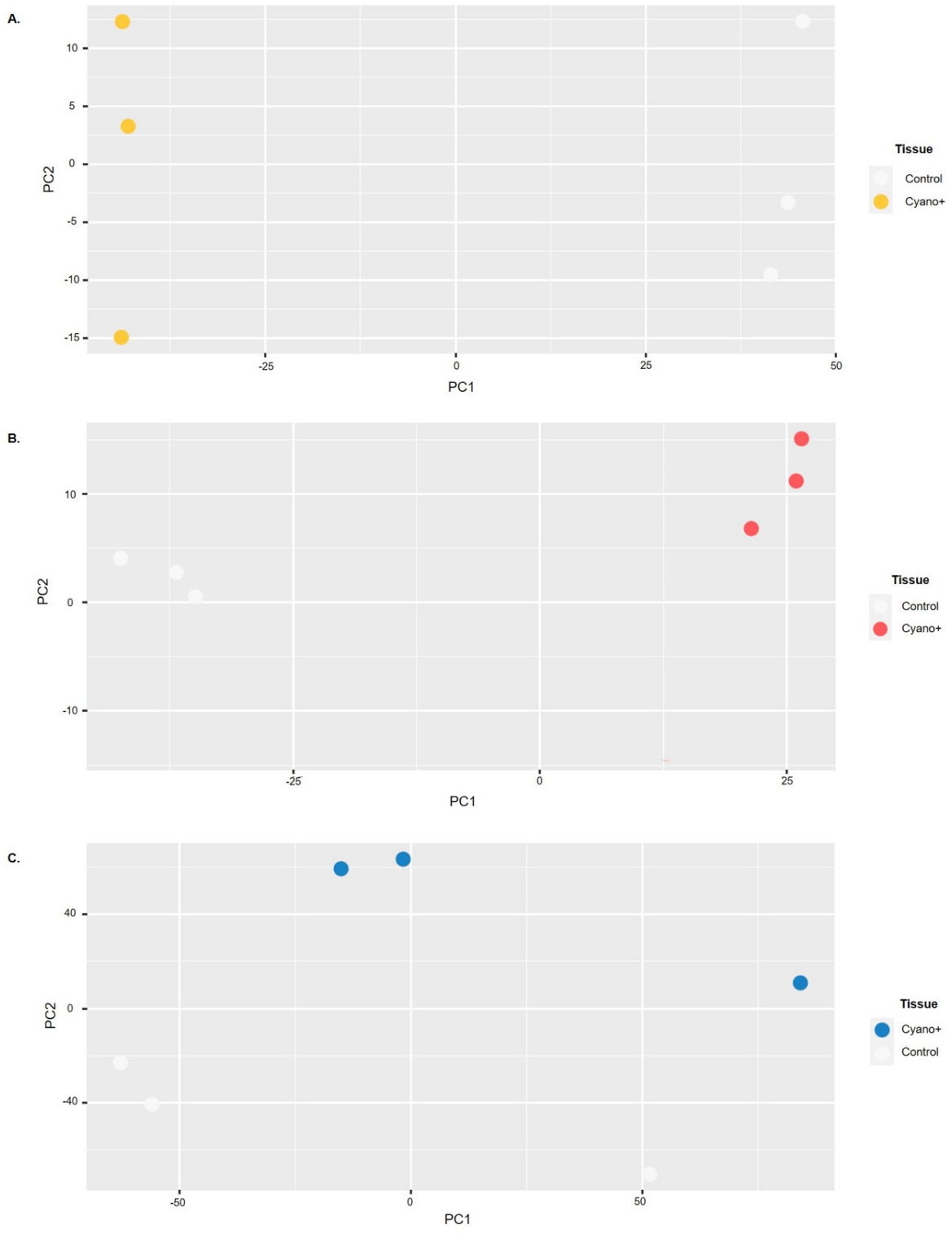


Figure S2.8. Transcriptomic data for *Anthoceros punctatus*, *Azolla pinnatta* spp. *africana* and *Gunnera perpensa*. PCA plots of genes differentially expressed in cyanobacterial containing (symbiotic) and control tissue for *An. punctatus* (A), *Az. filiculoides* (B) and *G. perpensa* (C). Cyanobacterial containing (symbiotic) tissue groups distinctly from control tissue in all species. Tissues are denoted by colours indicated in the legends.

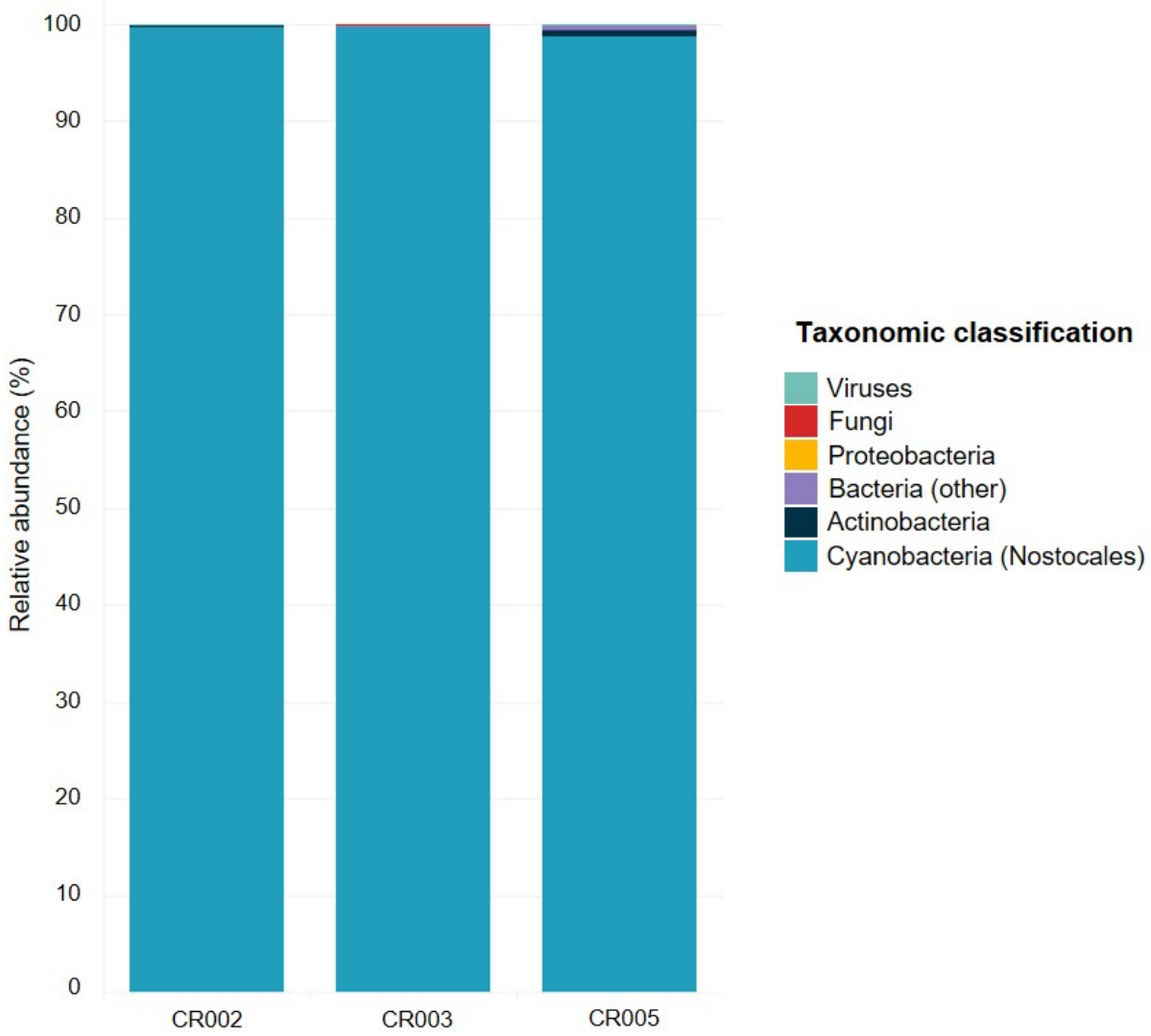


Figure S2.9. Taxonomic composition of endophytes in *Encephalartos natalensis* coralloid roots. Phylum and domain categories are listed in the legend to the right. Taxonomic classifications are based on reads that could be mapped to family level.

A.

B.

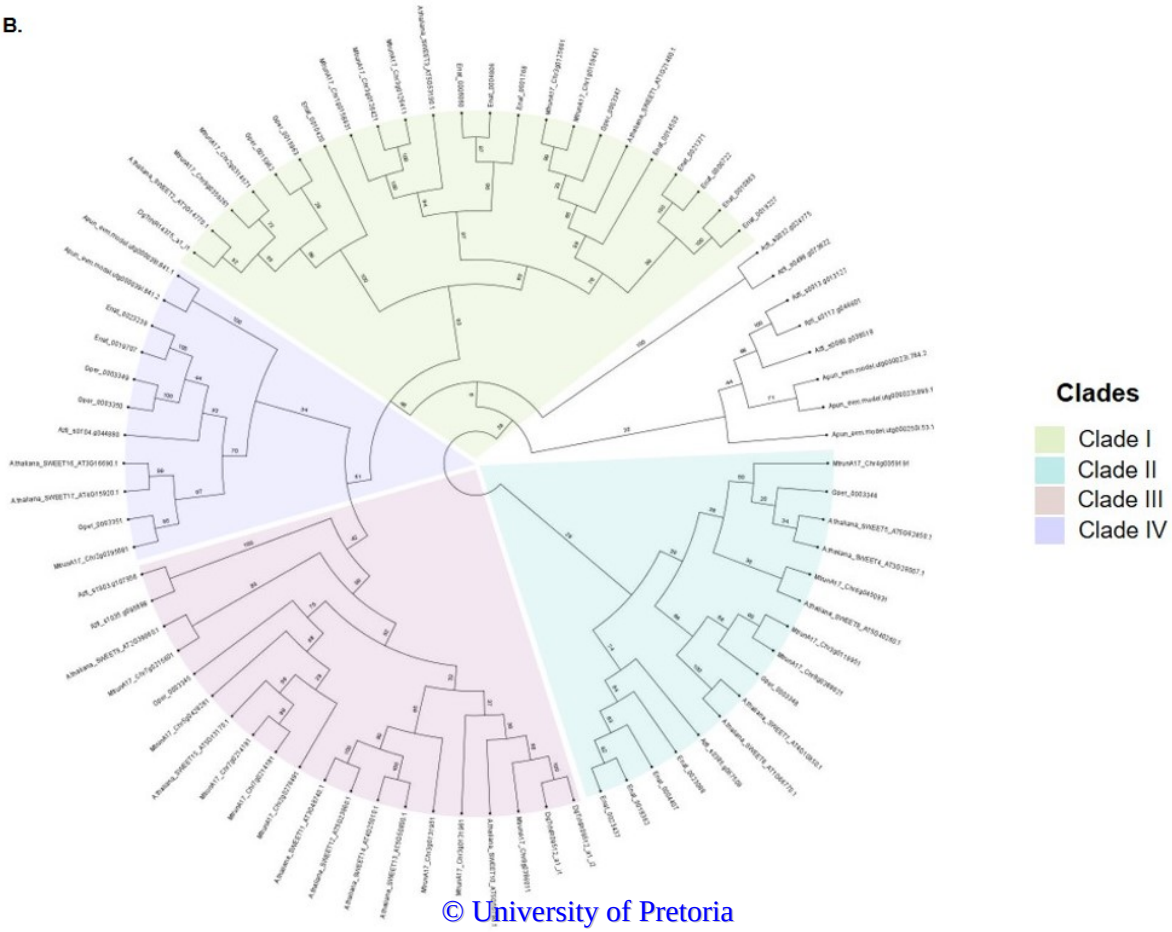


Figure S2.10. Trimmed phylogenetic analyses of protein sequences from the Sugars Will Eventually be Exported Transporters (SWEETs) proteins. *(A)* The *Arabidopsis thaliana* SWEET sequences formed four clades; SWEET1-3 (Clade I, green shading), SWEET4-8 (Clade II, blue shading), SWEET9-15 (Clade III, pink shading) and SWEET16-17 (Clade IV, purple shading) indicated by the colour legend. *(B)* The cyanobiosis lineage representative homologous sequences from *An. punctatus*, *Az. filiculoides*, *E. natalensis* and *G. perpensa* are scattered among the Arabidopsis SWEET sequences in all clades. Homologous sequences are included for *A. thaliana* and best BLASTp hits from *An. punctatus*, *Az. filiculoides*, *E. natalensis*, *G. perpensa*, *M. truncatula* and *D. glomerata*. Bootstrap values are indicated on branches. Species and protein IDs are included in each sequence label. Expression in symbiotic tissues are indicated for *An. punctatus* (orange), *Az. filiculoides* (red), *E. natalensis* (green), *G. perpensa* (blue), *M. truncatula* and *D. glomerata* (purple) next to each sequence label.

Table S2.1. Absolute counts and vegetative to heterocyst cell ratio from three biological replicate plants for *Azolla pinnatta* spp. *africana*, *Encephalartos natalensis* and *Gunnera perpensa*.

Species	Tissue type	Sample name	Vegetative cell count	Heterocyst count	Vegetative/heterocyst cell ratio	Percentage heterocysts (%)
Az. pinnatta spp. *africana*	Dorsal leaf lobe	AP_DL_4	7	–	–	0
			7	3	2.33	30
			4	–	–	0
			7	–	–	0
			4	–	–	0
		AP_DL_5	10	2	5.00	17
			8	2	4.00	20
			7	4	1.75	36
			9	–	–	0
			1	1	–	0
E. natalensis	Coralloid root	EN_CR_18	7	6	1.00	50
			3	4	2.00	33
			4	11	3.50	22
			1	–	0.75	57
			7	2	3.50	22
			2	1	1.17	46
		EN_CR_20	14	4	0.36	73
			8	18	0.44	69
			17	21	1.00	50
			18	21	0.10	91
			2	2	0.60	63
			12	19	0.63	61
			12	40	0.81	55
			9	15	0.86	54
			1	10	0.30	77
		EN_CR_26	7	25	0.80	56
			10	32	0.50	67
			5	42	0.50	67
			15	16	0.94	52
			6	12	0.37	73
			16	23	0.70	59
			7	19	0.28	78
			4	5	0.31	76
			7	14	0.12	89
G. perpensa	Stem gland	GP_SGC_29	4	3	1.33	43
			45	22	2.05	33

G. perpensa	Stem gland	GP_SGC_29	14	9	1.56	39
			42	26	1.62	38
			58	35	1.66	38
			10	5	2.00	33
			8	3	2.67	27
		GP_SGC_35	9	2	4.50	18
			8	3	2.67	27
			71	30	2.37	30
			42	17	2.47	29
			48	18	2.67	27
			49	25	1.96	34
			81	28	2.89	26
			50	23	2.17	32
			68	25	2.72	27
		GP_SGC_38	15	10	1.50	40
			12	4	3.00	25
			39	18	2.17	32
			31	28	1.11	47
			12	4	3.00	25
			14	5	2.80	26
			–	3	0.00	100
			4	2	2.00	33
			4	1	4.00	20
			39	15	2.60	28

Table S2.2. Summary of the mRNA-seq data obtained for each of the three biological replicate libraries for *Encephalartos natalensis* and *Gunnera perpensa.*

Species	Tissue type	Library name	Number of clean reads	Number of clean bases (Gb)	Clean data (Gb) per tissue	GC content (%)
E. natalensis	Mature leaflet	EN_ML_1_FW	35,793,487	11.0	59	47
		EN_ML_1_RW	35,793,487	11.0		47
		EN_ML_2_FW	33,760,464	9.5		47
		EN_ML_2_RW	33,760,464	9.5		47
		EN_ML_3_FW	31,937,859	9.0		46
		EN_ML_3_RW	31,937,859	9.0		46
	Immature leaflet	EN_IL_2_FW	34,804,196	9.8	37.4	46

E. natalensis	Immature leaflet	EN_IL_2_RW	34,804,196	9.8	37.4	46
		EN_IL_3_FW	31,532,536	8.9		47
		EN_IL_3_RW	31,532,536	8.9		47
	Mature rachis	EN_MR_1_FW	44,346,439	13.0	43.8	47
		EN_MR_1_RW	44,346,439	13.0		47
		EN_MR_3_FW	31,757,895	8.9		46
		EN_MR_3_RW	31,757,895	8.9		46
	Immature rachis	EN_IR_1_FW	35,176,829	9.9	63.8	46
		EN_IR_1_RW	35,176,829	9.9		46
		EN_IR_2_FW	36,329,652	11.0		47
		EN_IR_2_RW	36,329,652	11.0		47
		EN_IR_3_FW	37,607,945	11.0		48
		EN_IR_3_RW	37,607,945	11.0		49
	Leaf base	EN_LB_2_FW	23,707,724	6.7	13.4	47
		EN_LB_2_RW	23,707,724	6.7		47
	Stem	EN_ST_1_FW	31,385,089	8.8	37.2	46
		EN_ST_1_RW	31,385,089	8.8		46
		EN_ST_2_FW	34,678,896	9.8		47
		EN_ST_2_RW	34,678,896	9.8		47
	Upper primary root	EN_UPR_1_FW	32,266,928	9.1	56	45
		EN_UPR_1_RW	32,266,928	9.1		45
		EN_UPR_2_FW	31,926,465	9.0		44
		EN_UPR_2_RW	31,926,465	9.0		44
		EN_UPR_3_FW	35,048,567	9.9		44
		EN_UPR_3_RW	35,048,567	9.9		45
	Lower primary root	EN_LPR_1_FW	33,190,480	9.3	38	45
		EN_LPR_1_RW	33,190,480	9.3		45
		EN_LPR_2_FW	34,557,013	9.7		46
		EN_LPR_2_RW	34,557,013	9.7		46
	Coralloid root	EN_CR_2_FW	35,569,263	10.0	42	44
		EN_CR_2_RW	35,569,263	10.0		44
		EN_CR_3_FW	36,192,829	11.0		44
		EN_CR_3_RW	36,192,829	11.0		45
	Germinating seed	EN_GS_1_FW	36,478,147	11.0	39.4	46
		EN_GS_1_RW	36,478,147	11.0		46
		EN_GS_2_FW	31,035,126	8.7		45
		EN_GS_2_RW	31,035,126	8.7		45
G. perpensa	Stem gland	GP_SGC_1_CS04_FW	22,426,211	7.3	46.4	45
		GP_SGC_1_CS04_RW	22,426,211	7.3		46
		GP_SGC_2_CS05_FW	23,075,504	7.5		46
		GP_SGC_2_CS05_RW	23,075,504	7.5		46
		GP_SGC_3_CS06_FW	25,717,277	8.4		45

G. perpensa	Stem gland	GP_SGC_3_CS06_RW	25,717,277	8.4	46.4	46
	Stem	GP_ST_1_CS01_FW	30,049,839	9.8	49.6	45
		GP_ST_1_CS01_RW	30,049,839	9.8		46
		GP_ST_2_CS02_FW	22,071,733	7.2		45
		GP_ST_2_CS02_RW	22,071,733	7.2		46
		GP_ST_3_CS03_FW	23,772,373	7.8		46
		GP_ST_3_CS03_RW	23,772,373	7.8		46

Table S2.3. Summary of the *Encephalartos natalensis* and *Gunnera perpensa* transcriptome assemblies. The table summarises several statistics including the number of contigs and total annotations.

Species	Number of contigs	Number of bases	Mean contig length (bp)	N50 (bp)	N75 (bp)	Number of annotated contigs	BUSCO (complete)	BUSCO (fragmented)	BUSCO (missing)
E. natalensis	34,713	47,588,784	1,371	2,222	3,247	24,996	1,238	47	155
G. perpensa	33,346	37,344,818	1,120	1,714	2,572	23,101	1,301	40	99

Table S2.4. Summary of the tissue types collected from three biological replicate plants for *Encephalartos natalensis* and *Gunnera perpensa.*

Species	Tissue type	Sample name	Extraction method	PolyA RNA sequencing	Total RNA sequencing
E. natalensis	Mature leaflet	EN_ML_1	Plant/Fungi RNA Purification Kit	Yes	No
		EN_ML_2	Plant/Fungi RNA Purification Kit	Yes	No
		EN_ML_3	Plant/Fungi RNA Purification Kit	Yes	No
	Immature leaflet	EN_IL_1	Plant/Fungi RNA Purification Kit	Yes	No
		EN_IL_2	Plant/Fungi RNA Purification Kit	Yes	No
		EN_IL_3	Plant/Fungi RNA Purification Kit	Yes	No
	Mature rachis	EN_MR_1	Plant/Fungi RNA Purification Kit	Yes	No
		EN_MR_2	Plant/Fungi RNA Purification Kit	Yes	No
		EN_MR_3	Plant/Fungi RNA Purification Kit	Yes	No
	Immature rachis	EN_IR_1	Plant/Fungi RNA Purification Kit	Yes	No
		EN_IR_2	Plant/Fungi RNA Purification Kit	Yes	No
		EN_IR_3	Plant/Fungi RNA Purification Kit	Yes	No
	Leaf base	EN_LB_1	CTAB	Yes	No
		EN_LB_2	CTAB	Yes	No
		EN_LB_3	CTAB	Yes	No
	Stem	EN_ST_1	CTAB	Yes	No

E. natalensis	Stem	EN_ST_2	CTAB	Yes	No
		EN_ST_3	CTAB	Yes	No
	Upper primary root	EN_UPR_1	CTAB	Yes	Yes
		EN_UPR_2	CTAB	Yes	Yes
		EN_UPR_3	CTAB	Yes	Yes
	Lower primary root	EN_LPR_1	Plant/Fungi RNA Purification Kit	Yes	No
		EN_LPR_2	Plant/Fungi RNA Purification Kit	Yes	No
		EN_LPR_3	Plant/Fungi RNA Purification Kit	Yes	No
	Lateral root	EN_SR_1	Plant/Fungi RNA Purification Kit	Yes	Yes
		EN_SR_2	Plant/Fungi RNA Purification Kit	Yes	Yes
		EN_SR_3	Plant/Fungi RNA Purification Kit	Yes	Yes
	Precoralloid root	EN_PCR_1	Plant/Fungi RNA Purification Kit	Yes	Yes
		EN_PCR_2	Plant/Fungi RNA Purification Kit	Yes	Yes
		EN_PCR_3	Plant/Fungi RNA Purification Kit	Yes	Yes
	Coralloid root	EN_CR_1	Plant/Fungi RNA Purification Kit	Yes	Yes
		EN_CR_2	Plant/Fungi RNA Purification Kit	Yes	Yes
		EN_CR_3	Plant/Fungi RNA Purification Kit	Yes	Yes
G. perpensa	Leaf	GP_LF_1	Plant/Fungi RNA Purification Kit	Yes	No
		GP_LF_2	Plant/Fungi RNA Purification Kit	Yes	No
		GP_LF_3	Plant/Fungi RNA Purification Kit	Yes	No
	Stem gland	GP_SGC_1	CTAB	Yes	Yes
		GP_SGC_2	CTAB	Yes	Yes
		GP_SGC_3	CTAB	Yes	Yes
	Stem	GP_ST_1	CTAB	Yes	Yes
		GP_ST_2	CTAB	Yes	Yes
		GP_ST_3	CTAB	Yes	Yes
	Root	GP_RT_1	Plant/Fungi RNA Purification Kit	Yes	No
		GP_RT_2	Plant/Fungi RNA Purification Kit	Yes	No
		GP_RT_3	Plant/Fungi RNA Purification Kit	Yes	No

Table S2.5. Summary of the DNA-seq data obtained for each of the three biological replicate libraries for *Encephalartos natalensis* coralloid root tissue.

Sample name	Number of untrimmed reads (paired)	Number of trimmed reads (paired)	Number of unclassified reads	Number of classified reads	Reads mapped to Viridiplantae
CR002	35,329,749	31,550,291	27,132,270 (86%)	4,417,867 (14%)	84,529
CR003	4,723,941	4,234,083	3,614,933 (85.4%)	619,129 (14.6%)	12,915
CR005	49,984,602	47,524,148	40,066,829 (84.3%)	7,457,068 (15.7%)	230,997

Supplementary Data Set 2.1: Comprehensive candidate gene list.

Supplementary Data Set 2.2: Relative expression levels of significantly differentially expressed genes from *Anthoceros punctatus*, *Azolla filiculoides*, *Encephalartos natalensis*, *Gunnera perpensa*, *Medicago tuncatula* and *Dactylis glomerata*.

Supplementary Data Set 2.3: Taxonomic classification to family level of *Encephalartos natalensis* coralloid roots from all biological repeats.

CHAPTER 3

CONCLUDING REMARKS

Plant diversification and evolution was dependent on plant-microbe symbioses, which not only arose during, and likely facilitated land plant terrestrialisation around ~450 million years ago (Pirozynski and Malloch, 1975; Kenrick and Crane, 1997), but also has an ongoing impact on nutrient acquisition (van der Heijden et al., 2008; van der Heijden et al., 2016). The earliest fossil evidence for plant symbioses dates to the early-Devonian era (400 mya; Remy et al., 1994; Taylor et al., 1995), the forerunners of today's well-known mycorrhizae and root-nodule symbioses. It was not until the discovery of the common symbiotic pathway (CSP) and its conservation in intracellular symbioses that the molecular mechanisms governing these partnerships were uncovered (Oldroyd, 2013; Radhakrishnan et al., 2020). Well-known mycorrhizae and root-nodule symbioses are but two evolutionary innovations linked to the acquisition of nitrogen via partnerships with microorganisms. Plant-cyanobacterial symbiosis albeit a rare trait is another, but information on these partnerships have remained inadequate. Cyanobacterial symbiosis evolved independently across unrelated lineages of land plants (Raven, 2002), unlike those involving symbioses with arbuscular mycorrhizae (reviewed in Brundrett, 2002; Brundrett and Tedersoo, 2018), and rhizobia or *Frankia* diazotrophic bacteria and flowering plants (Soltis et al., 1995; Swensen, 1996). Most plant-microbe studies focus mainly on the angiosperms because of their species richness as well as their importance as crops, however, the genetic mechanisms underlying cyanobacterial partnerships emerged long before its partnership with the angiosperm *Gunnera* spp. and were already present in the early-diverging land plants (Raven, 2002). To truly understand plant symbioses for future synthetic biology, in particular the cyanobacterial partnerships, all lineages of plants need to be considered.

Nitrogen (N), among other nutrients, is an increasingly important resource globally. Not only has it been, and continues to be, important in both natural and agricultural ecosystems, but its limited availability in most soils is becoming an increasing concern as global populations and climate change are on the rise (reviewed in Guignard et al., 2017). Nitrogen is an essential element in plant metabolism and development and consequently a limiting factor in plant growth (Guignard et al., 2017). Therefore, synthetic nitrogen fertilisers are extensively used on agricultural crops each year to increase agricultural productivity around the world (Stewart et al., 2005). Unfortunately, their manufacturing and subsequent use is ineffective, since as much as half of the nitrogen applied to agricultural crops are not used by the plants instead causing a surplus that is lost to aquatic and terrestrial ecosystems (reviewed in Conant et al., 2013; Stevens, 2019). Ultimately contributing to various environmental problems, including biodiversity loss and further climate change (Conant et al., 2013; Stevens, 2019). Because of these environmental and ecological concerns and costs, there is a need to reduce the use of nitrogen fertiliser in agricultural systems. In both early land plant evolution and today's barren landscapes, the need to access nitrogen was and remains a major driver for the emergence of symbiotic interactions (reviewed in van der Heijden et al., 2015). Understanding the emergence of plant mutualistic symbioses as adaptations for nitrogen acquisition and the underlying genes and pathways involved are one of the first steps towards engineering these complex traits into crops of interest.

The work conducted in this MSc dissertation furthers our understanding of the evolution and development of plant-cyanobacterial partnerships for nitrogen acquisition, particularly in the cycad, *Encephalartos natalensis*. In Chapter 1, I provided a comprehensive summary of the emergence of plant mutualistic symbioses as an adaptation for nitrogen acquisition following land plant terrestrialisation. Special emphasis was placed on the evolution and establishment of mycorrhizae and root nodule symbioses, and their associated genes and proteins. This review highlighted that despite being important for nitrogen acquisition, not all plant symbioses have been afforded the same attention with cyanobacterial partnerships remaining largely elusive. Therefore, in chapter 2, I specifically focused on publicly available genomes and transcriptomes, as well as high confidence gene catalogues generated in this study, from representative cyanobiosis-forming plant species to investigate the unique biology of this partnership. A multi-disciplinary approach was utilised, encompassing light and transmission microscopy, transcriptomics and phylogenetics to study the underlying genetic basis for cyanobacterial partnerships in the cycad *E. natalensis* with other representative species serving as comparative analysis, including the hornwort *Anthoceros punctatus*, the fern *Azolla filiculoides* and the angiosperm *G. perpensa*. In addition, the designed methodology was applied to two root nodule hosts *Medicago truncatula* (legume) and *Dactylis glomerata* (grass) which served as additional comparative analyses.

The genetic mechanisms required for the development and maintenance of plant mutualistic symbioses predate the emergence of the first land plants and their subsequent improvement has since led to a variety of symbiotic associations present today (reviewed in Delaux and Schornack, 2021). I, therefore, hypothesised that despite plant-cyanobacterial partnerships having evolved independently (Raven, 2002) and also differing anatomically, they have drawn on pre-existing genes and pathways utilised by mycorrhizal and rhizobial interactions for their functioning, but to what extent was unclear (Figure 3.1). My investigations lend support to this hypothesis. The results of the study revealed homologs of key common symbiotic pathway genes to be highly conserved in at least two cyanobacterial symbioses, the hornwort *A. punctatus* and cycad *E. natalensis*. While our research provided candidate common symbiotic pathway genes potentially involved in the cyanobacterial partnership with cycad that can be the subject of further investigations, a key limitation was the inability to conclusively attribute these same homologs to all cyanobiosis-forming plant species investigated. Although available studies on the biology of *A. punctatus* and *Az. filiculoides* has highlighted the presence and absence of not only common symbiotic pathway orthologs but also those evolutionary linked to arbuscular mycorrhizae in these respective hosts (Li et al., 2018; Li et al., 2020), further investigations need to be done for the angiosperm *G. perpensa*. Dated studies from the late 1900s and early 2000s imply that all cyanobiosis lineage representatives aside from *Azolla* species retained arbuscular mycorrhizal symbiosis (Harley and Harley, 1987a, b; Gemma et al., 1992; Wang and Qiu, 2006). While this might be the case, no morphological or metagenomic evidence was found in our study to support this at least in the coralloid roots of *E. natalensis*. Together our results seem to suggest that cyanobacterial symbioses (in hosts that have retained mycorrhizal symbioses) recruited arbuscular mycorrhizae genes (CSP genes) in a similar way as root nodule symbioses have. Future investigations to fully resolve the possible neofunctionalisation of common

symbiotic pathway genes from pre-existing arbuscular mycorrhizae symbiosis and the evolutionary path taken by these distinct plant lineages for subsequent cyanobacterial partnerships will benefit from high-quality genomes, time-based colonising and cell level 'omics' as well as full microbiome genome sequencing.

The comprehensive, tissue-specific transcriptomic dataset from cycad along with the in-depth anatomical analysis of its various root tissues gave us incredible insight into the unique biology of the *E. natalensis*-cyanobacterial partnership (Figure 3.2). While it has long since been speculated that putative carbon and nitrogen sources involved in cyanobacterial partnerships, in particular with cycad, are sucrose, fixed carbon and amino acids (Pate et al., 1988; Rai et al., 2000; Rai et al., 2002; Rai et al., 2019). Our multi-disciplinary approach uncovered (potentially for the first time) carbon and nitrogen metabolising enzymes pointing to potential carbon and nitrogen sources. Metabolic pathway enrichments performed in chapter 2 linked genes involved in starch and sucrose catabolism, and citrulline and ornithine biosynthesis to *E. natalensis* coralloid roots, findings that together with anatomical investigations and a previous study done by Pate et al. (1988) greatly add to the growing body of knowledge of this partnership. Of particular note, our results suggest an independent acquisition of Sugars Will Eventually be Exported Transporters (SWEETs) in cyanobiosis-forming plant lineages, utilising, like the root nodule angiosperms, genes from an existing and ancestral family. While we could not resolve whether the cyanobiont produces these amino acids and subsequently provides it to the cycad or the cyanobiont provides ammonia (NH_3) and the cycad uses this precursor to synthesise the amino acids, our data along with the canonical though in other species seems to supports the latter. Future investigations to resolve the true carbon and nitrogen sources provided by the cyanobiont will benefit from total RNA sequencing and in-depth metabolomics studies.

This study opens subsequent areas of interest particularly associated with cyanobacterial-plant partnerships which have not yet enjoyed much attention. As mentioned above, future research will benefit from further and high-resolution 'omics' studies (genomics, metabolomics, cell-level and total RNA transcriptomics) in understudied plant lineages to resolve and identify additional genes involved in cyanobacterial symbioses. Similarly, as with the comprehensive, tissue-specific transcriptomic dataset we generated here for the cycad, *E. natalensis*, further symbiosis research would not only benefit from tissue-specific but time-scale data as well. Tissue-specific and time-scale data will shed light on the roles of early signalling genes involved in the development and establishment of this unique symbiosis. In addition, the function of candidate genes identified here and from future studies can be investigated to determine whether or not they can rescue *M. truncatula* mutant plants. This study has highlighted many additional questions and new directions of research into cyanobacterial symbioses that will contribute to our knowledge of mutualistic plant symbioses even further. This study has not only added fundamental knowledge about the last remaining critically under-studied plant-microbe symbiosis for nitrogen acquisition, not only providing invaluable resources but also sets the groundwork for future advanced synthetic biology applications in engineering plant-cyanobacterial symbioses for agriculture.

3.1. References

Ahern CP, Staff IA (1994) Symbiosis in cycads: the origin and development of coralloid roots in *Macrozamia communis* (Cycadaceae). Am J Bot **81**: 1559-1570.

Brundrett MC (2002) Coevolution of roots and mycorrhizas of land plants. New Phytol **154**: 275-304.

Brundrett MC, Tedersoo L (2018) Evolutionary history of mycorrhizal symbioses and global host plant diversity. New Phytol **220**: 1108-1115.

Conant RT, Berdanier AB, Grace PR (2013) Patterns and trends in nitrogen use and nitrogen recovery efficiency in world agriculture. Glob Biogeochem Cycles **27**: 558-566.

Delaux PM, Schornack S (2021) Plant evolution driven by interactions with symbiotic and pathogenic microbes. Science **371**: eaba6605.

Gemma JN, Koske RE, Flynn T (1992) Mycorrhizae in Hawaiian pteridophytes: occurrence and evolutionary significance. Am J Bot **79**: 843-852.

Guignard MS, Leitch AR, Acquisti C, Eizaguirre C, Elser JJ et al. (2017) Impacts of nitrogen and phosphorus: from genomes to natural ecosystems and agriculture. Front Ecol Evol **5**: 677-679

Harley JL, Harley EL (1987a) A check-list of mycorrhiza in the British flora. New Phytol **105**: 1-102.

Harley JL, Harley EL (1987b). A check-list of mycorrhiza in the British flora-addenda, errata and index. New Phytol **107**: 741-749.

Kenrick P, Crane PR (1997) The origin and early evolution of plants on land. Nature **389**: 33-39.

Li FW, Nishiyama T, Waller M, Frangedakis E, Keller J et al. (2020) *Anthoceros* genomes illuminate the origin of land plants and the unique biology of hornworts. Nat Plants **6**: 259-272.

Li FW, Brouwer P, Carretero-Paulet L, Cheng S, de Vries J et al. (2018) Fern genomes elucidate land plant evolution and cyanobacterial symbioses. Nat Plants **4**: 460-472.

Oldroyd GED (2013) Speak, friend, and enter: signalling systems that promote beneficial symbiotic associations in plants. Nat Rev Microbiol **11**: 252-263.

Pate JS, Lindblad P, Atkins CA (1988) Pathways of assimilation and transfer of fixed nitrogen in coralloid roots of cycad-*Nostoc* symbioses. Planta **176**: 461-471.

Pirozynski KA, Malloch DW (1975) The origin of land plants: a matter of mycotrophism. BioSyst **6**: 153-164.

Radhakrishnan GV, Keller J, Rich MK, Vernié T, Mbadinga Mbadinga DL et al. (2020) An ancestral signalling pathway is conserved in intracellular symbioses-forming plant lineages. Nat Plants **6**: 280-289.

Rai AN, Söderbäck E, Bergman B (2000) Tansley Review No. 116 Cyanobacterium–plant symbioses. New Phytol **147**: 449-481.

Rai AN, Bergman B, Rasmussen U (2002) Cyanobacteria in symbiosis. (Springer).

Rai AN, Singh AK, Syiem MB (2019) Plant growth-promoting abilities in cyanobacteria. In Cyanobacteria from Basic Science to Applications, AK Mishra, DN Tiwari, and AN Rai, eds (Academic Press), pp. 459-476.

Raven JA (2002) Evolution of cyanobacterial symbioses. In Cyanobacteria in Symbiosis, AN Rai, B Bergman, and U Rasmussen, eds (Dordrecht: Springer Netherlands), pp. 329-346.

Remy W, Taylor TN, Hass H, Kerp H (1994) Four hundred-million-year-old vesicular arbuscular mycorrhizae. Proc Natl Acad Sci USA **91**: 11841-11843.

Soltis DE, Soltis PS, Morgan DR, Swensen SM, Mullin BC et al. (1995) Chloroplast gene sequence data suggest a single origin of the predisposition for symbiotic nitrogen fixation in angiosperms. Proc Natl Acad Sci USA **92**: 2647-2651.

Stevens CJ (2019) Nitrogen in the environment. Science **363**: 578-580.

Stewart WM, Dibb DW, Johnston AE, Smyth TJ (2005) The contribution of commercial fertilizer nutrients to food production. Agron. J **97**: 1-6.

Swensen SM (1996) The evolution of actinorhizal symbioses: evidence for multiple origins of the symbiotic association. Am J Bot **83**: 1503-1512.

Taylor TN, Remy W, Hass H, Kerp H (1995) Fossil arbuscular mycorrhizae from the early Devonian. Mycologia **87**: 560-573.

van der Heijden MG, Martin FM, Selosse MA, Sanders IR (2015) Mycorrhizal ecology and evolution: the past, the present, and the future. New Phytol **205**: 1406-1423.

van der Heijden MGA, Bardgett RD, van Straalen NM (2008) The unseen majority: soil microbes as drivers of plant diversity and productivity in terrestrial ecosystems. Ecol Lett **11**: 296-310.

van der Heijden MGA, Bruin SD, Luckerhoff L, van Logtestijn RSP, Schlaeppi K (2016) A widespread plant-fungal-bacterial symbiosis promotes plant biodiversity, plant nutrition and seedling recruitment. ISME J. **10**: 389-399.

Wang B, Qiu YL (2006) Phylogenetic distribution and evolution of mycorrhizas in land plants. Mycorrhiza **16**: 299-363.

3.2. Figures

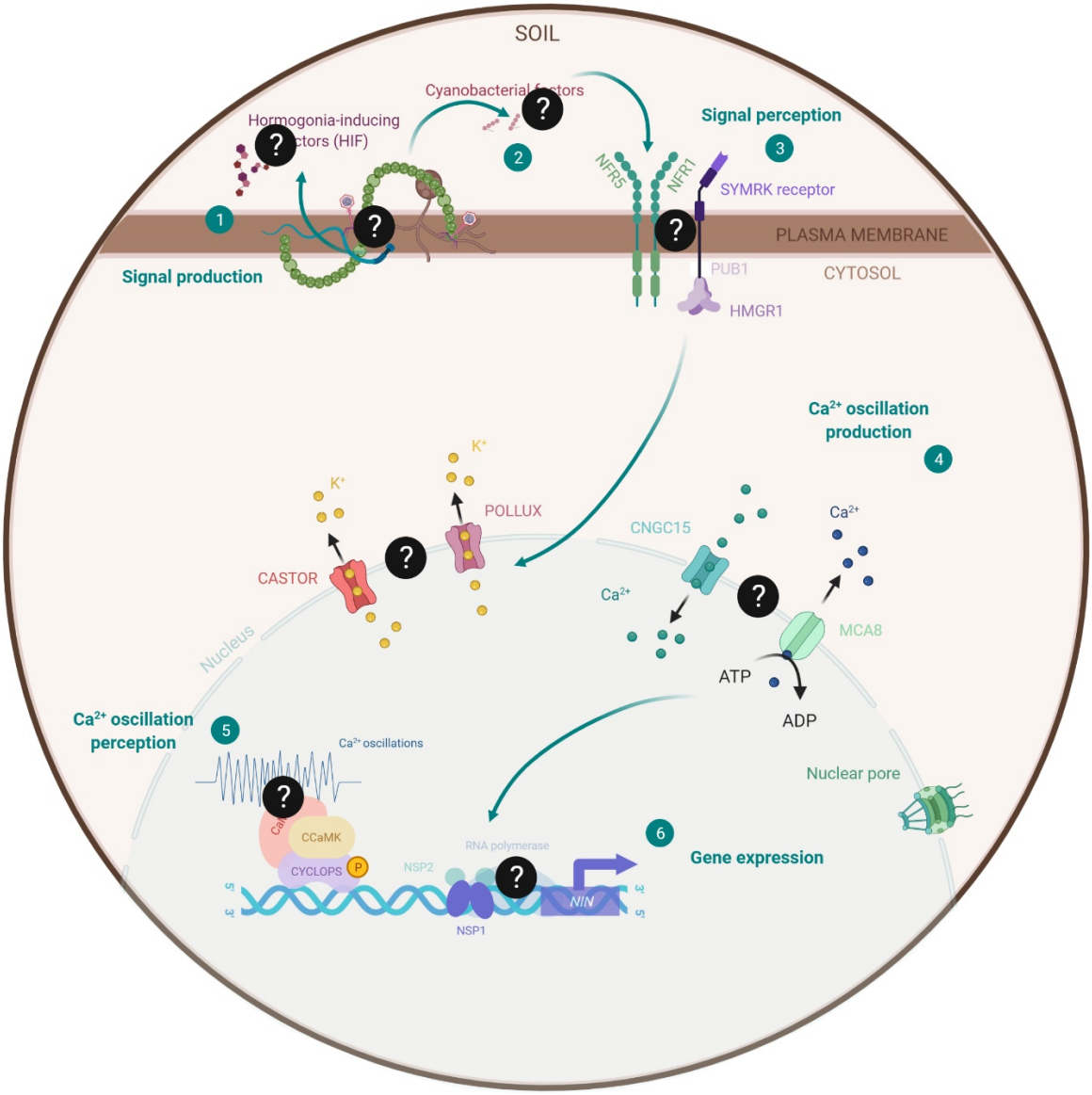


Figure 3.1. Common symbiosis pathway genes for cyanobacterial symbioses and the cycad, *Encephalartos natalensis*. (1, 2) Signal production: The production of hormogonia-inducing factors (HIFs) by plants and the subsequent signal production by cyanobacteria (if any) is unknown. The presence of microbes other than cyanobacteria is unclear, but metagenomic data seems to suggest that cyanobacteria dominate the coralloid roots. (3) Signal perception: The perception of hormogonia (if any) by unknown receptors. (4) Calcium oscillation production: downstream of unknown symbiotic receptors, nuclear calcium oscillation is caused by the movement of calcium into the nucleoplasm. Potassium-permeable channels, CASTOR and POLLUX, compensate for the resulting depolarisation while an unknown calcium symporter (if present) pumps calcium back into the nuclear envelope. (5) Calcium oscillation perception: CCaMK forms a complex with phosphorylated CYCLOPS to decode calcium oscillations. (6) Gene expression: Downstream transcription factors and genes are still unknown. Created with Biorender.com.

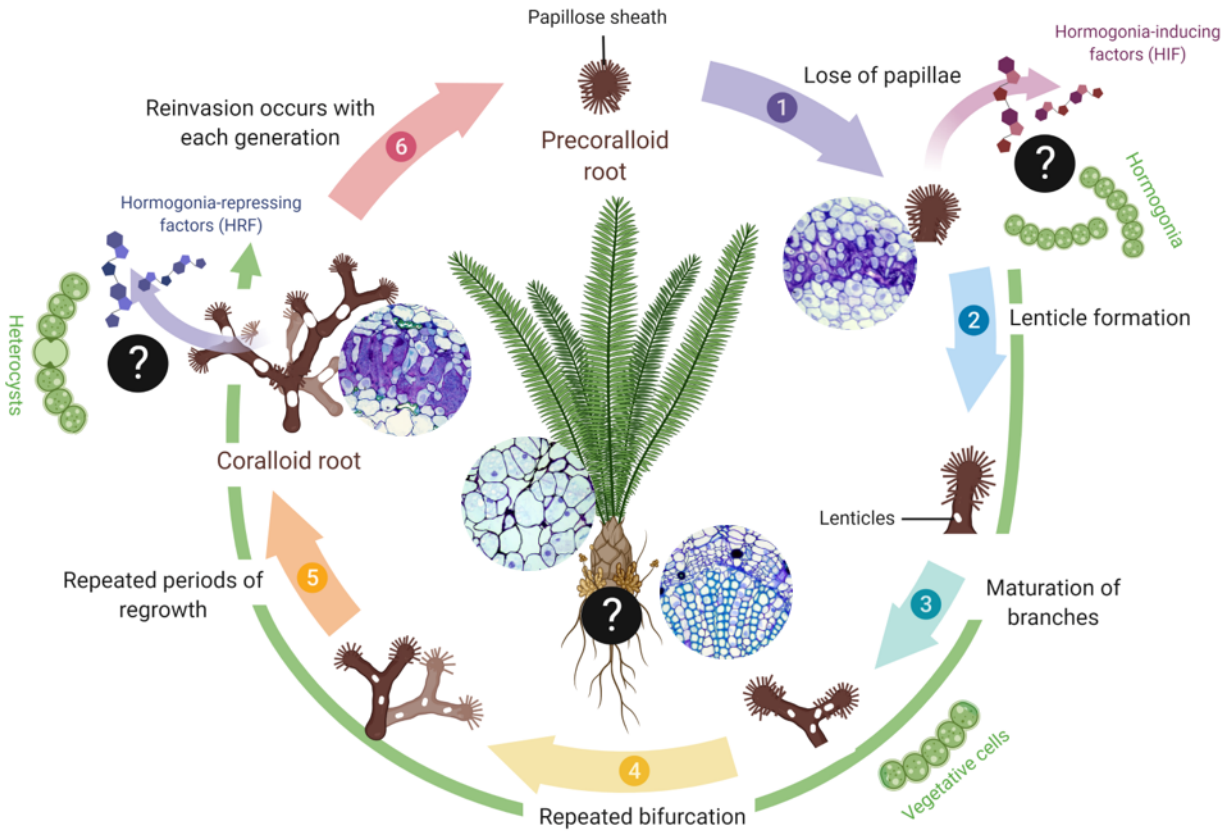


Figure 3.2. Precoralloid maturation and coralloid root development. (1) Precoralloid roots arise as apogeotrophic, papillose root which becomes apapillose as a thin dermal layer replaces the papillose sheath. (2) A transition stage as apogeotrophic precoralloid roots are invaded by hormogonia after which apical lenticles form. Hormogonia invade these precoralloid roots due to the release of an as yet unknown hormogonia-inducing factor (HIF). (3) Coralloid roots grow exhibiting a change to geotropic growth. (4) Coralloid roots mature exhibiting dichotomy. (5) Development of mature coralloid roots with well-developed, dichotomously branched, geotrophic root structure. Once mature, coralloid roots induce heterocyst formation by releasing an unknown hormogonia-repressing factor (HRF). Precoralloid and coralloid roots are thought to be highly specialised types of lateral roots emerging as "transformed normal, lateral root apices" but our morphological investigation seems to suggest otherwise. Adapted from Ahern et al. (1994) and created with Biorender.com.