**UNIVERSITY OF EAST ANGLIA**
**JOHN INNES CENTRE**

# The conservation of arbuscular mycorrhizal symbiosis between a liverwort and an angiosperm

AISLING LOUISE COOKE

Thesis submitted to the University of East Anglia
for the degree of Doctor of Philosophy

Research conducted at the John Innes Centre

2018

“Mīlle viae dūcunt hominēs per saecula ~~Rōmam~~ [calcium]”

*Liber Parabolarum*, 591 (1175), Alain de Lille.

# Abstract

Approximately 80% of all land plants form mutually beneficial interactions with soil fungi from the phylum Glomeromycota, a relationship called arbuscular mycorrhizal (AM) symbiosis. AM symbiosis enhances the uptake of mineral nutrients from the soil by these plants and positively affects resistance to disease and abiotic stress. AM symbiosis is predicted to have evolved in the first plants that colonized land, over 450 million years ago based on evidence from the fossil record and from the presence of AM symbiosis in the earliest diverging land plant clades. The molecular mechanisms of AM symbiosis have been elucidated in detail in the 21$^{st}$ century using studies in angiosperm model species, but it is not clear that these discoveries also apply to the species from other land plant clades where AM symbiosis is present.

In this study, I assess the conservation of known molecular mechanisms of AM symbiosis between the model angiosperm, *Medicago truncatula* and the liverwort, *Marchantia paleacea*. Putative AM symbiosis genes were identified by phylogenetic and sequence analysis and the function of these genes in AM symbiosis was supported by expression analysis during colonization of *M. paleacea* with AM fungi. Functional conservation of ion channels in symbiosis signalling was demonstrated through the generation of mutants in *M. paleacea* by CRISPR/Cas9 genome editing, and quantification of the AM symbiosis phenotype. The conservation of molecular function was demonstrated by complementation of orthologous mutants in *M. truncatula*.

This thesis provides insight into the evolution of the molecular mechanisms of AM symbiosis since the liverwort and angiosperm lineages diverged and provide further insights into the generation of nuclear $Ca^{2+}$ oscillations during symbiosis signalling.

# Acknowledgements

To Myriam. I could not have done it without you. From the bottom of my heart, thank you!

To Giles, who welcomed me into his lab and gave me the opportunity to work with Marchantia, and thanks also to Jeremy Murray and Allan Downie for their insights over the past four years.

To the Charpentier lab, current and former: Pierre Dangeville, Pablo del Cerro, Fran Robson, Rik Huisman and Nuno Leitao.

To the Symbiosis Super Group, who welcomed me in and have taught me so much; Aaron Thomas, Akira Akamatsu Andy Breakspear, Chengdu Liu, Donna Cousins, Doreen Feike, Eleni Soumpourou, Feng Feng, Fran Robson, Guru Radhakrishnan, Ioannis Tamvakis, Jian Feng, Jo Harrison, Jodi Lilley, Jongho Sun, Katharina Scheissl, Leoni Luginbuehl and Ram Karunakaran.

To my fellow members of Team Marchantia, Pierre-Marc Delaux and Guru Radhakrishnan.

To my PhD student comrades in arms: Leonie Luginbuehl, Guru Radhakrishnan an Nuno Leitao. It wouldn’t have been nearly as much fun without you.

To my supervisory committee, past and present: Diane Saunders, Sophien Kamoun and Christine Faulkner.

To the John Innes Foundation, for funding the PhD Rotation Programme, and t Steph Bornemann and Nick Brewin for keeping this programme running, not to mention my fellow 2014 cohort: Billy Aldridge, Cathy Mansfield, Javi Galdon-Armero and Sue Kuhaudomlarp.

To Mark Buttner, Alison Smith and Xiaoqi Feng, for my wonderful rotations.

To all the support services at JIC, particularly the Media Kitchen and Horticultural Services, who have helped make everything possible.

To my Hawaiian Red Shrimp, for keeping me company on my desk and persistently not-dying.

To all my Norwich friends; Alex, Becky, Billy, Cass, Cat (n) and Cat (2n), Cathy, Erin, Jo, Marie, Millie, Mikki, Nicola, Tom and Thom.

To my family, Mam, Dad, Izzy, Leo and Des and my boyfriend Alex. Thank you for supporting and humouring me through all the PhD madness.

# Table of Contents

# List of Figures

# List of Tables

# Chapter I – Introduction

Symbiosis, defined as a close and prolonged association between two or more organisms of different species, is a ubiquitous feature of life on earth. These associations, whether parasitic, commensal or mutualistic, are a driving force behind evolution, precipitating major evolutionary events such as the origin of eukaryotes, the Cretaceous radiations of insects and angiosperms, and the formation of coral reefs. Arbuscular mycorrhizal (AM) symbiosis, the association between land plants and fungi of the phylum Glomeromycota, is possibly the most widespread symbiotic association on earth. AM symbiosis is an ancient association, present in the common ancestor of the extant land plants and likely to be instrumental in the colonization of the terrestrial environment by the first land plants. While the molecular mechanisms of AM symbiosis have been dissected in angiosperm model plant systems, angiosperms a highly derived land plant lineage in terms of morphology and genome structure and are not necessarily representative of the entire land plant clade. These molecular mechanisms have not been studied in detail in the non-flowering plants. This thesis is dedicated to studying the molecular mechanisms of AM symbiosis in one of the first plant lineages to diverge, the liverworts (see Figure 1.1), and further develops the liverwort species *Marchantia paleacea* as a model for studying AM symbiosis in a basal land plant.

## 1.1 Arbuscular mycorrhizal symbiosis

In AM symbiosis, Glomeromycete fungi invade and colonize land plants, forming intracellular structures that function as an interface for nutrient exchange (Schüβler et al., 2001). These often form highly branched, tree-like structures called arbuscules, but other structures such as intracellular hyphal coils may also be formed, with or without the presence of arbuscules (Dickson, 2004). About 80% of AM fungi also form storage vesicles within the root. At the arbuscule, the fungus provides mineral nutrients scavenged from the soil in exchange for fixed carbon from the plant. Glomeromycetes are obligate biotrophs and with few exceptions, depend entirely on plants for fixed carbon (Gianinazzi-Pearson, 1996).

Mycorrhizal symbiosis is found in 80% of land plant species and 92% of land plant families, with AM symbiosis being the most common (Smith and Read, 2008; Trappe, 1987; Wang and Qiu, 2006). The hyphal network of the fungus acts as an extension of the plant rooting system, allowing it to functionally explore a larger soil volume. This is particularly important to the uptake of phosphate, which tends to be patchily distributed and poorly available in soil (Javot et al., 2007b; Schachtman et al., 1998). AM symbiosis also contributes to the acquisition of other essential nutrients including potassium, nitrogen and sulphur (Garcia et al., 2016).

AM symbiosis also affects the plant response to other abiotic stresses. AM symbiosis can sometimes improve drought resistance and tolerance and mitigate the negative effects of salinity on at least some crop plants, potentially by reducing mineral deficiencies in salt stressed plants. (Augé, 2001; Fileccia et al., 2017; Hameed et al., 2014). In heavy metal contaminated soils, AM symbiosis can have a protective effect on the host plant by immobilising and effectively filtering out heavy metals from the rhizosphere (Hildebrandt et al., 2007).

Symbiosis with mycorrhizal fungi affects interactions between plants and other organisms. AM symbiosis is widely reported to reduce the damage caused by parasitic nematodes and soil borne diseases, such as *Fusarium* and *Phytophthora* but may increase susceptibility to some shoot borne diseases and sap-feeding insects (de la Pena et al., 2006; Gange, 2007; Li et al., 2006; Shaul et al., 1999; Whipps, 2004). These conflicting effects may be because AM symbiosis can modulate plant defence responses through multiple mechanisms. These include changing nutrient status, altered defence related gene transcription, accumulation of insect antifeedant compounds, altered release of volatiles in the shoot and low level accumulation of defensive compounds in mycorrhized roots (Pozo and Azcón-Aguilar, 2007)

Although mycorrhiza literally means fungus-root, fungal symbionts occur in non-vascular plants without rooting systems, where they colonize the thallus and rhizoids (see Figure 1.2) (Read et al., 2000). I will refer to all symbioses between land plants and Glomeromycete fungi as arbuscular mycorrhizal symbiosis, as they share the same structural and functional characteristics, regardless of the plant organ colonized (Humphreys et al., 2010).

## 1.2 The common symbiosis signalling pathway

Arbuscular mycorrhizal symbiosis requires reprogramming of plant cell structure, metabolism and defence responses to allow the fungal symbiont to enter and colonize the root, or rooting system. The transcriptional changes required for these processes are induced by the common symbiosis signalling pathway, a mechanism common to AM symbiosis and to the root nodule symbiosis formed between nitrogen fixing bacteria and several members of the Eurosid I clade, including Fabales, Fagales, Cucurbitales and Rosales (Gualtieri and Bisseling, 2000; Oldroyd, 2013). Root nodule

symbiosis involves the development of specialized organs, called nodules, that accommodate nitrogen fixing bacteria in a very low oxygen environment required for nitrogenase activity (Shah and Brill, 1977). Root nodule symbiosis is believed to have evolved by co-opting the more ancient signalling machinery of AM symbiosis, with nitrogen fixing bacteria using chitin based compounds to signal to the host plant (Parniske, 2008).

The common symbiosis signalling pathway (CSSP) is a plant response to these beneficial microbial partners that activates transcriptional reprogramming through characteristic calcium ($Ca^{2+}$) oscillations in the nucleus. These characteristic oscillations of $Ca^{2+}$ concentration associated with the plant nucleus are necessary to initiate both AM and root nodule endosymbioses (Cárdenas et al., 1999; Shaw and Long, 2003). As many of the molecular mechanisms of the CSSP were first elucidated in root nodule symbiosis and then found to also function in AM symbiosis, this signalling pathway will be described in the context of both symbioses.

### 1.2.1 Exchange of signal molecules

Both root nodule symbiosis and AM symbiosis are initiated by the exchange of signal molecules in the rhizosphere, between the endosymbiont and the host plant. In AM symbiosis, the plant hormone strigolactone promotes hyphal branching, mitochondrial activation, spore germination and colonization of the root by AM fungi (Akiyama et al., 2005; Besserer et al., 2008; Besserer et al., 2006; Gomez-Roldan et al., 2008). In legume-rhizobia symbiosis, the best described form of root nodule symbiosis, flavonoids released by the plant activate symbiotic gene expression in rhizobial bacteria (Carl and Phillips, 1990).

This symbiotic gene expression in rhizobia leads to production and release of lipochitooligosaccharides (LCOs) so called Nod factors, which induce nuclear associated $Ca^{2+}$ oscillations in the host plant (Ehrhardt et al., 1996). Nod factors are made up of four or five N-acetylglucosamine (GlcNAc) residues and an N-acyl group attached to the non-reducing terminal sugar (Denarie et al., 1996). Nod factors vary between rhizobial species, in the length and degree of saturation in the N-acyl group and in the methyl, fucosyl, acetyl and sulphate groups that can decorate the GlcNAc residues. Legumes also vary in which Nod factors they will respond to, allowing rhizobial bacteria and legumes to form species specific symbioses (Denarie et al., 1996; Lerouge et al., 1990).

AM fungi also secrete signal molecules that activate symbiotic gene expression (Kosuta et al., 2003). Those include a variety of chitin based Myc factors: LCOs that are chemically similar to Nod factors, and short chain chitin oligomers (COs) such as CO4, CO8 (Maillet et al., 2011; Sun et al., 2015). Both AM LCOs and COs elicit nuclear associated $Ca^{2+}$ oscillations (Genre et al., 2013; Maillet et al., 2011; Sun et al., 2015; Walker et al., 2000). It is unclear how the host distinguishes between Myc factor and chitin molecules associated with pathogenic fungi but exudates from plant pathogens have not been found to induce nuclear associated $Ca^{2+}$ oscillations in legume root cells.

Plants detect microbial glycan molecules, including chitin based compounds, with complexes of receptor kinases and receptor like proteins with extracellular domains containing lysin motifs (LysM receptor kinases and LysM RLKs) (Gust et al., 2012). Nod factors are perceived by a complex made up of LysM RLK 3 (LYK3) and Nod Factor Perception (NFP), a non- functional kinase, in *Medicago truncatula*, and by their orthologues Nod Factor Receptor (NFR)1 and NFR5, respectively in *Lotus*

*japonicus* (Arrighi et al., 2006; Broghammer et al., 2012; Limpens et al., 2003; Madsen et al., 2003).

However, in AM symbiosis, the "Myc factor receptor" is not clearly defined, probably due to the diversity of Myc factor and to the functional redundancy in the LysM RLK families. In tomato, virus induced gene silencing of the NFP orthologue, *SlLYK10*, greatly reduces AM colonization (Buendia et al., 2016). In the nodulating non-legume *Parasponia andersonii*, RNAi targeting of the *NFP* homologue *PaNFP1* blocked infection by AM fungi but it is unclear whether a second receptor, *PaNFP2* was also affected in this experiment (Op den Camp et al., 2011; van Velzen et al., 2018). In *M. truncatula, NFP* is involved in the root branching and gene expression responses to Myc LCOs (Czaja et al., 2012; Genre et al., 2013; Maillet et al., 2011; Sun et al., 2015). However, *MtNFP* is not necessary for AM symbiosis or $Ca^{2+}$ oscillations induced by Myc COs and *MtLYK3* is not necessary for AM colonisation or the root branching response to LCOs (Amor et al., 2003; Maillet et al., 2011). Although legumes contain a closely related paralogue of *MtNFP*/*LjNFR5*, called *MtLYK1*/*LjLYS1*, this finding that *NFP* is not necessary for normal AM colonization is unlikely to be due to functional redundancy between the two paralogues. In the model legume *L. japonicus*, a triple mutant of *nfr5lys11nfr1* had normal colonisation by AM fungi (Rasmussen et al., 2016).

In legumes, a second LysM-RLK functions in recognizing chitin oligomers in both symbiotic and pathogenic contexts. RNAi of *LYK9* in pea (*Pisum sativum*) reduces transcription of symbiosis genes when CO5 is applied and reduces the defence response to pathogenic fungi (Leppyanen et al., 2018). The mutants *Mtlyk9* and *Ljlys6*, homologues of *PsLYK9*, are also more susceptible to pathogenic fungi and while AM colonization is normal, the symbiotic CO response has not been assessed in

these mutants (Bozsoki et al., 2017). In rice, a dual function receptor, *OsCERK*, also detects both pathogens and AM fungi (Miyata et al., 2014). Unlike legume species, this receptor is essential to the formation of AM symbiosis, most likely because rice does not have a symbiotic $Ca^{2+}$ response to Myc-LCOs (Sun et al., 2015).

In addition to LysM containing receptors, the receptor complex that recognises symbiotic signals includes a leucine rich repeat receptor kinase called *MtDMI2*/*LjSYMRK* that is necessary for both AM and root nodule symbiosis signalling (Endre et al., 2002; Miwa et al., 2006; Stracke et al., 2002). In *L. japonicus*, overexpression of *SYMRK* is sufficient to induce AM symbiosis related gene expression and nodule formation (Ried et al., 2014). Although SYMRK associates with NFR1 and NFR5 at the plasma membrane, overexpression of *SYMRK* can induce nodulation in the absence of these LysM RLKs, while *NFR1* and *NFR5* overexpression does not lead to nodulation in a *symrk* mutant background (Ried et al., 2014). This indicates that SYMRK is necessary and sufficient to initiate symbiotic signalling downstream of NFR1 and NFR5.

### 1.2.3 Activating the calcium machinery

There are several potential mechanisms of signal transduction that may connect signal molecule detection at the plasma membrane to nuclear associated $Ca^{2+}$ oscillations, based on both genetic and pharmacological data. SYMRK interacts with 3-Hydroxy-3-Methylglutaryl CoA Reductase 1 (HMGR1), an enzyme in the mevalonate synthesis pathway (Kevei et al., 2007b). Pharmacological inhibition or RNAi knockdown of *HGMR1* impairs both root nodule and AM symbioses, and RNAi knockdown impairs both the nuclear $Ca^{2+}$ oscillation and transcriptional responses to Nod factor applied to *M. truncatula* roots (Kevei et al., 2007a; Venkateshwaran et al., 2015). This suggests that component of the mevalonate pathway might be involved in the

generation of a secondary messenger to activate the nuclear $Ca^{2+}$ oscillation. Interestingly, exogenous application of mevalonate to *M. truncatula* roots induces $Ca^{2+}$ oscillations and symbiotic marker gene expression but this effect is dependent on NFP and partially dependent on DMI2 (SYMRK) (Venkateshwaran et al., 2015). This indicates that the receptor like kinases are required for the activation of the $Ca^{2+}$ oscillations in addition to the mevalonate pathway.

$Ca^{2+}$ oscillations induced by the G-protein agonist, Mastoparan, or its synthetic analogue, Mas7, activate symbiosis marker genes, dependent on the symbiotic $Ca^{2+}$ oscillation decoder, the $Ca^{2+}$ and calmodulin protein kinase CCaMK (Charron et al., 2004; Pingret et al., 1998; Sun et al., 2007). These $Ca^{2+}$ oscillations can be triggered in *dmi1* knockout mutants, suggesting that G-protein signalling may trigger nuclear $Ca^{2+}$ oscillations by activating directly the $Ca^{2+}$ channel. In this scenario, the counter ion balance movement would be sustained by another cation channel activated by Mastoparan signalling (Peiter et al., 2007; Sun et al., 2007). In soybean, the regulator of G-protein signalling (RGS) proteins are phosphorylated by NFR1 and heterotrimeric G proteins positively regulate nodulation, while active Gα proteins negatively regulate nodulation (Choudhury and Pandey, 2013, 2015; Choudhury and Pandey, 2016). Altogether, this suggest that G-protein signalling might be required in the generation of a secondary messenger.

Further potential mechanisms of signal transduction include the mitogen activated protein kinase (MAPK) cascade, as RNAi knockdowns of *SIP2*, a MAPK kinase that interacts with SYMRK in *L. japonicus*, reduces nodulation and the expression of root nodule symbiosis genes (Chen et al., 2012). In rice, a component of the intracellular receptor complex that detects karrikin molecules found in smoke, *Dwarf 14 Like* (*D14L*) is necessary for AM symbiosis development and the gene expression response

to mycorrhizal exudates (Gutjahr et al., 2015). However, the ligand of D14L is unknown and it is not clear how it might interact with other members of the CSSP. Finally, identification of *CNGC15* as a symbiosis associated $Ca^{2+}$ channel suggests that cyclic nucleotide monophosphate (cNMP) could regulate nuclear $Ca^{2+}$ oscillations (Charpentier et al., 2016). The production of cNMPs could be activated by one or more of the mevalonate, G protein, the karrikin receptor complex or the MAPK kinase cascade.

### 1.2.4 Decoding calcium oscillations

Symbiotic $Ca^{2+}$ oscillations in the nucleus are decoded by the nuclear-localized $Ca^{2+}$ and calmodulin-dependent protein kinase CCAMK/DMI3 (Levy et al., 2004; Mitra et al., 2004). Mutants of *ccamk* failed to form nodules or AM symbioses but nevertheless responded to Nod factor with nuclear $Ca^{2+}$ oscillations (Ehrhardt et al., 1996; Wais et al., 2000). CCaMK is unique among $Ca^{2+}$ binding proteins as it binds $Ca^{2+}$ by two different mechanisms; directly at three EF hand domains and indirectly through calmodulin (Miller et al., 2013; Patil et al., 1995; Takeda et al., 2012). Direct $Ca^{2+}$ binding occurs at lower $Ca^{2+}$ concentrations, and $Ca^{2+}$ binding at even one EF hand promotes the auto-kinase activity of CCaMK. This phosphorylates a specific threonine residue (Thr-271 in *M. truncatula*), which alters the hydrogen bonding of the protein so its conformation does not allow substrate phosphorylation (Miller et al., 2013). Phosphorylation at this residue is essential to maintaining inactive CCaMK, as a single amino acid substitution results in constitutively active CCaMK and spontaneous nodule formation in the absence of rhizobia (Tirichine et al., 2006).

During nuclear $Ca^{2+}$ oscillations, $Ca^{2+}$ binds calmodulin in the nucleus, which then binds to CCaMK. Calmodulin binds preferentially to autophosphorylated CCaMK, suggesting that direct $Ca^{2+}$ binding also primes the protein for activation during

symbiotic signalling (Sathyanarayanan et al., 2001). Bound calmodulin overrides the effect on the hydrogen bond network of threonine phosphorylation, leading to activation of CCaMK and substrate phosphorylation. This mechanism allows CCaMK to be activated temporarily and specifically by the $Ca^{2+}$ oscillations characteristic of symbiosis signalling (Miller et al., 2013).

CCaMK phosphorylates and forms a complex with CYCLOPS/IPD3, a nuclear localized DNA binding protein, necessary for AM and rhizobial infection and symbiotic gene expression (Messinese et al., 2007; Singh et al., 2014; Yano et al., 2008). CYCLOPS must be phosphorylated at specific serine residues to function in symbiotic transcriptional reprogramming, with phosphomimic mutations at these residues leading to spontaneous symbiotic gene expression, and nodule development independently of CCaMK (Singh et al., 2014). However, *cyclops* mutants in an autoactive CCaMK background have reduced but not abolished nodulation, suggesting that CCaMK may also activate transcription necessary for nodule organogenesis, but not endosymbiont infection, independently of CYCLOPS (Yano et al., 2008).

### 1.2.2 Nuclear calcium machinery

The nuclear $Ca^{2+}$ oscillations induced by Nod and Myc factors are generated by a complex of ion channels located at the nuclear envelope membrane that release $Ca^{2+}$ ions from the nuclear envelope lumen (Capoen, 2011). In *M. truncatula*, these are composed of a potassium permeable channel, Does not Make Infections 1 (DMI1, also known as POLLUX in other species) (Ane et al., 2004; Charpentier et al., 2008; Venkateshwaran et al., 2012) and the Class III of cyclic nucleotide gated channel, CNGC15 (Charpentier et al., 2016). Other angiosperms outside the Vicoid legumes also require a paralogue of *DMI1*/*POLLUX*, called *CASTOR*, for Nod factor-induced

nuclear $Ca^{2+}$ oscillations (Ane et al., 2004; Charpentier et al., 2008; Imaizumi-Anraku et al., 2005; Venkateshwaran et al., 2012).

DMI1 is essential for activating the $Ca^{2+}$ response to symbiosis signal molecules, as Myc factor and Nod factor induced nuclear $Ca^{2+}$ oscillations are abolished in knockout *dmi1* mutants (Ane et al., 2004; Charpentier et al., 2008; Imaizumi-Anraku et al., 2005; Sun et al., 2015). DMI1 is also essential for infection by AM fungi and rhizobial bacteria, and for the development of nodules (Ane et al., 2004; Catoira et al., 2000). However, DMI1 is not necessary for Mastoparan-induced nuclear $Ca^{2+}$ oscillations that structurally mimic Nod and Myc factors-induced nuclear $Ca^{2+}$ oscillations (Peiter et al., 2007; Sun et al., 2007). This suggests that DMI1 modulates nuclear $Ca^{2+}$ oscillations associated with symbioses signalling.

This is consistent with the data concerning the only symbiotic $Ca^{2+}$ channel that has been identified. *CNGC15* was identified through a bioinformatic screen of the *M. truncatula* genome and confirmed with reverse genetics using the *M. truncatula* transposon tagged mutant collection (Charpentier et al., 2016; Pislariu et al., 2012; Young et al., 2011). *M. truncatula* contains three *CNGC15* homologues, *CNGC15a*, *CNGC15b* and *CNGC15c*. Mutations in any of these genes leads to symbiotic defects including reduced nodule number, fewer rhizobial infections and lower AM colonization (Charpentier et al., 2016). These mutants also had fewer root cells respond to Nod factor with $Ca^{2+}$ oscillations, while many of the cells that did respond did not sustain these $Ca^{2+}$ oscillations and had altered oscillation frequency (Charpentier et al., 2016). CNGC15 proteins localize to the nuclear membrane, where they physically interact with DMI1 (Charpentier et al., 2016).This analysis indicates that all three homologues are required for efficient activation of nuclear $Ca^{2+}$ oscillations by Nod factor and Myc factor, possibly forming a tetrameric complex.

A third component of the nuclear $Ca^{2+}$ generation machinery, MCA8, a sarco/endoplasmic reticulum (SERCA) $Ca^{2+}$-ATPase, localised to both the endoplasmic reticulum and unclear envelope, has also been identified via reverse genetic screening (Capoen, 2011). Pharmacological inhibition of SERCA-type $Ca^{2+}$-ATPases inhibit Nod factor-induced nuclear $Ca^{2+}$ oscillations and symbiosis marker gene expression, and RNAi knockdown of *MCA8* blocks Nod factor-induced nuclear $Ca^{2+}$ oscillations and leads to reduced AM colonization and nodulation (Capoen, 2011; Capoen et al., 2009; Charron et al., 2004; Engstrom et al., 2002).

Mathematical modelling demonstrated that $Ca^{2+}$ oscillations can be generated through the interplay of the three components described; DMI1, CNGC15 and MCA8 (Capoen, 2011; Charpentier and Oldroyd, 2013). This model requires that DMI1 and CNGC15 are activated simultaneously, with CNGC15 releasing $Ca^{2+}$ from the nuclear envelope lumen, counterbalanced by potassium flow through DMI1. This is consistent with yeast two hybrid assay and in *planta* bimolecular fluorescence data that indicates that DMI1 physically interacts with all three CNGC15 homologues in *M. truncatula* (Charpentier et al., 2016). Hypothetically, a secondary messenger could activate DMI1 or CNGC15, leading to channels conformational change and simultaneous opening of both DMI1 and CNGC15. Mathematical modelling also predicts that regulation of DMI1 by $Ca^{2+}$ is necessary to maintain $Ca^{2+}$ oscillations after activation (Charpentier and Oldroyd, 2013). In this model, MCA8 would be required to replenish the $Ca^{2+}$ store formed by the lumen of the nuclear envelope contiguous with the lumen of the endoplasmic reticulum.

In addition to the ion channels and transporters that directly generate nuclear $Ca^{2+}$ oscillations, several components of the nuclear pore complex, specifically the Nup84 subcomplex, are necessary for symbiotic $Ca^{2+}$ oscillations. These include Nucleoporin

133 (NUP133), NUP85 and NENA (related to nuclear pore components Sec13 in humans and Seh1 in yeast) (Groth et al., 2010; Kanamori et al., 2006; Saito et al., 2007). In *nup133* and *nup85* mutants, nodule formation, nuclear $Ca^{2+}$ oscillations and AM colonisation are impaired (Kanamori et al., 2006; Saito et al., 2007). In *nena* mutants, AM colonization and nuclear $Ca^{2+}$ oscillations are impaired but while rhizobial infection does not occur through root hairs, nodules may develop and rhizobia can enter the root by crack entry (Groth et al., 2010). Nucleoporins may function in nuclear $Ca^{2+}$ oscillation generation through trafficking of secondary messengers and modulators of the $Ca^{2+}$ machinery, or in localizing the nuclear $Ca^{2+}$ machinery components DMI1, CNGC15 and MCA8 in the nuclear envelope membrane. Alternatively, mathematical modelling suggests that nuclear pores play a role in coupling $Ca^{2+}$ oscillations in the nucleoplasm and the cytoplasm, which may be important in sustaining oscillations over longer periods (Martins et al., 2016).

## 1.3 Arbuscule development

After initial symbiosis signalling, AM symbiosis develops in two distinct stages. In angiosperm model species, the fungal hypha forms an appressorium like structure on the plant epidermis, then crosses the epidermis though an intracellular accommodation structure called the pre-penetration apparatus. The pre-penetration apparatus is formed by reorganization of the cytoskeleton and ER, with the cell nucleus first migrating towards the fungal appressorium, then migrating towards the next cell layer ahead of a cytoplasmic column (Genre et al., 2005). Secretion within this column generates a hollow tube continuous with the apoplast, through which the fungal hypha can cross the epidermis. In *Arum*-type fungal morphology, the fungus grows through the apoplast of the root cortex, with side branches entering cortical cells and branching dichotomously to form arbuscules, while in *Paris*-type morphology the fungal hyphae

form intracellular coils, with occasional branches, and grow intracellularly across the cortex. Intra-radicle hyphal growth can also form intermediate morphology (Dickson, 2004). This depends on both the plant colonised and the fungal inoculum; with the Arum-type found in the legume model species *M. truncatula* and *L. japonicus* and Paris type colonization seen in in many woodland species, while both types of morphology are formed by different fungi in tomato (Smith and Smith, 1997; Cavagnaro et al., 2001). Although Paris-type morphology may be more common than Arum-type, the molecular mechanisms of intracellular nutrient exchange structures has been described mostly in the context of Arum-type arbuscules.

Arbuscules are composed of finely branched fungal hyphae surrounded by the plant derived peri-arbuscular membrane (PAM), providing a huge surface area for nutrient exchange. The generation of the PAM is essential to allow the arbuscule to form, as plant mutants affecting membrane generation, symbiotic transcriptional reprogramming and nutrient exchange result in absent or stunted arbuscules. The PAM contains a specific suite of proteins found only on this membrane, targeted there by precise temporal regulation of transcription that couples expression of arbuscule specific genes to the transient re-orientation of polar secretion that occurs during arbuscule development (Pumplin et al., 2012).

The cytoplasmic protein VAPYRIN is necessary for arbuscule development, as well as initial infection by AM fungi, and is also required for infection in root nodule symbiosis (Murray et al., 2011b; Pumplin et al., 2010a). In RNAi *VAPYRIN* knockdowns and in *vapyrin* mutants, most fungal hyphopodia formed at the epidermis fail to penetrate the root and those that do are incapable of developing arbuscules in cortical cells (Pumplin et al., 2010b). VAPYRIN is made up of two protein-protein interacting domains, an N-terminal Major Sperm Protein (MSP) domain and a C-

terminal domain containing several ankyrin repeats and is thought to function by bridging proteins that remodel the structure of the plant cell to accommodate AM fungi.

One of these proteins is Exo70I, a component of the exocyst complex involved in vesicle trafficking that facilitates polarised exocytosis (Eckardt, 2008). In *M. truncatula*, *Exo70I* is specifically expressed in arbuscule containing cells where it co-localizes with VAPYRIN in punctate structures at the tips of the growing PAM (Zhang et al., 2015). In *exo70i* mutants, AM fungi can invade cortical cells but do not form mature arbuscules with fine branching or incorporate characteristic PAM proteins (Zhang et al., 2015). *VAPYRIN* is likely to have other interacting partners involved in fungal entry to cells and rhizobial infection of root hairs, possibly including additional members of the Exo70 family also upregulated in AM colonized roots (Harrison and Ivanov, 2017).

In addition to the exocyst, there are also symbiosis specific components of the SNARE (soluble N-ethylmaleimide-factor attachment protein receptors) complex that mediates vesicle fusion during exocytosis. The v-SNAREs, vesicle-associated membrane proteins VAMP721d and VAMP721e are necessary for arbuscule development and for symbiosome formation in root nodule symbiosis (Ivanov et al., 2012). The t-SNARE, *Syntaxin of Plants 132* (*SYP132*) is also involved in forming the PAM, with RNAi of *SYP132α* spliceform leading to lower mycorrhizal colonization and impaired differentiation of rhizobia in nodules (Huisman et al., 2016). However, the requirement of these SNARE components in symbiosis seems to due to their specific expression during symbiosis, rather than specificity in vesicle targeting, as RNAi knockdowns of *VAMP721d/VAMP721e* and *SYP132α* can be rescued by

expression of an alternative syntaxin or VAMP gene under the *VAMP721e* or *SYP132* promoter, respectively (Huisman, 2018).

## 1.4 Nutrient exchange

Functional nutrient exchange is necessary for the development and maintenance of arbuscules in AM symbiosis. AM symbiosis is suppressed by high phosphate conditions, indicating that the symbiosis is only allowed to proceed when providing a limiting nutrient to the host plant (Breuillin et al., 2010; Gu et al., 2011). The symbiosis specific PHT1 family phosphate transporter, *PT4* (in *M. truncatula* and *L. japonicus*, *PT11* in rice) is necessary for phosphate uptake by the plant during AM symbiosis (Harrison et al., 2002; Javot et al., 2007a; Paszkowski et al., 2002; Volpe et al., 2016). In *M. truncatula*, *pt4* mutants and RNAi knockdowns accumulate polyphosphate (polyP) in the arbuscules, arbuscules senesce prematurely and fungal infection is compromised (Javot et al., 2007a). *PT4* is likely to function in phosphate sensing as well as in nutrient uptake, as *pt4* mutant roots in *M. truncatula* and *L. japonicus* have impaired Pi responses in the absence of endosymbionts, suggesting that *PT4* also mediates early root responses to phosphate status in non-mycorrhizal roots (Volpe et al., 2016). Phosphate transport by *PT4*, a Pi/H+ symporter, is driven by the acidic environment of the peri-arbuscular space. The pH gradient between the peri-arbuscular space and the cytoplasm is generated by the proton pump ATPase, *HA1*, with *ha1* mutants having a similar phenotype to that seen in *pt4* mutants (Krajinski et al., 2014; Wang et al., 2014).

While plants can regulate arbuscule development and prevent it from proceeding in the absence of beneficial mineral uptake from AM fungi, this mineral nutrient does not have to be phosphate. In nitrogen limiting conditions, AM symbiosis develops

normally in *pt4* mutants without premature arbuscule degeneration (Breuillin-Sessoms et al., 2015). This effect is dependent on *AMT2;3,* an ammonium transporter specifically upregulated during AM symbiosis. As *AMT2;3* expressed in yeast does not rescue ammonium transport mutants, this gene may function in nutrient sensing, while ammonium uptake from the peri-arbuscular space is mediated by the closely related ammonium transporters, *AMT2;4* and *AMT2;5*, which are also upregulated in AM colonized roots (Breuillin-Sessoms et al., 2015).

Nutrient transport from the plant to the fungus is also necessary for the development and maintenance of AM symbiosis. As obligate biotrophs, AM fungi depend on host plants for a supply of fixed carbon. This fixed carbon is likely to be transported in the form of lipids. Genomic analysis of the AM fungi *Rhizophagus irregularis* and *Gigaspora rosea* indicates that these species do not contain the genetic machinery required for de novo synthesis of fatty acids and, although AM fungi store carbon predominantly as triacylglycerols, little to no lipid synthesis takes place in the extraradical mycelium (Bécard et al., 1991; Bonfante et al., 1994; Pfeffer et al., 1999; Tang et al., 2016; Tisserant et al., 2013; Wewer et al., 2014).

During AM colonization, multiple fatty acid metabolism genes are upregulated by the host plant. These include the glycerol-3-phosphate acyltransferase (GPAT), *Reduced Arbuscular Mycorrhization 2* (*RAM2*), the mycorrhization specific palmitoyl-ACP thioesterase *FatM,* and multiple *Apetala* 2-domain transcription factors homologous to the *WRINKLED* (*WRI*) transcription factors that regulate fatty acid production in *A. thaliana* (Luginbuehl et al., 2017). Both *FatM* and *RAM2* are necessary for normal arbuscule development, with *ram2* and *fatm* mutants unable to develop mature finely branched arbuscules (Bravo et al., 2017; Bravo et al., 2016; Gobbato et al., 2012; Gobbato et al., 2013; Luginbuehl et al., 2017). In the current model for mycorrhizal

fatty acid biosynthesis, FatM preferentially releases 16:0 fatty acids from acyl carrier proteins in the chloroplast, which are later converted to 16:0 β-monoacylglycerol (β-MAG) by RAM2. The 16:0 β-MAG or a derivative is exported at the peri-arbuscular membrane and taken up by AM fungi, which can then convert it to other fatty acids found in AM fungi (Bravo et al., 2017; Luginbuehl et al., 2017). Although some changes in lipid metabolism during AM symbiosis may be involved in generating the surface area of the PAM, this specific biosynthetic pathway is more likely to function in fungal nutrition. The arbuscule phenotype of *ram2* mutants can be complemented by co-culture with a WT plant, indicating that *ram2* mutant cells are still capable of generating the PAM (Luginbuehl et al., 2017).

Although the transporter that exports fatty acids from the plant cell has not been identified, current evidence points towards two half ATP-binding cassette (ABC) transporters from the G sub family, *Stunted Arbuscule* (*STR*) and *STR2*. *STR* and *STR2* are expressed during AM colonization and form a complex on the PAM with both essential to normal arbuscule development (Zhang et al., 2010). Although the substrate of the complex has not been identified, *str* accumulates 16:0 βMAG to almost wild type levels despite low fungal colonisation (Bravo et al., 2017). Other ABCG transporters function in lipid transport or in depositing lipid-based biopolymers, making a lipid based substrate plausible for the STR/STR2 complex (Fedi et al., 2017; Hwang et al., 2016; Yadav et al., 2014; Zhang et al., 2010).

## 1.5 Transcription during AM symbiosis

Root colonization and arbuscule development in AM symbiosis are controlled by transcription factors downstream of symbiotic $Ca^{2+}$ spiking (Floss et al., 2013). These

transcription factors include CYCLOPS and GRAS domain transcription factors (Singh et al., 2014).

GRAS-domain transcription factors are a plant specific family of proteins with diverse roles in plant development and signalling named after the first members of this family discovered: *GIBBERELLIC-ACID INSENSITIVE* (*GAI*), *REPRESSOR of GAI* (*RGA*) and *SCARECROW* (*SCR*) (Pysh et al., 1999). Several members are required for AM symbiosis including REQUIRED FOR ARBUSCULAR MYCORRHIZAE 1 (RAM1), REQUIRED FOR ARBUSCULE DEVELOPMENT1 (RAD1), DELLA proteins and NODULATION SIGNALLING PATHWAY 1 (NSP1) and NSP2 (Delaux et al., 2013a; Floss et al., 2013; Gobbato et al., 2012; Kaló et al., 2005; Maillet et al., 2011; Park et al., 2015; Smit et al., 2005; Xue et al., 2015).

RAM1 is necessary for arbuscule development and appears to function by driving the arbuscule specific transcriptional program, including genes involved in nutrient transport (e.g. *PT4*), in arbuscule development (e.g. *Exo70I*) and in the arbuscule fatty acid biosynthesis pathway (Gobbato et al., 2012; Luginbuehl et al., 2017). It also drives expression of another GRAS-domain transcription factor, *RAD1*. *rad1* mutant has a similar but less severe phenotype to *ram1* (Pimprikar et al., 2016; Rich et al., 2015; Xue et al., 2015).

NSP1 and NSP2, are essential to root nodule symbiosis but have relatively moderate mutant phenotypes in AM symbiosis. Fungal colonisation is reduced in *nsp1* and *nsp2*, and *nsp2* is also impaired in the root branching response to Myc LCOs (Delaux et al., 2013a; Kalo et al., 2005; Maillet et al., 2011; Smit et al., 2005). It is likely that NSP1 and NSP2 function earlier in AM symbiosis signalling, as both transcription factors

regulate the synthesis of strigolactone, the phytohormone that activates symbiotic responses in AM fungi (Liu et al., 2011).

The DELLA motif containing proteins are also positive regulators of AM symbiosis. As DELLA proteins function in perceiving the phytohormone gibberellic acid (GA), their function integrates symbiosis regulation with plant responses to light, nutrient status, abiotic stress and interacting phytohormone pathways (Achard et al., 2007; Brian, 1959; Floss et al., 2013; Navarro et al., 2008). The DELLA proteins bind to DNA and repress the plant response to GA, while in the presence of GA, DELLA proteins are degraded, allowing GA induced gene expression. Wild type *M. truncatula* treated with GA, the double mutant (*Mtdella1/Mtdella2*), and the triple mutant (*Mtdella1/Mtdella2Mtdella3*), are still colonized by AM fungi but form much fewer arbuscules than untreated WT roots (Floss et al., 2017; Floss et al., 2016; Floss et al., 2013; Foo et al., 2013; Yu et al., 2014). GA treated roots and *della* mutants are also impaired in root nodule development and rhizobial infection (Jin et al., 2016; Maekawa et al., 2009; Tatsukami and Ueda, 2016). Surprisingly, DELLA proteins and NSP1 also seem to function late in AM symbiosis development, regulating the senescence of arbuscules, possibly by mediating the MYB family transcription factor, *MYB1*, that promotes arbuscule degeneration (Floss et al., 2017).

## 1.6 Evolution of AM symbiosis

### 1.6.1 origins of AM symbiosis

Molecular phylogeny and the fossil record indicate that AM symbiosis evolved at the same time as the first land plants, or shortly afterwards. Land plant spores appear in the fossil record just over 470 million years ago, in the early middle Ordovician period, while Glomeromycete spore appear 455-460 million years ago in the late middle

Ordovician (Redecker et al., 2000; Rubinstein et al., 2010) . Unambiguous evidence of AM symbiosis first appears in the fossil record in the early Devonian, just over 400 million years ago. In fossils of the early plant *Aglaophyton*, fungi colonise the growth axes and form characteristic dichotomously branched arbuscules in the parenchymous tissue. These fossils preserve exceptional intracellular details and the arbuscules seen in *Aglaophyton* are comparable to those characteristic of contemporary symbiosis with Glomeromycete fungi (Remy et al., 1994).

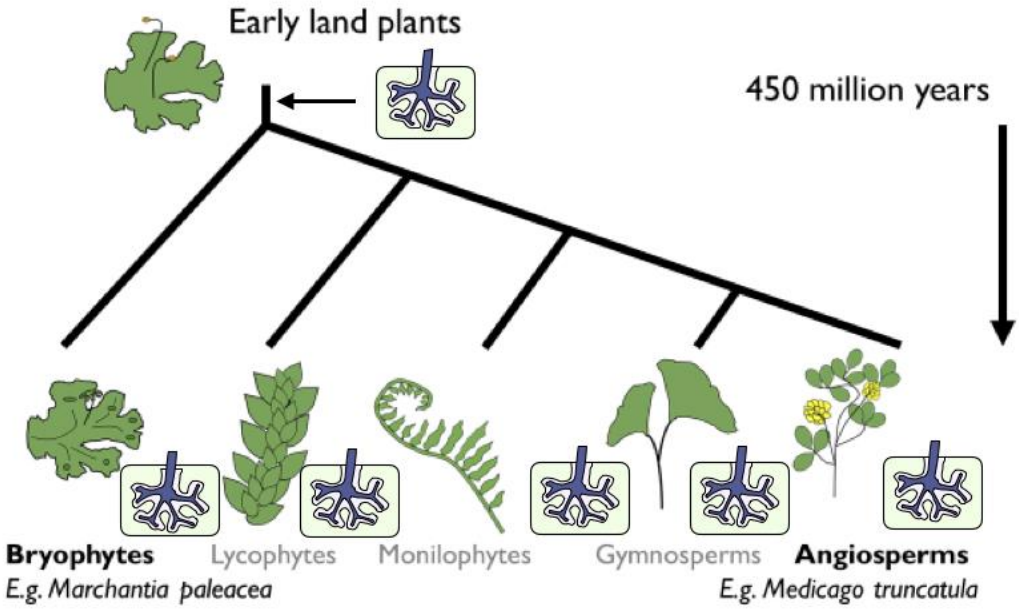



**Figure 1.1 | Phylogram of the major land plant clades, with the bryophyte/tracheophyte node left unresolved.** Arbuscular mycorrhizal symbiosis was present in the last common ancestor of the extant land plants and is present in at least some species of all the major land plant clades, with the exception of the mosses.

AM symbiosis occurs in all major land plant clades, including two of the three bryophyte lineages. This places a single origin of AM symbiosis more than 450 million years ago in the last common ancestor of the land plants (Figure 1.1). The evolution

of AM symbiosis in the land plants can be studied by comparing the molecular mechanisms of AM symbiosis between liverwort and angiosperm model species. As such, molecular phylogenetics also places the evolution of AM symbiosis more than 450 million years ago, before bryophytes diverged from vascular plants, or tracheophytes (Magallon et al., 2013).

AM symbiosis was most likely a feature of the last common ancestor of the extant land plants, as it is present in the tracheophytes and two of the three bryophyte clades, liverworts and hornworts (Read et al., 2000; Schüßler, 2000). Despite reports of fungal colonization of mosses, fungal structures are typically found in dead cells where it is unlikely that a functional symbiosis can take place (Parke and Linderman, 1980; Rabatin, 1980; Zhang and Guo, 2007). Although early land plant phylogeny has proven difficult to resolve, none of the possible topologies of hornworts, liverworts, mosses and tracheophytes supported by current methods place mosses as earliest diverging land plant lineage, suggesting that there has been a secondary loss of AM symbiosis in the moss lineage (Harrison, 2017; Rensing, 2018).

It is possible that early AM symbiosis was crucial in the colonization of land by plants, as the earliest land plants lacked true roots and soils were poorly developed (Nicolson, 1967; Pirozynski and Malloch, 1975). As the fungus fossil record is poor, it has proven difficult to date the colonization of the terrestrial environment by fungi but some phylogenetic studies indicate that Glomeromycete fungi evolved before the evolution of land plants (Lucking et al., 2009). Early Glomeromycete fungi may have acquired fixed carbon through alternative symbioses, as the basal Glomeromycete, *Geosiphon pyriformis*, does not form AM symbioses with land plants but instead forms a symbiosis with otherwise free-living soil cyanobacteria, usually *Nostoc punctiforme* (Schüßler, 2012). Early plant-fungal symbioses may have included other fungal

lineages as well as Glomeromycota. Liverworts from the early diverging lineage Haplomitriopsida do not form AM symbioses but form mutually beneficial associations with Mucoromycotina fungi (Field et al., 2015). Symbiosis with Mucoromycotina is found in most of the major land plant clades and this association has an equally ancient fossil record, as plant-Mucoromycotina associations are seen in 407 million year old fossils of the ancient plant *Horneophyton lignieri* (Strullu-Derrien et al., 2014).

Some of the strongest evidence for AM symbiosis in the last common ancestor of the extant land plants comes from genomic comparisons between bryophytes and angiosperms in which the molecular mechanisms of AM symbiosis have been characterized. Liverwort and hornwort genomes contain the genetic machinery for generating and decoding symbiosis associated $Ca^{2+}$ oscillations, and DMI3/CCaMK is functionally conserved between these lineages and angiosperms (Delaux et al., 2015; Wang et al., 2010). Further AM symbiosis specific genes have been identified by comparing the genomes of angiosperms that do or do not form AM symbioses, and many of these are retained in bryophyte species (Bravo et al., 2016; Delaux et al., 2014; Favre et al., 2014). Many AM symbiosis specific genes found in the liverwort *Lunularia cruciata* are also specifically upregulated by AM symbiosis (Delaux et al., 2015). Surprisingly, many of these genes and gene families are also found in the sister group to land plants, the charophycean algae, suggesting that the raw material for evolving symbiosis was present before the colonization of land (Delaux et al., 2015).

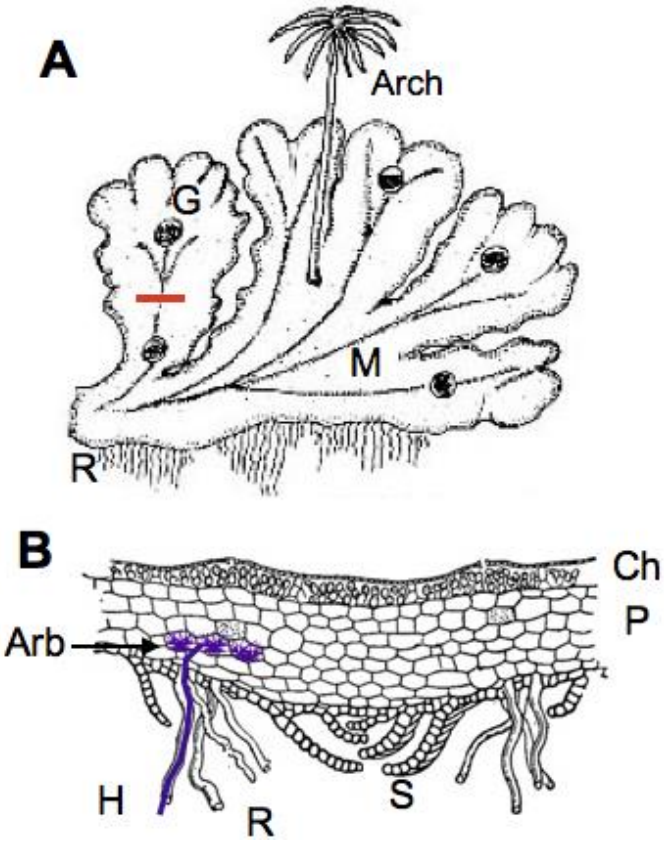



**Figure 1.2 | Anatomy of *Marchantia sp.*.** A. The gametophyte and the female reproductive structure, the archegonium (Arch) of *Marchantia sp.*, including single celled rhizoids (R) extending from the ventral surface, the midrib at the centre of the dorsal surface and asexual reproductive structures, gemma cups (G). B. Cross section of the gametophyte thallus indicated by the red line in A. The thallus is comprised of a photosynthetic chorenchymous layer (Ch) and a parenchymous layer (P), with scales (S) and rhizoids (R) on the ventral surface. During AM colonization, AM fungal hyphae (H) colonise the gametophyte via the rhizoids and arbuscules (Arb) develop in the parenchymous layer. Figure adapted from (Dittmer, 1964) A and Alamy stock photo B.

### 1.6.2 Recent evolution of AM symbiosis

The molecular mechanisms of AM symbiosis are largely conserved in angiosperms (Bravo et al., 2016). However, there has been some evolution of symbiosis genes within different angiosperm clades. While *PT4* is necessary for AM symbiosis and phosphate uptake in both monocots and dicots, additional phosphate transporters that

function in AM symbiosis have been independently recruited during angiosperm evolution. In rice, two PHT1s are required for normal colonization and arbuscule development. *OsPT11* is the orthologue of *MtPT4* and is necessary for symbiotic phosphate uptake, while OsPT13 is not required for phosphate transport but is necessary for fungal colonisation and arbuscule development (Yang et al., 2012). In the dicots, the legume *Astragalus sinicus* contains a *PT4* orthologue and requires a second transporter, *PT1*, for normal levels of root colonisation and arbuscule development (Xie et al., 2013).

Like OsPT13, AsPT1 is not required for symbiotic phosphate uptake suggesting that it functions in phosphate sensing as a transceptor (Xie et al., 2013). *OsPT13* is found in a monocot specific expansion of the PHT1 family, while *AsPT1* is part of a clade containing dicots and the basal angiosperm *Amborella trichopoda*, including the *M. truncatula* orthologue of *AsPT1*, *MtPT8*. Although *MtPT8* is upregulated during AM symbiosis, the *Mtpt8* mutant does not have an obvious symbiotic phenotype, indicating that the symbiotic function of this phosphate transporter has evolved within the legume family (Breuillin-Sessoms et al., 2015). Similarly, the ammonium transporter necessary for maintaining arbuscule development in the absence of phosphate uptake, *MtAMT2;3*, does not have a direct orthologue in monocots. *MtAMT2;3* is found in a clade with *MtAMT2;4* and *MtAMT2;5*, which are also upregulated during AM symbiosis, as well as a single copy monocot *AMT2* (Breuillin-Sessoms et al., 2015). Duplication and subfunctionalization, or recruitment of additional nutrient transporters during AM symbiosis appears to have continued over in multiple angiosperm clades.

### 1.6.3 Evolution of root nodule symbiosis

Root nodule symbiosis between plants and nitrogen fixing rhizobial bacteria (legumes and *Parasponia sp*.) or nitrogen fixing Actinorhizal bacteria is thought to have evolved from AM symbiosis in the clade within the Eurosid dicots made up of Fabales, Fagales, Cucurbitales, and Rosales, called the nitrogen-fixing root nodule (NFN) clade (Kistner and Parniske, 2002; Soltis et al., 1995). However, of the 28 families that form this clade, only one (Fabaceae) is predominantly nodulating, nine have one or a few nodulating genera and the remaining eighteen do not form root nodule symbioses (Doyle, 2011). Nodulation was thought to have evolved multiple times after a "predisposition event" in the common ancestor of this clade, based on maximum parsimony and phylogenetic modelling (Doyle, 2011; Kistner and Parniske, 2002; Soltis et al., 1995; Swensen, 1996; Werner et al., 2014).

Recent genomic comparisons between nodulating and non-nodulating species in this clade found evidence for widespread independent secondary losses of genes specific to root nodule symbiosis (Griesmann et al., 2018; van Velzen et al., 2018). This indicates that root nodule symbiosis used to be a more widespread trait in this clade and supports the multiple secondary losses hypothesis. These analyses found evidence of parallel loss or pseudogenization of three genes essential for root nodule symbiosis in non-nodulating species in the NFN; *NODULE INCEPTION* (*NIN*), *RHIZOBIUM-DIRECTED POLAR GROWTH* (*RPG*) and *NFP/NFR5* (Griesmann et al., 2018; van Velzen et al., 2018). These genes are retained in species outside the NFN clade, suggesting that evolutionary pressure to lose the capacity to form root nodule symbiosis has driven the fixation of the loss of function of mutants in these genes (Griesmann et al., 2018; van Velzen et al., 2018). This suggests that the current

distribution of the root nodule symbiosis trait in plants has been strongly affected by selection against forming associations with nitrogen fixing bacteria.

Some aspects of symbiosis signalling appear to be conserved within the Eurosid angiosperms, but not with other angiosperms. *SYMRK* from other nodulating Eurosids, and *SYMRK* from *Tropaeolum majus*, a mycorrhized Eurosid from outside the nodule forming clade, can complement the nodule symbiosis phenotype of *L. japonicus symrk* mutant (Gherbi et al., 2008; Markmann et al., 2008). However, *SymRK* genes from rice and the Asterid tomato, (*Solanum lycopersicum*), rescue AM symbiosis but do not fully restore the nodulation phenotype of *Ljsymrk* (Markmann et al., 2008). This indicates that the function of SYMRK may have acquired functions in root nodule symbiosis in specific clades within the angiosperms. This change in function may be due to the composition of the extracellular domains of SYMRK in these species; while Eurosid SYMRKs have three LRR domains, only two are found in tomato and rice (Markmann et al., 2008).

Downstream of *SYMRK*, the cation channel *POLLUX/DMI1* from rice does not fully complement either AM or root nodule symbiosis in *Ljpollux*. Overexpression of *OsPollux* restored AM symbiosis and nodule development in fewer than 20% of transformed root systems and did not restore infection by rhizobia in any root systems (Banba et al., 2008). Due to a mutation in the smallest parts of the channel pore, selectivity filter only a single cation channel, *MtDMI1*, is required to generate calcium signals in *M. truncatula*, while two *CASTOR* and *POLLUX* are required in most other angiosperms (Venkateshwaran et al., 2012). However, by modifying a single amino acid, *POLLUX* from the legume *L. japonicus*, as well as *POLLUX* from the Eurosid *A. thaliana*, can rescue rhizobial infection and nodulation in *Mtdmi* mutants even though *A. thaliana* does not form either AM or root nodule symbiosis

(Venkateshwaran et al., 2012). This indicates that at least two components of symbiosis, *SYMRK* and *DMI1*/*POLLUX*, had become functionally distinct before the evolution of nodulation in Fabales, Fagales, Cucurbitales, and Rosales.

### 1.6.4 Conservation of AM symbiosis between angiosperms and liverworts

AM symbiosis is structurally and functionally conserved between liverworts and angiosperms. Associations between fungi and liverworts have been observed since the 19$^{th}$ century (Schacht, 1854). Several Glomeromycete-liverwort associations have been identified more recently, based on ultrastructural details of intracellular fungi, showing hyphal coils and arbuscules, and sequencing of fungal symbionts (Fonseca et al., 2006; Ligrone et al., 2007; Ligrone and Lopes, 1989; Russell and Bulman, 2005). AM symbiosis is also functional in liverworts, as (Humphreys et al., 2010) found that arbuscular mycorrhizal fungi enhance uptake of mineral nutrients, primary production and asexual reproduction in *Marchantia paleacea*. The colonisation pattern of AM fungi described in *M. paleacea* is shown in figure 1.2.

Orthologues of many previously identified symbiosis genes, including the transcription factors *NSP1*, *NSP2*, *RAD1* and *RAM1*, arbuscule associated genes *RAM2*, *VAPYRIN*, *STR* and *STR2*, and the CSSP components *DMI1/Pollux*, *DMI2/SYMRK*, *DMI3/CCamK* and *IPD3/CYCLOPS* are present in liverwort transcriptomes (Delaux et al., 2015). The LysM RLK family has been expanded in the angiosperms but liverwort transcriptomes contain a homologue to the clade of LysM RLKs containing *MtNFP*/*LjNFR5* and other LYR type receptors, and a homologue to the clade containing *MtLYK9*/*LjNFR1*, *OsCERK* and the LYK type receptors (Delaux et al., 2015). In the liverwort *Lunularia cruciata*, the transcription factors *LcNSP1* and *LcRAD1*, two *RAM2* paralogues, *LcSTR* and *LcSTR2* are upregulated by AM colonization. Paralogues of the symbiotic nutrient transporters

*MtPT4* and *MtHA1* are also upregulated, but direct orthologues are not present in liverwort transcriptomes (Delaux et al., 2015).

The $Ca^{2+}$ oscillation signal decoder, *CCamK*, is functionally conserved between the *Closterium peracerosum–strigosum‐littorale* complex, a charophyte alga and *M. truncatula*, with *CpCCamK* fully rescuing the AM phenotype of *Mtccamk-1* (Delaux et al., 2015). It is likely that *CCamK* function is also conserved between liverworts and angiosperms, given that liverworts are more closely related to angiosperms than charophyte algae, but the functional conservation of other AM symbiosis genes has not been assessed outside of the angiosperms (Delaux et al., 2015).

Assessing the functional conservation of additional AM symbiosis genes between angiosperms and liverworts is an opportunity to study the evolution of this plant-fungal association over 450 million years of evolution. The molecular mechanisms of AM symbiosis in liverworts are likely to be representative of the common ancestor of the land plants, as liverworts have not undergone the repeated rounds of whole genome duplication seen in other land plant lineages, particularly the angiosperms (Berrie, 1960; Bowman et al., 2016b; Fritsch, 1991; Ren et al., 2018). The genetic mechanisms of AM symbiosis in liverworts, particularly if functionally conserved with angiosperm orthologues, are therefore likely to represent the core genetic toolkit required for AM symbiosis. The developmental context of AM symbiosis in liverworts, the gametophyte thallus, is also more representative of the first land plants than the sporophyte root system studied in angiosperms (Jones and Dolan, 2012; Kenrick and Strullu-Derrien, 2014).

## 1.7 Thesis outline and objectives

Where the genetic mechanisms of AM symbiosis differ between angiosperms and liverworts, it presents an opportunity to see how they have evolved since these lineages diverged. The life history, morphology and nutritional demands of land plants have changed massively in this time, with the evolution of a dominant sporophyte, vasculature, true roots and greater photosynthetic capacity in higher plants, any of which could place a distinct evolutionary pressure on AM symbiosis (Brodribb and Field, 2010).

Although *Marchantia polymorpha* has been developed as a model bryophyte, AM symbiosis is absent from the subspecies developed as a model organism (Bowman et al., 2016a; Fonseca et al., 2006; Ligrone et al., 2007). This project uses the closely related species *Marchantia paleacea* as it has documented AM symbiosis that nutritionally benefits the host plant (Humphreys et al., 2010).

First, putative AM symbiosis genes in *M. paleacea* will be identified by phylogenetic analysis (Chapter III). The symbiotic function of these genes will be verified where possible by assessing their regulation in response to AM fungi and comparing this to known patterns of gene expression in angiosperms (Chapter IV). Finally, the functional conservation of $Ca^{2+}$ signalling genes in the common symbiotic signalling

pathway will be assessed by transcomplementation experiments and the generation of mutant lines in *M. paleacea* (Chapter V).

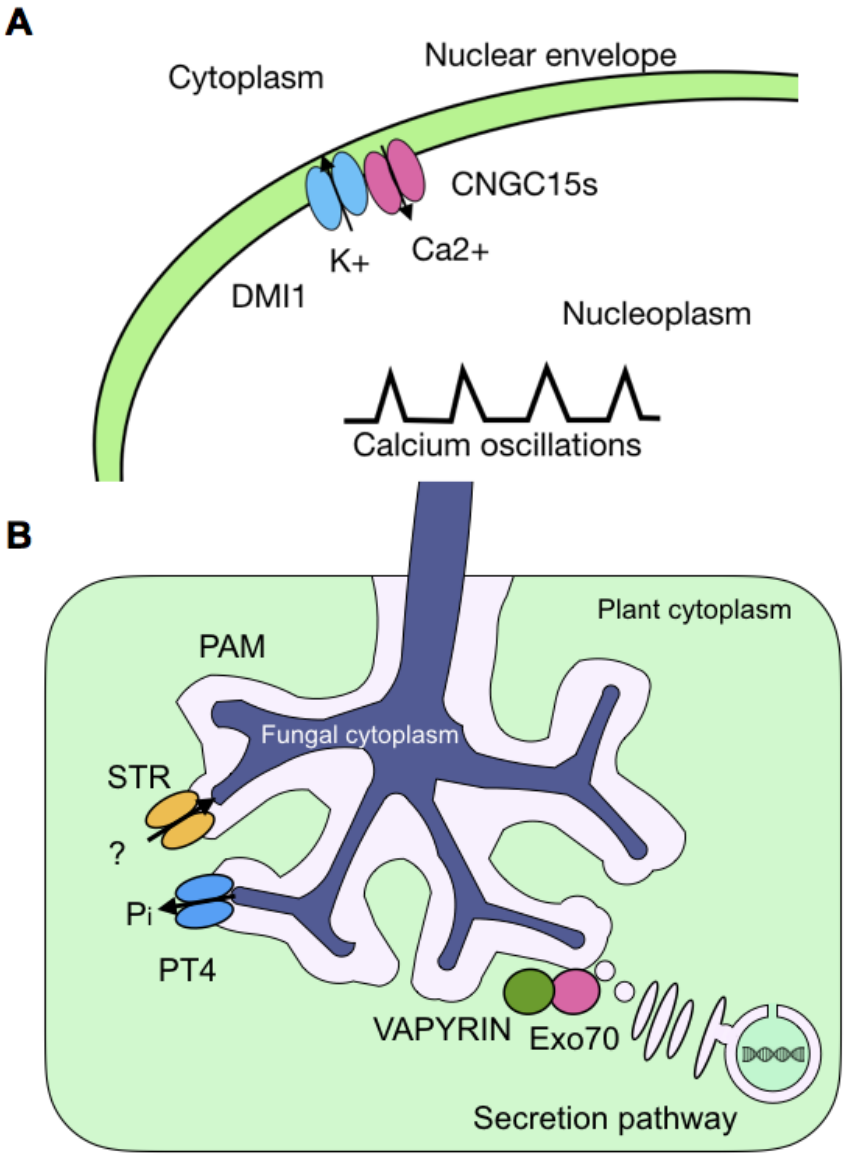



**Figure 1.3 | AM symbiosis genes identified in *M. truncatula* investigated in *M. paleacea* in this thesis.** A. Calcium signalling machinery in the common symbiosis including the cation channel DMI1 and the cyclic nucleotide gated calcium channels, the CNGC15s. B. Proteins associated with nutrient transport across the peri-arbuscular membrane (PAM), STR and VAPYRIN and proteins associated with generation of the PAM, VAPYRIN and Exo70.

# Chapter II – Materials and Methods

## 2.1 Generating the *M. paleacea* genome

### 2.1.1 Genomic DNA extraction from *M. paleacea*

1 g of 8 week old *M. polymorpha* gametophyte tissue grown in sterile conditions on ½ B5 Gamborg's media was flash frozen and finely ground in liquid nitrogen. Genomic DNA was extracted by the modified CTAB method with chloroform/isoamyl alcohol clean up as described in (Healey et al., 2014), without using polyvinylpyrrolidone (PVP). DNA quality was assessed by gel-electrophoresis for 20 minutes at 20V and 60 minutes at 30V in a 1% agarose tris-acetate-EDTA (ethylenediaminetetraacetic acid) (TAE) buffer and DNA concentration was analysed by spectrophotometrically by NanoDrop machine followed by Qubit analysis, which indicated that the DNA concentration was 61-64.2 ng/uL..

### 2.1.2 Genome sequencing

The library preparation and sequencing were carried out by GENEWIZ (South Plainfield, NJ). Both short-insert paired-end and long insert mate-pair libraries were produced for genome sequencing. For the paired-end library, DNA fragmentation was performed using the Adaptive Focused Acoustics shearing method by Covaris (Covaris Inc., Woburn, MA). Library preparation was performed using the NEBNext Ultra DNA Library Prep Kit (New England Biolabs, Ipswich, MA) with bead size selection for 300-400 bp fragments, based on the manufacturer's protocol, resulting in an average library insert size of 336 bp. The mate-pair library was prepared using the Nextera Mate-Pair DNA library prep kit (Illumina, San Diego, CA) according to the manufacturer's protocol. Following DNA fragmentation, tagging and purification, gel

size selection was used to obtain 3-5 kb fragments with an average size of 4,311 bp. Sequencing was performed on 2x100bp Rapid Run mode on the Illumina HiSeq2500 sequencer. The data files of the reads from the sequenced paired-end and mate-pair libraries were then provided by GENEWIZ as FASTQ files.

### 2.1.3 Genome assembly

Trimming of adapter and low-quality sequences of the paired-end library was performed using Trimmomatic v0.33 with the following parameters (ILLUMINACLIP:TruSeq2-PE.fa:2:30:10 LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:12). The NextClip pipeline was used to process the mate-pair library to retain predicted genuine long insert mate-pairs (Leggett et al., 2014). Genome assembly and scaffolding of contigs was performed by a colleague, Guru Radhakrishnan, using the following assemblers: ABySS(Simpson et al., 2009), SOAPdenovo2 (Luo et al., 2012), CLC Genomics Workbench (https://www.qiagenbioinformatics.com/products/clc- genomics-workbench/), IDBA-UD (Peng et al., 2012), MEGAHIT (Peng et al., 2012), SPAdes (Bankevich et al., 2012), Platanus (Kajitani et al., 2014), MaSuRCa (Zimin et al., 2013), Meraculous (Chapman et al., 2011), Velvet (Zerbino and Birney, 2008) and Minia (Chikhi and Rizk, 2012). Scaffolding of contigs recovered using the assembly tools used was performed using the scaffolder of SOAPdenovo2 (Luo et al., 2012). Assembly statistics were calculated using the ABySS-fac tool from ABySS (Simpson et al., 2009). The final assembly was 238.61 Mb, made up of 22,669 contigs, with an N50 length of 77.78 Kb and 112x sequencing depth.

## 2.2 Bioinformatics

### 2.2.1 Sequence search

The assembled *M. paleacea* genome was searched using Basic Local Alignment Search Tool (BLAST) for the genes analysed in Chapter III using the amino acid sequence of known *M. truncatula* symbiosis genes as a query. Where possible, members of these gene families were also recovered from publicly available angiosperm, lycophyte, gymnosperm, bryophyte and chlorophyte genomes by protein BLAST search of the Phytozome 12 database (https://phytozome.jgi.doe.gov/), the NCBI database (https://blast.ncbi.nlm.nih.gov/), the Congenie (http://congenie.org) database, The Arabidopsis Information Resource (TAIR) (www.arabidopsis.org), and the *Klebsorbidium nitens* genome project (http://www.plant morphogenesis.bio.titech.ac.jp/~algae_genome_project/klebsormidium/) (Goodstein et al., 2012; Hori et al., 2014; Sundell et al., 2015). For clades in the green lineage poorly represented by publicly available sequenced genomes, the gymnosperms and the charophyte algae, sequenced transcriptomes from these clades were searched using BLAST using the 1000 Plants database (Johnson et al., 2012; Matasci et al., 2014; Wickett et al., 2014; Xie et al., 2014). Results from these BLAST searches were then used as queries to BLAST search the *A. thaliana* genome (TAIR), to confirm the sequences retrieved from each species were members of the correct gene family. The sequences used for phylogenetic analyses in Chapter III are listed in Appendix A.

### 2.2.2 Sequence alignment

Amino acid sequences were aligned using the multiple sequence alignment tool ClustalX 2.1 (Larkin et al., 2007). Alignments were manually edited in MEGA7 (Molecular Evolutionary Genetics Analysis) to remove sequences with >50% of the

*M. truncatula* query gene length absent and to remove highly divergent regions and regions where gaps were present in >50% of sequences within the alignment (Kumar et al., 2016). As per (Cvrčková et al., 2012), N-terminal part of up to 300 amino acids in Exo70 sequences was removed due to highly divergent sequences. Exo70 sequences were also aligned to a guide tree generated from *A. thaliana* and *M. truncatula* sequences only rather than generated *de novo*, due to the presence of multiple highly divergent sequences. The N-terminal region of DMI1 sequences up to the second transmembrane domain was also removed due to highly divergent sequences.

### 2.2.3 Tree construction

Maximum likelihood phylogenetic trees were generated using PhyML 3.0 with automatic model selection by Smart Model Selection (SMS) selected by Akaike Information Criterion (Guindon et al., 2010; Lefort et al., 2017). Branch support was calculated by aBayes (Bayesian-like transformation of approximate likelihood ratio test) analysis (Anisimova et al., 2011). Trees were viewed and edited for clarity in Figtree v1.4.3 (http://tree.bio.ed.ac.uk/software/figtree/).

### 2.2.4 Intron-exon structure analysis

The genes were extracted from the genome of *M. paleacea* using the *M. truncatula* protein sequences as queries for BLAST searches. The coding sequences were retrieved using the "Find Open Reading Frames" tool in Geneious v10.1 (http://www.geneious.com/). The coding sequence obtained was aligned to the region of the *M. paleacea* genome containing the ORF using MEGA7 to deduce the intron-exon structures for the genes and verified using the *M. paleacea* transcriptome available in the 1000 Plants database (Matasci et al., 2014; Wickett et al., 2014; Xie et al., 2014; Johnson et a(Kumar et al., 2016)l., 2012). The intron-exon structures of genes analysed were plotted using the Exon-Intron Graphic Maker version 4

(http://wormweb.org/exonintron). The intron-exon structure of *M. truncatula* genes were taken from the currently available annotated genome sequences.

### 2.2.5 Protein sequence similarity analyses

Protein sequence were aligned using the SIM local similarity program and these alignments were visualised using LALNVIEW (Duret et al., 1996; Huang and Miller, 1991).

### 2.2.6 Identification of nuclear localisation signal motifs

Potential nuclear localisation signals were identified by LOCALIZER Plant and SeqNLS prediction tools (Lin and Hu, 2013; Sperschneider et al., 2017).

## 2.3 Plant material and growth conditions

### 2.3.1 Marchantia paleacea

*M. paleacea* thalli were kindly provided by Katie Field and David Beerling (University of Sheffield) (Humphreys et al., 2010). Plants were grown and maintained on soil (10% peat mixed with sand) in controlled environment rooms (16/8-hour day-night cycle at $22^0$C, light intensity of 100 µmol/$m^2$s). Gemma from these thalli were collected and sterilised based on the methods described in (Vujičić et al., 2011). Gemma cups were collected using a micropipette tip and placed into an Eppendorf tube. Gemmae were surface sterilised with 5% sodium hypochlorite solution for 30 seconds, which was removed with 5 rinses with water. The sterilised gemmae were grown on Gamborg's half-strength B5 medium without sucrose (Table 2.1) on sterile tissue culture plates under control conditions (16/8-hour day-night cycle at $22^0$C, light intensity of 100 µmol/$m^2$s). Gemma cup production was induced by transferring 4-week-old thalli onto Gamborg's half-strength B5 medium with 1% sucrose.

For *Agrobacterium* mediated transformation of *M. paleacea*, gametophyte fragments were co-cultured in liquid Gamborg's half-strength B5 medium with 1% sucrose with shaking in vented 50mL culture tubes (24 hour light at 22$^0$C, light intensity of 100 µmol/m$^2$s). Selection for transformants was carried out on solid (1% agar) Gamborg's half-strength B5 medium without sucrose with hygromycin selection (concentration 10 µg/mL).

### 2.3.2 Medicago truncatula

*M. truncatula* cultivar R108 was used as wild type (Hoffmann et al., 1997). All mutant lines used in this study (*Mtcngc15a*, *Mtcngc15b*, *Mtcngc15c*) are derived from this cultivar (Charpentier et al., 2016). *M. truncatula* seeds were scarified with sandpaper and surface-sterilised in 10% sodium hypochlorite solution for 3 min. Seeds were washed 5 times with sterile water and imbibed in water with 5 µg/ml nystatin and 50 µg/ml augmentin (amoxicillin clavulanate) for 4 hours before plating on water agar (DWA). Seeds were stratified at 4$^0$C for 5 days in the dark and subsequently germinated overnight at room temperature in the dark.

For *M. truncatula Agrobacterium*-mediated transformations, plants were grown on plates containing modified Fahraeus agar medium (modFP; Table 2.1) in controlled environment room (16/8-hour day-night cycle at 23$^0$C, with 32% humidity, light intensity of 300 µmol/m$^2$s).

For mycorrhization and nodulation assays, *M. truncatula* germinated seedlings were first grown on modFP as described above. After transformation (Section 2.7), plants were transferred to trays containing a 1:1 mix of terragreen and sand, inoculated with

fungal or rhizobial inoculum (Section 2.8.2 and 2.9), and grown in a controlled environment (16/8-hour day-night cycle at $22^0$C, light intensity of 300 µmol/$m^2$s).

For mycorrhizal assay, *M. paleacea* gametophytes (Section 2.8.1) were grown for 8 weeks under 16/8-hour day-night cycle at $22^0$C, with light intensity of 100 µmol/$m^2$s in 1 L vented containers ¼ filled with a soil mix consisting of 80% base mix (comprised of 90% sharp sand, 10% John Innes No. 1 compost), 20% *R. irregularis* inoculum.

### 2.3.3 Generation of AM fungi inoculum (*Rhizophagus irregularis*)

AM inoculum was propagated by growing chive seeds on a soil mix of 45% terragreen, 45% sand, 10% John Innes No. 1 compost, inoculated with 5,000 sterile spores of *R. irregularis* for 8 weeks in controlled environment (16/8-hour day-night cycle at $22^0$C, light intensity of 300 µmol/$m^2$s). At 8 weeks, chive plants were removed and the soil mix containing hyphae and spores was stored at 4°C for up to 6 months. Further propagation was continued using chive plants as above method above, grown on a 20% colonised soil inoculum/80% fresh soil mix.

| Medium | Composition for 1 L |
|---|---|
| Gamborg's B5 | KNO3 2500 mg, CaCl2(2H2O) 150 mg, MgSO4(7H2O) 250 mg, (NH4)2SO4 134 mg, NaH2PO4(H2O) 150 mg, KI 0.75 mg, H3BO3 3.0 mg, MnSO4(H2O) 10 mg, ZnSO4(7H2O) 2.0 mg, Na2MoO4(2H2O) 0.25 mg, CuSO4(5H2O) 0.025 mg, CoCl2(6H2O) 0.025 mg, Ferric-EDTA 43 mg, sucrose 2%, pH 5.5, inositol 100 mg, nicotinic acid 1.0 mg, pyridoxine.HCl 1.0 mg, thiamine. HCl 10 mg, kinetin 0.1 mg, 2,4-D 1.0 mg, 10 g agar |
| Water agar (DWA) | Bacto agar 15 g |
| modFP | CaCl2(2H2O) 0.1 g, MgSO4 0.12 g, KHPO4 0.01 g, Na2HPO4(12H2O) 0.150 g, ferric citrate 5 mg, H3BO3 2.86 g, MnSO4 2.03 g, ZnSO4(7H2O) 0.22 g, CuSO4(5H2O) 0.08 g, H2MoO4(4H2O) 0.08 g, NH4NO3 0.5 mM, Formedium agar 8 g, pH 6.0. |
| SOC (Super Optimal broth with Catabolite repression) | Tryptone 20 g, yeast extract 5 g, NaCl 0.58 g, KCl 0.19 g, MgCl2 2.03 g, MgSO4(7H2O) 2.46 g, Glucose 3.6 g |
| LB (Luria-Bertani) | Tryptone 10 g, yeast extract 5 g, NaCl 5 g. To obtain agar medium, 10 g Lab M No.1 agar was added |
| TY (Rhizobium complete medium) | Tryptone 5 g, yeast extract 3 g, CaCl2(6H2O) 1.32 g. For solid medium 10 g Lab M No.1 agar was added. |
| **Soil medium** | **Composition** |
| John Innes potting compost no. 1 (Levington) | 7 parts loam, 3 parts peat, 2 parts sand. 0.6kg ground limestone, 1.2 kg hoof and horn meal, 1.2 kg superphosphate, 0.6 kg potassium sulphate added per cubic metre. |
| Terragreen and sand mix | 1:1 mix of terragreen (Oil-dry UK ltd) and sharp sand (BB Minerals). |

**Table 2.1 | Composition of media used for the growth of plants and bacteria.**

## 2.4 Media and antibiotics

The composition of the media used for plant and bacterial growth in the current study are detailed in the Table 2.1. Antibiotics (Table 2.2) were added to the media where necessary for the selection of transgenic plants and bacteria.

| **Antibiotic/selection marker** | **Dissolved in** | **Final concentration liquid culture (**µg/mL) | **Final concentration solid media (µg/mL)2** |
|---|---|---|---|
| Rifampicin | DMSO | 20 | 100 |
| Spectinomycin | sterile water | 200 | 400 |
| Kanamycin | sterile water | 20 | 20 |
| Carbenicillin | sterile water | 50 | 100 |
| Hygromycin | sterile water | - | 10 |
| Cefotaxime | sterile water | - | 100 |
| X-Gal (5-Bromo-4-chloro-3-indolyl β-D-galactopyranoside) | DMF | - | 40 |
| Nystatin | DMSO | 5 | 2 |
| Augmentin (amoxicillin and clavulanate potassium) | sterile water | 100 | - |

**Table 2.2 | Antibiotics and visual selection markers used for the selection of transformed plants and bacteria.**

## 2.5 Molecular cloning

All vectors used in this study were cloned using the GoldenGate cloning strategy (Engler et al., 2014; Weber et al., 2011). The DNA components were synthesised by

GenArt (Life Technologies). DNA sequences used in molecular cloning that were not synthesised by GenArt included the *MpaU6-1* promoter and the *MpaTUBa1* promoter, which were amplified from *M. paleacea* genomic DNA in a 50 µL reaction containing 10 µL HF buffer, 1 µL dNTPs, 2.5 µL 10 µM forward primer, 2.5 µL 10 µM reverse primer, 0.5 µL Phusion DNA Polymerase, 100 ng *M. paleacea* genomic DNA. Cycling conditions were as follows: initial melting at 98°C for 30 seconds followed by 30 cycles of 10 seconds at 98°C, 10 seconds at the annealing temperature specified in Table 2.3, 60 seconds at 72 °C, followed by a final extension of 2 minutes at 72 °C. Primers used included extensions compatible with the GoldenGate cloning system and are listed in Table 2.3. Where necessary, Golden Gate incompatible restriction sites were removed by site directed mutagenesis using the Q5® Site-Directed Mutagenesis Kit according to the manufacturers instructions with primers as described in Table 2.3 (New England Biolabs).

DNA sequences used in molecular cloning that were not synthesised by GenArt also included short sequences used to generate the targeting RNA sequence (crRNA) that together with a Cas9 nuclease-recruiting sequence (tracrRNA) comprises the single guide RNA (sgRNA) that directs Cas9 nuclease activity during CRSIPR/Cas9 genome editing. Sequences were selected using sgRNA Scorer 2.0 using the genomic DNA sequences of the genes targeted (Chari et al., 2017).

These were synthesised as DNA oligomers by Eurofins genomics (https://www.eurofinsgenomics.eu/) and annealed at equimolar concentrations to form double stranded DNA molecules (dsDNA) in a solution of 10mM Tris-HCl (pH 7.5), 50 mM NaCl, 1mM EDTA, 100 mg/ µL total DNA oligomers.

To generate level 1 vectors, a 15 µL reaction mix was set up containing 100 ng of each level 0 plasmid, 100 ng backbone plasmid, 0.15 µL 100 x BSA, 1.5 µL 10 x T4 buffer,

1 µL BsaI (New England Biolabs), 1 µL T4 DNA ligase (New England Biolabs) and water. Cycling conditions for digestion and ligation of the DNA components were as follows: 25 cycles of 3 min at 37°C and 4 min at 16°C, followed by 5 min at 50°C and 5 min at 80°C.

To generate level 1.5 vectors (insertion of short guide sequences to a level 1 vector for CRISPR/Cas9 genome editing), 15 µL reaction mix was set up containing 100 ng of the acceptor level 1 plasmid, a nanogram quantity of the annealed oligo mixture in a 3:1 molar ratio to the acceptor plasmid, 0.15 µL 100 x BSA, 1.5 µL 10 x T4 buffer, 1 µL BsmBI (New England Biolabs), 1 µL BsaI (New England Biolabs), 1 µL T4 DNA ligase (New England Biolabs) and water.

To generate level 2 vectors, a 15 µL reaction mix was set up containing 100 ng of each level 1 plasmid, 100 ng of the backbone plasmid pICH50505 (Icon-Genetics), 0.15 µL 100x BSA, 1 µL BpiI (Thermo Fisher Scientific), 1.5 µL 10 x T4 buffer, 1 µL T4 DNA ligase (New England Biolabs) and water. Cycling conditions for digestion and ligation of the DNA components were as follows: 25 cycles of 3 min at 37°C and 4 min at 16°C, followed by 5 min at 50°C and 5 min at 80°C. Level 2 vectors used in this study are listed in Table 2.4.

For amplification, level 0, level 1 and level 2 plasmids were transformed into chemically competent *E. coli* DH5α strain as follows: 2 µL of reaction was mixed with 20 µL of chemically competent *E. coli* DH5α (ThermoFisher Scientific) and incubated 30 minutes on ice, heat shocked for 30 seconds at 42°C and incubated for 5 minutes on ice, before being suspended in 200 µL SOC (Table 2.2). The resuspended *E. coli* was incubated for 1 hour at 37°C at 220 rpm before being plated on LB agar plates (Table 2.2) supplemented with the appropriate antibiotics and incubated overnight at 37°C.

Colony PCR was used to check whether the plasmid of interest was successfully transformed by using a micropipette tip touched to the colony of interest and dipped in a PCR reaction comprised of 10 μL 2X GoTaq® Green Master Mix, 1 μL 10 nM forward primer, 1 μL 10 nM reverse primer and 8 μL water. Cycling conditions were 2 minutes at 95°C, 30 cycles of 30 seconds at 95 °C, 30 seconds at the appropriate annealing temperature, 1 minute at 72°C, followed by an extension of 5 minutes at 72°C and incubation at 4°C. Primers and annealing temperatures used were as described in Table 2.3.

| Name | 3' to 5' sequence | Used in this study for | Tm |
|---|---|---|---|
| MpaDMI1gRNA1aF | cactctgtggtctcaTCTCGCAGTCACCGAGTGGAGCGTTTtgagaccacgaagtg | sgRNA guide synthesis | N/A |
| MpaDMI1gRNA1aR | CACTTCGTGGTCTCAAAACGCTCCACTCGGTGACTGCGAGATGAGACCACAGAGTG | sgRNA guide synthesis | N/A |
| MpaDMI1gRNA1bF | cactctgtggtctcaTCTCGAAGCAGCGGCATCAGGCGCGTTTtgagaccacgaagtg | sgRNA guide synthesis | N/A |
| MpaDMI1gRNA1bR | CACTTCGTGGTCTCAAAACGCGCCTGATGCCGCTGCTTCGAGATGAGACCACAGAGTG | sgRNA guide synthesis | N/A |
| MpaDMI1gRNA2aF | cactctgtggtctcaTCTCGTTGAGCAGGCTGGTAATGTTTtgagaccacgaagtg | sgRNA guide synthesis | N/A |
| MpaDMI1gRNA2aR | CACTTCGTGGTCTCAAAACATTACCAGCCTGCTCAACGAGATGAGACCACAGAGTG | sgRNA guide synthesis | N/A |
| MpaDMI1gRNA2bF | cactctgtggtctcaTCTCGCTCGCTTGAAAGTGAAAGGTTTtgagaccacgaagtg | sgRNA guide synthesis | N/A |
| MpaDMI1gRNA2bR | CACTTCGTGGTCTCAAAACCTTTCACTTTCAAGCGAGCGAGATGAGACCACAGAGTG | sgRNA guide synthesis | N/A |
| MpaCNGC1gRNA1aF | CACTCTGTGGTCTCATCTCGAGGCATTTACTTATCCAGCGTTTTGAGACCACGAAGTG | sgRNA guide synthesis | N/A |
| MpaCNGC1gRNA1aR | CACTTCGTGGTCTCAAAACGCTGGATAAGTAAATGCCTCGAGATGAGACCACAGAGTG | sgRNA guide synthesis | N/A |
| MpaCNGC1gRNA1bF | CACTCTGTGGTCTCATCTCGTCCAGCACAGATTCCAGGCGTTTTGAGACCACGAAGTG | sgRNA guide synthesis | N/A |
| MpaCNGC1gRNA1bR | CACTTCGTGGTCTCAAAACGCCTGGAATCTGTGCTGGACGAGATGAGACCACAGAGTG | sgRNA guide synthesis | N/A |
| MpaCNGC1gRNA2aF | CACTCTGTGGTCTCATCTCGCCTTCAAAGGCTATGTTCGTTTTGAGACCACGAA GTG | sgRNA guide synthesis | N/A |

| | | | |
|---|---|---|---|
| MpaCNGC1gRNA2aR | CACTTCGTGGTCTCAAAA CGAACATAGCCTTTGAAG GCGAGATGAGACCACAG AGTG | sgRNA guide synthesis | N/A |
| MpaCNGC1gRNA2bF | CACTCTGTGGTCTCATCT CGGCCCAGGGCCGAGG CCAGGGTTTTGAGACCA CGAAGTG | sgRNA guide synthesis | N/A |
| MpaCNGC1gRNA2bR | CACTTCGTGGTCTCAAAA CCCTGGCCTCGGCCCTG GGCCGAGATGAGACCAC AGAGTG | sgRNA guide synthesis | N/A |
| MpaCNGC1gRNA1aF | CACTCTGTGGTCTCATCT CGAGGCATTTACTTATCC AGCGTTTTGAGACCACGA AGTG | sgRNA guide synthesis | N/A |
| MpaCNGC1gRNA1aR | CACTTCGTGGTCTCAAAA CGCTGGATAAGTAAATGC CTCGAGATGAGACCACA GAGTG | sgRNA guide synthesis | N/A |
| MpaCNGC1gRNA1bF | CACTCTGTGGTCTCATCT CGTCCAGCACAGATTCCA GGCGTTTTGAGACCACG AAGTG | sgRNA guide synthesis | N/A |
| MpaCNGC1gRNA1bR | CACTTCGTGGTCTCAAAA CGCCTGGAATCTGTGCT GGACGAGATGAGACCAC AGAGTG | sgRNA guide synthesis | N/A |
| MpaCNGC1gRNA2aF | CACTCTGTGGTCTCATCT CGCCTTCAAAGGCTATGT TCGTTTTGAGACCACGAA GTG | sgRNA guide synthesis | N/A |
| MpaCNGC1gRNA2aR | CACTTCGTGGTCTCAAAA CGAACATAGCCTTTGAAG GCGAGATGAGACCACAG AGTG | sgRNA guide synthesis | N/A |
| MpaCNGC1gRNA2bF | CACTCTGTGGTCTCATCT CGGCCCAGGGCCGAGG CCAGGGTTTTGAGACCA CGAAGTG | sgRNA guide synthesis | N/A |
| MpaCNGC1gRNA2bR | CACTTCGTGGTCTCAAAA CCCTGGCCTCGGCCCTG GGCCGAGATGAGACCAC AGAGTG | sgRNA guide synthesis | N/A |
| MpaCNGC1gRNA3aF | cactctgtggtctcaTCTCGTAGAGTT ATTCAATCAGAGTTTtgagaccacg aagtg | sgRNA guide synthesis | N/A |
| MpaCNGC1gRNA3aR | CACTTCGTGGTCTCAAAACTCT GATTGAATAACTCTACGAGATG AGACCACAGAGTG | sgRNA guide synthesis | N/A |
| MpaCNGC1gRNA3bF | cactctgtggtctcaTCTCGCAACCGG CGCTGCACACCGGTTTtgagacca cgaagtg | sgRNA guide synthesis | N/A |
| MpaCNGC1gRNA3bR | CACTTCGTGGTCTCAAAACCGG TGTGCAGCGCCGGTTGCGAGA TGAGACCACAGAGTG | sgRNA guide synthesis | N/A |
| MpaCNGC1gRNA4aF | cactctgtggtctcaTCTCGCTAGATG CAATTTGCGGTTTtgagaccacgaa gtg | sgRNA guide synthesis | N/A |
| MpaCNGC1gRNA4aR | CACTTCGTGGTCTCAAAACCGC AAATTGCATCTAGCGAGATGAG ACCACAGAGTG | sgRNA guide synthesis | N/A |
| MpaCNGC1gRNA4bF | cactctgtggtctcaTCTCGGAAAGCA TGACGACCAAGTTTtgagaccacga agtg | sgRNA guide synthesis | N/A |
| MpaCNGC1gRNA4bR | CACTTCGTGGTCTCAAAACTTG GTCGTCATGCTTTCCGAGATGA GACCACAGAGTG | sgRNA guide synthesis | N/A |

| | | | |
|---|---|---|---|
| MpaCNGC1gRNA5aF | cactctgtggtctcaTCTCGGACCTCAAATTCGTGGCGTTTtgagaccacgaagtg | sgRNA guide synthesis | N/A |
| MpaCNGC1gRNA5aR | CACTTCGTGGTCTCAAAACGCCACGAATTTGAGGTCCGAGATGAGACCACAGAGTG | sgRNA guide synthesis | N/A |
| MpaCNGC1gRNA5bF | cactctgtggtctcaTCTCGAAGCTTTCTGAACTGTCGTTTtgagaccacgaagtg | sgRNA guide synthesis | N/A |
| MpaCNGC1gRNA5bR | CACTTCGTGGTCTCAAAACGACAGTTCAGAAAGCTTCGAGATGAGACCACAGAGTG | sgRNA guide synthesis | N/A |
| MpaCNGC1gRNA6aF | cactctgtggtctcaTCTCGTCATAGAAGTAGAAAGCGTTTtgagaccacgaagtg | sgRNA guide synthesis | N/A |
| MpaCNGC1gRNA6aR | CACTTCGTGGTCTCAAAACGCTTTCTACTTCTATGACGAGATGAGACCACAGAGTG | sgRNA guide synthesis | N/A |
| MpaCNGC1gRNA6bF | cactctgtggtctcaTCTCGCATACATGCAAGCAAATGAGTTTtgagaccacgaagtg | sgRNA guide synthesis | N/A |
| MpaCNGC1gRNA6bR | CACTTCGTGGTCTCAAAACTCATTTGCTTGCATGTATGCGAGATGAGACCACAGAGTG | sgRNA guide synthesis | N/A |
| MpaCNGC3gRNA1aF | cactctgtggtctcaTCTCGGTTGGGAGAAGAGCTCCAAGTTTtgagaccacgaagtg | sgRNA guide synthesis | N/A |
| MpaCNGC3gRNA1aR | CACTTCGTGGTCTCAAAACTTGGAGCTCTTCTCCCAACCGAGATGAGACCACAGAGTG | sgRNA guide synthesis | N/A |
| MpaCNGC3gRNA1bF | cactctgtggtctcaTCTCGCTTGAAATCAGGTGTTTCCAGTTTtgagaccacgaagtg | sgRNA guide synthesis | N/A |
| MpaCNGC3gRNA1bR | CACTTCGTGGTCTCAAAACTTGGAGCTCTTCTCCCAACCGAGATGAGACCACAGAGTG | sgRNA guide synthesis | N/A |
| MpaCNGC3gRNA2aF | cactctgtggtctcaTCTCGACATCTGCATTCGGATCGAAGTTTtgagaccacgaagtg | sgRNA guide synthesis | N/A |
| MpaCNGC3gRNA2aR | CACTTCGTGGTCTCAAAACTTCGATCCGAATGCAGATGTCGAGATGAGACCACAGAGTG | sgRNA guide synthesis | N/A |
| MpaCNGC3gRNA2bF | cactctgtggtctcaTCTCGAAGTGGAATAAGTTCTTCGGTTTtgagaccacgaagtg | sgRNA guide synthesis | N/A |
| MpaCNGC3gRNA2bR | CACTTCGTGGTCTCAAAACCGAAGAACTTATTCCACTTCGAGATGAGACCACAGAGTG | sgRNA guide synthesis | N/A |
| MpaCNGC3gRNA3aF | cactctgtggtctcaTCTCGATGGCATCTAGTAGTGTTTtgagaccacgaagtg | sgRNA guide synthesis | N/A |
| MpaCNGC3gRNA3aR | CACTTCGTGGTCTCAAAACACTACTAGATGCCATCGAGATGAGACCACAGAGTG | sgRNA guide synthesis | N/A |
| MpaCNGC3gRNA3bF | cactctgtggtctcaTCTCGCAAACATACATAGTGCGTGGTTTtgagaccacgaagtg | sgRNA guide synthesis | N/A |
| MpaCNGC3gRNA3bR | CACTTCGTGGTCTCAAAACCACGCACTATGTATGTTTGCGAGATGAGACCACAGAGTG | sgRNA guide synthesis | N/A |
| MpaCNGC3gRNA4aF | cactctgtggtctcaTCTCGTTTGCTCTGTGTTCAGAGCGTTTtgagaccacgaagtg | sgRNA guide synthesis | N/A |
| MpaCNGC3gRNA4aR | CACTTCGTGGTCTCAAAACGCTCTGAACACAGAGCAAACGAGATGAGACCACAGAGTG | sgRNA guide synthesis | N/A |

| | | | |
|---|---|---|---|
| MpaCNGC3gRNA4bF | cactctgtggtctcaTCTCGGAGAGTGTGACGACAAAGTTTtgagaccacgaagtg | sgRNA guide synthesis | N/A |
| MpaCNGC3gRNA4bR | CACTTCGTGGTCTCAAAACTTTGTCGTCACACTCTCCGAGATGAGACCACAGAGTG | sgRNA guide synthesis | N/A |
| MpaCNGC3gRNA5aF | cactctgtggtctcaTCTCGATACTACTCGCACCAGGTTTtgagaccacgaagtg | sgRNA guide synthesis | N/A |
| MpaCNGC3gRNA5aR | CACTTCGTGGTCTCAAAACCTGGTGCGAGTAGTATCGAGATGAGACCACAGAGTG | sgRNA guide synthesis | N/A |
| MpaCNGC3gRNA5bF | cactctgtggtctcaTCTCGAAGTTATTAGAACTACGAGTTTtgagaccacgaagtg | sgRNA guide synthesis | N/A |
| MpaCNGC3gRNA5bR | CACTTCGTGGTCTCAAAACTCGTAGTTCTAATAACTTCGAGATGAGACCACAGAGTG | sgRNA guide synthesis | N/A |
| MpaCNGC3gRNA6aF | cactctgtggtctcaTCTCGTTCGTAGCCAGTCAGTTCGTTTtgagaccacgaagtg | sgRNA guide synthesis | N/A |
| MpaCNGC3gRNA6aR | CACTTCGTGGTCTCAAAACGAACTGACTGGCTACGAACGAGATGAGACCACAGAGTG | sgRNA guide synthesis | N/A |
| MpaCNGC3gRNA6bF | cactctgtggtctcaTCTCGCATGCTTCATCCAAGCCGTTTtgagaccacgaagtg | sgRNA guide synthesis | N/A |
| MpaCNGC3gRNA6bR | CACTTCGTGGTCTCAAAACGGCTTGGATGAAGCATGCGAGATGAGACCACAGAGTG | sgRNA guide synthesis | N/A |
| MpaACT_1F | CATGGATACCGGATGCCTCC | quantitative RT PCR | 62°C |
| MpaACT_1R | GGTGCCCAGAGGTTCTGTTC | quantitative RT PCR | 62°C |
| MpaAPT_2F | GGGTACACTTGCTGCAGGAA | quantitative RT PCR | 62°C |
| MpaAPT_2R | CTCACGGCCCTTTAGATCCG | quantitative RT PCR | 62°C |
| MpaVPY_5F | CCGTGGCGTACAAGATTCAG | quantitative RT PCR | 60°C |
| MpaVPY_5R | GGTTCATGGCAATCTCGACC | quantitative RT PCR | 60°C |
| MpaPTD_1F | TCCATGATCCTTGCTTGCCGT | quantitative RT PCR | 60°C |
| MpaPTD_1R | GGAACTCTTCCCTGCCAGAC | quantitative RT PCR | 60°C |
| MpaSTR_3F | TGCTCGTCTATCTCCCGTTC | quantitative RT PCR | 60°C |
| MpaSTR_3R | ACGAAGGAGTTGGTGGTG | quantitative RT PCR | 60°C |
| MpaExo70GI_2F | AGCTACATCCAGACCTTCCG | quantitative RT PCR | 60°C |
| MpaExo70GI_2R | TTCCGTTGCGTGGTCTACTA | quantitative RT PCR | 60°C |
| DMI1-qPCR-46619F | CGTCGACACCTTGAGAGCTT | quantitative RT PCR | 60°C |
| DMI1-qPCR-47470R | TGGTACGGCATTCGCTTTGA | quantitative RT PCR | 60°C |
| MpaCNG1-qPCR-1427F | AAACTCAGGGCGTGGATGAG | quantitative RT PCR | 60°C |
| MpaCNG1-qPCR-1582R | GCAAGACCTCGCAAATTGCA | quantitative RT PCR | 60°C |
| MpaCNG2-qPCR-1534F | ATTTGCGAACGGTTGAAGCC | quantitative RT PCR | 60°C |
| MpaCNG2-qPCR-1725R | ACACCACGACAACAGTTCGT | quantitative RT PCR | 60°C |
| MpaCNG3-qPCR-9014F | ACGACGCACACGATCTTTCA | quantitative RT PCR | 60°C |
| MpaCNG3-qPCR-10133R | TCGCACTTCTCATCGGCAAT | quantitative RT PCR | 60°C |
| MpaCNG4-qPCR-1439F | TTTGGATGAAGCGACGTCGA | quantitative RT PCR | 60°C |
| MpaCNG4-qPCR-1512R | ACCCATCTGAATCGCCCAG | quantitative RT PCR | 60°C |
| Genotype_DM I1_rev2 | ATCTTTGGAATCCGGACC CC | Genotyping | 59°C |
| Genotype_CNGC1cut_rev2 | GGGTCCGAGCTTATTGCAGT | Genotyping | 59°C |
| Genotype_CNGC1cut_fwd2 | CCGGACAGTTCAGAAAGCTT | Genotyping | 57°C |
| Genotype_CNGC3cut_fwd2 | TCGGGTGGTAGTTTGCTGTT | Genotyping | 57°C |
| Genotype_CNGC3cut_rev2 | AAGACGGAAGCGTTTTGCTG | Genotyping | 57°C |
| Genotype_CNGC3cut_fwd3 | AGCTTCTTACTTGGGCGTTG | Genotyping | 57°C |
| Genotype_CNGC3KO_fwd2 | CAGATAAAGGTTCAGAG GCC | Genotyping | 57°C |
| Genotype_CNGC3KO_rev2 | CTTCTGTACGAAGTCGTT GC | Genotyping | 57°C |
| DMI1_genotyp e_fwd | GCGAGAGTCAGCTTTCTT CA | Genotyping | 57°C |

| | | | |
|---|---|---|---|
| DMI1_genotyp e_rev | TGCCACTGGACATATGG CCA | Genotyping | 59°C |
| CNGC1cut_ge notype_fwd | GAGATACGGAAGAGTGGATG | Genotyping | 57°C |
| CNGC1_cutge notype_rev | TGTGGCTGTCATCTCGGAAG | Genotyping | 59°C |
| CNGC3cut_ge notype_fwd | GTGCGGAGGGTAAGAATCACA | Genotyping | 59°C |
| CNGC3cut_ge notype_rev | TGGCTCAGAAGGCTTCGGAA | Genotyping | 59°C |
| CNGC1_genot ype_fwd | GATCATCAGTGGTGGTTGGG | Genotyping | 59°C |
| CNGC1_genot ype_rev | ATCACCGTTCGGAGGACCGT | Genotyping | 61°C |
| CNGC3_genot ype_fwd | TCCGTACGAGTACCAGATTG | Genotyping | 57°C |
| CNGC3_genot ype_rev | AGAGGTAGTACAGCCACCAA | Genotyping | 57°C |
| SDMMpau6Node606_FWD | AAAAGAGAGGCGAATACAATCT TG | Molecular Cloning | 58°C |
| SDMMpau6Node606_FWD | GCGTGACAATTTACAAGAAAAT AATTTC | Molecular Cloning | 58°C |
| MpaPU U6-1 fwd | CGAAGACAAGAGGAGACTGGC TTCAATCAAAGC | Molecular Cloning | 72+°C |
| MpaPU U6-1 rev | TGCCGTCTCCGAGATAGGGTG CAGCTGCG | Molecular Cloning | 72+°C |
| SgRNA acceptor FWD | CGCAGCTGCACCCTATCTCGG AGACGGCA | Molecular Cloning | 72+°C |
| SgRNA acceptor REV | CGCAGCTGCACCCTATCTCGG AGACGGCA | Molecular Cloning | 72+°C |
| cloningMpaPUTUA5_FWD | CACTCTGTGGTCTCAGGAGTC TCGAGTAACAGTTTTTC | Molecular Cloning | 72+°C |
| cloningMpaPUTUA5_REV | CACTTCGTG GTC TCACATTCT CTTCCTTCTCCTGAAAC | Molecular Cloning | 72+°C |
| FOR-KAN | ATGATTGAACAAGATGGATTG | colony PCR | 53°C |
| REV-KAN | TCAGAAGAACTCGTCAAGAAG | colony PCR | 56°C |
| FOR-HYG | ATGAAAAAGCCTGAACTCACC | colony PCR | 58°C |
| REV-HYG | CTATTCCTTTGCCCTCGGA | colony PCR | 58°C |

**Table 2.3 | Primers used in this study.**

| Name | PR 1 | PR 2 | PR 3 | PR 4 | PR 5 | Use in the study |
|---|---|---|---|---|---|---|
| AC218 | p35S::HYG::tNOS | pMpaDMI1::GUS::t35S | p35S::mCherry:tNOS | | | GUS reporter |
| AC219 | p35S::HYG::tNOS | pMpaCNGC1::GUS::t35S | p35S::mCherry:tNOS | | | GUS reporter |
| AC220 | p35S::HYG::tNOS | pMpaCNGC2::GUS::t35S | p35S::mCherry:tNOS | | | GUS reporter |
| AC221 | p35S::HYG::tNOS | pMpaCNGC3::GUS::t35S | p35S::mCherry:tNOS | | | GUS reporter |
| AC222 | p35S::HYG::tNOS | pMpaCNGC4::GUS::t35S | p35S::mCherry:tNOS | | | GUS reporter |
| AC265 | p35S::HYG::tNOS | pMpoEF1a::AtCas9::t35S | pMpaU6-1::sgRNADMI1-1a | pMpaU6-1::sgRNADMI1-1b | p35S::mCherry:tNOS | Genome editing: *MpaDMI1* |
| AC266 | p35S::HYG::tNOS | pMpoEF1a::AtCas9::t35S | pMpaU6-1::sgRNADMI1-2a | pMpaU6-1::sgRNADMI1-2b | p35S::mCherry:tNOS | Genome editing: *MpaDMI1* |
| AC267 | p35S::HYG::tNOS | pMpoEF1a::AtCas9::t35S | pMpaU6-1::sgRNACNGC1-1a | pMpaU6-1::sgRNACNGC1-1b | p35S::mCherry:tNOS | Genome editing: *MpaCNGC1* |
| AC268 | p35S::HYG::tNOS | pMpoEF1a::AtCas9::t35S | pMpaU6-1::sgRNACNGC1-2a | pMpaU6-1::sgRNACNGC1-2b | p35S::mCherry:tNOS | Genome editing: *MpaCNGC1* |
| AC269 | p35S::HYG::tNOS | pMpoEF1a::AtCas9::t35S | pMpaU6-1::sgRNACNGC1-3a | pMpaU6-1::sgRNACNGC1-3b | p35S::mCherry:tNOS | Genome editing: *MpaCNGC1* |
| AC270 | p35S::HYG::tNOS | pMpoEF1a::AtCas9::t35S | pMpaU6-1::sgRNACNGC1-4a | pMpaU6-1::sgRNACNGC1-4b | p35S::mCherry:tNOS | Genome editing: *MpaCNGC1* |
| AC271 | p35S::HYG::tNOS | pMpoEF1a::AtCas9::t35S | pMpaU6-1::sgRNACNGC1-5a | pMpaU6-1::sgRNACNGC1-5b | p35S::mCherry:tNOS | Genome editing: *MpaCNGC1* |
| AC272 | p35S::HYG::tNOS | pMpoEF1a::AtCas9::t35S | pMpaU6-1::sgRNACNGC1-6a | pMpaU6-1::sgRNACNGC1-6b | p35S::mCherry:tNOS | Genome editing: *MpaCNGC1* |
| AC273 | p35S::HYG::tNOS | pMpoEF1a::AtCas9::t35S | pMpaU6-1::sgRNACNGC3-1a | pMpaU6-1::sgRNACNGC3-1b | p35S::mCherry:tNOS | Genome editing: *MpaCNGC3* |
| AC274 | p35S::HYG::tNOS | pMpoEF1a::AtCas9::t35S | pMpaU6-1::sgRNACNGC3-2a | pMpaU6-1::sgRNACNGC3-2b | p35S::mCherry:tNOS | Genome editing: *MpaCNGC3* |
| AC275 | p35S::HYG::tNOS | pMpoEF1a::AtCas9::t35S | pMpaU6-1::sgRNACNGC3-3a | pMpaU6-1::sgRNACNGC3-3b | p35S::mCherry:tNOS | Genome editing: *MpaCNGC3* |
| AC276 | p35S::HYG::tNOS | pMpoEF1a::AtCas9::t35S | pMpaU6-1::sgRNACNGC3-4a | pMpaU6-1::sgRNACNGC3-4b | p35S::mCherry:tNOS | Genome editing: *MpaCNGC3* |
| AC277 | p35S::HYG::tNOS | pMpoEF1a::AtCas9::t35S | pMpaU6-1::sgRNACNGC3-5a | pMpaU6-1::sgRNACNGC3-5b | p35S::mCherry:tNOS | Genome editing: *MpaCNGC3* |
| AC278 | p35S::HYG::tNOS | pMpoEF1a::AtCas9::t35S | pMpaU6-1::sgRNACNGC3-6a | pMpaU6-1::sgRNACNGC3-6b | p35S::mCherry:tNOS | Genome editing: *MpaCNGC3* |
| AC283 | p35S::HYG::tNOS | pMpaTUa1::AtCas9::t35S | pMpaU6-1::sgRNADMI1-1a | pMpaU6-1::sgRNADMI1-1b | p35S::mCherry:tNOS | Genome editing: *MpaDMI1* |
| AC284 | p35S::HYG::tNOS | pMpaTUa1::AtCas9::t35S | pMpaU6-1::sgRNADMI1-2a | pMpaU6-1::sgRNADMI1-2b | p35S::mCherry:tNOS | Genome editing: *MpaDMI1* |
| AC285 | p35S::HYG::tNOS | pMpaTUa1::AtCas9::t35S | pMpaU6-1::sgRNACNGC1-1a | pMpaU6-1::sgRNACNGC1-1b | p35S::mCherry:tNOS | Genome editing: *MpaCNGC1* |
| AC286 | p35S::HYG::tNOS | pMpaTUa1::AtCas9::t35S | pMpaU6-1::sgRNACNGC1-2a | pMpaU6-1::sgRNACNGC1-2b | p35S::mCherry:tNOS | Genome editing: *MpaCNGC1* |
| AC287 | p35S::HYG::tNOS | pMpaTUa1::AtCas9::t35S | pMpaU6-1::sgRNACNGC1-3a | pMpaU6-1::sgRNACNGC1-3b | p35S::mCherry:tNOS | Genome editing: *MpaCNGC1* |
| AC288 | p35S::HYG::tNOS | pMpaTUa1::AtCas9::t35S | pMpaU6-1::sgRNACNGC1-4a | pMpaU6-1::sgRNACNGC1-4b | p35S::mCherry:tNOS | Genome editing: *MpaCNGC1* |
| AC289 | p35S::HYG::tNOS | pMpaTUa1::AtCas9::t35S | pMpaU6-1::sgRNACNGC1-5a | pMpaU6-1::sgRNACNGC1-5b | p35S::mCherry:tNOS | Genome editing: *MpaCNGC1* |
| AC290 | p35S::HYG::tNOS | pMpaTUa1::AtCas9::t35S | pMpaU6-1::sgRNACNGC1-6a | pMpaU6-1::sgRNACNGC1-6b | p35S::mCherry:tNOS | Genome editing: *MpaCNGC1* |
| AC291 | p35S::HYG::tNOS | pMpaTUa1::AtCas9::t35S | pMpaU6-1::sgRNACNGC3-1a | pMpaU6-1::sgRNACNGC3-1b | p35S::mCherry:tNOS | Genome editing: *MpaCNGC3* |
| AC292 | p35S::HYG::tNOS | pMpaTUa1::AtCas9::t35S | pMpaU6-1::sgRNACNGC3-2a | pMpaU6-1::sgRNACNGC3-2b | p35S::mCherry:tNOS | Genome editing: *MpaCNGC3* |

| | | | | | | |
|---|---|---|---|---|---|---|
| AC293 | p35S::HYG::tNOS | pMpaTUa1::AtCas9::t35S | pMpaU6-1::sgRNACNGC3-3a | pMpaU6-1::sgRNACNGC3-3b | p35S::mCherry:tNOS | Genome editing: *MpaCNGC3* |
| AC294 | p35S::HYG::tNOS | pMpaTUa1::AtCas9::t35S | pMpaU6-1::sgRNACNGC3-4a | pMpaU6-1::sgRNACNGC3-4b | p35S::mCherry:tNOS | Genome editing: *MpaCNGC3* |
| AC295 | p35S::HYG::tNOS | pMpaTUa1::AtCas9::t35S | pMpaU6-1::sgRNACNGC3-5a | pMpaU6-1::sgRNACNGC3-5b | p35S::mCherry:tNOS | Genome editing: *MpaCNGC3* |
| AC296 | p35S::HYG::tNOS | pMpaTUa1::AtCas9::t35S | pMpaU6-1::sgRNACNGC3-6a | pMpaU6-1::sgRNACNGC3-6b | p35S::mCherry:tNOS | Genome editing: *MpaCNGC3* |
| AC305 | pAtUbi10::mCherry::tNOS | | | | | *M. truncatula* complementation |
| AC306 | pAtUbi10::mCherry::tNOS | pLjUbi1::MpaCNGC1::t35S | | | | *M. truncatula* complementation |
| AC307 | pAtUbi10::mCherry::tNOS | pLjUbi1::MpaCNGC3::t35S | | | | *M. truncatula* complementation |

**Table 2.4. Design of GoldenGate level 2 binary expression vectors.**

## 2.6 *Agrobacterium tumefaciens* mediated transformation of *M. paleacea*

Chemically competent *Agrobacterium tumefaciens* AGL1 (Lazo et al., 1991) was prepared as described in (Lin, 1995) and transformed with level 2 plasmids by mixing 200-400 ng of plasmid with 20 µL of competent cells in an eppendorf tube. This mixture was added to an electroporation cuvette and electroporated using the Gene-Pulser (BioRad) with resistance at 200 Ω, voltage at 1.25 V and capacitance at 25 µFd. 500 µ of SOC media was added and the agrobacterium suspension was incubated at 30 at 28ºC for 1 hour at 220 rpm. The cells were spread on to LB plates with the appropriate antibiotics and were incubated for 2-3 days at 28ºC. Colony PCR to confirm the presence of the plasmid of interest in positive clones was performed as described in Section 2.4.

For the *M. paleacea* transformation, thalli were grown for 6 weeks under control conditions in a Percival CU36L4-LED growth cabinet (12/12-hour day-night cycle at $22^0$C, light intensity of 100 µmol/$m^2$s, 50% relative humidity). Transformation of *M. paleacea* was performed according to (Kubota et al., 2013), with the following modification; approximately 2g of *M. paleacea* thalli were added to 100 mL of liquid

Gamborg's half-strength B5 with 1% sucrose liquid medium (Table 2.1) and blended for 30 seconds at 6000 rpm, the co-culture was performed in sterile vented 50 mL falcon tubes containing ½ Gamborg's B5 medium with 1% sucrose and 100 mM acetosyringone, shaken at 30 rpm grown under continuous light (60 μmol/$m^2$s) at 22ºC for four days. The plant co-culture was washed twice with sterile water and transferred to plates containing ½ Gamborg's B5 medium without sucrose, with hygromycin 10 μg/mL and cefotaxime 100 μg/mL.

The plants were grown under continuous light (60 μmol/$m^2$s) at 22ºC for 3-4 weeks at which point putative transformants could be observed. Transformation efficiency was measured using a DMR/MZFLIII microscope (Leica Microsystems, UK) to visualise the transformation marker mCherry included in the L2 vectors transformed (Table 2.4).

## 2.7 *Agrobacterium rhizogenes* mediated transformation of *M. truncatula*

Electro-competent *A. rhizogenes* AR1193 were transformed with the GoldenGate level 2 vector as follows; 20 μL of competent cells were mixed with 200-400 ng of plasmid and incubated on ice for 10 minutes. The cell mixtures were then transformed by electroporation as described. 100 μL of cell suspension was spread on LB plates containing the appropriate antibiotics and grown at 28ºC for 2-3 days. Colony PCR to confirm the presence of the plasmid of interest in positive clones was performed as described in Section 2.4.

For *M. truncatula* transformation, positive colonies were grown in liquid LB medium (Table 2.2) with antibiotics for 2 days, spun down and resuspended in water to an OD600 of 2. The *A. rhizogenes*-mediated transformation was performed according to (Boisson-Dernier et al., 2001) with the following modifications. *M. truncatula* seedlings were germinated overnight as described in Section 2.3 and approximately 3

mm of the root tips were cut off. The cut root of the seedling was dipped into the cell suspension of *A. rhizogenes* AR1193 carrying the appropriate GoldenGate level 2 vector. After dipping, seedlings were placed onto modFP plates (Table 2.1) and grown in controlled environment as described in Section 2.3. Four weeks after transformation, plants were screened for transformed roots using a DMR/MZFLIII microscope (Leica) to visualise the transformation marker mCherry. Plants with transformed roots were used for mycorrhization or nodulation assays.

## 2.8 Mycorrhization assays

### 2.8.1 Inoculation of *M. paleacea*

*M. paleacea* gemmae were grown in vitro on 1% agar ½ B5 medium without sucrose for 4 weeks in a Snijders ECDO1E growth cabinet at 22°C, continuous light and 50% relative humidity. Thalli were then removed from the medium and placed in 1L vented containers with 250 mL soil comprising 20% chive root inoculum prepared as described in Section 2.3.3, as well as 72% sand, 8% John Innes No. 1 compost, or without inoculum in pots containing 9% terragreen, 81% sand, 10% John Innes No. 1 compost for non-mycorrhized controls. Plants were removed and cleaned at 42 dpi (gene expression assay) and 56 dpi (mutant phenotyping mycorrhization assay).

### 2.8.2 Inoculation of *M. truncatula*

*M. truncatula* plants producing transformed roots were selected for arbuscular mycorrhization assays. Before inoculation the untransformed roots were cut off the root system. Plants were transferred to a soil mix comprising 8-% 1:1 terragreen and sand mix (Table 2.2) and 20% chive root inoculum. Plants were grown for 5 weeks in

controlled environment (16/8-hour day-night cycle at $22^0$C, light intensity of 300 µmol/$m^2$s).

### 2.8.3 Ink staining and quantification of fungal infection structures in *M. truncatula*

*M. truncatula* roots were washed in water and pre-cleared by incubation in 10% (w/v) KOH at 95 ºC for 18 min. Roots were rinsed 3 times in water and stained at 95ºC for 7 minutes in a solution of 5% Waterman black ink and 5% acetic acid (Vierheilig et al., 1998). Roots were de-stained in water for 2 days. Mycorrhizal infection structures were quantified using the gridline intersect method (Giovannetti and Mosse, 1980) by using a DFC420 colour camera (Leica).

### 2.8.4 Aniline blue staining of fungal infection structures in *M. paleacea*

*M. paleacea* thalli were cleared for 24 hours in 10% KOH solution and subsequently washed three times in water. The thalli were bleached in a solution of hydrogen peroxide (7.5%) and ammonium hydroxide (0.7%) for 3 minutes, washed several times in water and then acidified in 5% lactic acid for 1 hour. The fungal structures were then stained in a solution of 0.01% aniline blue in lactic acid for 24 hours, per (Upson et al., 2007). The thalli were destained for at least 24 hours in 80% lactic acid, with at least 3 changes of destaining solution during that time. Fungal colonization was imaged using the Leica DM6000 light microscope.

## 2.9 Nodulation assay

*M. truncatula* transformed plants were transferred to a 1:1 mix of terragreen and sand and grown in controlled environment (Section 2.3) for 1 week before inoculation. For the inoculation, *Sinorhizobium meliloti* strain 2011 was grown in liquid YT medium

for 2 days at 28ºC. The culture was spun down by centrifugation (10 minutes at 4000 rpm) and resuspended in water to a final OD600 of 0.01. Each plant was inoculated with 5 mL of *S. meliloti* suspension.

## 2.10 Histochemical GUS staining of *M. paleacea*

To visualize GUS activity, *M. paleacea* tissue were incubated in a solution containing 50 mM sodium phosphate buffer, 0.5 mM $K_3Fe(CN)_6$, 0.5 mM $K_4Fe(CN)_6$, 0.3% (v/v) Triton$^{TM}$ X-100, supplemented with 0.6 mg/mL β-$_D$-glucopyranosiduronic acid (X-gluc) dissolved in N,N-dimethylformamide. The tissues were vacuum infiltrated in the staining solution for 10 minutes and incubated 1h, 6h or 24h at 37 ºC in the dark. The reaction was stopped by removing the staining solution and incubating the tissues in 70% ethanol. Ethanol was exchanged until the tissues were cleared. GUS staining was imaged using the Leica M205FA stereoscope or the Leica DM6000 light microscope.

## 2.11 Quantification of gene expression

### 2.11.1 RNA isolation

RNA extraction was carried out using the RNeasy mini plant kit (Qiagen, USA) following the manufacturer's protocols, including on column DNase I treatment. Samples were prepared by washing soil particles (where necessary) from plants with dH2O, placing on tissue paper to absorb excess water, flash freezing and grinding 100 mg samples in liquid nitrogen. RNA quality was assessed by gel electrophoresis and nanodrop analysis. PCR screening for genomic DNA contamination was carried out in a 20 μL reaction comprised of 10 μL 2X GoTaq® Green Master Mix, 1 μL 10 nM forward primer, 1 μL 10 nM reverse primer and 7.5 μL water and 0.5 μL RNA extraction. Primers used were as described in Table 23.

### 2.11.2 Quantitative real-time PCR (qRT-PCR)

cDNA synthesis was performed using the Bio-Rad iScript cDNA Synthesis Kit. Quantitative real time PCR (qPCR) was performed using the Bio-Rad one-step RT-PCR SYBR Green kit. Values were normalized by expression of the *M. paleacea* housekeeping genes *MpaAPT* and *MpaACT*, as orthologues of these genes in *M. polymorpha* are stably expressed across plant development, hormone treatments, nutrient stress and cold stress (Saint-Marcoux et al., 2015).

## 2.12 Screening of *M. paleacea* mutant lines

Regenerating plants transformed with CRISPR/Cas9 genome editing vectors (as seen in Table 2.4) were to screened to visualise the transformation marker mCherry using a DMR/MZFLIII microscope (Leica Microsystems, UK). Genomic DNA was extracted by placing approximately 10 mg of tissue in 100 µL quick DNA extraction buffer containing 1M KCl, 100 mM Tris-HCl (pH 7.5), 10 mM EDTA. This was incubated at 95°C for 5 minutes, followed by cooling to 4°C. 400 µL of water was added to the solution and 1 µL was used as a template for PCR analysis. PCR reactions were carried out in a 20 µL reaction comprised of 10 µL 2X GoTaq® Green Master Mix, 1 µL 10 nM forward primer, 1 µL 10 nM reverse primer, 1 µL template and 7 µL water. Cycling conditions were 2 minutes at 95°C, 30 cycles of 30 seconds at 95 °C, 30 seconds at the appropriate annealing temperature, 1 minute at 72°C, followed by an extension of 5 minutes at 72°C and incubation at 4°C. Primers and annealing temperatures used were as described in Table 2.3.

PCR products were visualized by gel electrophoresis and mutants were identified based on the presence of amplicon size that was smaller than that expected from wild

type genomic DNA by the length predicted by the two sgRNAs used in CRISPR/Cas9 genome editing. Deletions predicted by the CRISPR target sites were confirmed by Sanger sequencing performed by Eurofins Genomics (https://www.eurofinsgenomics.eu/).

# Chapter III - Phylogenetic and protein sequence analyses of putative symbiosis genes in *M. paleacea*

## 3.1 Introduction

Phylogenetics and phylogenomics have been useful tools for studying the molecular mechanisms of AM symbiosis. Phylogenomics indicates that the loss of AM symbiosis correlates with the loss of a specific set of genes with symbiotic function in independent losses of AM symbiosis (Delaux et al., 2014). These phylogenetic patterns have been used to develop pipelines to identify genes with symbiotic function (Bravo et al., 2016; Favre et al., 2014). Genes identified by these pipelines have been further investigated by molecular biology methods, leading to the discovery of *FatM*, an acyl-ACP thioesterase that functions specifically in fatty acid biosynthesis in arbuscule containing cells (Bravo et al., 2017; Luginbuehl et al., 2017).

Phylogenomics of root nodule symbiosis in angiosperms has been instrumental in understanding when and how this trait was evolved from AM symbiosis (Griesmann et al., 2018). Phylogenomics of basal land plants and green algae has additionally been used to show that a number genes that function specifically in AM symbiosis or root nodule symbiosis in angiosperms are also present in charophyte algae, at least one of which, *CCaMK* is functionally conserved between angiosperms and charophyte algae closely related to the land plants (Delaux et al., 2015). Despite this, fungal symbionts

have not been documented in any charophyte algae and AM symbiosis is thought to have evolved in the earliest land plants (Redecker et al., 2000; Rubinstein et al., 2010). Many symbiosis specific genes are found in the basal land plants, including the liverworts (Delaux et al., 2015). Due to the nature of liverwort genomes, these represent an excellent potential model for studying the molecular mechanisms of AM symbiosis. Unlike most land plant lineages, liverworts have not undergone the same repeated whole genome duplications (WGDs) and show little evidence of polyploidy (Berrie, 1960; Fritsch, 1991). Recent analysis indicates that most regulatory genes exist as single paralogues in the *Marchantia polymorpha* genome, although some gene families, such as phosphate and ammonium transporters, are relatively over represented (Bowman et al., 2017).

In contrast, all angiosperms are paleopolyploids and WGDs occur frequently in the angiosperm lineage (Albert et al., 2013; Bowers et al., 2003; Cui et al., 2006; De Bodt et al., 2005). Model plant species such as *Arabidopsis thaliana*, rice (*Oryza sativa*) and *Medicago truncatula* have genomes that indicate that multiple whole genome duplications took place in these lineages since the evolution of the angiosperms (Ren et al., 2018). These WGDs lead to both neofunctionalization and genetic redundancy of duplicate genes, leading to difficulty in identifying some molecular mechanisms by forward genetic screens in angiosperms (Wang et al., 2011).

Using liverworts as a model for plant molecular genetics presents two major advantages. First, the simpler genome structure seen in *M. polymorpha* should result in reduced functional redundancy, and observable phenotypes in single mutants where double, triple or even higher orders of mutants are necessary to see phenotypes in orthologous angiosperm genes. Secondly, where a single paralogue is retained, this portion of a liverwort genome more closely resembles the predicted genome of the

common ancestor of the land plants than other model species (Bowman et al., 2017). The major alternative bryophyte model species, the moss, *Physcomitrella patens*, has undergone a whole genome duplication 30-60 MYA and polyploidy is extremely common in the moss lineage (Rensing et al., 2007).

*M. polymorpha* is not used as model plant to study AM symbiosis in this project as the symbiosis is absent from the subspecies *M. polymorpha ruderalis* (the subspecies developed as a model organism) and *M. polymorpha polymorpha*, although fungal colonization of parenchymous cells has been documented in *M. polymorpha montivagans* (Bowman et al., 2016a; Fonseca et al., 2006; Ligrone et al., 2007). This suggests that the ability to form AM symbiosis has recently been lost in these subspecies. This project uses *M. paleacea* as a model organism, as it has documented functional AM symbiosis and is closely related to *M. polymorpha* (Humphreys et al., 2010). *M. polymorpha* and *M. paleacea* diverged approximately 44 MYA but the genomes are relatively similar due to the low rate of DNA substitution in the Marchantiopsida class (Villarreal et al., 2016). This means that in addition to the advantages of working with a liverwort genome, there is also the potential for genomic comparisons between closely related species, one of which retains AM symbiosis, while the other has undergone recent loss of this association. Evidence of recent gene loss in *M. polymorpha* is documented in the PhD thesis of (Radhakrishnan, 2017).

The aim of the research presented in this chapter was to identify putative symbiosis genes in *M. paleacea* and assess their relationship to known symbiosis genes documented in angiosperm species. Putative symbiosis genes with roles in arbuscule development, nutrient transport and symbiosis signalling could be identified by orthology to those genes studied in angiosperms. A putative symbiotic phosphate transporter and signalling $Ca^{2+}$ channel could not be identified based on orthology

with angiosperm genes but candidate symbiotic genes could be identified from gene sequence features and comparison to putative symbiosis genes identified in other liverwort species. Further symbiosis genes in *M. paleacea* could be identified using a similar approach.

## 3.2 Results

### 3.2.1 *Marchantia paleacea* genome sequencing

The genomic DNA extracted from axenic culture was sequenced to generate 218 million 100bp paired-end reads with an average insert size of 336 bp, and 74 million 100bp mate-pair reads with an average insert size of 4311bp. These reads were assembled by a colleague (Radhakrishnan, 2017) to generate an assembly 238.61 Mb in size, comprising 22,669 scaffolds, with a 40.3% G+C content. The N50 length of this genome was 77.78 kb, in contrast to the recently released version 2.0 *M. polymorpha* genome, where an N50 of 256.9 kb was achieved (Bowman et al., 2017).

| ***M. truncatula*** | ***M. paleacea*** | **Phylogenetic relationship** |
|---|---|---|
| *MtDMI1* | *MpaDMI1* | Single *M. paleacea* gene with multiple co-orthologues in *M. truncatula* |
| *MtCNGC15s* | *MpaCNGC3* | Multiple co-orthologues in both *M. truncatula* and *M. paleacea* |
| *MtVAPYRIN* | *MpaVAPYRIN* | Direct orthologue present |
| *MtEXO70I* | *MpaEXO70GI* | Single *M. paleacea* gene with multiple co-orthologues in *M. truncatula* |
| *MtPT4* | *MpaPHT1D* | Gene family present, no orthologue to Mt Sym gene |
| *MtSTR* | *MpaSTR* | Direct orthologue present |

Table 3.1 | Summary of findings for the gene families investigated in Chapter III

### 3.2.2 Identification of putative symbiosis genes by phylogeny: *DMI1*

Putative symbiosis genes were identified by Basic Local Alignment Search Tool (BLAST) analysis of the *M. paleacea* genome using identified symbiosis genes from the *M. truncatula* as a query and CLC Main workbench software. The *M. paleacea* transcriptome, publicly available at the 1000 Plants project online portal (https://sites.google.com/a/ualberta.ca/onekp/) was also BLAST searched using the same queries in case of gaps in our genome assembly (Johnson et al., 2012; Matasci et al., 2014; Wickett et al., 2014; Xie et al., 2014). Amino acid sequences were aligned by CLUSTAL to representative amino acid sequences of the same genome family from sequenced genomes of land plants and chlorophyte algae, see Appendix 1 (Larkin et al., 2007). Where no genome sequences are available, or poor sequenced genome quality resulted in no hits in sequences genomes, transcriptome data was used. This was essential for the charophyte algae, for which no published genome sequences are available apart from the basal charophyte, *Klebsormidium nitens* (Hori et al., 2014).

Two genes involved in symbiosis signalling were analysed phylogenetically; the potassium permeable cation channel *DMI1* and the cyclic nucleotide gated $Ca^{2+}$ channel, *CNGC15*. DMI1 and CNGC15 form a complex in the membrane of the nuclear envelope and generate the nuclear calcium oscillations characteristic of symbiosis signalling (Ane et al., 2004; Charpentier et al., 2016; Kistner et al., 2005). Although these proteins interact and function in generating the same signal, these genes have contrasting patterns of evolution.

### 3.2.2.1 Calcium signalling: *DMI1*

*DMI1* is present as a single copy in *M. paleacea* and as a single copy and a second sequence homologous to a short fragment (named *MapoDMI1f*, reference to published genome in Appendix 2) of the cytoplasmic C-terminal domain in *M. polymorpha*, which could not be used in phylogenetic analysis due to its short length. Although multiple copies of *DMI1* are found in other clades of land plants, a single paralogue is likely to be representative of the common ancestor of the land plants. *DMI1* was duplicated independently in the lycophytes (as seen in *Selaginella moellendorfii*), the mosses, and in the common ancestor of angiosperms and gymnosperms, as seen in Figure 3.1. and the table of DMI1 sequences used in Appendix 2. In addition, a POLLUX-like cation channel of unknown function is also found in land plants and charophytes. Although *DMI1* has been duplicated twice in mosses, once early in the evolution of this clade and again more recently in *P. patens*, *S. fallax* contains only one full length gene, *SfDMI1a*, while *SfDMI1c* and both *P. patens DMI1* paralogues form a clade of truncated genes lacking all transmembrane domains. *DMI1* is also present in the charophyte algae and *DMI1* like sequences are present in chlorophyte algae, consistent with (Delaux et al., 2015).

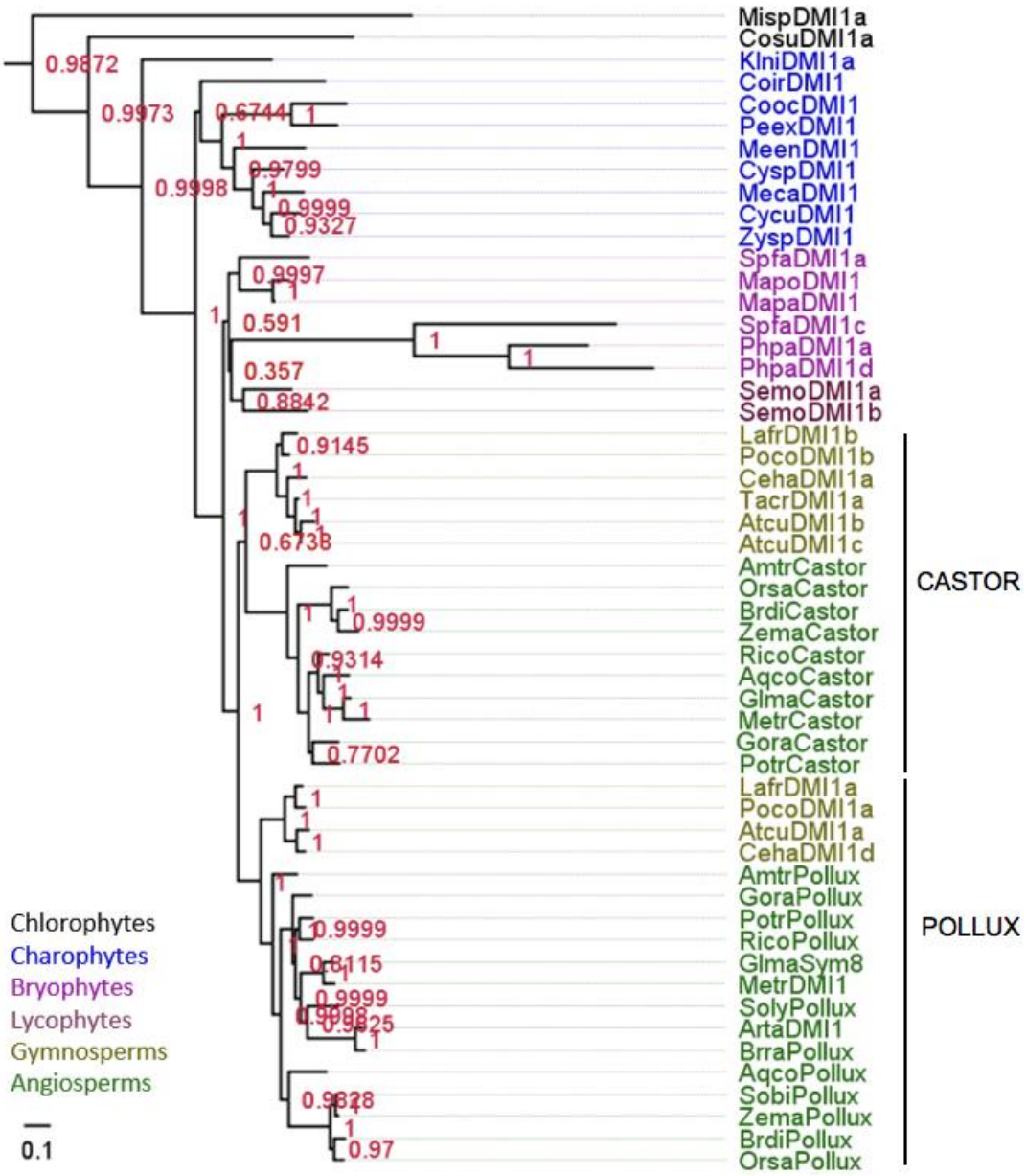



Figure 3.1 | Maximum likelihood phylogeny of *DMI1* type cation channels in the green lineage.

Maximum likelihood phylogeny of *DMI1* in the green lineage, rooted by *Pollux-Like* sequences from the green lineage. *DMI1* was likely present as a single copy in the common ancestor of the land plants. This gene underwent duplication to form the *Castor* and *Pollux* clades in the common ancestor of the angiosperms and the gymnosperms. Both *Castor* and *Pollux* are required for symbiosis signalling in the angiosperms, apart from in the Vicoid legumes, represented here by “Metr”, Medicago truntactula. The model lycophyte *S. moellendorfii* contains two copies of *DMI1, SemoDMI1a* and *SemoDMI1b* from a duplication specific to the lycophyte lineage. Scale bar represents substitutions per site. Key for abbreviations of species used are in Table 3.5.

Both *CASTOR* and *POLLUX* are necessary for symbiotic signalling in model angiosperms such as *L. japonica* and rice (*Oryza sativa*) but in *M. truncatula*, only the *POLLUX* homologue is required (Ane et al., 2004; Banba et al., 2008; Charpentier et al., 2008; Chen et al., 2009; Gutjahr et al., 2008; Imaizumi-Anraku et al., 2005; Peiter et al., 2007; Venkateshwaran et al., 2012). This change in function is associated with an amino acid substitution in the selectivity filter of the cation channel, the narrowest part of the channel pore. The selectivity filter determines the permeability of the channel to specific ions, leading to alterations in function. Two channels, *CASTOR* and *POLLUX* are required to generate calcium oscillations in most angiosperms and these have a selectivity filter composed of the amino acids ADSGNH, however a single cation channel, *MtDMI1*, with a selectivity filter composed of ADAGNH seen, is sufficient to generate symbiotic calcium oscillations in *M. truncatula* (Venkateshwaran et al., 2012). (Delaux et al., 2015) observed that the ancestral state in the common ancestor of the land plants was likely a single *DMI1*-like cation channel, as the ADAGNH type selectivity filter is seen in liverworts and in most charophytes (Table 3.2).

This analysis found that the selectively filter changed from ADAGNH to ADSGNH on two separate occasions where there was a duplication of *DMI1* in the land plant lineage. As well as the seed plants, which contain the two paralogues *CASTOR* and *POLLUX*, the lycophyte *S. moellendorfii* also contains two paralogues of *DMI1* from a separate duplication event (Figure 3.1), both of which also have the ADSGNH selectivity filter (Table 3.2).

| Clade | Species | Gene name | Selection filter |
|---|---|---|---|
| Chlorophyte | *Ostreococcus lucimarinus* | *OlDMI1* | LGADSTA |
| Chlorophyte | *Coccomyxa subellipsoidea* | *CosuDMI1* | VTFAGEA |
| Charophyte | *Klebsorbidium nitens* | *KnDMI1a* | ADAGNH |
| Charophyte | *Penium exiguum* | *PeDMI1* | ADAGNH |
| Charophyte | *Coleochaete irregularis* | *CiDMI1* | ADAGNH |
| Bryophyte | *Marchantia paleacea* | *MpaDMI1* | ADAGNH |
| Bryophyte | *Marchantia polymorphha* | *MpDMI1* | ADAGNH |
| Bryophyte | *Sphagnum fallax* | *SfDMI1a* | ADSGNH |
| Bryophyte | *Sphagnum fallax* | *SfDMI1c* | Absent |
| Lycophyte | *Selaginella moellendorfii* | *SmDMI1a* | ADSGNH |
| Lycophyte | *Selaginella moellendorfii* | *SmDMI1b* | ADSGNH |
| Angiosperm | *Oryza sativa* | *OsPOLLUX* | ADSGNH |
| Angiosperm | *Oryza sativa* | *OsCASTOR* | ADSGNH |
| Angiosperm | *Arabidopsis thaliana* | *AtDMI1* | ADSGSH |
| Angiosperm | *Medicago truncatula* | *MtDMI1* | ADAGNH |
| Angiosperm | *Medicago truncatula* | *MtCASTOR* | ADSGNH |

Table 3.2. Amino acid sequence of the selectivity filter of DMI1.

The selectivity filter of DMI1 affects the conductance of the channel when activated. Most angiosperms have a serine in the third amino acid position, while an alanine is found in most charophytes and in liverworts

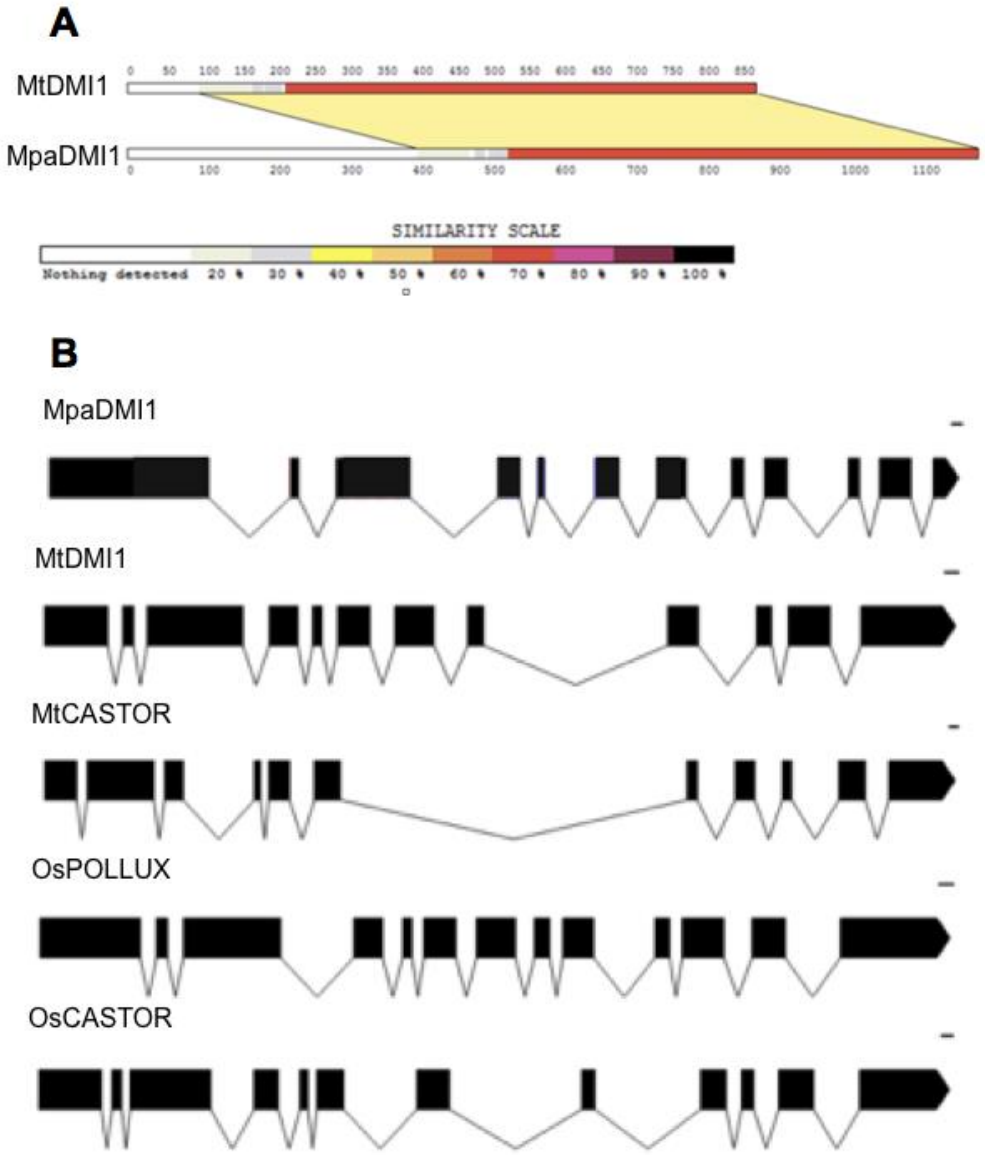



Figure 3.2 | Sequence analysis of DMI1 in the land plant lineage, showing conservation of amino acid sequence and intron exon structure.

A. Graphical view of protein sequence similarity generated by LALNVIEW (Duret et al., 1996). The C terminal region of DMI1 is highly similar between *M. truncatula* and *M. paleacea*, while the N terminal region is highly divergent. B. In the angiosperms *M. truncatula* and rice, the intron-exon structure of *DMI1*/*POLLUX* and *CASTOR* are similar but not identical, with 12 exons in *MtDMI1* and 13 in *OsPOLLUX*, and 11 exons in *MtCASTOR* and 12 in *OsCASTOR* *MpaDMI1* is also made up of 12 exons. Scale bar represents 100 bp.

Although MtDMI1 and MpaDMI1 have 76% protein sequence identity and 87% similarity from the second transmembrane domain onwards, the N terminal region of these proteins is highly divergent and 305 amino acids longer in MpaDMI1 (Figure 3.2 A). Intron-exon structure is also conserved between *MpaDMI1* and *MtDMI1* (Figure 3.2 B suggesting conservation in transcriptional regulation of this gene (Heyn et al., 2015; Li, Chen, and Zhang, 2012).

### 3.2.3 Identification of putative symbiosis genes by phylogeny: *CNGCs*

The second member of the cation complex that generates symbiotic calcium oscillations, CNGC15, is a member of the Group III cyclic nucleotide gated cation channels. This subsection of a large gene family in the angiosperms also contains *AtCNGC14*, *AtCNGC16*, *AtCNGC17* and *AtCNGC18* in *A. thaliana*. In contrast to the angiosperms, basal land plants such as bryophytes and lycophytes do not have Group III CNGCs, but CNGCs from these plants form a sister group to a clade containing the angiosperm Group I, Group II and Group III CNGCs (Figure 3.3). In *M. paleacea* there are two CNGCs in this sister group, *MpaCNGC1 and MpaCNGC3*. The phylogeny in Figure 3.3 indicates that the Group I/II/III CNGCs were represented by a single gene in the common ancestor of the land plants. The Group IVa and Group IVb CNGCs appear to be ancient gene subfamilies, as they are represented by charophytes, bryophytes, lycophytes and angiosperms.

The Group I/II/III clade of CNGCs may represent a more recent evolution, as there is not a clearly defined charophyte CNGCs group in this clade. Instead, charophyte CNGCs closest to Group I/II/III diverge at several poorly resolved nodes (node support shown in the extended tree in figure A.1). However, the amino acid sequence of the selection filter, GN[GNST], is conserved between many charophyte Group

I/II/III like CNGCs and angiosperm Group I/II/III CNGCs (Table 3.3). The selection filter of CNGCs is a triplet of amino acids present in the pore region, identified on the basis of homology to the bacterial (*Streptomyces lividans*) $K^+$-selective channel KcsA (Zelman et al., 2012). This suggests that the relative conductance of this group of CNGCs to $Na^+$, $K^+$, and $Ca^{2+}$ ions is similar in both charophyte algae and land plants. None of the charophyte algae transcriptomes in this analysis contained all three subfamilies of CNGC, with Group IVb and either Group IVa (in most Zygnematophyceae) or Group I/II/III (Coleochaetophyceae, Klebsorbidiophyceae and *Spirogyra sp*.) represented (examples in Table 3.3).

| Clade | Species | Group I/II/III | Group IVa | Group IVb |
|---|---|---|---|---|
| Charophyte | *C. irregularis* | 2 **(GQN, GNA)** | - | 1 (GNA) |
| Charophyte | *Spirotaena sp.* | 3 **(GQG)** | - | 2 (ANA) |
| Charophyte | *P. exiguum* | - | 2 (AGN) | 2 (GNA) |
| Charophyte | *C. cushleckae* | - | 1 (AGN) | 1 (GNA) |
| Bryophyte | *M. paleacea* | 2 (**GQG, GQT**) | 1 (AGN) | 1 (GNA) |
| Bryophyte | *P. patens* | 4 (**SQT, TQT**, AQT) | 1 (**GGN**) | 3 (GNA) |
| Lycophyte | *S. moellendorfii* | 4 (**GQN**, **GQG**) | 4 (AGN, **GGN**) | 4 (GNA, GNL) |
| Angiosperm | *O. sativa* | 11 (**GQN, GQG, GQT,** GNP) | 2 (AGN) | 3 (GND, GNL) |
| Angiosperm | *A. thaliana* | 16 (**GQN, GQG, GQS**) | 2 (AGN) | 2 (AND, GNL) |
| Angiosperm | *M. truncatula* | 12 (**GQN, GQG, GQS**, GEN) | 6 (AGN, AGG) | 3 (GND, GNL) |

Table 3.3 | Amino acid sequence of the selectivity filter in the CNGC gene family.

The amino acid triplet defining the selectivity filter was identified by homology to the known structure of the *Streptomyces lividans* K+-selective channel (PDB 1BL8) (Kaplan et al., 2007).

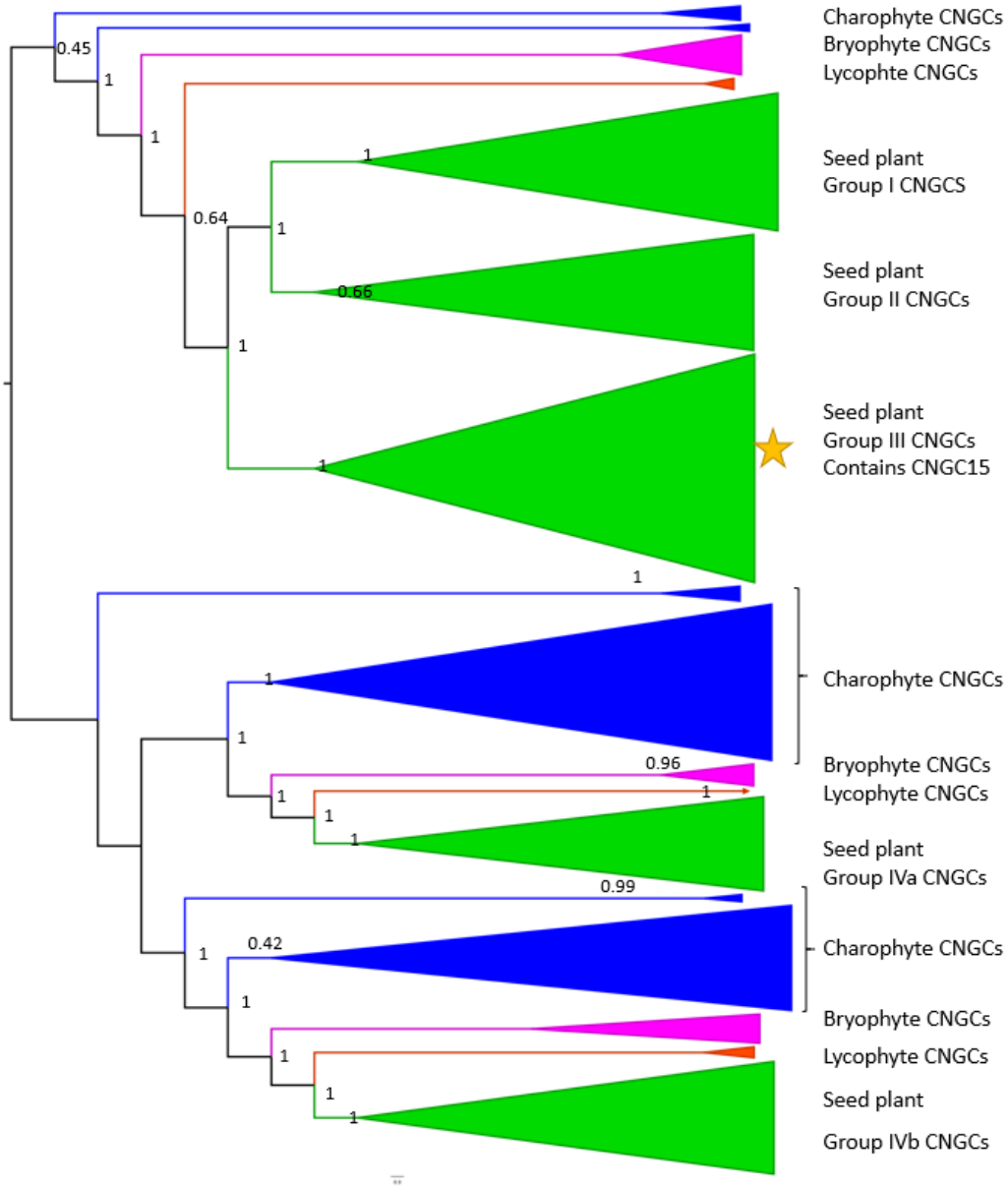



Figure 3.3 | Maximum likelihood cladogram of the CNGCs in the Streptophytes

A maximum likelihood cladogram of streptophyte CNGCs, rooted with the chlorophyte CNGC like sequence, *VcCNGC*. In the land plants, the CNGCs fall into three major clades; the Group IVa CNGCs, the Group IVb CNGCs and a clade containing angiosperm and gymnosperm Groups I, II and III CNGCs as well as a single orthologous clade of bryophyte CNGCs and a single orthologous clade of lycophyte CNGCs. Stars indicate the presence of symbiosis associated CNGCs. Branch support by aBayes analysis.

BLAST searches of the publicly available chlorophyte genomes *Volvox carteri*, *Chlamydomonas reinhardtii*, *Dunaliella salina*, *Coccomyxa subellipsoidea*, *Micromonas pusilla*, *Micromonas sp.* and *Ostreococcus lucimarinus* found no ion channels with both CNG binding and CaM binding domains characteristic of CNGCs (Zelman et al., 2012).

As the Group I/II/III CNGCs expanded independently in the angiosperm and bryophyte lineages, there is no direct CNGC15 orthologue in the basal land plants. The Group I/II/III CNGCs, MpaCNGC1 and MpaCNGC3, both contain the cyclic nucleotide binding and calmodulin binding domains seen in MtCNGC15a/b/c. However, MpaCNGC3 has higher protein sequence identity to MtCNGC15a/b/c (56%, 65% and 63%, respectively) than MpaCNGC1 (46%, 55% and 52%, respectively) (Figure 3.4). MpaCNGC3 contains a predicted nuclear localisation signal (NLS), as is also seen in MtCNGC15a/b/c, however, putative NLS sequences are seen in all members of the CNGC family in *M. paleacea* (Table 3.4), so either of the Group I/II/III CNGCs, MpaCNGC1 or MpaCNGC3 could localise to the nuclear envelope, as well as the more divergent Group IVa CNGC, MpaCNGC4 or the Group IVb CNGC MpaCNGC2.

| Gene | Position | Sequence | Score | Motif match |
|---|---|---|---|---|
| *MpaCNGC1* | 443-451 | RV**KRRD**TEE | 0.829 | yes |
| *MpaCNGC1* | 463-468 | RNRVRR | 0.528 | no |
| *MpaCNGC1* | 659-668 | RRHRSRKLGP | 0.856 | no |
| *MpaCNGC2* | 442-446 | M**RRRQ** | | yes |
| *MpaCNGC2* | 647-661 | RRFKKQRDERKEREK | | no |
| *MpaCNGC3* | 656-665 | RWRRYQ**RRKL** | 0.53 | yes |
| *MpaCNGC3* | 730-733 | PKPS | | no |
| *MpaCNGC4* | 460-475 | FLMALG**RRKL**EMQLKR | 0.815 | yes |
| *MpaCNGC4* | 483-497 | **KRRRL**PVLL**RRRV**RQ | 0.824 | yes |
| *MpaCNGC4* | 692-701 | QKAWRARRAR | | no |

Table 3.4 | Potential nuclear localization signals in the CNGC gene family in *M. paleacea*.

All members of the CNGC family in *M. Paleacea* contain potential NLS sequences. Potential NLS sequences were identified by LOCALIZER Plant and SeqNLS (shaded) (Lin and Hu, 2013; Sperschneider et al., 2017). Motifs that matched the basic canonical motif conserved in *M. truncatula* nuclear localization signals (K/R, K/R, K/R, X) are in bold.

Additionally, the intron-exon structure is conserved between angiosperm *CNGC15* genes and *MpaCNGC3* but not with *MpaCNGC1* (Figure 3.5). Although the length of the entire genes has changed, this is mostly due to changes in the intron lengths and not in the exon lengths. These genes were composed of seven exons and six introns, however, *MpaCNGC1* was composed of a single exon and contained no introns. All together, these analyses suggest that MpaCNGC3 could be the closest functional analogue to MtCNGC15a,b,c.

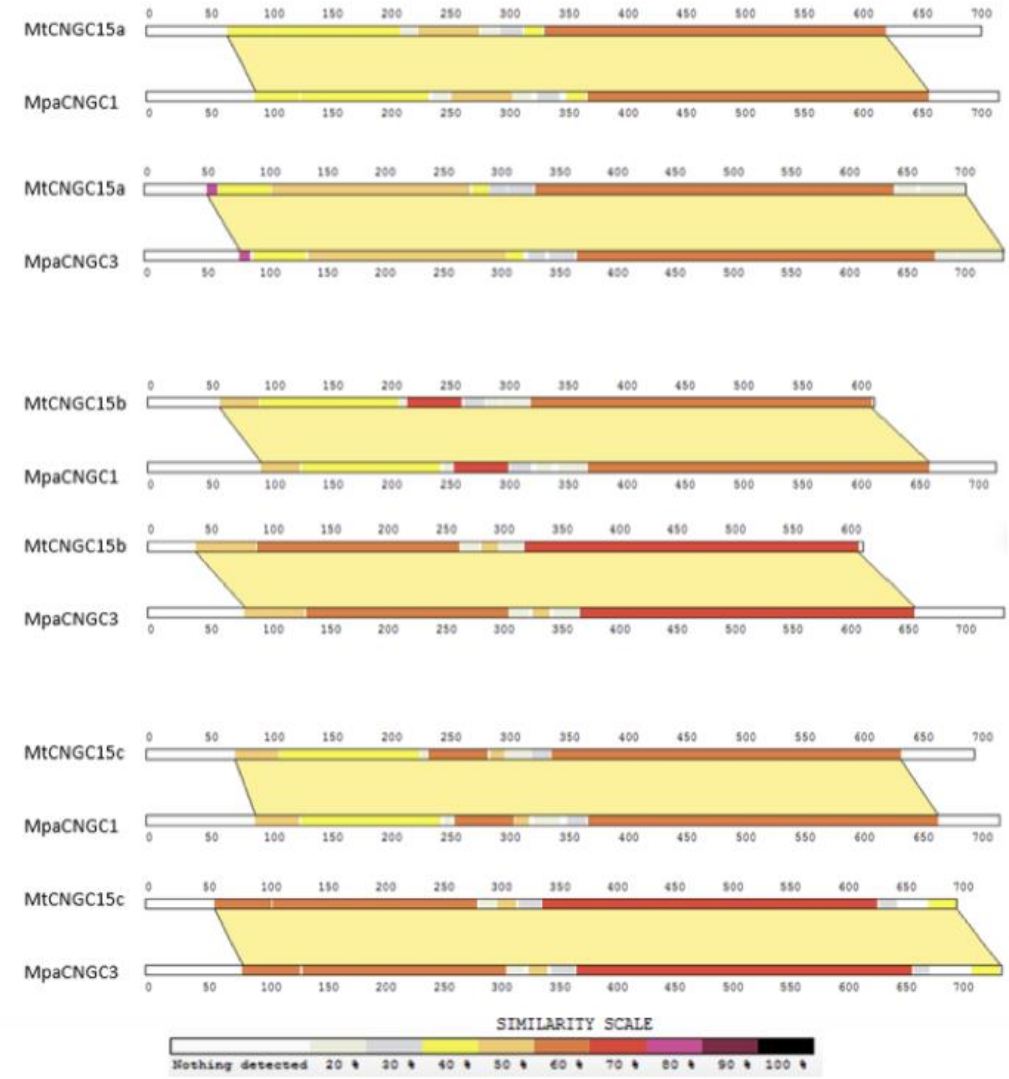



**Figure 3.4 | Amino acid sequence evolution in CNGC15 in the land plant lineage.** MtCNGC15s are more similar to MpaCNGC3 than MpaCNGC1. The greatest sequence similarity is between MtCNGC15c and MpaCNGC3.


**Figure 3.5 | Evolution of intron exon structure in CNGC15 in the land plant lineage**. Intron-exon structure is conserved between *MtCNGC15a*, *OsCNGC15* and *MpaCNGC3*, but not *MpaCNGC1,* despite phylogenetic analysis indicating that *MpaCNGC3* and *MpaCNGC1* are the products of a gene duplication in the liverwort lineage. Scale bar represents 100 bp.

### 3.2.4 Identification of putative symbiosis genes by phylogeny: arbuscule development

The molecular mechanisms of arbuscule development involves the reorientation of secretion in the plant cell to generate the peri-arbuscular membrane. This project focused on two components of this machinery with specific roles in generating the symbiotic membrane interface; the exocyst component Exo70I and VAPYRIN, a

protein that in vascular plants is found only in species forming AM or root nodule symbioses.

Exo70 is one of 8 components of the exocyst complex that functions in vesicle tethering and targeting to the plasma membrane prior to vesicle fusion, necessary for functions as diverse as polarized secretion, neuron development, cytokinesis, cell migration morphogenesis, organogenesis and polarized cell growth (Hala et al., 2008; Li et al., 2013; Lukáš et al., 2006; Rawat et al., 2017; Tu et al., 2015; Zhu et al., 2017). Although *Exo70* is present as a single locus in mammals and in yeast (*Saccharmomyces cerevisiae*), this gene family is expanded in the land plants. While other members of the exocyst complex are present as one or a few loci in land plants, the copy number of *Exo70* is 23 in *A. thaliana* and 47 in rice (Cvrčková et al., 2012). Phylogenetic analysis of the *Exo70* gene family in the green lineage (Figure 3.6) suggests that much of the expansion in the *Exo70* gene family took place in the angiosperm lineage and there are three *Exo70* subfamilies in the land plants, which is consistent with the findings of (Cvrčková et al., 2012; Lukáš et al., 2006). These subfamilies are Exo70.1, which contains the *Exo70A* group in angiosperms, Exo70.3, which contains the angiosperm *Exo70G* and *Exo70I* groups, and Exo70.2, which contains the remaining angiosperm groups. These three ancestral paralogues also appear to be present in at least some charophyte algae, implying that they are not an innovation of the land plants (Figure 3.6). Each of these *Exo70* subfamilies is represented by a single gene in the liverworts *M. polymorpha* and *M. paleacea*, which is likely to be representative of the common ancestor of the land plants.

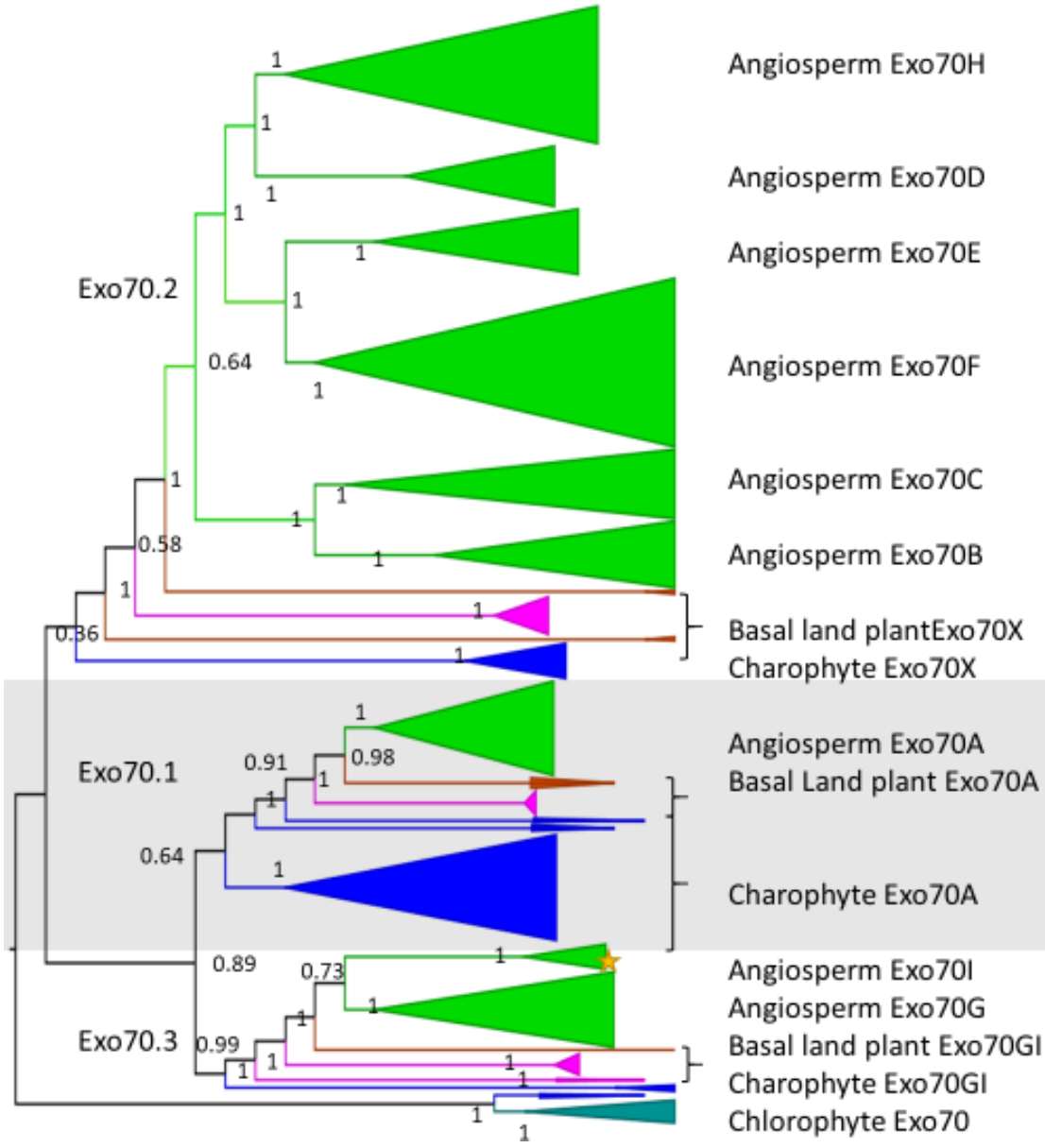



Figure 3.6 | Maximum likelihood cladogram of Exo70 in the green lineage, rooted by chlorophyte Exo70 sequences.

There are three major subfamilies of Exo70 in the streptophytes. The Exo70.2 subfamily has expanded to form Exo70C, Exo70D, Exo70E, Exo70F and Exo70H in the angiosperms. The Exo70.3 gene present in the common ancestor of the land plants likely had both developmental and symbiotic functions. Branch support shown is aBayes analysis. The star indicates the presence of identified symbiotic Exo70 in *M. truncatula*.

In angiosperms, the *Exo70I* group is found only in plants that form AM symbioses, expressed specifically in arbuscule containing tissue and is necessary for the secretion

of the PAM during arbuscule formation, (Zhang et al., 2015). The role of the Exo70G group in angiosperms is unknown, as the *Atexo70g1* mutant has no clear phenotype and no *Atexo70g2* mutants have been described (Lukáš et al., 2006). In the moss, *P. patens*, the mutant *Ppexo70.3d* has pleiotropic defects in cell elongation, cell differentiation, reproductive development, and cell wall and cuticle deposition (Rawat et al., 2017). It is possible that the single *Marchantia sp.* Exo70.3 paralogue has functions in both symbiosis and other aspects of plant development, as it is retained in the *M. polymorpha* genome, and a completely symbiosis specific Exo70 evolved only in the angiosperm lineage.

VAPYRIN is made up of two protein-protein interaction domains, a Major Sperm Protein (MSP) domain and multiple ankyrin repeats. Although the specific function of *Vapyrin* is unknown, it is likely to be involved in protein-protein interactions in the cellular machinery that accommodates both AM fungi and rhizobia (Feddermann et al., 2010; Murray et al., 2011a; Pumplin et al., 2010a). *VAPYRIN* is typically found as a single copy in endosymbiotic plant species, although there are exceptions, as *S. moellendorfii* has two copies of *VAPYRIN* and *M. polymorpha* retains a copy of *VAPYRIN* despite having lost AM symbiosis (Figure 3.7). A paralogue of *VAPYRIN*, VAPYRIN Like (VPYL) with both MSP and ankyrin repeat domains is found in the symbiotic eudicots but also in the mosses *S. fallax* and *P.patens*. A paralogue of *VAPYRIN* lacking a recognizable MSP domain is also found in monocots, *M. paleacea* and *M. polymorpha* (seen in the extended VAPYRIN like phylogeny in Figure A.4). *S. moellendorfii* appears to lack a *VPYL* in either clade. The function of *VPYL* is unknown but its presence in mosses indicates that it may have non-symbiotic functions. This analysis found that *VAPYRIN* is unique to the land plants, with its closest relatives in the charophyte algae containing the ankyrin repeat domain only.

Although *VAPYRIN* is likely to have conserved function in AM symbiosis since the earliest land plants, *VPYL* genes arose at a similar time and are found in moss genomes, where AM symbiosis is not present. As *VAPYRIN* is also found in *M. polymorpha*, this phylogenetic analysis suggests that there may additional non-symbiotic functions of members of the *VAPYRIN* gene family

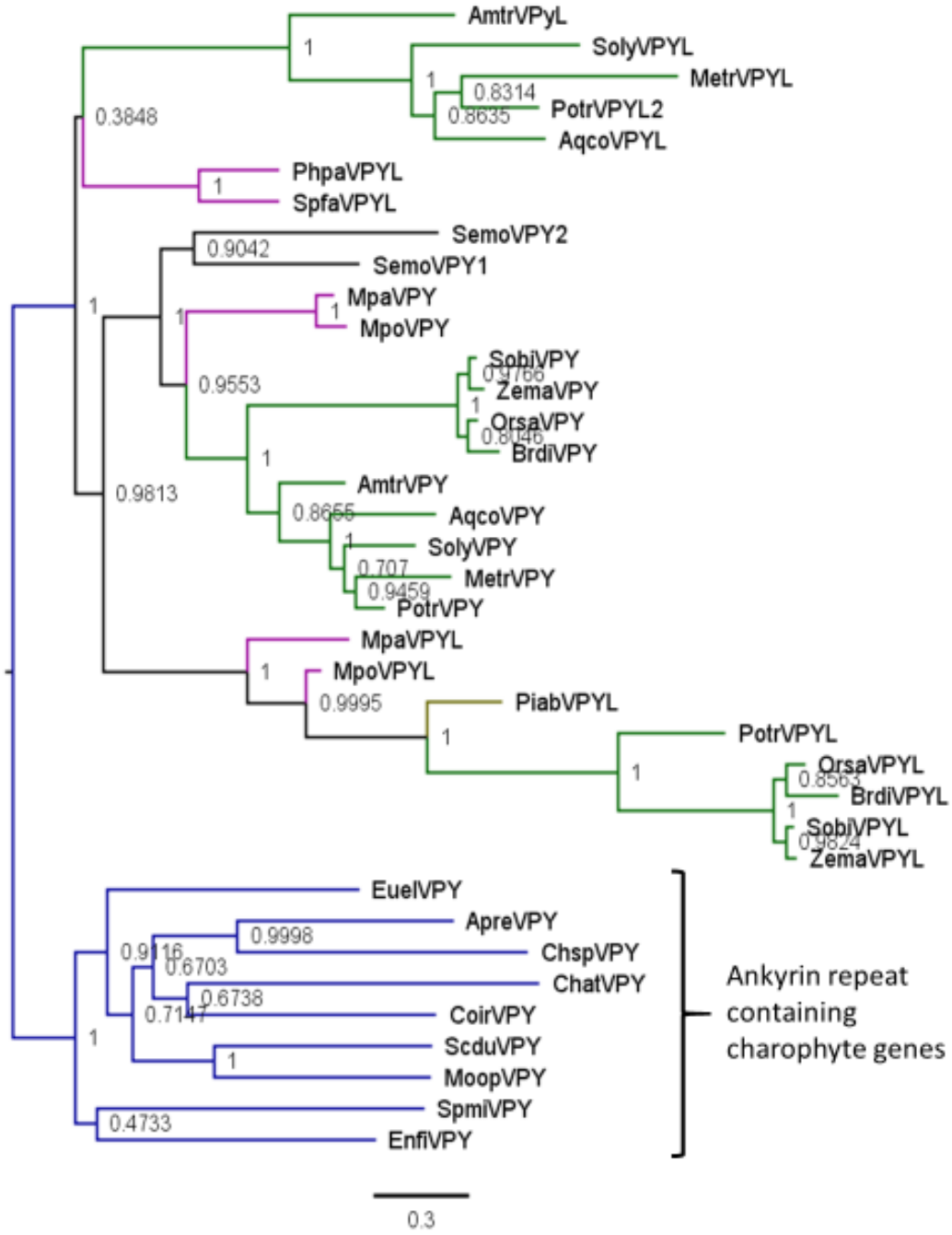



Figure 3.7 | Maximum likelihood phylogeny of VAPYRIN in the green lineage.

*Vapyrin* and *VPYL*, containing both MSP and ankyrin repeat domains, are present only in land plants. Mosses do not contain *Vapyrin* orthologues but do contain *VPYL*, suggesting non-symbiotic functions for this gene. Phylogeny rooted using chlorophyte ankyrin repeat containing genes. Key for abbreviations of species used are in Table 3.5.

| Chlorophyte algae | |
|---|---|
| Euel | *Eudorina elegans* |
| Voca | *Volvox carteri* |
| Apre | *Aphanochaete repens* |
| Chre | *Chlamydomonas reinhardtii* |
| Chsp | *Chlamydomonas sp.* |
| Scdu | *Scherffelia dubia* |
| Moop | *Monomastix opisthostigma* |
| Mipu | *Micromonas pusilla CCMP1545* |
| Misp | *Micromonas sp. RCC299* |
| Cosu | *Coccomyxa subellipsoide* |
| **Charophyte algae** | |
| Klni | *Klebsorbidium nitens* |
| Klsu | *Klebsormidium subtile* |
| Stom | *Staurodesmus omearii* |
| Coir | *Coleochaete irregularis* |
| Chat | *Chlorokybus atmophyticus* |
| Spmi | *Spirotaenia minuta* |
| Enfi | *Entransia fimbriat* |
| Cooc | *Cosmarium ochthodes* |
| Peex | *Penium exiguum* |
| Meen | *Mesotaenium endlicherianum* |
| Mekr | *Mesotaenium kramstei* |
| Meca | *Mesotaenium caldariorum* |
| Cybr | *Cylindrocystis brebissonii* |
| Cysp | *Cylindrocystis sp.* |
| Cycu | *Cylindrocystis cushleckae* |
| Spsp | *Spirotaenia sp.* |
| Zysp | *Zygnemopsis sp.* |

| Bryophytes | |
|---|---|
| Spfa | *Sphagnum fallax* |
| Phpa | *Physcomitrella patens* |
| Mapo | *Marchantia polymorpha* |
| Mapa | *Marchantia paleacea* |
| Lucr | *Lunularia cruciata* |
| **Lycophytes** | |
| Semo | *Selaginella moellendorffii* |
| **Gymnosperms** | |
| Lafr | *Lagarostrobos franklinii* |
| Poco | *Podocarpus coriaceus* |
| Ceha | *Cephalotaxus harringtonia* |
| Tacr | *Taiwania cryptomerioides* |
| Atcu | *Athrotaxis cupressoides* |
| **Angiosperms** | |
| Amtr | *Amborella trichopoda* |
| Orsa | *Oryza sativa* |
| Brdi | *Brachypodium distachyon* |
| Zema | *Zea mays* |
| Rico | *Ricinus communis* |
| Aqco | *Aquilegia coerulea* |
| Glma | *Glycine max* |
| Metr | *Medicago truncatula* |
| Gora | *Gossypium raimondii* |
| Potr | *Populus trichosarpa* |
| Soly | *Solanum lycopersicum* |
| Arta | *Arabidopsis thaliana* |
| Sobi | *Sorghum bicolor* |
| Piab | *Picea abies* |

Table 3.5 | Species abbreviations used in Figures 3.1, 3.6, 3.7 and 3.8.

### 3.2.5 Identification of putative symbiosis genes by phylogeny: nutrient transport

STR and STR2 are two closely related half ABCG transporters AM that function specifically in AM symbiosis (Zhang et al., 2010). The substrate of STR and STR2 is unknown but they have been proposed to export lipids across the periarbuscular membrane to be taken up by the AM fungi, most likely in the form of 16:0 β - monoacylglycerol (Bravo et al., 2017; Roth and Paszkowski, 2017). STR and STR2 are closely related to a group of half ABCG transporters found in non-symbiotic plants also, some of which function in root suberin deposition, as well as forming the desiccation tolerant pollen wall and maintaining dormancy in the seed coat (Fedi et al., 2017; Yadav et al., 2014).

This analysis found that both *STR* and *STR2* are present in *M. paleacea*, and these were likely present in the common ancestor of the land plants (Figure 3.8). This suggests that the function of these genes in AM symbiosis has been conserved since that time. This analysis also found a subset of charophyte algae have an ABCG transporter gene closely related to the *STR* and *STR2* clade, although this node has low statistical support. A similar group was found in the larger dataset used in (Delaux et al., 2015). This clade includes sequences from *Coleochaete irregularis* and in *Mesotaenium endlicherianum*, which are species of charophyte algae that molecular phylogenetics place close to the land plant lineage in the streptophyte clade, and like the land plants, form a phragmoplast during cell division. This data suggests that a proto-*STR* may have been present in the common ancestor of land plants, Mesotanieaceae and Coleochaetaceae, and that this gene underwent duplication to form *STR* and *STR2* in the land plants, while a single paralogue was subject to different selection pressure in charophyte species.

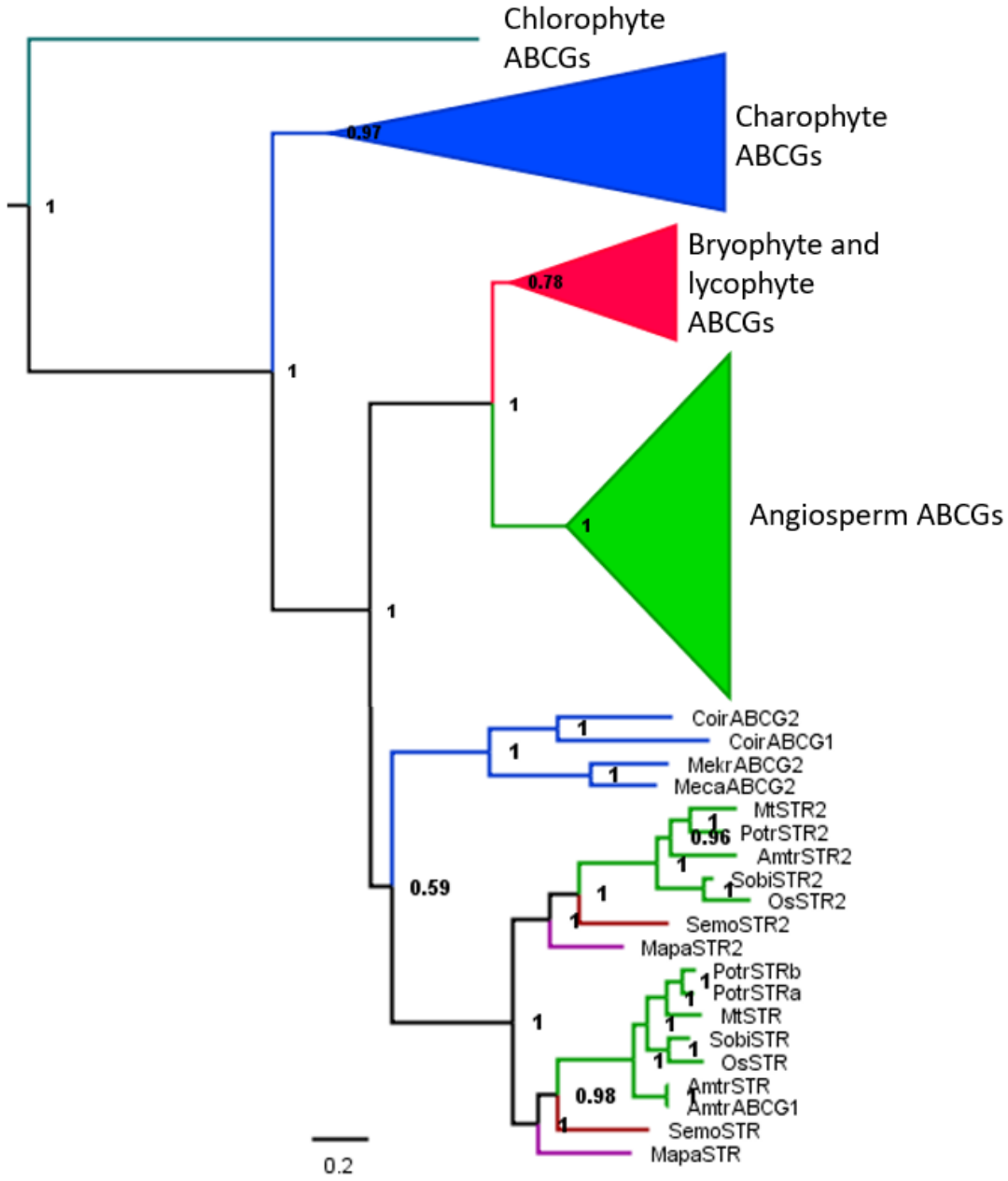



Figure 3.8 | Maximum likelihood phylogeny of STR and related ABCG transporters in the streptophytes.

*STR* and *STR2* were present in the common ancestor of the land plants. Some charophyte algae contain ABCG sequences similar to *STR* and *STR2*. *STR* and *STR2* are closely related to further ABCG transporters in the angiosperms. Key for abbreviations of species used are in Table 3.5. Charophytes branches are shown in blue, bryophyte branches are magenta, angiosperm branches are green. The tree was rooted using charophyte sequences (turquoise). Scale bar shows substitutions per site. Branch support calculated by aBayes analysis.

As AM symbiosis in *M. paleacea* functions in phosphate uptake, this implies that the genome must contain a phosphate transporter gene that functions at the arbuscule (Humphreys et al., 2010). Unlike angiosperms, bryophytes contain both Na+/Pi symporter PTBs and the H+/Pi PHT1 symporters (Bonnot et al., 2017). It seems likely that a PHT1 transporter will function in the arbuscule, given that in angiosperms the periarbuscular space forms an acid environment that drives PHT1 transport (Guttenberger, 2000). In the land plants, the PHT1 gene family is characterized by repeated gene duplication and gene loss. This phylogenetic analysis (Figure 3.9) found three distinct angiosperm clades; the highly divergent monocot and dicot clade , named clade II, a symbiosis specific clade containing *MtPT4* and *OsPT11*, named clade I and group formed by the dicot clade III and monocot clade IV, which have expanded independently in the angiosperms, named using the convention established in the literature (Nagy et al., 2005). In this analysis, the monocot specific AMF inducible clade V, proposed in (Ceasar et al., 2014), fell inside the monocot clade IV, although long branches in the phylogram indicate that these genes are quite divergent compared to the rest of clade IV.

The phylogeny presented in Figure 3.9 suggests that a single PHT1 gene was present in the common ancestor of the land plants, and that the PHT1 gene family expanded independently in the liverwort, moss, lycophyte and angiosperm lineages and is similar to that published in (Delaux et al., 2015). Although some basal land plant PHT1s appear to be more closely related to the angiosperm PHT1 I clade, this may be an artefact caused by long branch attraction and may not indicate a true phylogenetic relationship.

(Delaux et al., 2015) identified a PHT1 phosphate transporter specifically upregulated during AM symbiosis in the liverwort *Lunularia cruciata (LucrPTMYyc)*. Notably, this PHT1 transporter is not part of the bryophyte clade more closely related to the angiosperm PHT1 I group. Although there is not a clear orthologue to this phosphate transporter present in *M. paleacea*, the closest paralogue to *LucrPTMyc* may be *MpaPHT1D* (Figure 3.10). *MpaPHT1D* also lacks a corresponding orthologue in the *M. polymorpha* genome, further suggesting that this phosphate transporter has a function specific to AM symbiosis lost in the non -symbiotic *M. polymorpha* lineage. It is worth noting that *M. polymorpha* contains three additional PHT1 genes generated by a tandem duplication since *M. paleacea* and *M. polymorpha* diverged (Figure 3.10). This suggests a recent adaptation in phosphate uptake in *M. polymorpha* that may compensate for the recent loss of nutrient uptake via AM symbiosis since this species diverged from *M. paleacea*.

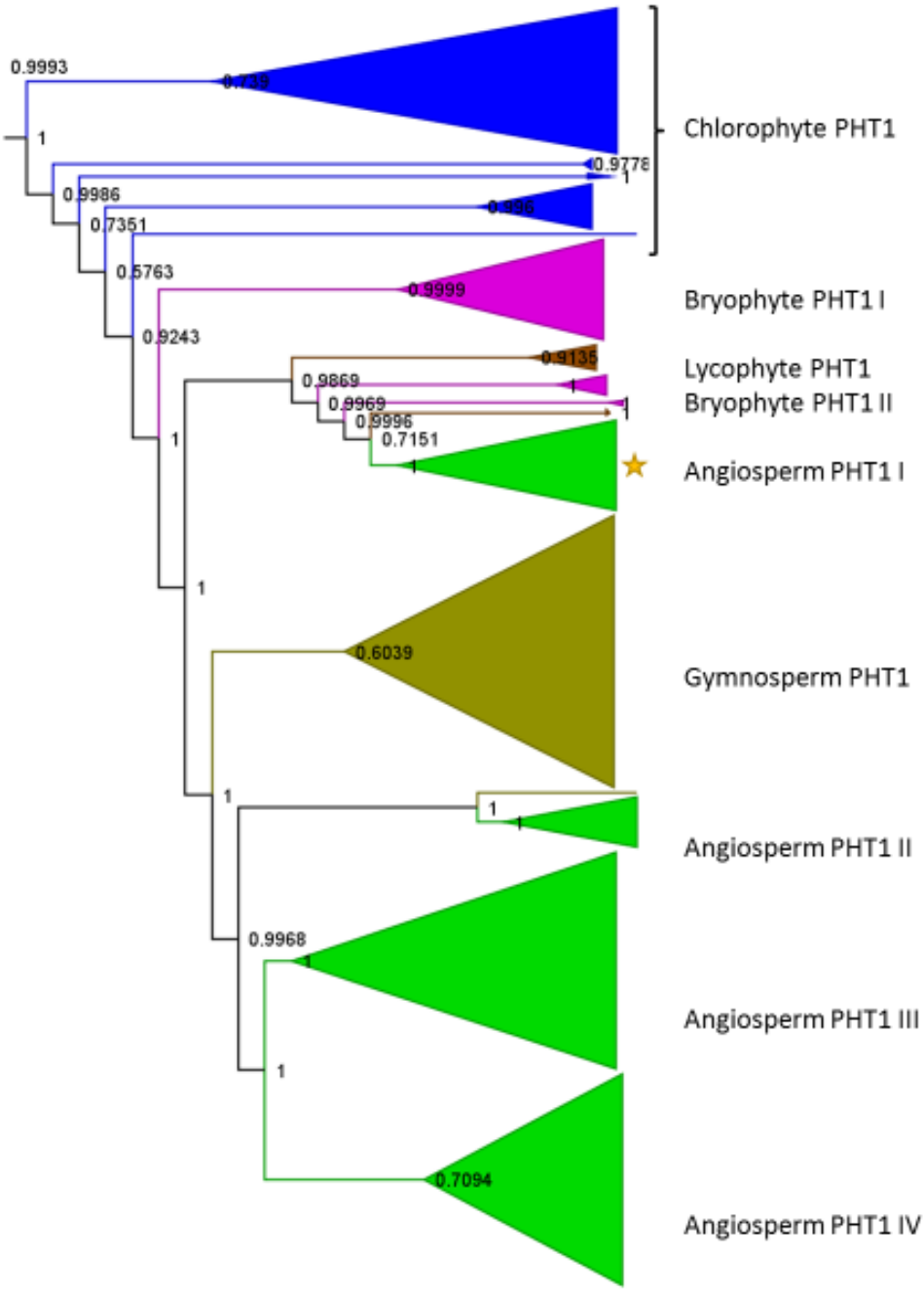



Figure 3.9 | Maximum likelihood cladogram of PHT1 transporters in the green lineage.

The PHT1 phosphate transporter family was probably represented by a single gene in the common ancestor of the land plants. Independent expansions of PHT1 gene have taken place in the angiosperms, gymnosperms, lycophytes and bryophytes. Some lycophyte and bryophyte PHT1 sequence appear to be more closely related to the angiosperm PHT1 I group than to other angiosperm PHT1s.

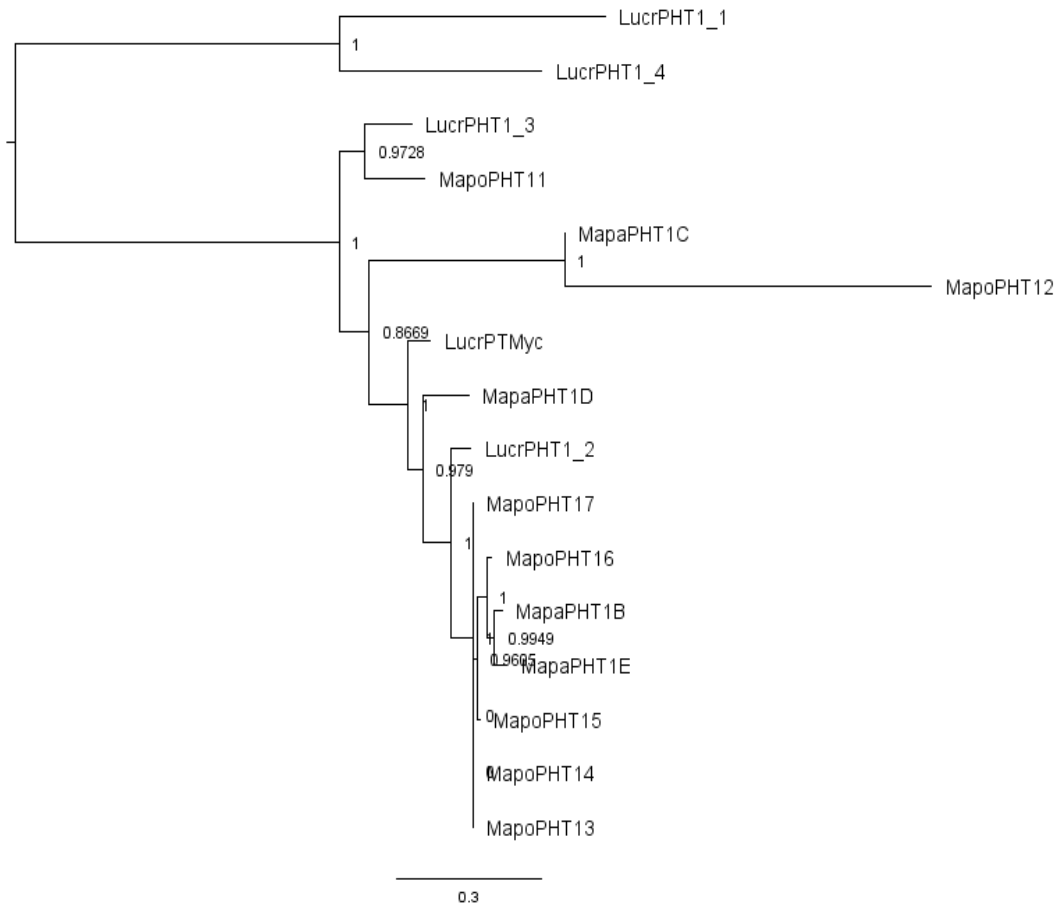



Figure 3.10 | Unrooted maximum likelihood phylogeny of PHT1 phosphate transporters in Marchantiopsida.

Gene loss and gain in the PHT1 family has continued during liverwort evolution. An orthologue of the symbiosis associated PHT1 in *L. cruciata*, *LucrPTMyc* is present in *M. paleacea* but not *M. polymorpha*. The PHT1 family has expanded recently in the *Marchantia* genus, with tandem duplications increasing the number of PHT1s in *M. polymorpha* (*MapoPHT1;3-MapoPHT1;6* are adjacent to each other on scaffold 0195 in the current *M. polymorpha* genome assembly).

## 3.3 Discussion

Phylogenetic analysis of symbiosis genes in land plants has recently provided insight into the evolution of plant endosymbiosis and the molecular mechanisms necessary for AM symbiosis. This phylogenetic analysis has shown that many molecular mechanisms thought to be symbiosis specific evolved in an algal ancestor of land plants before the transition to land and led to the discovery of a symbiosis specific fatty acid metabolic pathway. The power of phylogenetic analysis to answer biological questions about AM symbiosis continues to increase as more plant genomes and transcriptomes are published.

In this chapter, I conducted phylogenetic analysis of six symbiosis genes involved in different stages of symbiosis establishment. This analysis focused on asking how these genes had evolved since the bryophytes and angiosperms diverged and how these gene families would have looked in the common ancestor of the land plants. This analysis was then ultimately used in selecting gene families to study the functional conservation of symbiosis genes.

Some phylogenetic analysis has been conducted previously on all the gene families evaluated. A *DMI1* orthologue in a non-symbiotic plant (*A. thaliana*) was discovered when *DMI1* was first shown to function in root nodule and AM symbiosis in *M. truncatula*, suggesting that it may have non-symbiotic functions (Ane et al., 2004). The presence of a paralogue that is also necessary for symbiosis signaling was discovered in *L. japonicus* and found to also be necessary in other angiosperm species such as rice (Charpentier et al., 2008; Imaizumi-Anraku et al., 2005). Only *MtDMI1* is necessary for AM symbiosis but both *LjPOLLUX* and *LjCASTOR* are indispensable

due to recent evolution of the selection filter of the ion channel in the legumes, affecting cation permeability and therefore signalling function (Venkateshwaran et al., 2012). The *M. truncatula*-type selection filter, ADAGNH, appears to be the ancestral sequence, as it is seen ion liverworts and in charophyte algae, but the *L. japonicus*-type selection filter is also seen in mosses and lycophytes.

The three main subgroups of land plant CNGCs found in this analysis were also recovered in (Charpentier et al., 2016; Saand et al., 2015; Zelman et al., 2012). A putative symbiotic CNGC could not be identified on the basis of phylogeny as this gene family has expanded independently in bryophyte and angiosperm lineages. However, comparisons of intron-exon structure and protein similarity provides evidence that one of the two Group I/II/III CNGCs in *M. paleacea*, *MpaCNGC3,* could have the same function as *CNGC15a,b,c* and generate nuclear $Ca^{2+}$ signals in AM symbiosis.

The analysis in this chapter indicates that the three main subgroups are also present in in charophyte algae, although no transcriptome had all three subgroups represented. This suggests that the algal ancestor of the land plants had all three CNGC subfamilies represented, including a Group I/II/III CNGC but one or more of these subfamilies was lost in different charophyte lineages. Several charophyte Group IVa and Group IVb CNGCs have NLS sequences (see appendix 7.1.1), suggesting that they could function in nuclear $Ca^{2+}$ signalling in concert with *DMI1* and *CCaMK* genes found in these species.

Some arbuscule related genes have direct orthologues in *M. paleacea*, with single orthologues of *VAPYRIN*, *STR* and *STR2* present. This suggests that these genes have had an unchanged function in AM symbiosis since the common ancestor of the land plants. However, most bryophyte PHT1 phosphate transporters do not fall into the

same sub-families as angiosperm PHT1s. Transcriptional analysis of another liverwort in the class Marchantiopsida, *L. cruciata*, has previously identified a PHT1 specifically upregulated in AM symbiosis (Delaux et al., 2015). Although there is an orthologue of this gene in *M. paleacea*, it is absent from *M. polymorpha*, despite expansion of the PHT1 family in *M. polymorpha* as a whole. This is the strongest candidate for a PHT1 gene that functions in AM symbiosis in *M. paleacea*.

The Exo70.3 subfamily is represented by a single paralogue in *M. paleacea*. This is likely to function in generating the PAM during symbiosis, but it may also function in plant development. Although the function of the non-symbiotic Exo70.3, *Exo70G*, is unknown in the angiosperms, two *Exo70G* paralogues are expressed in multiple plant organs during development in *A. thaliana* and one Exo70.3 paralogue is essential to normal development in the moss, *P. patens* (Li et al., 2010).

The putative symbiosis related genes identified in this chapter will be further investigated by expression analysis in Chapter IV. The functional conservation of *DMI1* and *CNGC* will also be analysed. While the function of *VAPYRIN*, *Exo70I* (or *MpaExo70GI*), *STR* and *STR2* are very likely to be conserved based on phylogenetic analysis, this is less clear for *DMI1* and *CNGC* analysed here. In addition, these components of symbiotic $Ca^{2+}$ signalling form a functional unit in the angiosperms, despite differing evolutionary histories. The potential symbiotic PHT1 identified in this chapter may be an interesting area for further research. As the impact of AM symbiosis on phosphate uptake in *M. paleacea* has been described, the tools required to dissect function of PHT1s in AM symbiosis in this species already exist (Humphreys et al., 2010).

# Chapter IV – Gene expression analysis of putative symbiosis genes in *M. paleacea*

## 4.1 Introduction

Multiple genes with function in AM symbiosis and arbuscule development have been identified based on specific upregulation during colonization with Glomeromycete fungi. As such, the phosphate transporter *PT4*, which is specifically expressed in arbuscule containing cells, was demonstrated to be essential for normal arbuscule development and phosphate uptake (Harrison et al., 2002; Javot et al., 2007a; Nagy et al., 2005; Paszkowski et al., 2002). The proton pump ATPase, HA1, that generates the pH gradient across the periarbuscular membrane and that drives phosphate transport, was similarly identified on the basis of arbuscule specific expression, while transcriptional co-expression analysis was used to identify Exo70I, a component of the exocyst complex involved in secretion of the PAM (Krajinski et al., 2014; Krajinski et al., 2002; Zhang et al., 2015). Similarly, the ACP thioesterase, FatM, part of the lipid biosynthesis and export gene cluster is upregulated by the transcription factor *RAM1* during AM colonization (Luginbuehl et al., 2017). Expression analyses of the liverwort *Lunularia cruciata*, found that orthologues of several known AM genes were upregulated eight weeks post inoculation with AM (Delaux et al., 2015). These included the transcription factors *LcRAD1* and *LcNSP1*, a PHT1 phosphate transporter (*LcPTMyc*) and a proton pump ATPase (*LcHa*), and two *RAM2* paralogues.

In contrast, genes involved in the mycorrhizal signal perception and transduction, are not upregulated by AM fungi. As such, *DMI1* is constitutively expressed in the roots of *M. truncatu*la, and not induced by AM colonization (Ane et al., 2004; Limpens et al., 2003; Riely et al., 2007). The nucleoporin, *NUP85* is also expressed constitutively, unaffected by inoculation with AM fungi (Saito et al., 2007). In *M. truncatula*, all three *CNGC*15 paralogues are expressed in the roots outside of symbiotic colonization, as well as in various other tissues, with the highest levels of *MtCNGC15a* expression seen in the stem and leaf and the highest levels of *MtCNC15c* seen in the flower (Charpentier et al., 2016). Although *CCaMK* is slightly upregulated by inoculation with rhizobia in *M. truncatula*, it is not upregulated in rice by AM symbiosis (Chen et al., 2007; Chen et al., 2008; Gutjahr et al., 2008). *MtCCaMK* is expressed in the roots in the absence of symbionts, and weakly in flower tissues (Gutjahr et al., 2008; Poovaiah et al., 1999).

The aim of the research presented in this chapter was to investigate the expression of the AM related genes identified in *M. paleacea* genome by phylogenetic analysis in Chapter III, and to assess the spatial expression patterns of the ion channels for consistency with potential function in early AM signalling.

## 4.2 Results

### 4.2.1 Colonization by *Rhizophagus irregularis* induces expression of arbuscular mycorrhiza associated genes

To assess gene expression during AM symbiosis in *M. paleacea*, the conditions for AM colonization of *M. paleacea* were first tested with the homemade chive roots *R. irregularis* inoculum. Four-week-old *M. paleacea* thalli were grown from gemmae in vitro before transfer to soil. The soil was composed of 90% sharp sand, 10% John

Innes No. 1 compost to replicate the low nutrient supply conditions in (Humphreys et al., 2010) where AM symbiosis is shown to enhance *M. paleacea* primary production. This base mixture was combined in a 4:1 ratio with either chive roots *R. irregularis* inoculum in a soil medium made up of 45% terragreen, 45% sand and 10% John Innes No. 1 compost, or non-mycorrhizal control made of the same soil medium.

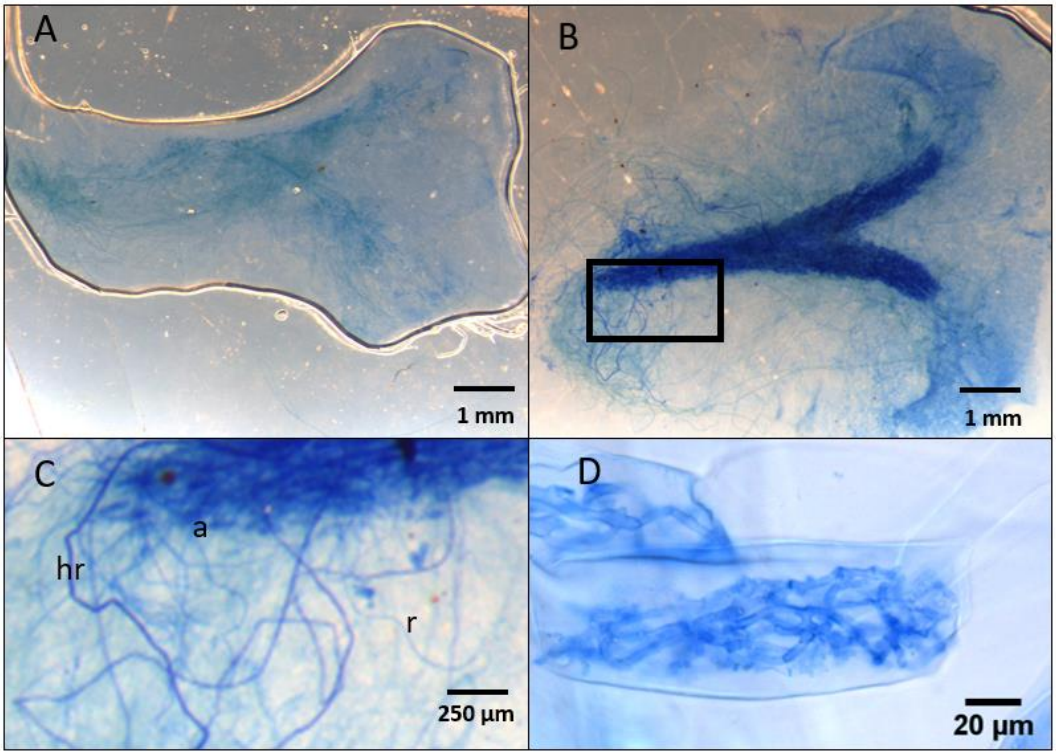



Figure 4.1 | Fungal infection structures in *M. paleacea* thalli at 6 weeks post inoculation with *R. irregularis*.

*M. paleacea* thallus non-colonized (A) and colonized (B-C). Fungal structures were stained with aniline blue and visualized in dark blue, while plant tissue appear light blue. C, Magnification of the boxed region in B showing rhizoids containing fungal hyphae (hr), light blue coloured rhizoids without hyphae (r) and intracellular fungal structures in the parenchymous tissue of the thallus (a). D. These fungal structures are highly branched, with an arbuscule like appearance.

A subset of plants was stained with aniline blue  per (Newsham et al., 2014) to check for fungal colonization at two weeks intervals until widespread colonisation of inoculated thalli by fungal structures had occurred. At 6 weeks post inoculation (w.p.i.), widespread colonization of the parenchymal region of the thallus along the midrib was observed (Figure 4.1. A, B). Colonization occurred as fungal hyphae grew through rhizoids to the ventral surface of the thallus (Figure 4.1 C) and intracellular fungal structures developed in the cells of the thallus parenchyma along the midrib (Figure 4.1 D). Fungal structures seen in these cells had highly branched arbuscule morphology (Figure 4.1 D).

The conditions for AM colonization of *M. paleacea* being successfully established, a mycorrhizal time course experiment at 8 dpi, 28 dpi and 42 dpi was performed to analyse the expression of candidate genes required for AM development. The Arbuscule development associated genes *MpaVAPYRIN*, *MpaExo70GI*, *MpaPHT1D* and *MpaSTR* were all significantly upregulated at 42 dpi (Figure 4.2A). Nutrient transport associated genes were highly expressed in AM colonized gametophytes, with a 19-fold increase in *MpaPHT1D* expression and a 23-fold increase in *MpaSTR* expression in inoculated gametophytes, relative to the uninoculated control (Figure 4.2A). A similar fold induction (25-fold) was also seen for *MpaVAPYRIN*. There was also a statistically significant upregulation of *MpaExo70GI* expression at 42dpi. In contrast to the other arbuscule associated genes measured here, fold induction of *MpaExo70GI* at 42 dpi was relatively low at 2.9-fold relative to the uninoculated control. Expression of arbuscule associated genes also increased in the uninoculated control grown on soil (Figure 4.2), suggesting that the expression of arbuscule

associated genes may be affected by the developmental and nutrient status of the plant, as well as AM colonization.

In contrast to the genes required for arbuscule development, there was no statistically significant difference in the expression of *MpaDMI1*, *MpaCNGC1*, *MpaCNGC2*, *MpaCNGC3* or *MpaCNGC4* between inoculated and control plants at 42 dpi (Figure 4.2B). This is consistent with findings in angiosperm models, where *MtDMI1* is not upregulated by AM inoculation and no symbiosis associated upregulation of a *CNGC* has been documented.

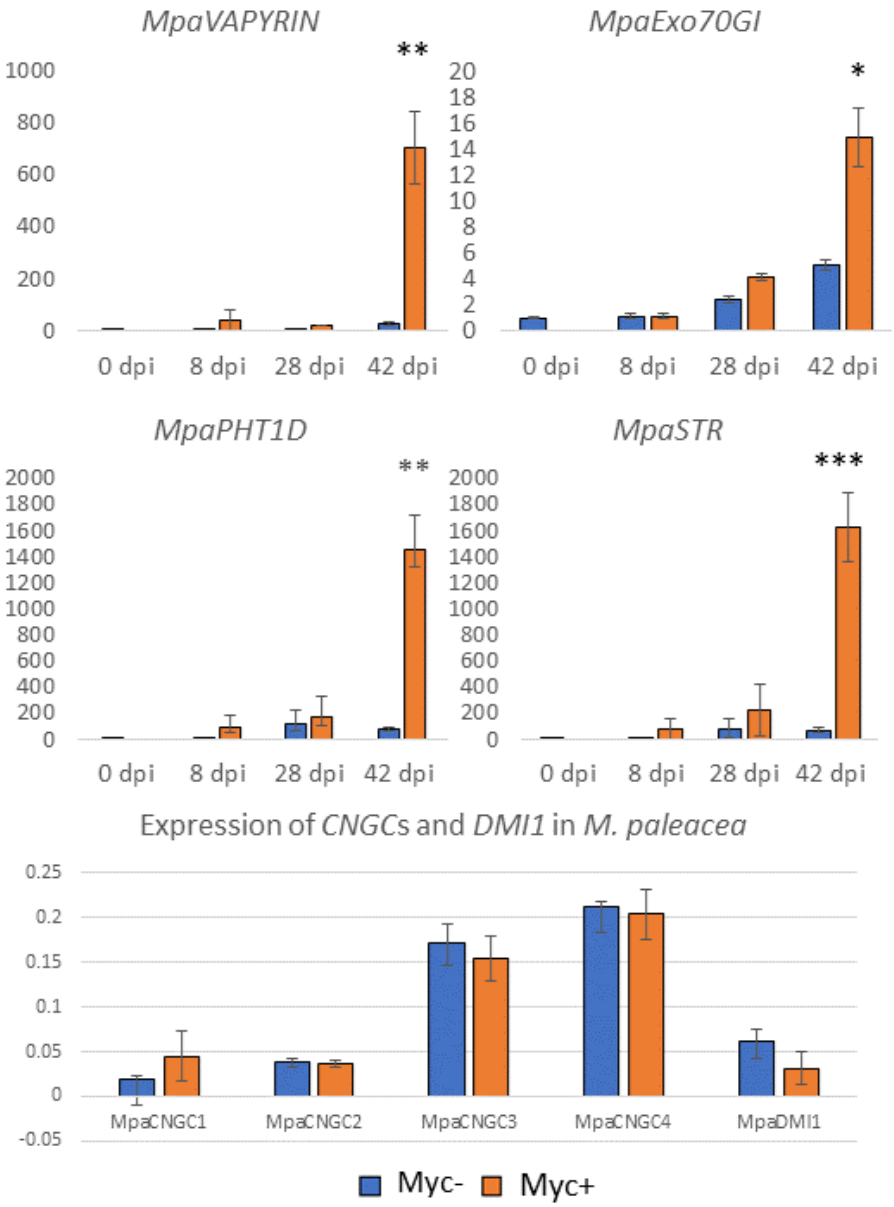



**Figure 4.2 | Quantification of transcript levels of target *M. paleacea* genes putatively involved in arbuscular symbiosis.**

Expression levels of *MpaVAPYRIN*, *MpaExo70GI*, *MpaPHT1D*, *MpaSTR MpaCNGC1*, *MpaCNGC2*, *MpaCNGC3*, *MpaCNGC4* and *MpaDMI1* were measured by qRT-PCR and normalized to MpaACTIN7 and M. paleacea Adenine phosphorylase transferase 2 (MpaAPT)(Saint Marcoux et al., 2015). Expression was monitored in gametophyte plants at 0, 8, 28 and 42 days post inoculation (dpi) with R. irregularis (Myc + plants) and in mock inoculated plants at the same time points (Myc-) for *MpaVAPYRIN*, *MpaExo70GI*, *MpaPHT1D* and *MpaSTR* **(A)**, and at 6 weeks post inoculation only for *MpaCNGC1*, *MpaCNGC2*, *MpaCNGC3*, *MpaCNGC4* and *MpaDMI1* (**B**). Gene expression at data point 0 day post inoculation (dpi) represents pre-grown gametophytes before transfer to low nutrient soil. Bar represent means of 4 biological replicates ± SEM. Asterisks indicate statistical differences between expression levels in mycorrhized and non-mycorrhized tissues at the corresponding time point (Student's T-test with *, P<0.05; **, P<0.01; ***, P<0.001).

### 4.2.2 Analyses of MpaCNGC1, MpaCNGC2, MpaCNGC3, MpaCNGC4 and MpaDMI1 promoter activity in M. paleacea.

To characterize the spatial expression pattern of the *M. paleacea CNGC*s and *DMI1*, the promoter region of each gene, defined between the ATG of each *MpaCNGC* and *MpaDMI1*, and the end of the next up-stream predicted transcript, was cloned in fusion with the coding sequence of the *ß-glucuronidase* (*GUS*) reporter gene using GoldenGate strategy. In the vectors constructed, *GUS* expression was driven by the 2 kb promoter of *MpaDMI1*, 1.5 kb, 1.5 kb, 2 kb and 1.5 kb of *MpaCNGC1*, *2*, *3*, and *4*, respectively. The fluorescent reporter mCherry was included in the level 2 construct to facilitate the detection of transformant. The fluorescent reporter (mCherry) was detected in 2, 1 and 6 of each transformants with *pMpaDMI1:GUS*, *pMpaCNGC3:GUS* and *pMpaCNGC1*,*2* and *4*, respectively. Those lines were selected for GUS activity analyses at different stages of *M. paleacea* development.

*pMpaDMI1:GUS* was first detected in apical notch (meristematic region) of gemma developing within the gemma cup (Figure 4.3A). Stronger staining reveals broader expression of *pMpaDMI1:GUS* throughout the gemma (Figure 4.3B and C), and over the 2 week old thallus and rhizoid (Figure 4.3D and F). Stronger expression of *pMpaDMI1:GUS* is observed in the region of the thallus along the midrib (Figure 4.3D and E), an area which is colonized by AM fungi (Figure 4.1B).

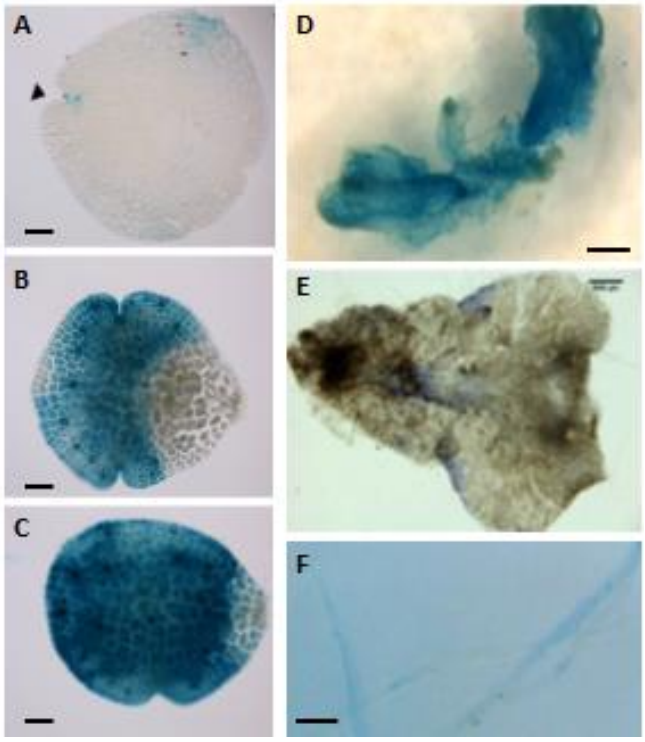



**Figure 4.3 | *MpaDMI1* promoter activity in *M. paleacea***

A-C, *MpaDMI1:GUS* expression in developing gemmae from a 1 week old gemma cup stained for 1h (A), 6h (B), 24h (C) of ß-Glucuonidase (GUS) activity. D, 2 week old thallus stained for 1h of GUS activity. E, 8 week old thallus stained for 24h of GUS activity. F, Rhizoid from 2 weeks old thallus stained for 2h. Scale bars indicate 50 µm (A-C, F) and 500 µm (D, E). Black arrowhead indicates one of the two apical notches.

In contrast to *MpaDMI1:GUS* expression, *pMpaCNGCs:GUS* expression is absent in the gemma developing from the gemma cup (Figure 4.4). *pMpaCNGC3:GUS*, *pMpaCNGC4:GUS* and *pMpaCNGC1:GUS* expressions are detected in the thalli including the parenchymal region of the thalli along the midrib and in the rhizoids (Figure 4.5 and Figure 4.6A and B). *pMpaCNGC2:GUS* expression was not detected in the rhizoid or in the midrib, but only in some air pores (Figure 4.6C). The absence of expression of *MpaCNGC2* from the tissue colonized by AM fungi (rhizoid and parenchymal region of the thalli along the midrib) suggest strongly that MpaCNGC2

is not required for AM symbioses in *M. paleacea*. In the other hand, *MpaCNGC1*, *MpaCNGC34* and *MpaCNGC4*are all expressed in the tissue colonized by AM fungi similarly to *pMpaDMI1*.

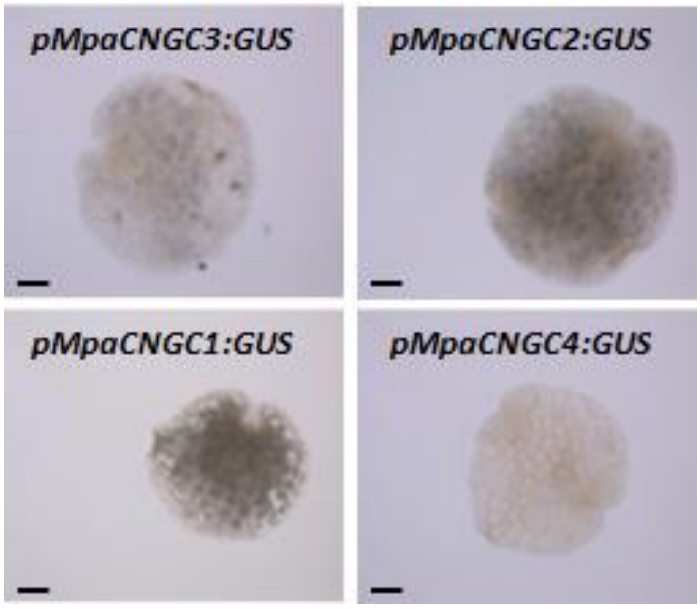



**Figure 4.4 | Promoter activity of the MpaCNGC1,2,3 and 4 in developing *M. paleacea* gemma.**

Gemma developing on 1-week old gemma cup were stained for 24h for GUS activity. Scale bar indicate 50 µm.

pMpaCNGC3:GUS pMpaCNGC4:GUS

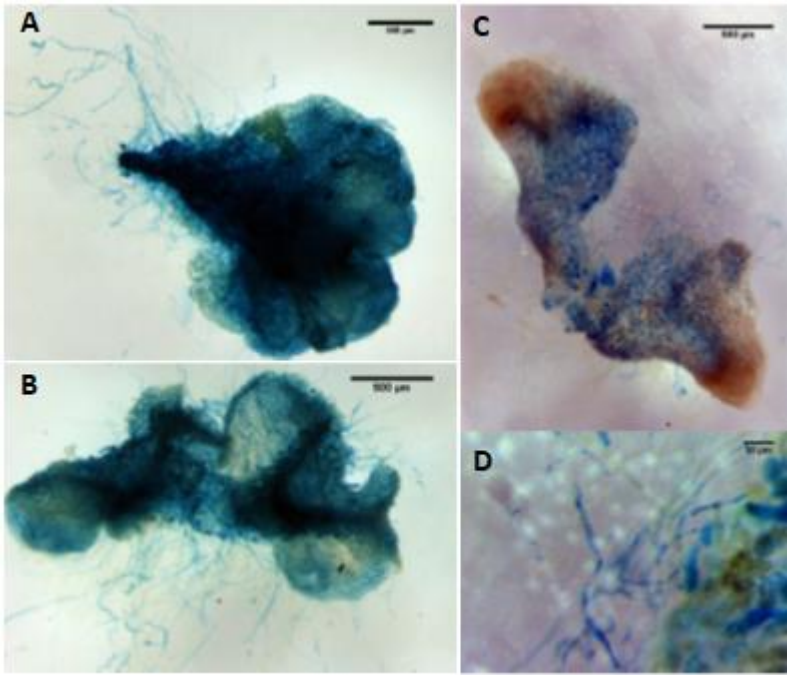



**Figure 4.5 | Promoter activity of the *MpaCNGC3* and *MpaCNGC4* in *M. paleacea*.**

2 weeks old thalli were stained 6h (A) and 24h (B, C, D) for GUS activity. Scale bars indicate 500 µm (A, B, C) and 50 µm (D).

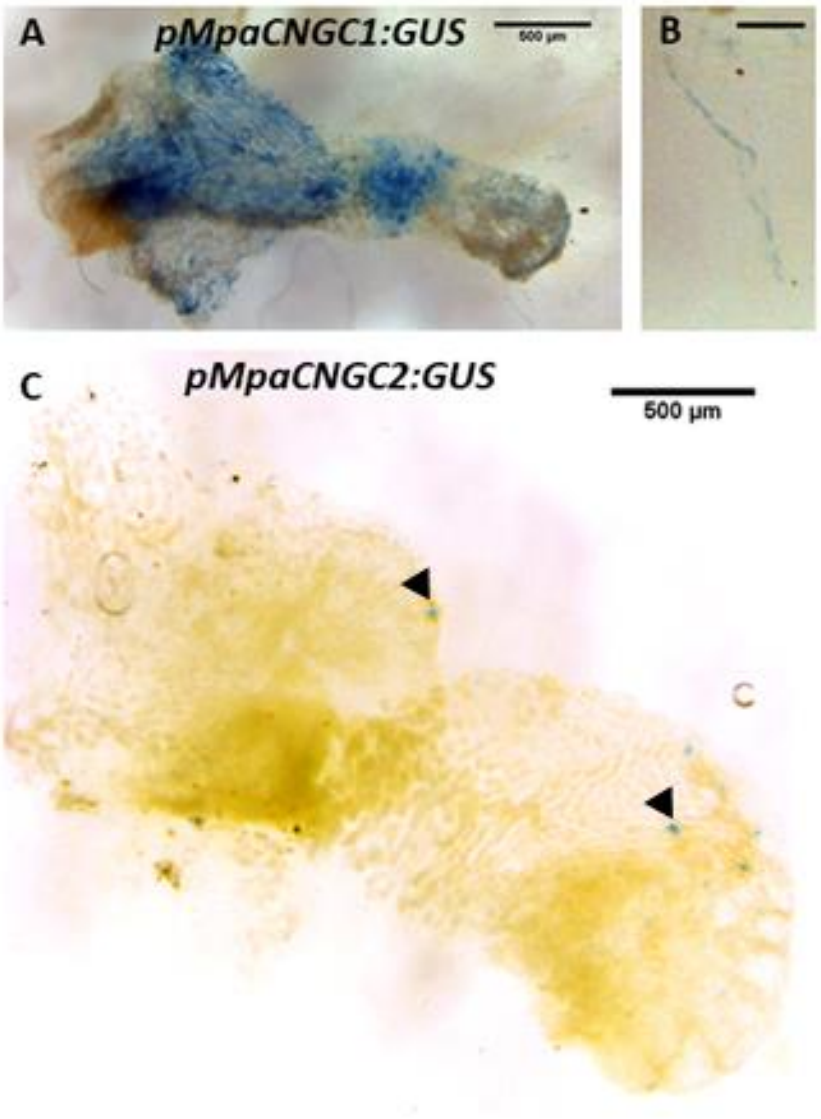



**Figure 4.6 | Promoter activity of the *MpaCNGC1* and *MpaCNGC2* in *M. paleacea*.**

2 weeks old thalli were stained 24h for GUS activity. B, Zoom on the rhizoid from picture (A). Scale bars indicate 500 µm (A,C) and 100 µm (B). Black arrowhead indicates air pores.

## 4.3 Discussion

Our understanding of AM symbiosis is thus far based almost entirely on studies in angiosperms (Delaux et al., 2015). To understand whether the insights gained into the AM symbioses in angiosperms are true for other land plants, it is important to extend our study to non-flowering plant such as liverworts, which represent one of the earliest diverging land plant lineages. In angiosperms, genes required for arbuscule development are upregulated after activation of AM signal transduction. In this study, I showed that AM related genes identified in *M. paleacea* genome by phylogenetic analysis and required for arbuscule development/function in angiosperm are induced by AM colonization in the liverwort *M. paleacea*. *MpaVAPYRIN*, *MpaExo70GI*, *MpaPHT1D* and *MpaSTR* are all significantly upregulated at 42 dpi. Similarly to the *MtVAPYRIN*, *MtPHT1D* and *MtSTR*, their *M. paleacea* related genes are highly upregulated by AM. However, in contrast to *MtExo70I*, which is upregulated up to 100-fold relative during AM colonization, (Zhang et al., 2015), fold induction of *MpaExo70GI* at 42 dpi was relatively low (2.9-fold relative to the control). As there is only one member of the Exo70.3 subfamily in *M. paleacea*, this gene may function in plant development outside of AM symbiosis. In angiosperms the Exo70.3 subfamily contains two clusters, the *Exo70I* cluster found only in species forming AM symbiosis, and the *Exo70G* cluster that is also found in non-AM plants (Zhang et al., 2015).

In angiosperms, genes required for AM signal transduction and notable for the activation of the nuclear $Ca^{2+}$ oscillation such as *MtDMI1* are not necessarily upregulated by AM fungi (Ane et al., 2004; Limpens et al., 2013). Similarly, I demonstrated that *MpaDMI1* and all *MpaCNGC*s are not upregulated by AM

colonization. Within the four *CNGC*s identified in *M. paleacea*, there is no clear functional homolog to *MtCNGC15a,b,c*. In order to gain insight into the putative *M. truncatula* functional homologue, I analysed the expression pattern of the *MpaCNGC1,2,3* and 4, along with *MpaDMI1*. *pMpaDMI1:GUS* is strongly expressed in rhizoid and more specifically in the parenchymal region of the thallus along the midrib, sites of AM colonization. Promoter GUS assays also indicated that *MpaDMI1* is expressed in the gemmae, while no GUS staining could be detected in the gemmae of *pMpaCNGC1:GUS*, *pMpaCNGC2:GUS*, *pMpaCNGC3:GUS* or *pMpaCNGC4:GUS* lines. The significance of *MpaDMI1* expression in the gemmae is unknown, as it is unclear what functions, if any, *MpaDMI1* might have at this stage in development. Gemmae may have the potential to respond to Myc factors with calcium oscillations, or *MpaDMI1* may have another function at this stage of development. Similarly to *pMpaDMI1:GUS*, *pMpaCNGC1:GUS*, *pMpaCNGC4:GUS* and *pMpaCNGC3:GUS* are all expressed in rhizoids and thalli, suggesting that they all potentially could function in AM symbiosis. In contrast, *pMpaCNGC2:GUS* was expressed specifically in the air pores, highlighting that *MpaCNGC2* is unlikely a player in AM symbiosis.

The expression analyses described in this chapter demonstrate that putative functional homologues for known AM genes in angiosperms are similarly upregulated in *M. paleacea* during AM colonization. Additionally, I demonstrated that the putative functional homologues of *MtDMI1* and *MtCNGC15s* are expressed in the tissue colonized by AM in *M. paleacea*. However, to assess whether *MpaDMI1* is the functional homologue of *MtDMI1* and whether *MpaCNGC3*, *MpaCNGC4* and/or *MpaCNGC1* are the functional homologue of *MtCNGC15a,b,c*, the identification and characterization of *M. paleacea* lines with mutations in those genes will be required.

# Chapter V – Functional conservation of calcium signalling components between *M. truncatula* and *M. paleacea*

## 5.1 Introduction

In angiosperms, signal molecules from endosymbionts such as AM fungi and rhizobia detected at the plasma membrane elicit nuclear $Ca^{2+}$ oscillations that are necessary for transcriptional reprogramming to establish functional symbioses (Oldroyd, 2013). In the model legume *Medicago truncatula*, these $Ca^{2+}$ oscillations are generated by a complex of ion channels localised to the nuclear membrane, comprising the potassium permeable channel, Does Not Make Infections 1 (DMI1) and the Group III cyclic nucleotide-gated channels CNGC15a, CNGC15b and CNGC15c (Ane et al., 2004; Charpentier et al., 2016; Riely et al., 2007). The sarco/endoplasmic reticulum calcium ATPase, MCA8, is also necessary for symbiotic $Ca^{2+}$ oscillations and is localised to the nuclear and endoplasmic reticulum membranes (Capoen, 2011). DMI1 and CNGC15 interact in *M. truncatula* and mathematical modelling predicts that these channels are activated simultaneously, with CNGC15 releasing $Ca^{2+}$ from the lumen of the nuclear envelope while DMI1 acts as a counter ion channel that balances the electrochemical gradient across the membrane (Charpentier and Oldroyd, 2013; Charpentier et al., 2016). MCA8 then acts to capture the released $Ca^{2+}$ ions to replenish the store in the lumen of the nuclear envelope (Capoen, 2011). The mechanism of

activation of CNGC15 is unknown but the protein contains a cyclic nucleotide binding domain, an IQ calmodulin (CaM) binding domain and may be modulated by its interaction with DMI1 (Charpentier et al., 2016)

Genes orthologous to DMI1, and CNGC15 paralogues, have been identified in basal land plants (Charpentier et al., 2016; Delaux et al., 2015; Saand et al., 2015). However, the Group I/II/III CNGCs have diversified independently in the liverwort and angiosperm lineages. *DMI1* has undergone duplication in the seed plants, with both paralogues necessary for symbiosis signalling, with the exception of *M. truncatula* (see Chapter III and (Venkateshwaran et al., 2012)). Given that functional AM symbiosis is conserved between *M. paleacea* and *M. truncatula*, this suggests that a CNGC paralogue with conserved symbiotic function is also present in the *M. paleacea* genome. *MpaCNGC1* and *MpaCNGC3* are the paralogues most closely related to *MtCNGC15a/b/c*. While both genes have potential symbiotic function based on phylogeny, MpaCNGC3 has higher protein similarity, conserved intron-exon structure and contains an NLS (Chapter III).

Transcomplementation assays have previously been used to assess the conservation of genetic function between land plants in multiple contexts and are useful in answering questions about the evolution of plant development. For example, the root hair development phenotype of the *A. thaliana* mutant, *Atrhd6* (*root hair defective 6*), can be rescued by a homologous gene from the moss *P. patens*, *RHD SIX-LIKE,* that promotes rhizoid development in moss gametophores (Menand et al., 2007). Transformation of root systems with symbiosis gene mutants in the model legumes *M. truncatula* and *L. japonicus* have shown that while the common symbiosis signalling pathway (CSSP) gene *CCaMK* is conserved with monocot angiosperms and even

charophyte algae, the function of *MtDMI1* has evolved more recently. Mtdmi1 can be complemented by its *Lotus japonicus* orthologue alone, *LjPOLLUX*, when carrying the substitution *S329A*, or with the co-expression of both *LjPOLLUX* and *LjCASTOR* (Delaux et al., 2015; Godfroy et al., 2006; Venkateshwaran et al., 2012). However, unlike *CCaMK*, there has been independent expansion of the CNGCs in angiosperms and to a lesser extent, liverworts (see Chapter III, (Charpentier et al., 2016)).

Although transcomplementation analysis indicates if a liverwort gene can still function in symbiosis in an angiosperm, a mutant in *M. paleacea* is required to definitively show that this specific gene also functions in AM symbiosis in a liverwort. As *M. paleacea* is not an established model species with existing mutant libraries, mutants were generated by targeted mutagenesis using the clustered regularly interspaced short palindromic repeat (CRISPR)/CRISPR-associated protein 9 (Cas9) system (Bortesi and Fischer, 2015). This system is adapted from the prokaryotic adaptive immune system and uses a single Cas9 endonuclease from *Streptococcus pyogenes* to cleave double strand breaks in DNA and a short guide RNA (gRNA) that targets a specific sequence for cleavage (Cong et al., 2013; Mali et al., 2013). The endogenous DNA repair systems, homology-directed repair (HDR) and non-homologous end joining (NHEJ), are exploited to induce specific gene editing or random micro-deletions, insertions or nucleotide substitutions (Cong et al., 2013). The CRISPR/Cas9 system has been used to generated knockout mutants in a variety of model angiosperms, as well as the model liverwort *Marchantia polymorpha*, however, CRISPR/Cas9 genome editing in the species *M. paleacea* has not previously been published (Albert et al., 2018; Liu et al., 2017; Sugano et al., 2018).

The aim of the research in this chapter is to investigate the function of *MpaDMI1*, *MpaCNGC1* and *MpaCNGC3* in a basal land plant that undergoes AM symbiosis

using CRISPR/Cas9 genome editing and transformation of *M. truncatula* mutants with putative homologues from the *M. paleacea* genome. The CNGC gene family in particular was selected because there was no one to one homology between a putative symbiosis in the *M. paleacea* genome and the identified angiosperm symbiosis genes, and a putative symbiosis gene could not be identified on the basis of gene expression patterns. Functional molecular genetics assays were required to identify any members of this gene family with a role in AM symbiosis in *M. paleacea*.

## 5.2 Results

### 5.2.1 Generating *M. paleacea* mutants

Initial attempts to generate CRISPR/Cas9 mutants used the methodology from previously published works in *M. polymorpha, ,* in which the *M. polymorpha* promoter of the U6 small nuclear RNA (*MpU6-1*) is used to drive the small guide RNA (sgRNA) transcription, and the promoter of *M. polymorpha Elongation Factor 1a* (*MpEF1a*) is used to drive the expression of an *A. thaliana* codon optimized *Cas9* (Sugano et al., 2018; Sugano et al., 2014). To identify the sgRNAs, I used sgRNA Scorer 2.0 (Chari et al., 2017). Two sgRNAs were used in each vector, both driven by *MpU6-1* to induce deletion of the region between the target sequences, thus resulting in a frame shift and loss of gene function (Figure 5.1 A). Dual sgRNA CRISPR/Cas9 mutagenesis was previously demonstrated to increase specificity and efficiency in generating reliable loss-of-function alleles in mice and *A. thaliana* (Pauwels et al., 2018; Ran et al., 2013; Zhou et al., 2014a; Zhou et al., 2014b). Mutants are also easily genotyped by polymerase chain reaction (PCR) as the predicted deletions mediated by dual sgRNA CRISPR/Cas9 generate shorter amplicons than the wild type.

For *MpaDMI1*, *MpaCNGC1* and *MpaCNGC3*, I designed sgRNAs to mediate knockout mutations by targeting the coding sequence upstream of transmembrane domains forming the ion channels (Figure 5.1 B). Two knockout vectors were designed for each desired mutation type, with different pairs of sgRNA sequences to allow for variation in sgRNA efficiency. For *MpaCNGC1* and *MpaCNGC3*, vectors were also designed with pairs of sgRNAs to target the coding sequence just upstream of the cyclic nucleotide binding domain and the IQ calmodulin binding domain. These mutants were designed to result in deregulated CNGCs, which could provide differing phenotypes to a loss of function mutants and in case *Mpacngc1* or *Mpacngc3* knockouts were lethal. No lethal CNGC mutant has been identified in plants thus far but to date CNGC mutants have all been evaluated in plants with larger CNGC families than *M. paleacea* (K Jha et al., 2016).

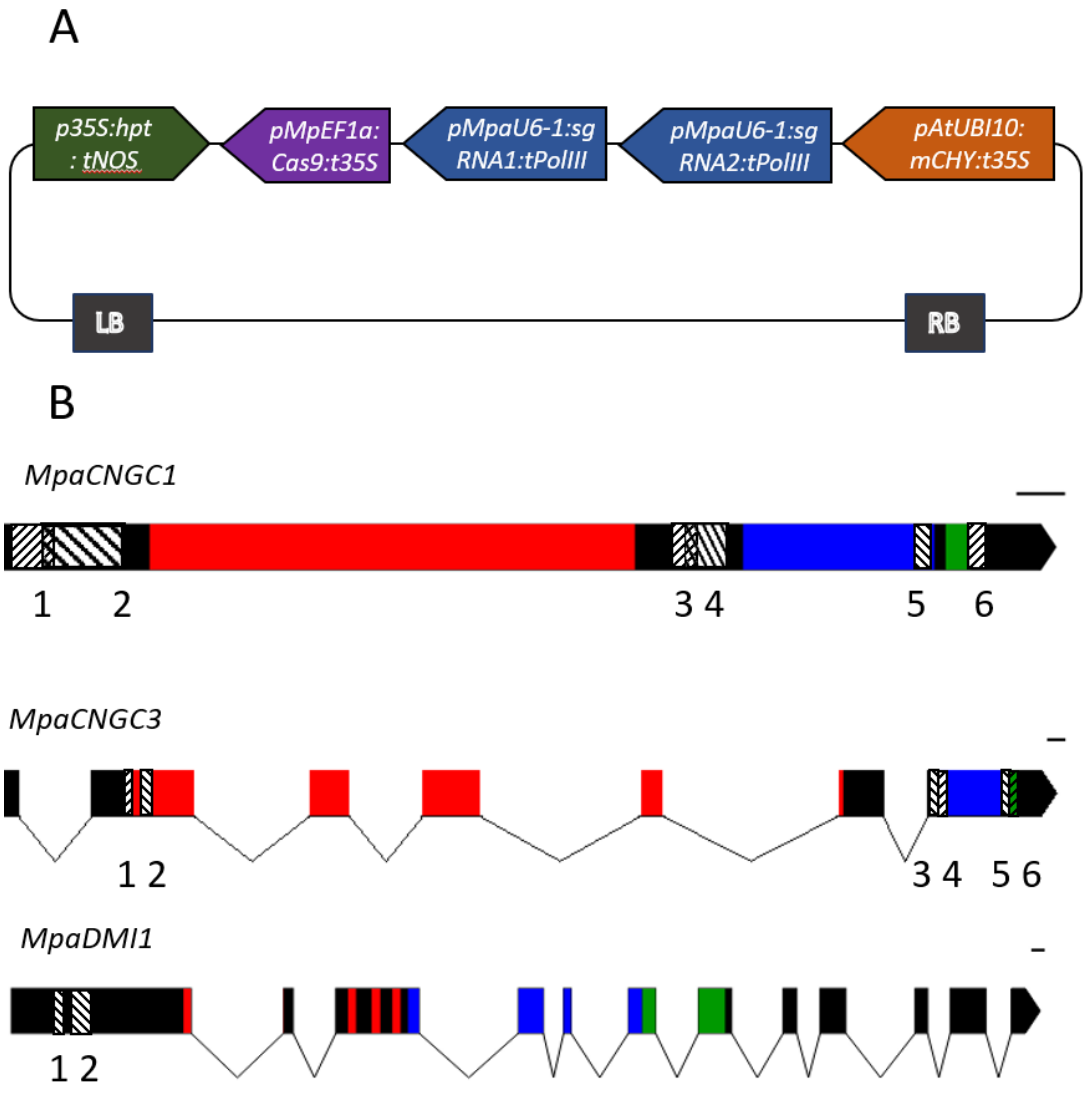



**Figure 5.1 | CRISPR/Cas9 genome editing in *M. paleacea.***

A. Construct design for CRISPR/Cas9 genome editing in *M. paleacea*. Two selection markers flanking the CRISPR/Cas9 machinery were used to select transformants containing the entire T-DNA. LB and RB are the left and right borders of the T-DNA respectively. B. Genes targeted for CRISPR/Cas9 mutagenesis. Diagonally striped boxes indicate predicted deletions induced by NHEJ using the double sgRNA approach. In the *MpaCNGC1* and *MpaCNGC3s*, red indicated the ion channel domain, blue indicates the CNG binding domain and green indicates the CaM binding domain. In *DMI1*, red indicated the transmembrane domains, blue indicates the NAD(P) binding domain and green indicates the *CASTOR/POLLUX/SYM8* domain. Numbers are used to refer to specific combinations of sgRNAs designed to induce the deletions shown and referenced in Table 5.2.

The vectors were constructed using Golden Gate cloning. I included two marker genes a hygromycin cassette and a fluorescent marker (mCherry), flanking the left border (LB) and the right border (RB) of the T-DNA, respectively (Figure 5.1A). Use of both markers allowed efficient selection of transformed lines. *M. paleacea* was transformed by co-culture with *A. tumefaciens* AGL1 containing the required binary vectors. The transformants were selected on hygromycin containing plates and subsequently genotyped by PCR.

Initial transformed lines designed to induce mutagenesis in *MpaDMI1*, *MpaCNGC1* and *MpaCNGC3* were all wild type. 1773 lines were screened in total. Using the same strategy, the CRISPR/Cas9 mediated mutagenesis conducted in *M. polymorpha* achieved over 70% genome editing efficiency (Sugano et al., 2014), indicating that *A. thaliana* codon optimized *Cas9* driven by the *MpEF1a* promoter was efficient. Additionally, the transformation of *M. paleacea* with a fluorescent marker under *MpEF1a* indicated strong expression in meristematic regions (Radhakrishnan, 2017), demonstrating that *MpEF1a* promoter from *M. polymorpha* is efficient in *M. paleacea*. I hypothesized that the *MpU6-1* used to drive sgRNA expression may be less efficient in *M. paleacea* than in *M. polymorpha*, as U6 promoters from evolutionarily distant plant species do not function as well as endogenous promoters in RNA mediated systems (Wang et al., 2008). Different U6 promoters within the same organisms can significantly affect CRISPR/Cas9 efficiency so selecting an optimal U6 promoter was necessary for generating mutant lines (Port et al., 2014). I identified seven conserved U6 sequences in the *M. paleacea* genome by BLAST searching using the conserved U6 motif from AtU6-1 as a query (Figure 5.2). The promoter region of the seven *M. paleacea* U6 was analysed for the presence or absence of restriction sites recognized

by the enzymes used in the GoldenGate assembly (Table 5.1). The promoter of the *MpaU6-1* contained a single GoldenGate compatible restriction site, and thus selected to be amplified using GoldenGate compatible and overlay extension primers. As such, 1.6kb upstream of the *MpaU6-1* was cloned from *M. paleacea* genomic DNA and inserted into the sgRNA acceptor GoldenGate module level 1. The Golden Gate compatible restriction site was removed by site directed mutagenesis using the Q5® Site-Directed Mutagenesis Kit (New England Biolabs). This promoter was used to drive the expression of the sgRNA as indicated in Figure 5.1A.

| | Restriction sites incompatible with GoldenGate cloning within 1.5kb of start | | | Substitutions from consensus sequence | Notes |
|---|---|---|---|---|---|
| | BsaI | BpiI | BsmBI | | |
| MpaU6;1 | - | - | 1,194 | 1 | |
| MpaU6;2 | - | 168 | 1,283 | 1 | |
| MpaU6;3 | 633 | 111, 1152 | - | 0 | |
| MpaU6;4 | 209, 471, 1245 | 1348 | 1348 | 0 | |
| MpaU6;5 | - | 210 | - | 6 | Contains insertions |
| MpaU6;6 | - | 219 | - | 3 | |
| MpaU6;7 | - | - | - | 4 | No G at transcription start site |

**Table 5.1 | Selection of endogenous U6 promoter in *M. paleacea* to drive sgRNA expression.**

The U6 promoter used must be compatible with GoldenGate cloning. Of the four most conserved U6 sequences in the *M. paleacea genome*, *MpaU6;1* had the fewest incompatible restriction sites in the promoter upstream of the conserved region. The remaining restriction site in the *MpaU6;1* promoter was removed by site directed mutagenesis after cloning this sequence from genomic DNA.

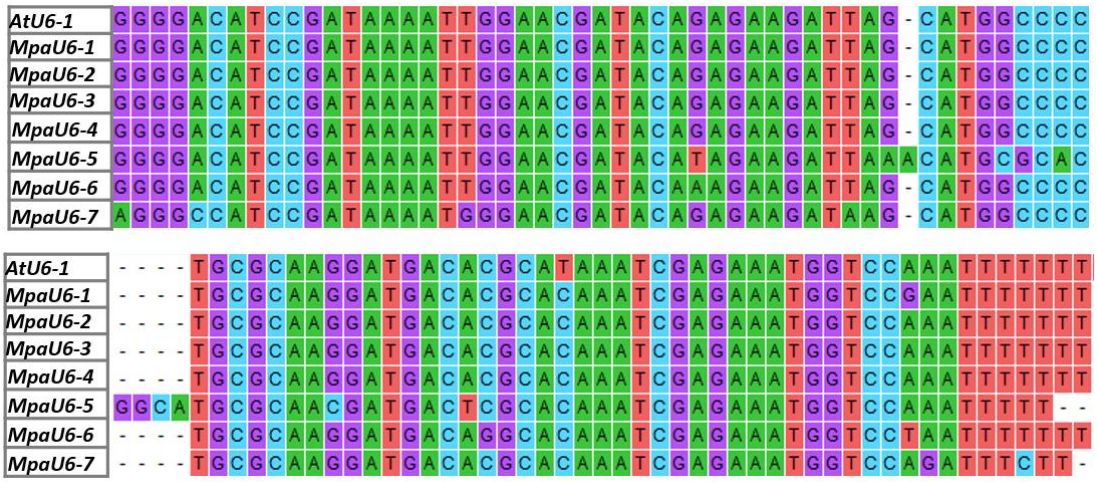



**Figure 5.2 | The conserved small nucleolar U6 RNA sequences in *M. paleacea*.**

The conserved region of U6 small nucleolar RNAs found by BLAST search of the *M. paleacea* genome were aligned to the BLAST query, AtU6-1 from *A. thaliana*. U6 sequences were found by BLAST searching the *M. paleacea* genome and aligned in MEGA7.The sequence of U6 small nucleolar RNA is typically highly conserved. Sequences that have diverged are less likely to be functional and are not preferred for driving sgRNA expression.

In order to potentially increase the efficiency of genome editing efficiency in *M. paleacea*, an alternative promoter to drive *Cas9* expression was also tested. In *M. polymorpha*, the promoter and 5' UTR of an α-Tubulin-like gene drives strong meristematic expression of fluorescent markers (Delmans et al., 2017). I PCR amplified the 1.4kb upstream of the orthologue to this gene in *M. paleacea* (named *pMpaTUBA1*) with GoldenGate compatible primer extensions and used this promoter to drive Cas9 expression in a second set of constructs for genome editing.

All the new constructs carrying the dual sgRNA driven by the *MpaU6* promoter, were transformed in *M. paleacea* via co-culture with *A. tumefaciens* AGL1. The primary transformant (T1 generation) were selected on hygromycin-containing plates and genotyped for the presence of the deletion expected. 25% of the lines transformed with the Cas9 driven by the *MpEF1a* promoter and expressing the dual *pMpaU6-1::sgRNA*

knocking out *MpaCNGC3* were mutants, and 2.5% of the line expressing the dual *pMpaU6-1::sgRNA* knocking out *MpaCNGC1* presented a deletion (Table 5.2). No mutants were obtained with the Cas9 driven by the promoter *MpaTUBa1* or targeting the regulatory motif of MpaCNGC1 and MpaCNGC3. Thus, I obtained two set of mutants; knock out for *MpaCNGC1* (2 lines) and knock out for *MpaCNGC3* (1 line).

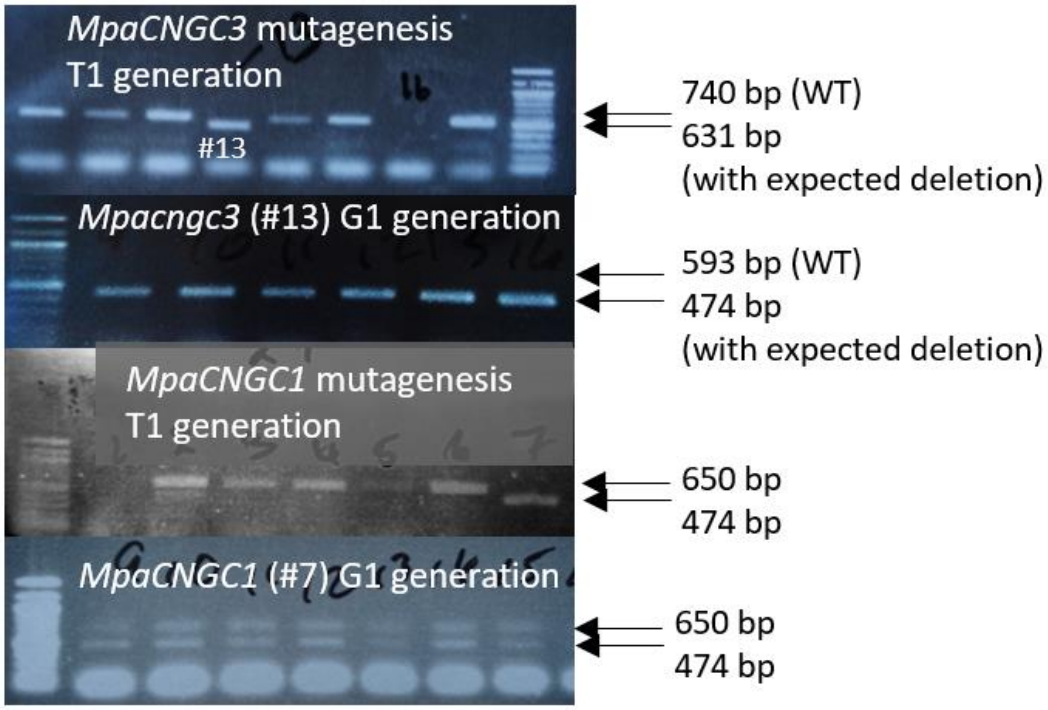



**Figure 5.3 | Genotyping by PCR of *M. paleacea* plants transformed with CRISPR/Cas9 genome editing vectors.**

PCR products amplified from the genomic region targeted that are shorter by the predicted deletion indicate successful genome editing. Ladder used in all gels is the NEB (New England Biolabs) 100 bp ladder.

The first asexual generation of those primary mutant, so called gemma (G), were further genotyped by PCR (Table 5.2). PCR products from all gemmalings grown from T1 Knock out *Mpacngc3* (#13) gemmae contained the shorter band indicating of the deletion. This band was sequenced to confirm a 109 bp deletion between bases 758 and 848 of the first exon of the coding sequence leading to a frameshift (sequence in

Appendix 3). PCR genotyping of gemmalings from T1 knock out *Mpacngc1* (#2 and #7) produced two amplicons, one indicating the presence of a deletion and a second longer amplicon characteristic of the WT gene (Figure 5.3). This indicate that the *Mpacngc1* mutants may not be stable. Both *Mpacngc1* #2 and #7 lines were maintained with the aim of finding an isogenic mutant line, while *Mpacngc3* #13 G1 gemmalings were propagated for arbuscular mycorrhiza phenotyping. Fewer lines transformed with the *MpaDMI1* CRISPR/Cas9 construct were obtained and these did not yield any mutant lines. This may have been due to non-optimal design of the guide RNAs in these constructs or simply due to the lower numbers of transformed lines obtained and screened, combined with generally low efficiency of genome editing.

| Promoter driving Cas9 | Gene targeted | sgRNA pair | Deletion type | Lines screened | Mutants found | Mutant names |
|---|---|---|---|---|---|---|
| MpoEFpro | MpaCNGC1 | 2 | Knockout | 8 | 2 | #2, #7 |
| MpoEFpro | MpaCNGC1 | 3 | CNGC binding | 32 | 0 | |
| MpoEFpro | MpaCNGC1 | 4 | CNGC binding | 24 | 0 | |
| MpoEFpro | MpaCNGC1 | 5 | CaM binding | 32 | 0 | |
| MpoEFpro | MpaCNGC3 | 1 | Knockout | 16 | 0 | |
| MpoEFpro | MpaCNGC3 | 2 | Knockout | 24 | 1 | #13 |
| MpoEFpro | MpaCNGC3 | 3 | CNGC binding | 72 | 0 | |
| MpoEFpro | MpaCNGC3 | 4 | CNGC binding | 76 | 0 | |
| MpoEFpro | MpaCNGC3 | 5 | CaM binding | 40 | 0 | |
| MpoEFpro | MpaCNGC3 | 6 | CaM binding | 24 | 0 | |
| pMpaTUa1 | MpaCNGC3 | 3 | CNGC binding | 2 | 0 | |
| pMpaTUa1 | MpaCNGC3 | 4 | CNGC binding | 6 | 0 | |
| pMpaTUa1 | MpaCNGC3 | 6 | CaM binding | 56 | 0 | |
| MpoEFpro | MpaDMI1 | 1 | Knockout | 1 | 0 | |
| MpoEFpro | MpaDMI1 | 2 | Knockout | 33 | 0 | |

**Table 5.2 | Lines screened for predicted deletions induced by CRISPR/Cas9 genome editing.**

To maximise the potential for generating mutants by CRISPR/Cas9 genome editing, multiple constructs were created for each gene and mutation type targeted. Two different promoters were used to drive Cas9 expression, the heterologous promoter previously used in *M. polymorpha*, *MpoEFpro* and an endogenous promoter, *pMpaTUa1.*

### 5.2.2 *Mpacngc3* mutants have defects in AM symbiosis

In vitro development of *Mpacngc1* chimaeras and *Mpacngc3* mutants was indistinguishable from wild type (Figure 5.4). This included asexual reproduction by gemma development and thallus development in gemmalings. There were no obvious rhizoid development phenotypes.

To assess any effect on AM symbiosis establishment, *Mpacngc3* and WT gemmalings were grown on soil in standard mycorrhization conditions, inoculated with *R. irregularis*, as described in Chapter IV. Four week old gemmalings were placed on a 90% horticultural sand/10% John Innes no. 1 compost mix with 20% inoculum grown with chive nurse plants. After 8 weeks, plants were removed from soil and washed with dH2O to remove soil particles, cleared and stained with aniline blue according to (Newsham et al., 2014). The AM fungi infect thalloid liverworts by growing through the rhizoid, possibly because this is the part of the plant in contact with the rhizosphere (Ligrone et al., 2007). Although the grid intersect method with whole root systems has been shown to be an accurate and reproducible way to quantify mycorrhizal colonisation in angiosperms, an assay that quantifies mycorrhizal colonisation has not been developed for non-vascular plants without rooting systems (Giovannetti and Mosse, 1980). Unlike roots, the area colonised by AM fungi in liverworts, the parenchymous tissue in the thallus midrib, is two dimensional and it is still unclear how much of the total thallus area can potentially be colonised by AM fungi. Therefore, instead of quantifying percent colonization of each plant, each thallus was examined under the Leica DM6000 light microscope and the approximate level of colonisation was determined. 40 WT thalli and 37 *Mpacngc3* thalli were classified as uninfected, rhizoid infection only, partially colonized or fully colonised following the criteria described in Figure 5.5. While 65% WT thalli were infected by AM fungi, only 27% of *Mpacngc3* thalli present some colonization (Figure 5.5 E). The colonization is limited to the rhizoid in 60% of the *Mpacngc3*. In 13% of the *Mpacngc3* no AM colonization was detected.

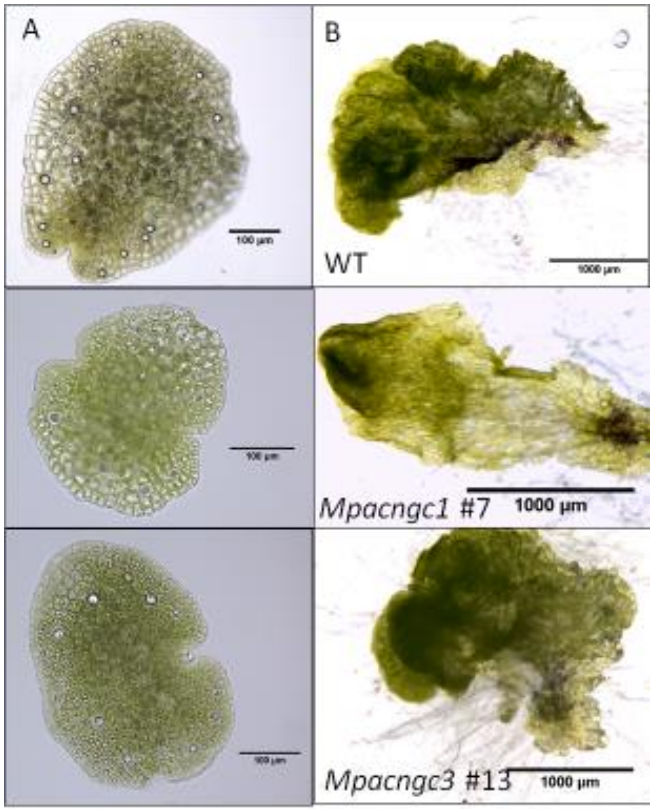



**Figure 5.4 | *Mpacngc3* and the chimeric *Mpacngc1* mutant line have normal gemma and thallus development.**

A. Gemmae taken from plants grown from the G1 generation of *Mpacngc3* #13, *Mpacngc1* #7 and WT plants. B. 15 day old gemmalings from the same grown in continuous light on sucrose free ½ B5 medium.

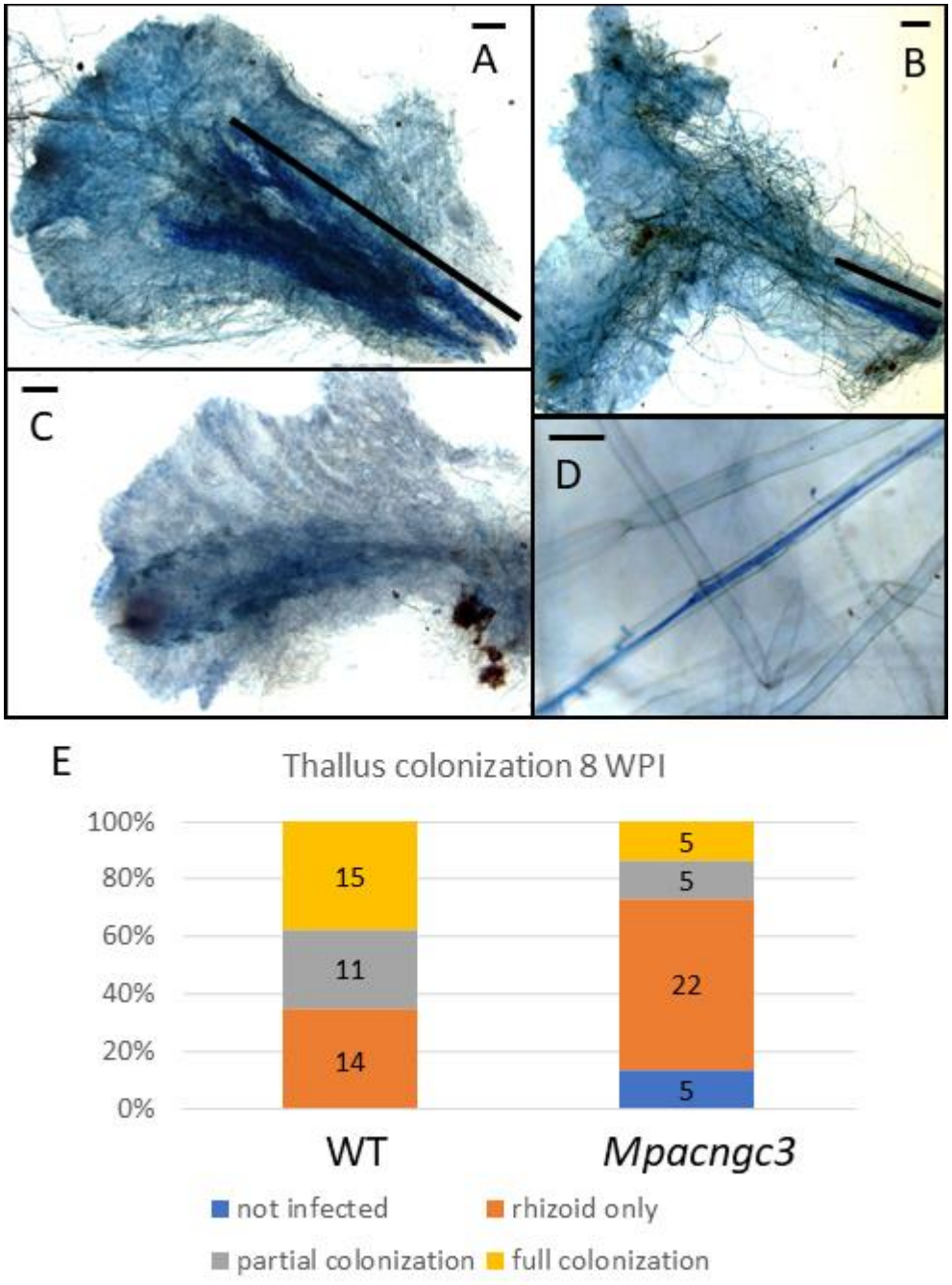



**Figure 5.5 | *Mpacngc3* is impaired in AM symbiosis.**

*M. paleacea* thalli were stained 8 weeks post inoculation with *R. irregularis*. A. Fully colonized thalli have dark blue staining indicating the presence of intracellular fungal structures covering more than half the length of the midrib (black line). B. Partially colonized thalli contain intracellular fungal structures coving less than half the length of the midrib. C. Uncolonized thalli do not have dark blue staining indicating the presence of AM fungi in the parenchymous tissue of the thallus. D. An *M. paleacea* rhizoid containing an AM fungal hypha. E. Proportion of thalli that are fully colonized, partially colonized, with rhizoid infection only and containing no fungal structures in WT (n = 40) and *Mpacngc3-1* (n = 37) plants. Statistical analysis by Fisher's Exact Test. Scale bars are 500µm in A, B and C and 50µm in D.

There was also a statistically significant reduction (p= 0.0013, Fisher's exact test) in the proportion of thalli where the parenchymous tissue of the thallus had been colonised by intra-cellular fungal structures, as seen in Figure 5.5 A and B. This indicates that AM symbiosis is impaired in *Mpacngc3* mutants, with a comparable phenotype to the *Mtcngc15abc* single mutants seen in (Charpentier et al., 2016). *Mpacngc3* plants are still capable of forming AM symbiosis, suggesting that these mutants are still capable of some symbiotic nuclear calcium signalling.

### 5.2.3 *MpaCNGC3* can complement the symbiosis phenotypes of *M. truncatula cngc15* mutants

To determine whether the molecular function of Group I/II/III CNGCs in symbiosis signalling has diverged during land plant evolution, I tested whether *MpaCNGC3,* or its paralogue *MpaCNGC1,* could rescue the defects in symbiosis seen in *M. truncatula* plants lacking one of the three *CNGC15* paralogues, *Mtcngc15a, Mtcngc15b* and *Mtcngc15c*. In these mutants nuclear associated calcium oscillations in response to symbiosis signalling molecules is impaired, with fewer cells responding and many of these cells have defects in oscillation frequency and in maintaining oscillations. As a lower proportion of cells still respond with WT nuclear associated calcium oscillations in *Mtcngc15a, Mtcngc15b* and *Mtcngc15c*, these mutants can still develop AM and root nodule symbioses but have lower colonisation levels, and form fewer nodules than wild type plants (Charpentier et al., 2016).

In this experiment, single *Mtcngc15a*, *Mtcngc15b* and *Mtcngc15c* mutants were transformed with *MpaCNGC1*, *MpaCNGC3* or an empty vector (EV) control containing only a fluorescent transformation marker (dsRED) and compared to wild type R108 plants also expressing the empty vector. Five weeks after inoculation with

*R. irregulars*, the root systems were screened for the dsRED expression and stained with the ink (Vierheilig et al., 1998). The frequency of intraradical fungal structures including arbuscules, intraradical hyphae and vesicles was quantified by grid intersect (Giovannetti and Mosse, 1980). The proportion of the rooting system containing arbuscules was significantly different between the wild type (R108) and mutants transformed with the empty vector (Figure 5.6, Appendix 2). *Mtcngc15b* mutants expressing *MpaCNGC3* had WT levels of arbuscules present in rooting systems. The arbuscule phenotype of *Mtcngc15a* was partially complemented by *MpaCNGC3* but *MpaCNGC3* did not complement *Mtcngc15c.* In all three single mutants, *MpaCNGC1* expression did not affect the symbiosis phenotype.

*MpaCNGC3* also recued the nodulation phenotypes of *Mtcngc15a* and *Mtcngc15c* (Figure 5.7). Transgenic hairy root systems were inoculated with *S. meliloti* and nodules were counted at 25 dpi. *MpaCNGC3* may also rescue the nodulation phenotype of *Mtcngc15b*, as average nodule numbers of *Mtcngc15b*/*MpaCNGC3* were higher than *Mtcngc15b*/EV but a slightly weaker mutant phenotype and small sample sizes means that this result is not statistically significant. *MpaCNGC1* did not rescue the nodulation phenotype of *Mtcngc15a*, *Mtcngc15b* or *Mtcngc15c*. These results indicate that the molecular function of Group I/II/III CNGCs in symbiosis signalling is partially conserved between *M. paleacea* and *M. truncatula*.

Percentage colonization of transgenic root systems by AM fungal structures.

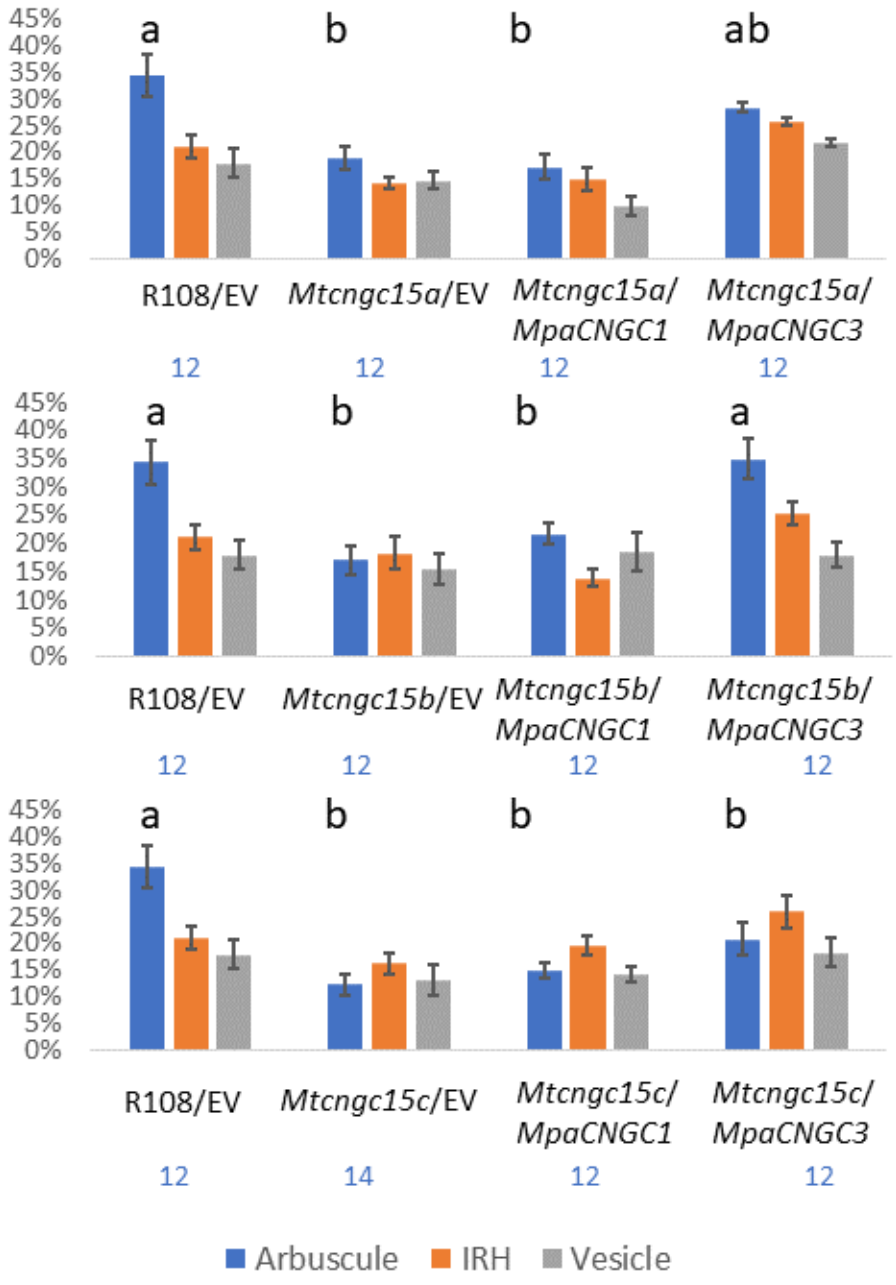



**Figure 5.6 | *MpaCNGC3* can complement the AM symbiosis phenotype of *Mtcngc15b* and partially complement *Mtcngc15a*.**

Average percentage of AM colonization in *M. truncatula* transgenic roots systems, including the wild type (R108) expressing the empty vector (EV), and single mutants *Mtcngc15a Mtcngc15b* and *Mtcngc15c* expressing empty vector (EV), *MpaCNGC1* or *MpaCNGC3* at 5 weeks post inoculation with *R. irregularis.* Blue bar: arbuscules, orange bar: intraradical hyphae, grey bar: vesicles. Letters indicate groups without statistically significant differences ($P<0.05$) according to Tukey's honest significance test after ANOVA. ANOVA found statistically significant differences only for arbuscules. Error bars represent standard error. Numbers in blue indicate sample size.

Average nodules per transformed root system.

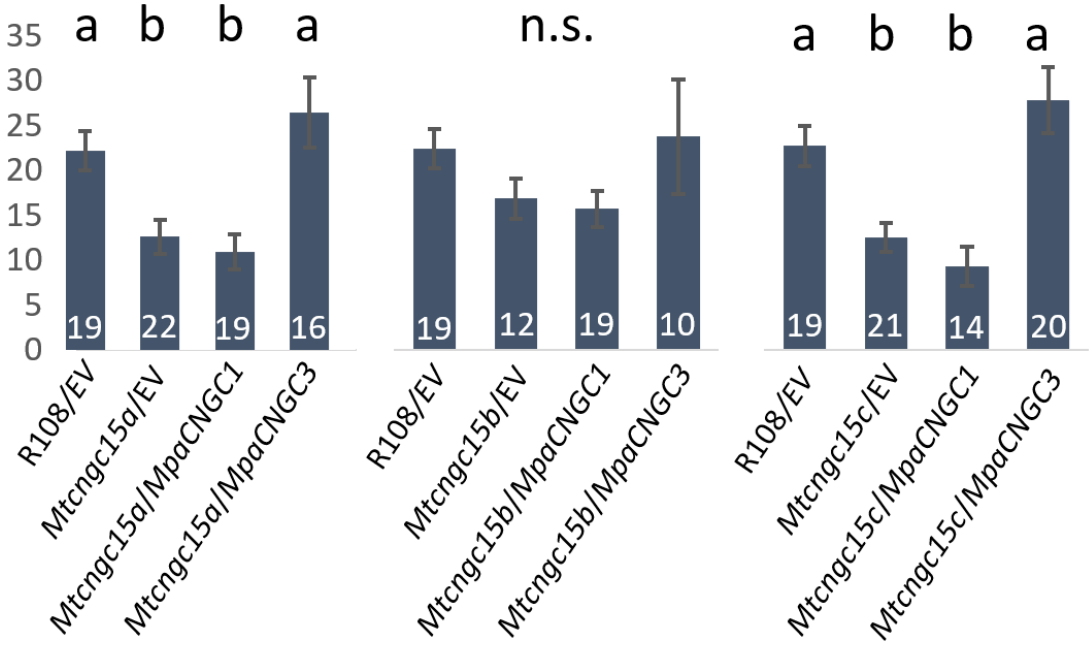



**Figure 5.7 | *MpaCNGC3* can complement the root nodule symbiosis phenotype of *Mtcngc15a* and *Mtcngc15c*.**

Average nodule number per transformed root system assessed at 25 days post inoculation with *S. meliloti.* Letters indicate groups without statistically significant differences (P<0.05) according to Tukey's honest significance test after ANOVA. Error bars represent standard error. Numbers in white indicate sample size.

## 5.3 Discussion

Although there has been rapid progress in elucidating the molecular mechanisms of AM symbiosis in the past 10-15 years, this is based entirely on studies in angiosperms, particularly the model legumes *M. truncatula* and *L. japonicus*, but also species such as the legume crops pea and soybean, rice, petunia and tomato (Delaux et al., 2013b). However, AM symbiosis is a feature of all major land plant clades apart from the mosses and was present in the common ancestor of the land plants and discoveries

made using angiosperm models may not apply to phylogenetically distant land plants (Delaux et al., 2015; Wang et al., 2010).

In this chapter, I show that the CRISPR/Cas9 system first developed in the widely used liverwort model species, *M. polymorpha*, can be adapted to induce targeted mutagenesis in a related species that forms AM symbiosis, *M. paleacea* (Sugano et al., 2018; Sugano et al., 2014). In addition to strong Cas9 expression, using a species-specific promoter to drive the expression of the dual sgRNA seems an important feature to obtain successful deletion. Thus, I could generate knock-out (KO) mutants in *M. paleacea CNGC3*. In order to measure whether the *Mpacngc3* KO mutants have mycorrhizal defects, I developed a method to quantify AM colonisation in *M. paleacea*. *M. paleacea* does not form multicellular rooting structures. Thus, AM colonize first the rhizoid and then the thallus, structures which due to a small size can be visualized using a light microscope to quantify AM. This method was used to demonstrate that the *M. paleacea* Group I/II/III CNGC identified by phylogeny and protein sequence analysis, MpaCNGC3, functions in AM symbiosis. Transformation of the model legume *M. truncatula* showed that *MpaCNGC3* could rescue the AM and nodulation phenotype of *Mtcngc15a*, but only the AM phenotype and nodulation phenotype of *Mtcngc15b* and *Mtcngc15c*, respectively. Although further replicates are required, this result suggests that other MpaCNGCs could be required to complement fully the *MtCNGC15*s, such as *MpaCNGC*4.

However, the AM symbiosis phenotype seen in *Mpacngc3* closely mirrored the phenotypes of *Mtcngc15a*, *Mtcngc15b* and *Mtcngc15c* (Charpentier et al., 2016). In *Mtcngc15a*, *Mtcngc15b* and *Mtcngc15c*, AM colonization is reduced but not abolished. This may be because some functional redundancy exists between these paralogues, as the copy number of *CNGC15* in angiosperms varies between one (rice,

*A. thaliana*) and five (*Glycine max*). In basal land plant such as *M. paleacea*, there is no direct CNGC15 orthologue and all four MpaCNGCs have potential predicted nuclear localization (Chapter III). This suggests that another MpaCNGC could contribute to generating symbiotic calcium signals the in *M. paleacea*.

*MpaCNGC3* at least partially complements the symbiosis phenotypes of all three *M. truncatula cngc15* single mutants. However, complementation of the AM symbiosis phenotype is not always associated with complementation in root nodule symbiosis. It is not clear why *MpaCNGC3* complements the AM phenotype of *Mtcngc15b* but does not complement nodulation (at least not significantly). Similarly, *MpaCNGC3* complements the nodulation phenotype of *Mtcngc15*c but not the AM phenotype. Variation in the degree to which *MpaCNGC3* could restore symbiosis in these mutants may be due to subfunctionalization or neofunctionalization among the three *CNGC15* paralogues in *M. truncatula*. A recent history of whole genome duplications has resulted in the expansion of many gene families in legumes, and widespread root nodule symbiosis in this clade may have provided evolutionary pressure for innovations in nuclear associated calcium signalling (Cannon, 2013; Cannon et al., 2006; Young and Bharti, 2012; Young et al., 2011).

# Chapter VI – General Discussion

Our understanding of the molecular mechanisms of AM symbiosis has advanced enormously over the past fifteen years. This is in part due to the development of legumes such as *M. truncatula* and *L. japonicus* as model species for molecular genetics, combined with advances in sequencing technologies. Reverse genetics approaches using transcriptome data and whole genome sequences of both mycorrhizal and non-mycorrhizal species have elucidated multiple genetic mechanisms necessary for AM symbiosis. As further plant and green lineage genomes are published, we have had greater insight into the evolution of AM symbiosis, but we still lack many of the molecular genetics resources of angiosperm model species for the functional analysis of molecular genetics in basal land plants.

The aim of this study was to develop the *M. paleacea* model system for studying AM symbiosis, to functionally analyse the conservation of putative symbiosis genes in this species and in doing so learn more about the evolution of AM symbiosis since the divergence of liverworts and angiosperms in the land plants.

In Chapter III, publicly available genome and transcriptome sequences are combined with the newly sequenced *M. paleacea* genome to identify putative symbiosis genes in a liverwort species for functional analysis. The expression of putative symbiotic gene expression, in Chapter IV, is then compared to expression patterns of known symbiotic genes from angiosperm model species.

Molecular genetics approaches are used to determine whether a potential symbiosis gene functions in AM symbiosis. In Chapter V, CRISPR- Cas9 genome editing is used to generate symbiosis mutants in a non-flowering plant system and the function of

symbiotic calcium signalling genes is shown to be conserved over 450 million years, based on complementation of *M. truncatula* mutants. The findings of this thesis are summarised in Table 6.1, with the specific findings for each gene family given in Table 6.2.

| | |
|---|---|
| **Chapter III** | Identified putative AM symbiosis genes by bioinformatic analyses. |
| **Chapter II, Chapter IV** | Established a staining protocol previously used to stain for ericoid mycorrhizal fungi in the liverwort *Cephaloziella varians* that could be used to demonstrate whole plant colonisation by AM fungi in *M. paleacea* (Upson et al., 2007). |
| **Chapter IV** | Demonstrated that genes with putative functions in arbuscule development and function in *M. paleacea* were upregulated during colonization with AM fungi. |
| **Chapter IV** | Identified spatial patterns of gene expression for *MpaDMI1* and the members of the CNGC gene family in *M. paleacea*. |
| **Chapter V** | Established functioning CRISPR-Cas9 genome editing in *M. paleacea*. |
| **Chapter V** | Generated a knockout mutant in *Mpacngc3* with impaired AM symbiosis. |
| **Chapter V** | Demonstrated partial conservation of function between *MpaCNGC3* and *MtCNGC15a*, *MtCNGC15b* and *MtCNGC15c*. |

**Table 6.1 | Summary of the major findings in this thesis.**

| *M. truncatula* | *M. paleacea* | Homology | Mpa gene expression | Molecular genetics |
|---|---|---|---|---|
| *MtDMI1* | *MpaDMI1* | 2 paralogues in flowering plants | Constitutive | |
| *MtCNGC15s* | *MpaCNGC3* Potentially other *MpaCNGC*s | Independent expansions of Group I/II/III CNGCs | Constitutive | Significant evidence for conserved function in calcium signalling |
| *MtVAPYRIN* | *MpaVAPYRIN* | Homologous | Upregulated by AM colonisation | |
| *MtEXO70I* | *MpaEXO70GI* | 2 paralogues in *M. truncatula* | Slightly upregulated by AM colonisation, higher basal expression | |
| *MtPT4* | *MpaPHT1D* | Same gene family only | Upregulated by AM colonisation | |
| *MtSTR* | *MpaSTR* | Homologous | Upregulated by AM colonisation | |

**Table 6.2 | Summary of the findings for genes with putative AM symbiosis function in *M. paleacea* identified in this thesis.**

## 6.1 On phylogenetic analysis of putative symbiotic genes in *M. paleacea*

Phylogenetic analysis of symbiosis genes in the green lineage has previously shown that symbiosis signalling genes were present in the common ancestor of the land plants and some components of the CSSP, namely the calcium signal decoder *CCaMK*, even predate the origin of land plants and AM symbiosis (Delaux et al., 2015; Wang et al., 2010). Phylogenetic analysis in Chapter III confirmed many of these previous findings.

### 6.1.1 Phylogenetic and sequence analysis of calcium signalling machinery in symbiosis

This thesis confirmed that *DMI1* like sequences are present in the charophyte algae, and that *DMI1* was likely to have been present as a single copy in the common ancestor of the land plants, as shown in (Delaux et al., 2015). By including gymnosperm sequences available from the 1KP database, this analysis also showed that the gene duplication event that led to the *CASTOR* and *POLLUX*/*DMI1* paralogues in angiosperms took place in the common ancestor of the seed plants, with independent duplications also occurring in the lycophyte and moss lineages, which were also seen in (Wang et al., 2010).

This duplication event may be associated with an amino acid transition in the selection filter region of the potassium channel. While most angiosperms require *CASTOR* and *POLLUX* to generate symbiotic nuclear associated calcium oscillations, only the *POLLUX* orthologue, *MtDMI1*, is required in *M. truncatula*, with RNAi of *MtCASTOR* having effect on symbiosis (Venkateshwaran et al., 2012). *LjPOLLUX* can gain a similar function to *MtDMI1* through a single amino acid transition in the selection filter, from ADSGNHA, which is seen in most angiosperms, to ADAGNHA. The ADAGNHA sequence is also retained in liverwort single copy *DMI1* sequences and most, but not all, charophyte algae sequences. Where *DMI1* duplications have occurred independently, both paralogues in the lycophyte *S. moellendorfii*, and a single paralogue each in *S. fallax* and *P. patens* also have the ADSGNHA type selection filter. However, the common ancestor of the land plants was likely to have had a single copy of *DMI1* with function in AM symbiosis signalling, so this gene is likely to have had the ADAGNHA sequence present in *M. paleacea*. It is unclear why

duplication of *DMI1* have evolved with the ADSGNHA selection filter on multiple occasions in the history of land plant evolution.

Although DMI1 and CNGC15 both physically interact and function in generating nuclear associated calcium signals, CNGCs have a very different evolutionary history in the land plants. The analysis in Chapter III found that there are three main subfamilies of CNGC in the land plants, with bryophyte, lycophyte and angiosperm representatives, with each family represented by a single gene in the common ancestor of the land plants. This is consistent with the phylogeny in (Charpentier et al., 2016) which was rooted using *Chlamydomonas reinhardtii* but differs slightly from the phylogeny in (Saand et al., 2015). As the analysis in this thesis used additional bryophyte genomes and charophyte transcriptomes, the phylogeny in Figure 3.1 is likely to be accurate.

In the angiosperms, Group IVa and Group IVb have relatively few members compared to the Group I, II and III CNGCs. In angiosperms, Group IVa CNGCs may be involved in mediating abiotic stress and defence responses. In *A. thaliana*, the IVa CNGCs *AtCNGC19* and *AtCNGC20* are upregulated in the shoot in response to salt stress, and in response to the bacterial elicitor, flg22 (Kugler et al., 2009; Moeder et al., 2011). In tomato, silencing of *SlCNGC15*, also a member of Group IVa, reduces the drought response, consistent with a role in osmotic stress responses (Saand et al., 2015). In the moss *P. patens*, two of the three IVa CNGCs have roles in thermosensing and activating heat shock transcription, which is a role not seen in Group IVa CNGCs in angiosperms to date (Finka and Goloubinoff, 2014). This may be due to acquisition of new functions during the expansion of Group IVa CNGCs in the moss lineage. The single member of the Group IVa CNGCs seen in *M. paleacea, MpaCNGC4*, is therefore likely to function in the response to biotic or abiotic stress.

In *A. thaliana*, mutants of the two IVb CNGCs, *Atcngc2* and *Atcngc4*, result in a constitutive immune response in leaves, impaired effector mediated hypersensitive response, dwarfism when grown on soil or media containing calcium and late flowering (Balagué et al., 2003; Chan et al., 2003; Jurkowski et al., 2004; Wang et al., 2017). This pleiotropic phenotype may occur because AtCNGC2 and AtCNGC4 form a complex that imports calcium from the apoplast to leaf cells, so mutants of either gene accumulate in toxic levels of calcium in the apoplast (Balagué et al., 2003; Chin et al., 2013). Impaired calcium import to the cytosol may affect use of calcium as a secondary messenger, resulting in defects in plant defence responses and developmental processes that rely on calcium signalling. If the Group IVb CNGC in *M. paleacea*, *MpaCNGC2*, has a similar function, *Mpacngc2* mutants may still present with a different phenotype as the role of calcium signalling in reproductive development and pathogen responses in liverworts is unknown.

A single CNGC underwent duplication to form the Group I, Group II and Group III CNGCs after the evolution of vascular plants but before the evolution of angiosperms. In angiosperms, this subfamily of CNGCs has diverse functions, and the function of the common ancestor of Group I, Group II and Group III CNGCs is unclear. The Group I CNGCs, *AtCNGC1*, *AtCNGC10* and *AtCNGC3* function in ion uptake for plant nutrition with *AtCNGC1* and *AtCNGC10* also involved in uptake of deleterious ions such as $Na^+$ and $Pb^{2+}$ (Borsics et al., 2007; Gobert et al., 2006; Guo et al., 2010; Jin et al., 2015; Ma et al., 2006; Sunkar et al., 2000). Two genes generated by a tandem repeat, *AtCNGC11* and *AtCNGC12*, function in regulating dark induced leaf senescence, root gravitropism and potentially in the hypersensitive response to R-gene containing pathogens (Urquhart et al., 2011; Urquhart et al., 2007).

Several Group II CNGCs have been described in *A. thaliana*. AtCNGC6 is activated by heat shock, leading to expression of heat shock proteins required for thermotolerance (Gao et al., 2012). A double mutant of *Atcngc5/6* was impaired in generating cation currents across the guard cell plasma membrane in response to externally applied membrane permeable cGMP (Wang et al., 2013). A double knockout of *Atcngc7/8* has impaired pollen tube growth with a high frequency of pollen grains bursting on germination (Tunc-Ozdemir et al., 2013).

Group III CNGCs form three distinct clusters in angiosperms. One cluster contains *CNGC15* genes, which function in symbiosis signalling in *M. truncatula* (Charpentier et al., 2016). In *A. thaliana* a second cluster is comprised of *AtCNGC14* and *AtCNGC17*, which function in root development and a third cluster contains two genes, *AtCNGC16* and *AtCNCG18*, necessary for male gametophyte development (Gao et al., 2016; Ladwig et al., 2015; Tunc-Ozdemir et al., 2013; Zhang et al., 2017). The common ancestor of the Group I, II and III CNGCs is likely to have functioned in regulation of development as multiple genes in this subfamily function in root and male gametophyte development, as well as the developmental response of accommodating endosymbionts. The expansion of this gene subfamily may be associated with developmental innovations in the seed plant lineage, such as the evolution of pollen tubes that deliver non-motile sperm to the egg, a feature of male gametophyte development in angiosperms, conifers and Gnetales (Friedman, 1993; Pettitt, 1977). Calcium signalling is essential to normal pollen germination and tube growth in angiosperms, with male fertility defects seen in multiple Group II and Group III CNGC mutants (Steinhorst and Kudla, 2013).

Although the intron-exon structure of Group I/II/III is variable in angiosperms, ranging from 3 introns (*ZmCNGC1* in *Zea mays*) to 13 introns (*MdCNGC24* in *Malus*

*domestica*), the intron exon structure of symbiosis associated *MtCNGC15a*, *MtCNGC15b* and *MtCNGC15c* in *M. truncatula* is conserved with one of the Group I/II/III CNGCs in *M. paleacea*, *MpaCNGC3* (Saand et al., 2015). Both *MtCNGC15a*, *MtCNGC15b* and *MtCNGC15c*, and *MpaCNGC3* are composed of 6 introns and 7 exons. In contrast, the other Group I/II/III CNGC in *M. paleacea*, *MpaCNGC1* does not contain any introns and has lower protein sequence similarity to *MtCNGC15a*, *MtCNGC15b* and *MtCNGC15c* than *MpaCNGC3*. Both genes contained a nuclear localization signal (NLS) indicating potential function in symbiosis. The symbiotic CNGCs in *M. truncatula* were initially identified by a bioinformatics approach that looked to identify transmembrane proteins containing motifs associated with $Ca^{2+}$ channels and ion channels, and NLS motifs (Charpentier et al., 2016). As *MpaCNGC3* also fits these criteria, it is the strongest candidate for a symbiotic CNGC in *M. paleacea*. This was confirmed by functional analysis in Chapter V, showing that a bioinformatics approach incorporating phylogeny, protein sequence similarity and localization motifs can identify genes with conserved function in phylogenetically distant land plant clades.

The phylogenetic analysis of the CNGC family in Chapter III also included charophyte algae. This analysis showed that the three main groups of CNGCs seen in the land plants (Group I/II/III, Group IVa and Group IVb) are also present in charophyte transcriptomes. Previous phylogenetic analyses of CNGCs have not included charophyte transcriptome data but by including this I was able to place the evolution of the three main CNGC groups to earlier in the green lineage, before the evolution of land plants. None of the transcriptomes analysed contained representatives of all three groups. Although transcriptome data typically will not cover all genes in a species genome, this may indicate that secondary gene loss within the CNGC gene family has

taken place in charophyte algae. Further analysis of CNGC function in charophyte algae is becoming possible, as emerging charophyte model species such as *Klebsorbidium*, *Micrasterias*, *Penium*, *Chara* and *Coleochaete* are developed for molecular biology.

I was unable to find CNGC channels containing CaM binding domains in any publicly available chlorophyte genome sequence (*Chlamydomonas reinhardtii* v5.5, *Dunaliella salina* v1.0, *Volvox carteri* v2.1, *Coccomyxa subellipsoidea* C-169 v2.0, *Micromonas pusilla* CCMP1545 v3.0, *Micromonas sp.* RCC299 v3.0.and *Ostreococcus lucimarinus* v2.0). It is possible that the weak hit to CNGC like channels previously identified in *C. reinhardtii* in (Verret et al., 2010) are more closely related to the structurally similar Shaker-like $K^+$ channels (Talke et al., 2003). Shaker-like $K^+$ channels differ functionally from CNGCs as they are primarily voltage gated, with activity modified by CNG binding (Ward et al., 2009).

### 6.1.2 Phylogenetic analysis of arbuscule development and nutrient transport associated genes

In Chapter III I also identified putative arbuscule associated genes in *M. paleacea*. I found single orthologues in the *M. paleacea genome* for one arbuscule development gene, *VAPYRIN*, and the half ABCG transporters *STR* and *STR2* that are thought to function in export of fatty acids from the plant cell for uptake by AM fungi.

This finding is consistent with (Delaux et al., 2015), which also identified orthologues to these genes in the transcriptome of the liverwort *Lunularia cruciata*. The presence of *VAPYRIN* and *STR*/*STR2* is significant because these genes have a phylogenetic pattern that includes most AM symbiosis forming plants but excludes most non-AM host species, and have been used as positive controls for AM symbiosis in

phylogenomics pipelines to identify AM symbiosis genes (Bravo et al., 2016; Pumplin et al., 2010a; Zhang et al., 2010).

This analysis also confirmed that *VAPYRIN* evolved in the common ancestor of the land plants but after the land plant lineage had diverged from extant charophyte algae species, as noted in (Delaux et al., 2015). This analysis also identified a group of charophyte algae ABCGs that may be more closely related to STR and STR2 than other half ABCG transporters in the land plants. Although the statistical support for this node in the phylogeny presented in Figure 3.7 is low (tree branch support from aBayes analysis = 0.5904), a similar cluster of charophyte half ABCG transporters were identified in (Delaux et al., 2015). These may represent a proto-*STR/STR2* that was recruited specifically for fungal nutrition during AM symbiosis.

Although the precise function of *VAPYRIN* is unknown, this gene is essential for both infection by endosymbionts (AM fungi and rhizobia in root nodule symbiosis) and for the development of intracellular structures required for functional symbiosis (arbuscules in AM symbiosis and symbiosomes in root nodule symbiosis) (Murray et al., 2011a; Pumplin et al., 2010a). *VAPYRIN* is composed of two protein-protein interacting domains, an MSP domain and an ankyrin repeat domain and VAPYRIN protein has been shown to interact with a symbiosis specific component of the exocyst complex, Exo70I (Pumplin et al., 2010a; Zhang et al., 2015). This suggests that VAPYRIN functions as a protein scaffold during secretion of a membrane by the plant cell to accommodate intracellular symbionts. STR and STR2 localise to the periarbuscular membrane (PAM), where they are required to maintain AM symbiosis (Zhang et al., 2010). The substrate that these transporters export to the periarbuscular space is unknown but they are thought to export lipids required for AM fungi nutrition

as part of the arbuscule specific lipid biosynthesis pathway (Bravo et al., 2017; Luginbuehl and Oldroyd, 2017).

*VAPYRIN*, *STR* and *STR2* are typically present as single copies in angiosperms forming AM symbioses. This pattern is not seen with all symbiosis associated genes. For example, the plasma membrane receptors that detect Myc factors are members of the LysM RLK family that has undergone repeated expansion and gene loss in the angiosperms, with no definitive "Myc factor receptor" gene that is required for AM symbiosis in all species (Gough et al., 2018). The PHT1 phosphate transporters required for AM symbiosis are another example; while a symbiosis specific clade containing *MtPT4* and *OsPT11* is conserved across angiosperms, several angiosperm species also require additional PHT1 phosphate transporters for functional AM symbiosis (Javot et al., 2007a; Karandashov et al., 2004; Paszkowski et al., 2002). The legume *Astragalus sinicus* contains an *MtPT4* orthologue but requires a second transporter, *AsPT1*, for normal levels of root colonisation and arbuscule development (Xie et al., 2013). As well as *OsPT11*, second phosphate transporter, *OsPT13* required for AM symbiosis in rice (*O. sativa*) (Paszkowski et al., 2002; Yang et al., 2012). *OsPT13* is a member of a monocot specific expansion of the PHT1 gene family (named clade IV in (Nagy et al., 2005)) which does not have a direct orthologue in other land plant clades (Nakagawa and Imaizumi-Anraku, 2015; Paszkowski et al., 2002; Yang et al., 2012).

While the molecular function of AM symbiosis genes from these gene families may not always be conserved, genes such as *VAPYRIN*, *STR* and *STR2* that are typically present as single copies are more likely to have a conserved role in AM symbiosis.

In contrast to *STR*, *STR2* and *VAPYRIN*, there is not one to one orthology between the phosphate transporter and Exo70 with putative AM function in the *M. paleacea* genome, and those identified to function in symbiosis in angiosperm model species.

*MtExo70I* is a component of the exocyst complex that is specifically upregulated during AM symbiosis and required for the development of fine branches in the arbuscule (Zhang et al., 2015). The *Exo70* gene family has expanded in the angiosperms but the *Exo70I* cluster is found only in species that form AM symbiosis (Zhang et al., 2015). *Exo70I* and *Exo70G* form one of three main *Exo70* groups in the angiosperms, named Exo70.3 (Cvrčková et al., 2012). This thesis showed that the Exo70.3 group is represented by a single member in *M. paleacea*, named *MpaExo70GI*. The gene duplication event that resulted in *Exo70G* and *Exo70I* may have taken place after the evolution of seed plants, as the Exo70.3 genes represented in available gymnosperm genomes form an independent expansion (see extended Exo70 phylogeny, appendix 2).

It is currently unknown whether *Exo70I* functions to define a symbiosis specific secretory pathway during arbuscule development, or if its function in AM symbiosis is due to dominant expression in arbuscule containing cells (Zhang et al., 2015). The SNARE protein complex, which also function in vesicle targeting during exocytosis, also contains symbiosis specific components that are specifically expressed in arbuscules, but these genes are are functionally redundant with non-symbiotic paralogues expressed under an arbuscule specific promoter (Huisman, 2018; Huisman et al., 2016; Ivanov et al., 2012). As the specific protein composition of the PAM is largely determined by changes in the transcription of cargo proteins during arbuscule development, a specifically defined symbiosis secretion pathway may not exist (Pumplin et al., 2012). However, the specific interaction of MtVAPYRIN with

MtExo70I, but not with MtExo70A1 or MtExo70B2, suggests that Exo70I has a unique role in generating the of finely branched architecture characteristic of arbuscules (Zhang et al., 2015).

Unlike *M. polymorpha* and *M. paleacea*, the genomes of moss species *P. patens* and *S. fallax* contain several Exo70.3 members, despite the absence of AM symbiosis from the moss lineage. Mutants of the *P. patens* gene *Ppexo70.3d* have severe defects in reproductive development, bud and gametophore development, cell elongation and differentiation in protonema and altered auxin production and responses (Rawat et al., 2017). The function of non-symbiotic Exo70.3 genes in angiosperms is not clear. While that *A. thaliana* mutant *Atexo70g1* has no obvious phenotype, silencing of *NbExo70G* in *Nicotiana benthamiana*, causes growth retardation during leaf development and reduces the callose deposition response to bacterial pathogens (Du et al., 2018; Lukáš et al., 2006). No Exo70.3 genes were found in the genome of the lycophyte *S. moellendorfii*, despite reports of AM symbiosis in multiple species of the *Selaginella* genus (Lee et al., 2001; Muthukumar and Prabha, 2013). However, the published *S. moellendorfii* genome contains gaps covering 1.9% of the estimated genome size so it is possible that an Exo70.3 is present in the genome but absent from the current assembly (Banks et al., 2011).

The members of the Exo70.1 and Exo70.2 subfamilies in the angiosperms have diverse roles in organogenesis, polarized cell growth, vascular development and pathogen defence (Hala et al., 2008; Koki et al., 2015; Li et al., 2013; Lukáš et al., 2006; Tu et al., 2015). As the composition of the Exo70 gene family in *M. paleacea* and *M. polymorpha* is likely to reflect that of the common ancestor of the land plants, functional analysis of the Exo70 gene family in these species would provide insight into the ancestral functions of these genes. The three Exo70 subfamiles in land plants

were also shown to be present in charophyte algae in Chapter III, suggesting that the functions of these subfamilies pre-dates the evolution of land plants.

Phylogenetic analysis of the PHT1 phosphate transporters could not identify a clear orthologue to the symbiosis specific phosphate transporters *MtPT4* and *OsPT11* in the *M. paleacea* genome. *MtPT4* and *OsPT11* represent a group of phosphate transporters conserved within the angiosperms that function in phosphate uptake at the PAM during AM symbiosis (Harrison et al., 2002; Javot et al., 2007a; Paszkowski et al., 2002; Yang et al., 2012). The PHT1 phosphate transporters have expanded independently in the angiosperms, gymnosperms and basal land plants, with potential recruitment of phosphate transporters for AM symbiosis function in each of these expansions. A homologue to PHT1 phosphate transporter upregulated during AM symbiosis in the liverwort *L. cruciata* was identified in the *M. paleacea* genome but not in the *M. polymorpha* genome, suggesting that this phosphate transporter also functions specifically in AM symbiosis. In *M. polymorpha*, the PHT1 gene family has undergone recent expansion with four PHT1 genes generated by a tandem repeat on scaffold 0195 of the current genome assembly (*Marchantia polymorpha v3.1*). This expansion of PHT1 phosphate transporters occurred after the divergence *of M. paleacea* and *M. polymorpha* and may represent an adaptation to non-AM mediated phosphate uptake by *M. polymorpha* as a weedy colonizer of disturbed and man-made environments, as well as stream and riverbanks, as suggested in (Bowman et al., 2017).

## 6.2 On the analysis of putative symbiosis gene expression in *M. paleacea*

In Chapter IV, I showed that the four putative arbuscule associated genes in *M. paleacea* identified in Chapter III are specifically upregulated during AM symbiosis. Specific upregulation during AM symbiosis is a hallmark of genes that function

specifically in the development and function of arbuscules and was instrumental in the identification of AM symbiosis genes including the nutrient uptake genes *PT4* and *HA1*, the exocyst component *Exo70I* and the fatty acid biosynthesis gene *FatM* (Harrison et al., 2002; Krajinski et al., 2014; Krajinski et al., 2002; Luginbuehl et al., 2017; Zhang et al., 2015).

In *M. paleacea*, the putative symbiosis genes *MpaPTD*, *MpaVAPYRIN* and *MpaSTR* all showed similar patterns of expression, with expression occurring specifically at 42 dpi with *R. irregularis*, when the fungus had colonized and developed intracellular structures in the parenchymous tissue across most of the thallus and, much lower levels of expression at other time points.

The fold change in expression for *MpaPTD* was not as dramatic as that seen in angiosperm model systems. Where *MtPT4* expression may be increased as much as 1038-fold *by R. irregularis* inoculation (Breuillin-Sessoms et al., 2015; Hogekamp et al., 2011; Luginbuehl et al., 2017).

It is possible that the difference in upregulation of symbiotic phosphate transporters is due to functional differences, or differences in regulation, between *MpaPTD* and *MtPT4*. This is further suggested by the phylogenetic analysis of Chapter III, which indicates that they were independently recruited for specific function in AM symbiosis. However, there is also a difference in methodology between the quantitative RT-PCR analysis in this thesis and methodology used in gene expression analysis during AM symbiosis in angiosperms. The RNA content of root systems only is extracted for quantitative RT-PCR analysis in angiosperm model plants, while in the liverwort *M. paleacea*, there are not distinct shoot and rooting systems, so RNA was extracted from whole plants. Instead, the gametophyte that forms AM symbiosis is comprised of a flat, prostrate thallus that accesses the rhizosphere through single

celled rhizoids growing from the ventral surface. Greater dilution of AM specific transcripts may therefore occur in in *M. paleacea* samples.

The expression pattern of *MpaExo70GI* during AM symbiosis also differed from that observed in *MtExo70I*. While *MtExo70I* is expressed specifically during AM symbiosis and upregulated approximately 18-fold by colonization with AM fungi, *MpaExo70GI* was upregulated 2.9-fold relative to the control in Chapter IV (Luginbuehl et al., 2017; Zhang et al., 2015). This may indicate that *MpaExo70GI* is constitutively expressed at a lower level due function outside of AM symbiosis, possibly resembling those of the *Exo70G* cluster, which like *Exo70I*, is a co-orthologue to *MpaExo70GI* in the angiosperms. Although the function of the Exo70G cluster is unknown, in *A. thaliana*, *AtExo70G1* is expressed in the QC, endodermis, and inner tissues of the root, and *AtExo70G2* is expressed in developing xylem elements, pollen and microspores, suggesting functions in cell differentiation (Li et al., 2010).

In general, the expression of putative arbuscule development and nutrient exchange genes identified by phylogenetic analysis was consistent with function in AM symbiosis, and in the case of *MpaExo70GI*, additional functions in *M. paleacea*. This validates the results of Chapter III and indicates conserved function of molecular mechanisms at the arbuscule for 450 million years of land plant evolution.

In contrast, the putative functional homologue for *MtDMI1* in *M. paleacea*, *MpaDMI1,* was not upregulated by AM colonization, and upregulation by AM symbiosis was not seen for any of the CNGC genes present in the *M. paleacea* genome that could act as functional homologues to *MtCNGC15s* that function in symbiotic signalling. This was consistent with expression analysis of *MtDMI1*, which is upregulated in the nodule meristem during root nodule symbiosis but is not

upregulated during AM symbiosis, with constitutive expression in the roots of both *M. truncatula* and the non-symbiotic plant *A. thaliana* (Ane et al., 2004; Limpens et al., 2013; Riely et al., 2007). Expression analysis of *MtCNGC15a*, *MtCNGC15b* and *MtCNGC15c* during AM symbiosis has not been published but these genes are expressed in the root of uninoculated plants (Charpentier et al., 2016).

The spatial expression patterns of *MpaDMI1* and of all CNGC genes in the *M. paleacea* genome (*MpaCNGC1*, *MpaCNGC2*, *MpaCNGC3* and *MpaCNGC4*) were analysed by staining for GUS activity in transgenic plants containing constructs in which the promoters of these genes were fused to the coding sequence of the *GUS* gene. *MpaDMI1* promoter activity was seen in gemmae, the asexual reproductive structures of *M. paleacea*, as well as globally in the developing thallus, including in the rhizoids. A distinct pattern of *MpaDMI1* promoter activity, localised to the thallus midrib and to the edges of the thallus between the apical area and the point of bifurcation during thallus development, was seen in 8 week old thalli. The expression pattern of *MpaDMI1* at the thallus midrib is consistent with a potential function in AM symbiosis. AM fungi infect through the rhizoids and colonise the parenchymous tissue of the thallus along the midrib, where they form arbuscule like intracellular structures. This suggests an expression pattern in tissues with potential for AM colonization, similar to the constitutive expression of *MtDMI1* in uninoculated root systems (Ane et al., 2004). The significance of *MpaDMI1* expression at the sides of the thallus, away from the midrib or growth apices is unclear, although this area of the thallus has been observed as a site of adventitious bud development in *M. polymorpha* in response to treatments such as UV-light, X-rays, drying and plasmolysing solution (Dickson, 1932).

*MpaCNGC1*, *MpaCNGC3* and *MpaCNGC4* promoter activity were not seen in gemmae but were seen throughout the developing thallus, including in the rhizoids. This expression pattern overlaps with that of *MpaDMI1*, indicating that these calcium channels could function with *MpaDMI1* in symbiotic nuclear associated calcium signalling. *MpaCNGC2* expression was restricted to the cells surrounding the air pores. Air pores in liverworts facilitate a balance between gas exchange and water loss, and to prevent liquid from outside penetrating the intercellular space of the air chambers at the dorsal surface of the thallus (Shimamura, 2016). Air pores have been recently shown to facilitate colonization of *M. polymorpha* by the oomycete pathogen *Phytophthora palmivora* (Carella et al., 2018). While stomata in other land plant clades are not developmentally homologous to air pores, they function similarly in gas exchange and are also a point of entry for pathogens, including oomycetes (Jones and Dolan, 2017; Kiefer et al., 2002; Zeng et al., 2010).

Overlapping expression of *MpaDMI1* and *MpaCNGC*s in the areas with potential for colonisation by AM function is consistent with a role in symbiont perception in the common symbiotic signalling pathway. However, unlike the specific expression of arbuscule associated genes during AM colonization, the expression patterns of *MpaDMI1* and *MpaCNGC*s do not provide direct evidence for function in AM symbiosis. The promoter sequences used here were also modified with nucleotide substitutions for compatibility with the GoldenGate modular cloning system, so the GUS activity pattern seen here may not faithfully recapitulate endogenous expression of these genes (Engler et al., 2014).

## 6.3 Functional conservation of symbiosis signalling between *M. paleacea* and *M. truncatula*

Through developing the CRISPR/Cas9 system for use in *M. paleacea* in Chapter V, I was able to generate a mutant line with measurable defects in AM symbiosis in a basal land plant.

In developing this system, I demonstrated the importance of promoter selection when driving expression of the sgRNA that directs Cas9 activity during genome editing. While in *M. polymorpha* sgRNA expression driven by the endogenous *MpU6-1* promoter can achieve up to 70% efficiency in genome editing, use of this promoter to drive sgRNA expression for genome editing in the closely related species *M. paleacea* was unsuccessful (Sugano et al., 2018). By instead using an endogenous *M. paleacea* U6 promoter sequence to drive sgRNA expression, I achieved genome editing efficiency of up to 25%. The U6 small nucleolar RNA promoters may not always function well in heterologous systems, as U6 promoters are less efficient when driving expression in distantly related angiosperm species (Wang et al., 2008). The efficiency of U6 and U3 promoters in CRISPR/Cas9 genome editing can vary even within the same species (Liang et al., 2016; Port et al., 2014). As the genome editing efficiency achieved in this study was still lower than that achieved using the same promoter and Cas9 compbinations (Sugano et al., 2018), analysis of the differential activity of all U6 promoters present in the *M. paleacea* genome may be required for optimization of CRISPR/Cas9 genome editing in this species.

The *Mpacngc3* knockout mutant generated by CRISPR/Cas9 genome editing was found to have a quantitative effect on AM symbiosis. While AM symbiosis was not abolished, the number of *Mpacngc3* mutant plants with colonisation of the thallus at 8 wpi with *R. irregularis* was significantly lower than in the wild type. Many *Mpacngc3* plants lacking colonisation of the thallus by AM fungi had fungal hyphae present in rhizoids, the site of initial infection by AM fungi in liverworts (Ligrone et

al., 2007; Ligrone and Lopes, 1989). However, as the arbuscule like structures where nutrient exchange takes place develop in the parenchymous tissue of the thallus, rhizoid infection alone is probably not sufficient for functional symbiosis. While AM fungal structures were present in 100% of wild type plants, there were no signs of fungal infection in 13% of *Mpacngc3* plants. This suggests that in *M. paleacea*, *MpaCNGC3* functions in the colonization of thallus tissue and the initial infection of rhizoids.

The molecular function of *MpaCNGC3*, but not its closest paralogue *MpaCNGC1*, is partially conserved with calcium channels that function in generating calcium oscillations in the common symbiosis signalling pathway in the angiosperm *M. truncatula*, *MtCNGC15a*, *MtCNGC15b* and *MtCNGC15c* (Chapter V). *MpaCNGC3* fully complements the AM symbiosis phenotype of *Mtcngc15b* and may partially complement the AM symbiosis phenotype of *Mtcngc15a*, but there is no statistically significant difference in the AM symbiosis phenotype of *Mtcngc15c* expressing *MpaCNGC3* and *Mtcngc15c* expressing the empty vector control. Surprisingly, *MpaCNGC3* could complement the nodulation phenotypes of both *Mtcngc15a* and *Mtcngc15c*, with inconclusive results in *Mtcngc15b*. This may be due to the evolution of *CNGC15* function in *M. truncatula*, as the nuclear associated calcium signalling pathway also functions in nodule symbiosis in this species. Duplication and subfunctionalization has taken place, as single mutants of *Mtcngc15a*, *Mtcngc15b* and *Mtcngc15c* all have defects in AM and root nodule symbiosis, indicating that they are not functionally redundant.

Alternatively, a second calcium channel may function in symbiosis signalling in *M. paleacea*, and may be required in addition to *MpaCNGC3* to complement the symbiosis phenotypes of *Mtcngc15* mutants. As AM symbiosis was not abolished in

*Mpacngc3*, it is likely that mutant plants can still generate the nuclear associated calcium oscillations required for symbiosis signalling. Another member of the CNGC gene family may have been recruited to function in generating nuclear calcium oscillations. This is unlikely to be *MpaCNGC1*, the second member of the Group I/II/III CNGCs in *M. paleacea*, as expression of *MpaCNGC1* did not complement any symbiosis phenotypes in *Mtcngc15a*, *Mtcngc15b* and *Mtcngc15c*. *MpaCNGC2* and *MpaCNGC4* must therefore be considered as possible mediators of symbiotic nuclear calcium oscillations, particularly *MpaCNGC4*, as the expression pattern of this gene overlaps significantly with that of *MpaCNGC3*.

In addition to the functional analysis of further members of the *M. paleacea* CNGC gene family, analysis of calcium signalling through genetically encoded fluorescent calcium reporters will allow further insights into the mechanisms of symbiotic calcium signal generation in this species.

## 6.4 Conclusions and outlook

The trait of AM symbiosis has been preserved in most lineages for over 450 million years of land plant evolution, including the evolution of vascular plants, rooting systems and the transition from gametophyte to sporophyte dominant generations in higher plants. The findings present in this thesis show that it is possible to predict gene function in AM symbiosis in basal land plant clades from phylogenetic and sequence analysis. This prediction can also be validated by studying the conservation of molecular function in AM symbiosis between an angiosperm and a basal land plant. Although previous studies have indicated that symbiosis gene orthologues from basal land plants and even algae can function in symbiosis signalling in an angiosperm, this study also shows that genome editing can be used to generate mutants with symbiosis phenotypes in a non-angiosperm model plant (Delaux et al., 2015; Wang et al., 2010).

Given the radically different developmental contexts of AM symbiosis in angiosperms and liverworts, where fungi colonize different plant organs and generations, it is possible that mutants of predicted symbiosis gene in *M. paleacea* may not exactly recapitulate those of their orthologues in angiosperms. For example, in liverworts the initial fungal infection and the development of intracellular nutrient structures take place in spatially separate tissues, the rhizoids and thallus parenchyma, respectively, while in angiosperms these processes are separated by a few cell layers in the root (Bonfante and Genre, 2010). Use of a liverwort model system for AM symbiosis may therefore allow dissection of the molecular mechanisms that control specific events in the development of AM symbiosis. The earliest land plants were likely to have been morphologically similar to the extant liverworts, so AM symbiosis in *M. paleacea* is far more representative of AM symbiosis in the earliest land plants than in any angiosperm model species (Taylor, 1995, 1997; Wellman et al., 2003). Further analysis of the molecular mechanisms of AM symbiosis in *M. paleacea* are likely to provide greater insight into the early evolution of the most widespread, and ancient, symbiotic association formed by land plants.

# References


Achard, P., Liao, L., Jiang, C., Desnos, T., Bartlett, J., Fu, X., and Harberd, N.P. (2007). DELLAs contribute to plant photomorphogenesis. Plant Physiol *143*, 1163-1172.

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

Albert, N.W., Thrimawithana, A.H., McGhie, T.K., Clayton, W.A., Deroles, S.C., Schwinn, K.E., Bowman, J.L., Jordan, B.R., and Davies, K.M. (2018). Genetic analysis of the liverwort Marchantia polymorpha reveals that R2R3MYB activation of flavonoid production in response to abiotic stress is an ancient character in land plants. New Phytol *218*, 554-566.

Albert, V.A., Barbazuk, W.B., Der, J.P., Leebens-Mack, J., Ma, H., Palmer, J.D., Rounsley, S., Sankoff, D., Schuster, S.C., and Soltis, D.E. (2013). The Amborella genome and the evolution of flowering plants. Science *342*, 1241089.

Amor, B.B., Shaw, S.L., Oldroyd, G.E., Maillet, F., Penmetsa, R.V., Cook, D., Long, S.R., Denarie, J., and Gough, C. (2003). The NFP locus of *Medicago truncatula* controls an early step of Nod factor signal transduction upstream of a rapid calcium flux and root hair deformation. The Plant Journal *34*, 495-506.

Ane, J.M., Kiss, G.B., Riely, B.K., Penmetsa, R.V., Oldroyd, G.E., Ayax, C., Levy, J., Debelle, F., Baek, J.M., Kalo, P*., et al.* (2004). *Medicago truncatula* DMI1 required for bacterial and fungal symbioses in legumes. Science *303*, 1364-1367.

Anisimova, M., Gil, M., Dufayard, J.-F., Dessimoz, C., and Gascuel, O. (2011). Survey of branch support methods demonstrates accuracy, power, and robustness of fast likelihood-based approximation schemes. Syst Biol *60*, 685-699.

Arrighi, J.-F., Barre, A., Ben Amor, B., Bersoult, A., Soriano, L.C., Mirabella, R., de Carvalho-Niebel, F., Journet, E.-P., Ghérardi, M., Huguet, T*., et al.* (2006). The Medicago truncatula Lysine Motif-Receptor-Like Kinase Gene Family Includes NFP and New Nodule-Expressed Genes. Plant Physiology *142*, 265-279.

Augé, R.M. (2001). Water relations, drought and vesicular-arbuscular mycorrhizal symbiosis. Mycorrhiza *11*, 3-42.

Balagué, C., Lin, B., Alcon, C., Flottes, G., Malmström, S., Köhler, C., Neuhaus, G., Pelletier, G., Gaymard, F., and Roby, D. (2003). HLM1, an essential signaling component in the hypersensitive response, is a member of the cyclic nucleotide–gated channel ion channel family. The Plant Cell *15*, 365-379.

Banba, M., Gutjahr, C., Miyao, A., Hirochika, H., Paszkowski, U., Kouchi, H., and Imaizumi-Anraku, H. (2008). Divergence of evolutionary ways among common sym genes: CASTOR and CCaMK show functional conservation between two symbiosis systems and constitute the root of a common signaling pathway. Plant Cell Physiol *49*, 1659-1671.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V.M., Nikolenko, S.I., Pham, S., Prjibelski, A.D*., et al.* (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol *19*, 455-477.

Banks, J.A., Nishiyama, T., Hasebe, M., Bowman, J.L., Gribskov, M., dePamphilis, C., Albert, V.A., Aono, N., Aoyama, T., Ambrose, B.A*., et al.* (2011). The Selaginella genome identifies genetic changes associated with the evolution of vascular plants. Science *332*, 960-963.

Bécard, G., Doner, L.W., Rolin, D.B., Douds, D.D., and Pfeffer, P.E. (1991). Identification and quantification of trehalose in vesicular-arbuscular mycorrhizal fungi by in vivo13C NMR and HPLC analyses*. New Phytol *118*, 547-552.
Berrie, G.K. (1960). The chromosome numbers of liverworts (Hepadcae and Anthocerotae). Transactions of the British Bryological Society *3*, 688-705.
Besserer, A., Bécard, G., Jauneau, A., Roux, C., and Séjalon-Delmas, N. (2008). GR24, a synthetic analog of strigolactones, stimulates the mitosis and growth of the arbuscular mycorrhizal fungus Gigaspora rosea by boosting its energy metabolism. Plant Physiol *148*, 402-413.
Besserer, A., Puech-Pages, V., Kiefer, P., Gomez-Roldan, V., Jauneau, A., Roy, S., Portais, J.C., Roux, C., Becard, G., and Sejalon-Delmas, N. (2006). Strigolactones stimulate arbuscular mycorrhizal fungi by activating mitochondria. PLoS Biol *4*, e226.
Boisson-Dernier, A., Chabaud, M., Garcia, F., Becard, G., Rosenberg, C., and Barker, D.G. (2001). Agrobacterium rhizogenes-transformed roots of Medicago truncatula for the study of nitrogen-fixing and endomycorrhizal symbiotic associations. Molecular plant-microbe interactions : MPMI *14*, 695-700.
Bonfante, P., Balestrini, R., and Mend Gen, K. (1994). Storage and secretion processes in the spore of Gigaspora margarita Becker &Hall as revealed by high-pressure freezing and freeze substitution. New Phytol *128*, 93-101.
Bonfante, P., and Genre, A. (2010). Mechanisms underlying beneficial plant–fungus interactions in mycorrhizal symbiosis. Nature communications *1*, 48.
Bonnot, C., Proust, H., Pinson, B., Colbalchini, F.P., Lesly-Veillard, A., Breuninger, H., Champion, C., Hetherington, A.J., Kelly, S., and Dolan, L. (2017). Functional PTB phosphate transporters are present in streptophyte algae and early diverging land plants. New Phytol *214*, 1158-1171.
Borsics, T., Webb, D., Andeme-Ondzighi, C., Staehelin, L.A., and Christopher, D.A. (2007). The cyclic nucleotide-gated calmodulin-binding channel AtCNGC10 localizes to the plasma membrane and influences numerous growth responses and starch accumulation in Arabidopsis thaliana. Planta *225*, 563-573.
Bortesi, L., and Fischer, R. (2015). The CRISPR/Cas9 system for plant genome editing and beyond. Biotechnol Adv *33*, 41-52.
Bowers, J.E., Chapman, B.A., Rong, J., and Paterson, A.H. (2003). Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. Nature *422*.
Bowman, J.L., Araki, T., Arteaga-Vazquez, M.A., Berger, F., Dolan, L., Haseloff, J., Ishizaki, K., Kyozuka, J., Lin, S.S., Nagasaki, H*., et al.* (2016a). The naming of names: Guidelines for gene nomenclature in Marchantia. Plant Cell Physiol *57*, 257-261.
Bowman, J.L., Araki, T., and Kohchi, T. (2016b). Marchantia : Past, present and future. Plant and Cell Physiology *57*, 205-209.
Bowman, J.L., Kohchi, T., Yamato, K.T., Jenkins, J., Shu, S., Ishizaki, K., Yamaoka, S., Nishihama, R., Nakamura, Y., Berger, F*., et al.* (2017). Insights into land plant evolution garnered from the Marchantia polymorpha genome. Cell *171*, 287-304.e215.
Bozsoki, Z., Cheng, J., Feng, F., Gysel, K., Vinther, M., Andersen, K.R., Oldroyd, G., Blaise, M., Radutoiu, S., and Stougaard, J. (2017). Receptor-mediated chitin perception in legume roots is functionally separable from Nod factor perception. Proceedings of the National Academy of Sciences *114*, E8118-E8127.
Bravo, A., Brands, M., Wewer, V., Dörmann, P., and Harrison, M.J. (2017). Arbuscular mycorrhiza-specific enzymes FatM and RAM2 fine-tune lipid biosynthesis to promote development of arbuscular mycorrhiza. New Phytol.

Bravo, A., York, T., Pumplin, N., Mueller, L.A., and Harrison, M.J. (2016). Genes conserved for arbuscular mycorrhizal symbiosis identified through phylogenomics. Nature plants *2*, 15208.
Breuillin-Sessoms, F., Floss, D.S., Gomez, S.K., Pumplin, N., Ding, Y., Levesque-Tremblay, V., Noar, R.D., Daniels, D.A., Bravo, A., Eaglesham, J.B.*, et al.* (2015). Suppression of arbuscule degeneration in Medicago truncatula phosphate transporter4 mutants is dependent on the ammonium transporter 2 family protein AMT2;3. Plant Cell *27*, 1352-1366.
Breuillin, F., Schramm, J., Hajirezaei, M., Ahkami, A., Favre, P., Druege, U., Hause, B., Bucher, M., Kretzschmar, T., Bossolini, E.*, et al.* (2010). Phosphate systemically inhibits development of arbuscular mycorrhiza in Petunia hybrida and represses genes involved in mycorrhizal functioning. Plant J *64*.
Brian, P. (1959). Effects of gibberellins on plant growth and development. Biological Reviews *34*, 37-77.
Brodribb, T.J., and Field, T.S. (2010). Leaf hydraulic evolution led a surge in leaf photosynthetic capacity during early angiosperm diversification. Ecol Lett *13*, 175-183.
Broghammer, A., Krusell, L., Blaise, M., Sauer, J., Sullivan, J.T., Maolanon, N., Vinther, M., Lorentzen, A., Madsen, E.B., Jensen, K.J.*, et al.* (2012). Legume receptors perceive the rhizobial lipochitin oligosaccharide signal molecules by direct binding. Proceedings of the National Academy of Sciences *109*, 13859-13864.
Buendia, L., Wang, T., Girardin, A., and Lefebvre, B. (2016). The LysM receptor-like kinase SlLYK10 regulates the arbuscular mycorrhizal symbiosis in tomato. New Phytol *210*, 184-195.
Cannon, S.B. (2013). The model legume genomes. Methods in molecular biology (Clifton, NJ) *1069*, 1-14.
Cannon, S.B., Sterck, L., Rombauts, S., Sato, S., Cheung, F., Gouzy, J., Wang, X., Mudge, J., Vasdewani, J., Schiex, T.*, et al.* (2006). Legume genome evolution viewed through the Medicago truncatula and Lotus japonicus genomes. Proceedings of the National Academy of Sciences *103*, 14959-14964.
Capoen, W. (2011). Nuclear membranes control symbiotic calcium signalling of legumes. Proc Natl Acad Sci U S A *108*, 14348-14353.
Capoen, W., Den Herder, J., Sun, J., Verplancke, C., De Keyser, A., De Rycke, R., Goormachtig, S., Oldroyd, G., and Holsters, M. (2009). Calcium spiking patterns and the role of the Calcium/Calmodulin-dependent Kinase CCamK in lateral root base nodulation of Sesbania rostrata. The Plant Cell *21*, 1526-1540.
Cárdenas, L., Feijó, J.A., Kunkel, J.G., Sánchez, F., Holdaway-Clarke, T., Hepler, P.K., and Quinto, C. (1999). Rhizobium Nod factors induce increases in intracellular free calcium and extracellular calcium influxes in bean root hairs. The Plant Journal *19*, 347-352.
Carella, P., Gogleva, A., Tomaselli, M., Alfs, C., and Schornack, S. (2018). Phytophthora palmivora establishes tissue-specific intracellular infection structures in the earliest divergent land plant lineage. Proceedings of the National Academy of Sciences.
Carl, A.M., and Phillips, D.A. (1990). Concurrent synthesis and release of nod-gene-inducing flavonoids from alfalfa roots. Plant Physiol *93*, 1552-1558.
Catoira, R., Galera, C., de Billy, F., Penmetsa, R.V., Journet, E.-P., Maillet, F., Rosenberg, C., Cook, D., Gough, C., and Dénarié, J. (2000). Four genes of Medicago truncatula controlling components of a nod factor transduction pathway. The Plant Cell *12*, 1647-1665.
Ceasar, S.A., Hodge, A., Baker, A., and Baldwin, S.A. (2014). Phosphate concentration and arbuscular mycorrhizal colonisation influence the growth, yield and expression of twelve PHT1 family phosphate transporters in foxtail millet (Setaria italica). PloS one *9*, e108459.
Chan, C.W.M., Schorrak, L.M., Smith, R.K., Bent, A.F., and Sussman, M.R. (2003). A cyclic nucleotide-gated ion channel, CNGC2, is crucial for plant development and adaptation to calcium stress. Plant Physiol *132*, 728.

Chari, R., Yeo, N.C., Chavez, A., and Church, G.M. (2017). sgRNA Scorer 2.0: A species-independent model to predict CRISPR/Cas9 activity. ACS synthetic biology *6*, 902-904.
Charpentier, M., Bredemeier, R., Wanner, G., Takeda, N., Schleiff, E., and Parniske, M. (2008). Lotus japonicus CASTOR and POLLUX are ion channels essential for perinuclear calcium spiking in legume root endosymbiosis. The Plant Cell *20*, 3467-3479.
Charpentier, M., and Oldroyd, G.E. (2013). Nuclear calcium signaling in plants. Plant Physiol *163*, 496-503.
Charpentier, M., Sun, J., Vaz Martins, T., Radhakrishnan, G.V., Findlay, K., Soumpourou, E., Thouin, J., Very, A.A., Sanders, D., Morris, R.J*., et al.* (2016). Nuclear-localized cyclic nucleotide-gated channels mediate symbiotic calcium oscillations. Science *352*, 1102-1105.
Charron, D., Pingret, J.L., Chabaud, M., Journet, E.P., and Barker, D.G. (2004). Pharmacological evidence that multiple phospholipid signaling pathways link Rhizobium nodulation factor perception in Medicago truncatula root hairs to intracellular responses, including Ca2+ spiking and specific ENOD gene expression. Plant Physiol *136*, 3582-3593.
Chen, C., Fan, C., Gao, M., and Zhu, H. (2009). Antiquity and function of castor and pollux, the twin ion channel-encoding genes key to the evolution of root symbioses in plants. Plant Physiol *149*, 306-317.
Chen, C., Gao, M., Liu, J., and Zhu, H. (2007). Fungal symbiosis in rice requires an ortholog of a legume common symbiosis gene encoding a Ca2+/calmodulin-dependent protein kinase. Plant Physiol *145*, 1619-1628.
Chen, C.Y., Ané, J.M., and Zhu, H.Y. (2008). OsIPD3, an ortholog of the Medicago truncatula DMI3 interacting protein IPD3, is required for mycorrhizal symbiosis in rice. New Phytol *180*.
Chen, T., Zhu, H., Ke, D., Cai, K., Wang, C., Gou, H., Hong, Z., and Zhang, Z. (2012). A MAP kinase kinase interacts with SymRK and regulates nodule organogenesis in Lotus japonicus. Plant Cell *24*, 823-838.
Chin, K., DeFalco, T.A., Moeder, W., and Yoshioka, K. (2013). The Arabidopsis cyclic nucleotide-gated ion channels AtCNGC2 and AtCNGC4 work in the same signaling pathway to regulate pathogen defense and floral transition. Plant Physiol *163*, 611-624.
Choudhury, S.R., and Pandey, S. (2013). Specific subunits of heterotrimeric G proteins play important roles during nodulation in soybean. Plant Physiol *162*, 522-533.
Choudhury, S.R., and Pandey, S. (2015). Phosphorylation-dependent regulation of G-protein cycle during nodule formation in soybean. Plant Cell *27*, 3260-3276.
Choudhury, S.R., and Pandey, S. (2016). Heterotrimeric G-protein complex and its role in regulation of nodule development. Endocytobiosis and Cell Research *27*, 29-35.
Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L*., et al.* (2013). Multiplex genome engineering using CRISPR/Cas systems. Science.
Cui, L., Wall, P.K., Leebens-Mack, J.H., Lindsay, B.G., Soltis, D.E., Doyle, J.J., Soltis, P.S., Carlson, J.E., Arumuganathan, K., Barakat, A*., et al.* (2006). Widespread genome duplications throughout the history of flowering plants. Genome Res *16*, 738-749.
Cvrčková, F., Grunt, M., Bezvoda, R., Hála, M., Kulich, I., Rawat, A., and Žárský, V. (2012). Evolution of the land plant exocyst complexes. Frontiers in Plant Science *3*, 159.
Czaja, L.F., Hogekamp, C., Lamm, P., Maillet, F., Martinez, E.A., Samain, E., Denarie, J., Kuster, H., and Hohnjec, N. (2012). Transcriptional responses toward diffusible signals from symbiotic microbes reveal MtNFP- and MtDMI3-dependent reprogramming of host gene expression by arbuscular mycorrhizal fungal lipochitooligosaccharides. Plant Physiol *159*, 1671-1685.
De Bodt, S., Maere, S., and Van de Peer, Y. (2005). Genome duplication and the origin of angiosperms. Trends Ecol Evol *20*, 591-597.

de la Pena, E., Echeverria, S.R., van der Putten, W.H., Freitas, H., and Moens, M. (2006). Mechanism of control of root-feeding nematodes by mycorrhizal fungi in the dune grass Ammophila arenaria. New Phytol *169*, 829-840.
Delaux, P.-M., Bécard, G., and Combier, J.-P. (2013a). NSP1 is a component of the Myc signaling pathway. New Phytol *199*, 59-65.
Delaux, P.-M., Radhakrishnan, G.V., Jayaraman, D., Cheema, J., Malbreil, M., Volkening, J.D., Sekimoto, H., Nishiyama, T., Melkonian, M., Pokorny, L*., et al.* (2015). Algal ancestor of land plants was preadapted for symbiosis. Proceedings of the National Academy of Sciences *112*, 13390-13395.
Delaux, P.M., Sejalon-Delmas, N., Bécard, G., and Ané, J.M. (2013b). Evolution of the plant-microbe symbiotic 'toolkit'. Trends Plant Sci *18*.
Delaux, P.M., Varala, K., Edger, P.P., Coruzzi, G.M., Pires, J.C., and Ane, J.M. (2014). Comparative phylogenomics uncovers the impact of symbiotic associations on host genome evolution. PLoS Genet *10*, e1004487.
Delmans, M., Pollak, B., and Haseloff, J. (2017). MarpoDB: An open registry for Marchantia polymorpha genetic parts. Plant and Cell Physiology *58*, e5-e5.
Denarie, J., Debelle, F., and Prome, J.C. (1996). Rhizobium lipo-chitooligosaccharide nodulation factors: signaling molecules mediating recognition and morphogenesis. Annu Rev Biochem *65*, 503-535.
Dickson, H. (1932). Polarity and the production of adventitious growing points in Marchantia polymorpha. Ann Bot *46*, 683-701.
Dickson, S. (2004). The Arum–Paris continuum of mycorrhizal symbioses. New Phytol *163*, 187-200.
Doyle, J.J. (2011). Phylogenetic perspectives on the origins of nodulation. Mol Plant-Microbe Interact *24*, 1289-1295.
Du, Y., Overdijk, E.J.R., Berg, J.A., Govers, F., and Bouwmeester, K. (2018). Solanaceous exocyst subunits are involved in immunity to diverse plant pathogens. J Exp Bot *69*, 655-666.
Duret, L., Gasteiger, E., and Perriere, G. (1996). LALNVIEW: a graphical viewer for pairwise sequence alignments. Comput Appl Biosci *12*, 507-510.
Eckardt, N.A. (2008). An Exocyst vesicle tethering complex in plants. The Plant Cell *20*, 1188-1188.
Ehrhardt, D.W., Wais, R., and Long, S.R. (1996). Calcium spiking in plant root hairs responding to Rhizobium nodulation signals. Cell *85*, 673-681.
Endre, G., Kereszt, A., Kevei, Z., Mihacea, S., Kalo, P., and Kiss, G.B. (2002). A receptor kinase gene regulating symbiotic nodule development. Nature *417*, 962-966.
Engler, C., Youles, M., Gruetzner, R., Ehnert, T.-M., Werner, S., Jones, J.D.G., Patron, N.J., and Marillonnet, S. (2014). A Golden Gate modular cloning toolbox for plants. ACS synthetic biology *3*, 839-843.
Engstrom, E.M., Ehrhardt, D.W., Mitra, R.M., and Long, S.R. (2002). Pharmacological analysis of nod factor-induced calcium spiking in Medicago truncatula. Evidence for the requirement of type IIA calcium pumps and phosphoinositide signaling. Plant Physiol *128*, 1390-1401.
Favre, P., Bapaume, L., Bossolini, E., Delorenzi, M., Falquet, L., and Reinhardt, D. (2014). A novel bioinformatics pipeline to discover genes related to arbuscular mycorrhizal symbiosis based on their evolutionary conservation pattern among higher plants. BMC Plant Biol *14*, 333.
Feddermann, N., Duvvuru Muni, R.R., Zeier, T., Stuurman, J., Ercolin, F., Schorderet, M., and Reinhardt, D. (2010). The PAM1gene of petunia, required for intracellular accommodation and morphogenesis of arbuscular mycorrhizal fungi, encodes a homologue of VAPYRIN. Plant J *64*.

Fedi, F., O'Neil, C., Menard, G., Trick, M., Corbineau, F., Bailly, C., Eastmond, P.J., and Penfield, S. (2017). Awake1, an ABC-type transporter, reveals an essential role for suberin in the control of seed dormancy. Plant Physiol.
Field, K.J., Rimington, W.R., Bidartondo, M.I., Allinson, K.E., Beerling, D.J., Cameron, D.D., Duckett, J.G., Leake, J.R., and Pressel, S. (2015). First evidence of mutualism between ancient plant lineages (Haplomitriopsida liverworts) and Mucoromycotina fungi and its response to simulated Palaeozoic changes in atmospheric CO2. New Phytol *205*, 743-756.
Fileccia, V., Ruisi, P., Ingraffia, R., Giambalvo, D., Frenda, A.S., and Martinelli, F. (2017). Arbuscular mycorrhizal symbiosis mitigates the negative effects of salinity on durum wheat. PloS one *12*, e0184158.
Finka, A., and Goloubinoff, P. (2014). The CNGCb and CNGCd genes from Physcomitrella patens moss encode for thermosensory calcium channels responding to fluidity changes in the plasma membrane. Cell Stress Chaperones *19*, 83-90.
Floss, D.S., Gomez, S.K., Park, H.-J., MacLean, A.M., Müller, L.M., Bhattarai, K.K., Lévesque-Tremblay, V., Maldonado-Mendoza, I.E., and Harrison, M.J. (2017). A transcriptional program for arbuscule degeneration during AM symbiosis is regulated by MYB1. Curr Biol *27*, 1206-1212.
Floss, D.S., Lévesque-Tremblay, V., Park, H.-J., and Harrison, M.J. (2016). DELLA proteins regulate expression of a subset of AM symbiosis-induced genes in Medicago truncatula. Plant Signaling & Behavior *11*, e1162369.
Floss, D.S., Levy, J.G., Lévesque-Tremblay, V., Pumplin, N., and Harrison, M.J. (2013). DELLA proteins regulate arbuscule formation in arbuscular mycorrhizal symbiosis. Proceedings of the National Academy of Sciences *110*, E5025-E5034.
Fonseca, H.M.A.C., Berbara, R.L.L., and Pereira, M.L. (2006). Lunularia cruciata, a potential in vitro host for Glomus proliferum and G. intraradices. Mycorrhiza *16*, 503.
Foo, E., Ross, J.J., Jones, W.T., and Reid, J.B. (2013). Plant hormones in arbuscular mycorrhizal symbioses: an emerging role for gibberellins. Ann Bot *111*, 769-779.
Friedman, W.E. (1993). The evolutionary history of the seed plant male gametophyte. Trends Ecol Evol *8*, 15-21.
Fritsch, R. (1991). Index to bryophyte chromosome counts, Bryophytorum Biblioteca. Bryophytorum Bibliotheca Science Publishers, Stuttgart, DE.
Gange, A.C. (2007). Insect–mycorrhizal interactions: patterns, processes, and consequences. In Ecological Communities: Plant Mediation in Indirect Interaction Webs, P.W. Price, T. Ohgushi, and T.P. Craig, eds. (Cambridge: Cambridge University Press), pp. 124-144.
Gao, F., Han, X., Wu, J., Zheng, S., Shang, Z., Sun, D., Zhou, R., and Li, B. (2012). A heat-activated calcium-permeable channel – Arabidopsis cyclic nucleotide-gated ion channel 6 – is involved in heat shock responses. The Plant Journal *70*, 1056-1069.
Gao, Q.-F., Gu, L.-L., Wang, H.-Q., Fei, C.-F., Fang, X., Hussain, J., Sun, S.-J., Dong, J.-Y., Liu, H., and Wang, Y.-F. (2016). Cyclic Nucleotide-Gated Channel 18 is an essential $Ca^{2+}$ channel in pollen tube tips for pollen tube guidance to ovules in Arabidopsis. Proceedings of the National Academy of Sciences *113*, 3096-3101.
Garcia, K., Doidy, J., Zimmermann, S.D., Wipf, D., and Courty, P.E. (2016). Take a trip through the plant and fungal transportome of mycorrhiza. Trends Plant Sci *21*, 937-950.
Genre, A., Chabaud, M., Balzergue, C., Puech-Pages, V., Novero, M., Rey, T., Fournier, J., Rochange, S., Becard, G., Bonfante, P.*, et al.* (2013). Short-chain chitin oligomers from arbuscular mycorrhizal fungi trigger nuclear Ca2+ spiking in Medicago truncatula roots and their production is enhanced by strigolactone. New Phytol *198*, 190-202.
Gherbi, H., Markmann, K., Svistoonoff, S., Estevan, J., Autran, D., Giczey, G., Auguy, F., Peret, B., Laplaze, L., Franche, C.*, et al.* (2008). SymRK defines a common genetic basis for

plant root endosymbioses with arbuscular mycorrhiza fungi, rhizobia, and Frankiabacteria. Proc Natl Acad Sci U S A *105*, 4928-4932.
Gianinazzi-Pearson, V. (1996). Plant cell responses to arbuscular mycorrhizal fungi: Getting to the roots of the symbiosis. The Plant Cell *8*, 1871-1883.
Giovannetti, M., and Mosse, B. (1980). An evaluation of techniques for measuring vesicular arbuscular mycorrhizal infection in roots. New Phytol *84*, 489-500.
Gobbato, E., Marsh, J.F., Vernie, T., Wang, E., Maillet, F., Kim, J., Miller, J.B., Sun, J., Bano, S.A., Ratet, P.*, et al.* (2012). A GRAS-type transcription factor with a specific function in mycorrhizal signaling. Curr Biol *22*, 2236-2241.
Gobbato, E., Wang, E., Higgins, G., Bano, S.A., Henry, C., Schultze, M., and Oldroyd, G.E. (2013). RAM1 and RAM2 function and expression during arbuscular mycorrhizal symbiosis and Aphanomyces euteiches colonization. Plant Signal Behav *8*.
Gobert, A., Park, G., Amtmann, A., Sanders, D., and Maathuis, F.J. (2006). Arabidopsis thaliana cyclic nucleotide gated channel 3 forms a non-selective ion transporter involved in germination and cation transport. J Exp Bot *57*, 791-800.
Godfroy, O., Debelle, F., Timmers, T., and Rosenberg, C. (2006). A rice calcium- and calmodulin-dependent protein kinase restores nodulation to a legume mutant. Molecular plant-microbe interactions : MPMI *19*, 495-501.
Gomez-Roldan, V., Fermas, S., Brewer, P.B., Puech-Pages, V., Dun, E.A., Pillot, J.-P., Letisse, F., Matusova, R., Danoun, S., Portais, J.-C.*, et al.* (2008). Strigolactone inhibition of shoot branching. Nature *455*, 189-194.
Goodstein, D.M., Shu, S., Howson, R., Neupane, R., Hayes, R.D., Fazo, J., Mitros, T., Dirks, W., Hellsten, U., Putnam, N.*, et al.* (2012). Phytozome: a comparative platform for green plant genomics. Nucleic Acids Res *40*, D1178-D1186.
Gough, C., Cottret, L., Lefebvre, B., and Bono, J.-J. (2018). Evolutionary history of plant lysm receptor proteins related to root endosymbiosis. Frontiers in Plant Science *9*.
Griesmann, M., Chang, Y., Liu, X., Song, Y., Haberer, G., Crook, M.B., Billault-Penneteau, B., Lauressergues, D., Keller, J., Imanishi, L.*, et al.* (2018). Phylogenomics reveals multiple losses of nitrogen-fixing root nodule symbiosis. Science *361*.
Groth, M., Takeda, N., Perry, J., Uchida, H., Draexl, S., Brachmann, A., Sato, S., Tabata, S., Kawaguchi, M., Wang, T.L.*, et al.* (2010). NENA, a Lotus japonicus homolog of Sec13, is required for rhizodermal infection by arbuscular mycorrhiza fungi and rhizobia but dispensable for cortical endosymbiotic development. Plant Cell *22*.
Gu, M., Chen, A., Dai, X., Liu, W., and Xu, G. (2011). How does phosphate status influence the development of the arbuscular mycorrhizal symbiosis? Plant Signal Behav *6*, 1300-1304.
Gualtieri, G., and Bisseling, T. (2000). The evolution of nodulation. Plant Mol Biol *42*, 181-194.
Guindon, S., Dufayard, J.F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol *59*, 307-321.
Guo, K.M., Babourina, O., Christopher, D.A., Borsic, T., and Rengel, Z. (2010). The cyclic nucleotide-gated channel AtCNGC10 transports Ca2+ and Mg2+ in Arabidopsis. Physiol Plant *139*, 303-312.
Gust, A.A., Willmann, R., Desaki, Y., Grabherr, H.M., and Nurnberger, T. (2012). Plant LysM proteins: modules mediating symbiosis and immunity. Trends Plant Sci *17*, 495-502.
Gutjahr, C., Banba, M., Croset, V., An, K., Miyao, A., An, G., Hirochika, H., Imaizumi-Anraku, H., and Paszkowski, U. (2008). Arbuscular mycorrhiza–specific signaling in rice transcends the common symbiosis signaling pathway. The Plant Cell *20*, 2989-3005.
Gutjahr, C., Gobbato, E., Choi, J., Riemann, M., Johnston, M.G., Summers, W., Carbonnel, S., Mansfield, C., Yang, S.Y., Nadal, M.*, et al.* (2015). Rice perception of symbiotic arbuscular mycorrhizal fungi requires the karrikin receptor complex. Science *350*, 1521-1524.

Guttenberger, M. (2000). Arbuscules of vesicular-arbuscular mycorrhizal fungi inhabit an acidic compartment within plant roots. Planta *211*, 299-304.
Hala, M., Cole, R., Synek, L., Drdova, E., Pecenkova, T., Nordheim, A., Lamkemeyer, T., Madlung, J., Hochholdinger, F., Fowler, J.E.*, et al.* (2008). An exocyst complex functions in plant cell growth in Arabidopsis and tobacco. Plant Cell *20*, 1330-1345.
Hameed, A., Dilfuza, E., Abd-Allah, E.F., Hashem, A., Kumar, A., and Ahmad, P. (2014). Salinity stress and arbuscular mycorrhizal symbiosis in plants. In Use of Microbes for the Alleviation of Soil Stresses, Volume 1 (Springer), pp. 139-159.
Harrison, C.J. (2017). Development and genetics in the evolution of land plant body plans. Phil Trans R Soc B *372*, 20150490.
Harrison, M.J., Dewbre, G.R., and Liu, J. (2002). A phosphate transporter from Medicago truncatula involved in the acquisition of phosphate released by arbuscular mycorrhizal fungi. The Plant Cell *14*, 2413-2429.
Harrison, M.J., and Ivanov, S. (2017). Exocytosis for endosymbiosis: membrane trafficking pathways for development of symbiotic membrane compartments. Curr Opin Plant Biol *38*, 101-108.
Healey, A., Furtado, A., Cooper, T., and Henry, R.J. (2014). Protocol: a simple method for extracting next-generation sequencing quality genomic DNA from recalcitrant plant species. Plant methods *10*, 1.
Hildebrandt, U., Regvar, M., and Bothe, H. (2007). Arbuscular mycorrhiza and heavy metal tolerance. Phytochemistry *68*, 139-146.
Hoffmann, B., Trinh, T.H., Leung, J., Kondorosi, A., and Kondorosi, E. (1997). A New Medicago truncatula Line with superior in vitro regeneration, transformation, and symbiotic properties isolated through cell culture selection. Mol Plant-Microbe Interact *10*, 307-315.
Hogekamp, C., Arndt, D., Pereira, P.A., Becker, J.D., Hohnjec, N., and Kuster, H. (2011). Laser microdissection unravels cell-type-specific transcription in arbuscular mycorrhizal roots, including CAAT-box transcription factor gene expression correlating with fungal contact and spread. Plant Physiol *157*, 2023-2043.
Hori, K., Maruyama, F., Fujisawa, T., Togashi, T., Yamamoto, N., Seo, M., Sato, S., Yamada, T., Mori, H., Tajima, N.*, et al.* (2014). Klebsormidium flaccidum genome reveals primary factors for plant terrestrial adaptation. Nature communications *5*, 3978.
Huang, X., and Miller, W. (1991). A time-efficient, linear-space local similarity algorithm. Advances in Applied Mathematics *12*, 337-357.
Huisman, R. (2018). Formation od a symbiotic host-microbe interface: the role of SNARE-mediated regulation of exocytosis. In Experimental Plant Sciences (Wageningen: Wageningen University).
Huisman, R., Hontelez, J., Mysore, K.S., Wen, J., Bisseling, T., and Limpens, E. (2016). A symbiosis-dedicated SYNTAXIN OF PLANTS 13II isoform controls the formation of a stable host-microbe interface in symbiosis. New Phytol *211*, 1338-1351.
Humphreys, C.P., Franks, P.J., Rees, M., Bidartondo, M.I., Leake, J.R., and Beerling, D.J. (2010). Mutualistic mycorrhiza-like symbiosis in the most ancient group of land plants. Nature communications *1*, 103.
Hwang, J.U., Song, W.Y., Hong, D., Ko, D., Yamaoka, Y., Jang, S., Yim, S., Lee, E., Khare, D., Kim, K.*, et al.* (2016). Plant ABC transporters enable many unique aspects of a terrestrial plant's lifestyle. Mol Plant *9*, 338-355.
Imaizumi-Anraku, H., Takeda, N., Charpentier, M., Perry, J., Miwa, H., Umehara, Y., Kouchi, H., Murakami, Y., Mulder, L., Vickers, K.*, et al.* (2005). Plastid proteins crucial for symbiotic fungal and bacterial entry into plant roots. Nature *433*, 527-531.

Ivanov, S., Fedorova, E.E., Limpens, E., De Mita, S., Genre, A., Bonfante, P., and Bisseling, T. (2012). Rhizobium–legume symbiosis shares an exocytotic pathway required for arbuscule formation. Proceedings of the National Academy of Sciences *109*, 8316-8321.
Javot, H., Penmetsa, R.V., Terzaghi, N., Cook, D.R., and Harrison, M.J. (2007a). A Medicago truncatula phosphate transporter indispensable for the arbuscular mycorrhizal symbiosis. Proceedings of the National Academy of Sciences *104*, 1720-1725.
Javot, H., Pumplin, N., and Harrison, M.J. (2007b). Phosphate in the arbuscular mycorrhizal symbiosis: transport properties and regulatory roles. Plant, Cell Environ *30*, 310-322.
Jin, Y., Jing, W., Zhang, Q., and Zhang, W. (2015). Cyclic nucleotide gated channel 10 negatively regulates salt tolerance by mediating Na+ transport in Arabidopsis. J Plant Res *128*, 211-220.
Jin, Y., Liu, H., Luo, D., Yu, N., Dong, W., Wang, C., Zhang, X., Dai, H., Yang, J., and Wang, E. (2016). DELLA proteins are common components of symbiotic rhizobial and mycorrhizal signalling pathways. Nature communications *7*, 12433.
Johnson, M.T.J., Carpenter, E.J., Tian, Z., Bruskiewich, R., Burris, J.N., Carrigan, C.T., Chase, M.W., Clarke, N.D., Covshoff, S., dePamphilis, C.W*., et al.* (2012). Evaluating methods for isolating total RNA and predicting the success of sequencing phylogenetically diverse plant transcriptomes. PloS one *7*, e50226.
Jones, V.A., and Dolan, L. (2012). The evolution of root hairs and rhizoids. Ann Bot *110*, 205-212.
Jones, V.A.S., and Dolan, L. (2017). MpWIP regulates air pore complex development in the liverwort Marchantia polymorpha. Development *144*, 1472-1476.
Jurkowski, G.I., Smith Jr, R.K., Yu, I.-c., Ham, J.H., Sharma, S.B., Klessig, D.F., Fengler, K.A., and Bent, A.F. (2004). Arabidopsis DND2, a second cyclic nucleotide-gated ion channel gene for which mutation causes the “defense, no death” phenotype. Mol Plant-Microbe Interact *17*, 511-520.
K Jha, S., Sharma, M., and K Pandey, G. (2016). Role of cyclic nucleotide gated channels in stress management in plants. Curr Genomics *17*, 315-329.
Kaló, P., Gleason, C., Edwards, A., Marsh, J., Mitra, R.M., Hirsch, S., Jakab, J., Sims, S., Long, S.R., and Rogers, J. (2005). Nodulation signaling in legumes requires NSP2, a member of the GRAS family of transcriptional regulators. Science *308*, 1786-1789.
Kalo, P., Gleason, C., Edwards, A., Marsh, J., Mitra, R.M., Hirsch, S., Jakab, J., Sims, S., Long, S.R., Rogers, J*., et al.* (2005). Nodulation signaling in legumes requires NSP2, a member of the GRAS family of transcriptional regulators. Science (New York, NY) *308*, 1786-1789.
Kanamori, N., Madsen, L.H., Radutoiu, S., Frantescu, M., Quistgaard, E.M., Miwa, H., Downie, J.A., James, E.K., Felle, H.H., Haaning, L.L*., et al.* (2006). A nucleoporin is required for induction of Ca2+ spiking in legume nodule development and essential for rhizobial and fungal symbiosis. Proc Natl Acad Sci U S A *103*, 359-364.
Kaplan, B., Sherman, T., and Fromm, H. (2007). Cyclic nucleotide-gated channels in plants. FEBS Lett *581*, 2237-2246.
Karandashov, V., Nagy, R., Wegmüller, S., Amrhein, N., and Bucher, M. (2004). Evolutionary conservation of a phosphate transporter in the arbuscular mycorrhizal symbiosis. P Natl Acad Sci USA *101*, 6285-6290.
Kenrick, P., and Strullu-Derrien, C. (2014). The Origin and Early Evolution of Roots. Plant Physiol *166*, 570-580.
Kevei, Z., Lougnon, G., Mergaert, P., Horváth, G.V., Kereszt, A., Jayaraman, D., Zaman, N., Marcel, F., Regulski, K., Kiss, G.B*., et al.* (2007a). 3-Hydroxy-3-Methylglutaryl Coenzyme A Reductase1 interacts with NORK and is crucial for nodulation in Medicago truncatula. The Plant Cell *19*, 3974-3989.
Kevei, Z., Lougnon, G., Mergaert, P., Horvath, G.V., Kereszt, A., Jayaraman, D., Zaman, N., Marcel, F., Regulski, K., Kiss, G.B*., et al.* (2007b). 3-hydroxy-3-methylglutaryl coenzyme a

reductase 1 interacts with NORK and is crucial for nodulation in Medicago truncatula. Plant Cell *19*, 3974-3989.
Kiefer, B., Riemann, M., Büche, C., Kassemeyer, H.-H., and Nick, P. (2002). The host guides morphogenesis and stomatal targeting in the grapevine pathogen Plasmopara viticola. Planta *215*, 387-393.
Kistner, C., and Parniske, M. (2002). Evolution of signal transduction in intracellular symbiosis. Trends Plant Sci *7*.
Kistner, C., Winzer, T., Pitzschke, A., Mulder, L., Sato, S., Kaneko, T., Tabata, S., Sandal, N., Stougaard, J., Webb, K.J.*, et al.* (2005). Seven Lotus japonicus genes required for transcriptional reprogramming of the root during fungal and bacterial symbiosis. Plant Cell *17*, 2217-2229.
Koki, F., Yoshiko, A., Akiko, I., Hiromasa, S., Kentaro, Y., Hiroyuki, K., Eiko, K., Hiroe, U., Tetsuro, Y., Sophien, K.*, et al.* (2015). Rice Exo70 interacts with a fungal effector, AVR-Pii, and is required for AVR-Pii-triggered immunity. The Plant Journal *83*, 875-887.
Kosuta, S., Chabaud, M., Lougnon, G., Gough, C., Denarie, J., Barker, D.G., and Becard, G. (2003). A diffusible factor from arbuscular mycorrhizal fungi induces symbiosis-specific MtENOD11 expression in roots of Medicago truncatula. Plant Physiol *131*, 952-962.
Krajinski, F., Courty, P.-E., Sieh, D., Franken, P., Zhang, H., Bucher, M., Gerlach, N., Kryvoruchko, I., Zoeller, D., Udvardi, M.*, et al.* (2014). The H(+)-ATPase HA1 of Medicago truncatula is essential for phosphate transport and plant growth during arbuscular mycorrhizal symbiosis. The Plant Cell *26*, 1808-1817.
Krajinski, F., Hause, B., Gianinazzi-Pearson, V., and Franken, P. (2002). Mtha1, a plasma membrane H+-ATPase gene from Medicago truncatula, shows arbuscule-specific induced expression in mycorrhizal tissue. Plant Biol *4*, 754-761.
Kubota, A., Ishizaki, K., Hosaka, M., and Kohchi, T. (2013). Efficient Agrobacterium-mediated transformation of the liverwort Marchantia polymorpha using regenerating thalli. Biosci Biotechnol Biochem *77*, 167-172.
Kugler, A., Kohler, B., Palme, K., Wolff, P., and Dietrich, P. (2009). Salt-dependent regulation of a CNG channel subfamily in Arabidopsis. BMC Plant Biol *9*, 140.
Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol Biol Evol *33*, 1870-1874.
Ladwig, F., Dahlke, R.I., Stührwohldt, N., Hartmann, J., Harter, K., and Sauter, M. (2015). Phytosulfokine regulates growth in Arabidopsis through a response module at the plasma membrane that includes CYCLIC NUCLEOTIDE-GATED CHANNEL17, $H^+$-ATPase, and BAK1. The Plant Cell *27*, 1718-1729.
Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R.*, et al.* (2007). Clustal W and Clustal X version 2.0. Bioinformatics *23*, 2947-2948.
Lazo, G.R., Stein, P.A., and Ludwig, R.A. (1991). A DNA transformation-competent Arabidopsis genomic library in Agrobacterium. Bio/technology (Nature Publishing Company) *9*, 963-967.
Lee, J.-K., Eom, A.-H., Lee, S.-S., and Lee, C.H. (2001). Mycorrhizal symbioses found in roots of fern and its relatives in Korea. Journal of Plant Biology *44*, 81.
Lefort, V., Longueville, J.-E., and Gascuel, O. (2017). SMS: Smart model selection in PhyML. Mol Biol Evol *34*, 2422-2424.
Leggett, R.M., Clavijo, B.J., Clissold, L., Clark, M.D., and Caccamo, M. (2014). NextClip: an analysis and read preparation tool for Nextera Long Mate Pair libraries. Bioinformatics *30*, 566-568.
Leppyanen, I., Shakhnazarova, V., Shtark, O., Vishnevskaya, N., Tikhonovich, I., and Dolgikh, E. (2018). Receptor-Like Kinase LYK9 in Pisum sativum L. is the CERK1-like receptor that

controls both plant immunity and am symbiosis development. International Journal of Molecular Sciences *19*, 8.
Lerouge, P., Roche, P., Faucher, C., Maillet, F., Truchet, G., Prome, J.C., and Denarie, J. (1990). Symbiotic host-specificity of Rhizobium meliloti is determined by a sulphated and acylated glucosamine oligosaccharide signal. Nature *344*, 781-784.
Levy, J., Bres, C., Geurts, R., Chalhoub, B., Kulikova, O., Duc, G., Journet, E.P., Ane, J.M., Lauber, E., Bisseling, T.*, et al.* (2004). A putative Ca2+ and calmodulin-dependent protein kinase required for bacterial and fungal symbioses. Science *303*, 1361-1364.
Li, H.-Y., Yang, G.-D., Shu, H.-R., Yang, Y.-T., Ye, B.-X., Nishida, I., and Zheng, C.-C. (2006). Colonization by the arbuscular mycorrhizal fungus Glomus versiforme induces a defense response against the root-knot nematode Meloidogyne incognita in the grapevine (Vitis amurensis Rupr.), which includes transcriptional activation of the class III chitinase gene VCH3. Plant and Cell Physiology *47*, 154-163.
Li, S., Chen, M., Yu, D., Ren, S., Sun, S., Liu, L., Ketelaar, T., Emons, A.-M.C., and Liu, C.-M. (2013). EXO70A1-mediated vesicle trafficking is critical for tracheary element development in Arabidopsis. The Plant Cell Online.
Li, S., van Os, G.M., Ren, S., Yu, D., Ketelaar, T., Emons, A.M., and Liu, C.M. (2010). Expression and functional analyses of EXO70 genes in Arabidopsis implicate their roles in regulating cell type-specific exocytosis. Plant Physiol *154*, 1819-1830.
Liang, G., Zhang, H., Lou, D., and Yu, D. (2016). Selection of highly efficient sgRNAs for CRISPR/Cas9-based plant genome editing. Scientific Reports *6*, 21451.
Ligrone, R., Carafa, A., Lumini, E., Bianciotto, V., Bonfante, P., and Duckett, J.G. (2007). Glomeromycotean associations in liverworts: a molecular, cellular, and taxonomic analysis. Am J Bot *94*, 1756-1777.
Ligrone, R., and Lopes, C. (1989). Cytology and development of a mycorrhiza-like infection in the gametophyte of Conocephalum conicum (L.) Dum. (Marchantiales, Hepatophyta). New Phytol *111*, 423-433.
Limpens, E., Franken, C., Smit, P., Willemse, J., Bisseling, T., and Geurts, R. (2003). LysM domain receptor kinases regulating rhizobial nod factor-induced infection. Science *302*, 630-633.
Limpens, E., Moling, S., Hooiveld, G., Pereira, P.A., Bisseling, T., Becker, J.D., and Küster, H. (2013). Cell- and tissue-specific transcriptome analyses of Medicago truncatula root nodules. PloS one *8*, e64377.
Lin, J.-J. (1995). Electrotransformation of Agrobacterium. In Electroporation Protocols for Microorganisms, J.A. Nickoloff, ed. (Totowa, NJ: Humana Press), pp. 171-178.
Lin, J.-r., and Hu, J. (2013). SeqNLS: Nuclear localization signal prediction based on frequent pattern mining and linear motif scoring. PloS one *8*, e76864.
Liu, W., Kohlen, W., Lillo, A., Op den Camp, R., Ivanov, S., Hartog, M., Limpens, E., Jamil, M., Smaczniak, C., Kaufmann, K.*, et al.* (2011). Strigolactone biosynthesis in Medicago truncatula and rice requires the symbiotic GRAS-type transcription factors NSP1 and NSP2. Plant Cell *23*, 3853-3865.
Liu, X., Wu, S., Xu, J., Sui, C., and Wei, J. (2017). Application of CRISPR/Cas9 in plant biology. Acta Pharmaceutica Sinica B *7*, 292-302.
Lucking, R., Huhndorf, S., Pfister, D.H., Plata, E.R., and Lumbsch, H.T. (2009). Fungi evolved right on track. Mycologia *101*, 810-822.
Luginbuehl, L.H., Menard, G.N., Kurup, S., Van Erp, H., Radhakrishnan, G.V., Breakspear, A., Oldroyd, G.E.D., and Eastmond, P.J. (2017). Fatty acids in arbuscular mycorrhizal fungi are synthesized by the host plant. Science.
Luginbuehl, L.H., and Oldroyd, G.E.D. (2017). Understanding the arbuscule at the heart of endomycorrhizal symbioses in plants. Curr Biol *27*, R952-r963.

Lukáš, S., Nicole, S., Marek, E., Michaël, Q., Marie-Theres, H., and Viktor, Ž. (2006). AtEXO70A1, a member of a family of putative exocyst subunits specifically expanded in land plants, is important for polar growth and plant development. The Plant Journal *48*, 54-72.
Luo, R., Liu, B., Xie, Y., Li, Z., Huang, W., Yuan, J., He, G., Chen, Y., Pan, Q., Liu, Y.*, et al.* (2012). SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. Gigascience *1*, 18.
Ma, W., Ali, R., and Berkowitz, G.A. (2006). Characterization of plant phenotypes associated with loss-of-function of AtCNGC1, a plant cyclic nucleotide gated cation channel. Plant Physiol Biochem *44*, 494-505.
Madsen, E.B., Madsen, L.H., Radutoiu, S., Olbryt, M., Rakwalska, M., Szczyglowski, K., Sato, S., Kaneko, T., Tabata, S., Sandal, N.*, et al.* (2003). A receptor kinase gene of the LysM type is involved in legumeperception of rhizobial signals. Nature *425*, 637-640.
Maekawa, T., Maekawa-Yoshikawa, M., Takeda, N., Imaizumi-Anraku, H., Murooka, Y., and Hayashi, M. (2009). Gibberellin controls the nodulation signaling pathway in Lotus japonicus. Plant J *58*, 183-194.
Magallon, S., Hilu, K.W., and Quandt, D. (2013). Land plant evolutionary timeline: gene effects are secondary to fossil constraints in relaxed clock estimation of age and substitution rates. Am J Bot *100*, 556-573.
Maillet, F., Poinsot, V., Andre, O., Puech-Pages, V., Haouy, A., Gueunier, M., Cromer, L., Giraudet, D., Formey, D., Niebel, A.*, et al.* (2011). Fungal lipochitooligosaccharide symbiotic signals in arbuscular mycorrhiza. Nature *469*, 58-63.
Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E., and Church, G.M. (2013). RNA-guided human genome engineering via Cas9. Science *339*, 823-826.
Markmann, K., Giczey, G., and Parniske, M. (2008). Functional adaptation of a plant receptor- kinase paved the way for the evolution of intracellular root symbioses with bacteria. PLoS Biol *6*, e68.
Martins, T.V., Evans, M.J., Wysham, D.B., and Morris, R.J. (2016). Nuclear pores enable sustained perinuclear calcium oscillations. BMC Systems Biology *10*, 55.
Matasci, N., Hung, L.H., Yan, Z., Carpenter, E.J., Wickett, N.J., Mirarab, S., Nguyen, N., Warnow, T., Ayyampalayam, S., Barker, M.*, et al.* (2014). Data access for the 1,000 Plants (1KP) project. Gigascience *3*, 17.
Menand, B., Yi, K., Jouannic, S., Hoffmann, L., Ryan, E., Linstead, P., Schaefer, D.G., and Dolan, L. (2007). An ancient mechanism controls the development of cells with a rooting function in land plants. Science *316*, 1477-1480.
Messinese, E., Mun, J.H., Yeun, L.H., Jayaraman, D., Rouge, P., Barre, A., Lougnon, G., Schornack, S., Bono, J.J., Cook, D.R.*, et al.* (2007). A novel nuclear protein interacts with the symbiotic DMI3 calcium- and calmodulin-dependent protein kinase of Medicago truncatula. Molecular plant-microbe interactions : MPMI *20*, 912-921.
Miller, J.B., Pratap, A., Miyahara, A., Zhou, L., Bornemann, S., Morris, R.J., and Oldroyd, G.E. (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.
Mitra, R.M., Gleason, C.A., Edwards, A., Hadfield, J., Downie, J.A., Oldroyd, G.E., and Long, S.R. (2004). A Ca2+/calmodulin-dependent protein kinase required for symbiotic nodule development: Gene identification by transcript-based cloning. Proc Natl Acad Sci U S A *101*, 4701-4705.
Miwa, H., Sun, J., Oldroyd, G.E., and Downie, J.A. (2006). Analysis of Nod-factor-induced calcium signaling in root hairs of symbiotically defective mutants of Lotus japonicus. Molecular plant-microbe interactions : MPMI *19*, 914-923.

Miyata, K., Kozaki, T., Kouzai, Y., Ozawa, K., Ishii, K., Asamizu, E., Okabe, Y., Umehara, Y., Miyamoto, A., Kobae, Y.*, 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.
Moeder, W., Urquhart, W., Ung, H., and Yoshioka, K. (2011). The role of cyclic nucleotide-gated ion channels in plant immunity. Molecular Plant *4*, 442-452.
Murray, J.D., Duvvuru Muni, R., Torres-Jerez, I., Tang, Y., Allen, S., Andriankaja, M., Li, G., Laxmi, A., Cheng, X., Wen, J.*, et al.* (2011a). Vapyrin, a gene essential for intracellular progression of arbuscular mycorrhizal symbiosis, is also essential for infection by rhizobia in the nodule symbiosis of Medicago truncatula. Plant J *65*.
Murray, J.D., Muni, R.R., Torres-Jerez, I., Tang, Y., Allen, S., Andriankaja, M., Li, G., Laxmi, A., Cheng, X., Wen, J.*, et al.* (2011b). Vapyrin, a gene essential for intracellular progression of arbuscular mycorrhizal symbiosis, is also essential for infection by rhizobia in the nodule symbiosis of Medicago truncatula. Plant J *65*, 244-252.
Muthukumar, T., and Prabha, K. (2013). Arbuscular mycorrhizal and septate endophyte fungal associations in lycophytes and ferns of south India. Symbiosis *59*, 15-33.
Nagy, F., Karandashov, V., Chague, W., Kalinkevich, K., Tamasloukht, M., Xu, G.H., Jakobsen, I., Levy, A.A., Amrhein, N., and Bucher, M. (2005). The characterization of novel mycorrhiza-specific phosphate transporters from Lycopersicon esculentum and Solanum tuberosumuncovers functional redundancy in symbiotic phosphate transport in solanaceous species. Plant J *42*.
Nakagawa, T., and Imaizumi-Anraku, H. (2015). Rice arbuscular mycorrhiza as a tool to study the molecular mechanisms of fungal symbiosis and a potential target to increase productivity. Rice *8*, 32.
Navarro, L., Bari, R., Achard, P., Lisón, P., Nemri, A., Harberd, N.P., and Jones, J.D. (2008). DELLAs control plant immune responses by modulating the balance of jasmonic acid and salicylic acid signaling. Curr Biol *18*, 650-655.
Newsham, K.K., Goodall-Copestake, W.P., Ochyra, R., and Váňa, J. (2014). Mycothalli of the hepatic Barbilophozia hatcheri in Antarctica: distribution and identities of mycobionts. Fungal Ecology *11*, 91-99.
Nicolson, T. (1967). Vesicular-arbuscular mycorrhiza—a universal plant symbiosis. Science Progress (1933-), 561-581.
Oldroyd, G.E. (2013). Speak, friend, and enter: signalling systems that promote beneficial symbiotic associations in plants. Nat Rev Microbiol *11*, 252-263.
Op den Camp, R., Streng, A., De Mita, S., Cao, Q., Polone, E., Liu, W., Ammiraju, J.S., Kudrna, D., Wing, R., Untergasser, A.*, et al.* (2011). LysM-type mycorrhizal receptor recruited for rhizobium symbiosis in nonlegume Parasponia. Science *331*, 909-912.
Park, H.-J., Floss, D.S., Levesque-Tremblay, V., Bravo, A., and Harrison, M.J. (2015). Hyphal branching during arbuscule development requires Reduced Arbuscular Mycorrhiza1. Plant Physiol *169*, 2774-2788.
Parke, J.L., and Linderman, R.G. (1980). Association of vesicular–arbuscular mycorrhizal fungi with the moss Funaria hygrometrica. Canadian Journal of Botany *58*, 1898-1904.
Parniske, M. (2008). Arbuscular mycorrhiza: the mother of plant root endosymbioses. Nat Rev Micro *6*, 763-775.
Paszkowski, U., Kroken, S., Roux, C., and Briggs, S.P. (2002). Rice phosphate transporters include an evolutionarily divergent gene specifically activated in arbuscular mycorrhizal symbiosis. Proceedings of the National Academy of Sciences *99*, 13324-13329.
Patil, S., Takezawa, D., and Poovaiah, B.W. (1995). Chimeric plant calcium/calmodulin-dependent protein kinase gene with a neural visinin-like calcium-binding domain. Proc Natl Acad Sci U S A *92*, 4897-4901.

Pauwels, L., De Clercq, R., Goossens, J., Inigo, S., Williams, C., Ron, M., Britt, A., and Goossens, A. (2018). A dual sgRNA approach for functional genomics in Arabidopsis thaliana. G3 (Bethesda, Md) *8*, 2603-2615.
Peiter, E., Sun, J., Heckmann, A.B., Venkateshwaran, M., Riely, B.K., Otegui, M.S., Edwards, A., Freshour, G., Hahn, M.G., Cook, D.R.*, et al.* (2007). The Medicago truncatula DMI1 protein modulates cytosolic calcium signaling. Plant Physiol *145*, 192-203.
Peng, Y., Leung, H.C., Yiu, S.M., and Chin, F.Y. (2012). IDBA-UD: a de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. Bioinformatics *28*, 1420-1428.
Pettitt, J. (1977). Detection in primitive gymnosperms of proteins and glycoproteins of possible significance in reproduction. Nature *266*, 530.
Pfeffer, P.E., Douds, D.D., Bécard, G., and 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., Bohmer, Monica J., Karl, L., Floss, Daniela S., Harrison, Maria J., Parniske, M.*, et al.* (2016). A CCaMK-CYCLOPS-DELLA complex activates transcription of RAM1 to regulate arbuscule branching. Curr Biol *26*, 987-998.
Pingret, J.L., Journet, E.P., and Barker, D.G. (1998). Rhizobium nod factor signaling. Evidence for a g protein-mediated transduction mechanism. Plant Cell *10*, 659-672.
Pirozynski, K.A., and Malloch, D.W. (1975). The origin of land plants: a matter of mycotrophism. BioSyst *6*, 153-164.
Pislariu, C.I., Murray, J.D., Wen, J., Cosson, V., Muni, R.R., Wang, M., Benedito, V.A., Andriankaja, A., Cheng, X., Jerez, I.T.*, et al.* (2012). A Medicago truncatula tobacco retrotransposon insertion mutant collection with defects in nodule development and symbiotic nitrogen fixation. Plant Physiol *159*, 1686-1699.
Poovaiah, B.W., Xia, M., Liu, Z., Wang, W., Yang, T., Sathyanarayanan, P.V., and Franceschi, V.R. (1999). Developmental regulation of the gene for chimeric calcium/calmodulin-dependent protein kinase in anthers. Planta *209*, 161-171.
Port, F., Chen, H.-M., Lee, T., and Bullock, S.L. (2014). Optimized CRISPR/Cas tools for efficient germline and somatic genome engineering in Drosophila. P Natl Acad Sci USA *111*, E2967-E2976.
Pozo, M.J., and Azcón-Aguilar, C. (2007). Unraveling mycorrhiza-induced resistance. Curr Opin Plant Biol *10*, 393-398.
Pumplin, N., Mondo, S.J., Topp, S., Starker, C.G., Gantt, J.S., and Harrison, M.J. (2010a). Medicago truncatula Vapyrin is a novel protein required for arbuscular mycorrhizal symbiosis. The Plant Journal *61*, 482-494.
Pumplin, N., Mondo, S.J., Topp, S., Starker, C.G., Gantt, J.S., and Harrison, M.J. (2010b). Medicago truncatulaVapyrin is a novel protein required for arbuscular mycorrhizal symbiosis. Plant J *61*.
Pumplin, N., Zhang, X., Noar, R.D., and Harrison, M.J. (2012). Polar localization of a symbiosis-specific phosphate transporter is mediated by a transient reorientation of secretion. Proceedings of the National Academy of Sciences *109*, E665–E672.
Pysh, L.D., Wysocka-Diller, J.W., Camilleri, C., Bouchez, D., and Benfey, P.N. (1999). The GRAS gene family in Arabidopsis: sequence characterization and basic expression analysis of the SCARECROW-LIKE genes. Plant J *18*, 111-119.
Rabatin, S.C. (1980). The occurrence of the vesicular-arbuscular-mycorrhizal fungus Glomus tenuis with moss. Mycologia *72*, 191-195.
Radhakrishnan, G.V. (2017). Tracing the evolution of the arbuscular mycorrhizal symbiosis in the plant lineage. In Cell and Developmental Biology, JIC (University of East Anglia).

Ran, F.A., Hsu, Patrick D., Lin, C.-Y., Gootenberg, Jonathan S., Konermann, S., Trevino, A.E., Scott, David A., Inoue, A., Matoba, S., Zhang, Y.*, et al.* (2013). Double Nicking by RNA-Guided CRISPR Cas9 for Enhanced Genome Editing Specificity. Cell *154*, 1380-1389.
Rasmussen, S.R., Füchtbauer, W., Novero, M., Volpe, V., Malkov, N., Genre, A., Bonfante, P., Stougaard, J., and Radutoiu, S. (2016). Intraradical colonization by arbuscular mycorrhizal fungi triggers induction of a lipochitooligosaccharide receptor. Scientific Reports *6*, 29733.
Rawat, A., Brejšková, L., Hála, M., Cvrčková, F., and Žárský, V. (2017). The Physcomitrella patens exocyst subunit EXO70. 3d has distinct roles in growth and development, and is essential for completion of the moss life cycle. New Phytol *216*, 438-454.
Read, D., Duckett, J., Francis, R., Ligrone, R., and Russell, A. (2000). Symbiotic fungal associations in 'lower' land plants. Philosophical Transactions of the Royal Society of London B: Biological Sciences *355*, 815-831.
Redecker, D., Kodner, R., and Graham, L.E. (2000). Glomalean fungi from the Ordovician. Science *289*.
Remy, W., Taylor, T.N., Hass, H., and Kerp, H. (1994). Four hundred-million-year-old vesicular arbuscular mycorrhizae. Proceedings of the National Academy of Sciences *91*, 11841-11843.
Ren, R., Wang, H., Guo, C., Zhang, N., Zeng, L., Chen, Y., Ma, H., and Qi, J. (2018). Widespread whole genome duplications contribute to genome complexity and species diversity in angiosperms. Molecular Plant *11*, 414-428.
Rensing, S.A. (2018). Plant evolution: phylogenetic relationships between the earliest land plants. Curr Biol *28*, R210-R213.
Rensing, S.A., Ick, J., Fawcett, J.A., Lang, D., Zimmer, A., Van de Peer, Y., and Reski, R. (2007). An ancient genome duplication contributed to the abundance of metabolic genes in the moss Physcomitrella patens. BMC Evol Biol *7*, 130.
Rich, M.K., Schorderet, M., Bapaume, L., Falquet, L., Morel, P., Vandenbussche, M., and Reinhardt, D. (2015). The Petunia GRAS transcription factor ATA/RAM1 regulates symbiotic gene expression and fungal morphogenesis in arbuscular mycorrhiza. Plant Physiol *168*, 788-797.
Ried, M.K., Antolin-Llovera, M., and Parniske, M. (2014). Spontaneous symbiotic reprogramming of plant roots triggered by receptor-like kinases. eLife *3*.
Riely, B.K., Lougnon, G., Ane, J.M., and Cook, D.R. (2007). The symbiotic ion channel homolog DMI1 is localized in the nuclear membrane of Medicago truncatula roots. Plant J *49*, 208-216.
Roth, R., and Paszkowski, U. (2017). Plant carbon nourishment of arbuscular mycorrhizal fungi. Curr Opin Plant Biol *39*, 50-56.
Rubinstein, C.V., Gerrienne, P., de la Puente, G.S., Astini, R.A., and Steemans, P. (2010). Early Middle Ordovician evidence for land plants in Argentina (eastern Gondwana). New Phytol *188*, 365-369.
Russell, J., and Bulman, S. (2005). The liverwort Marchantia foliacea forms a specialized symbiosis with arbuscular mycorrhizal fungi in the genus Glomus. New Phytol *165*, 567-579.
Saand, M.A., Xu, Y.P., Munyampundu, J.P., Li, W., Zhang, X.R., and Cai, X.Z. (2015). Phylogeny and evolution of plant cyclic nucleotide-gated ion channel (CNGC) gene family and functional analyses of tomato CNGCs. DNA Res *22*, 471-483.
Saint-Marcoux, D., Proust, H., Dolan, L., and Langdale, J.A. (2015). Identification of reference genes for real-time quantitative PCR experiments in the liverwort Marchantia polymorpha. PloS one *10*, e0118678.
Saito, K., Yoshikawa, M., Yano, K., Miwa, H., Uchida, H., Asamizu, E., Sato, S., Tabata, S., Imaizumi-Anraku, H., Umehara, Y.*, et al.* (2007). NUCLEOPORIN85 is required for calcium

spiking, fungal and bacterial symbioses, and seed production in Lotus japonicus. Plant Cell *19*, 610-624.
Sathyanarayanan, P.V., Siems, W.F., Jones, J.P., and Poovaiah, B.W. (2001). Calcium-stimulated autophosphorylation site of plant chimeric calcium/calmodulin-dependent protein kinase. J Biol Chem *276*, 32940-32947.
Schacht, H. (1854). Pilzfaden im Innern der Zellen und der

Starkmehlkorner vor. Flora *1854*, 618-624.
Schachtman, D.P., Reid, R.J., and Ayling, S.M. (1998). Phosphorus uptake by plants: from soil to cell. Plant Physiol *116*, 447-453.
Schüßler, A. (2000). Glomus claroideum forms an arbuscular mycorrhiza-like symbiosis with the hornwort Anthoceros punctatus. Mycorrhiza *10*, 15-21.
Schüßler, A. (2012). 5 The Geosiphon–Nostoc Endosymbiosis and Its Role as a Model for Arbuscular Mycorrhiza Research. In Fungal Associations (Springer), pp. 77-91.
Schüβler, A., Schwarzott, D., and Walker, C. (2001). A new fungal phylum, the Glomeromycota: phylogeny and evolution. Mycol Res *105*, 1413-1421.
Shah, V.K., and Brill, W.J. (1977). Isolation of an iron-molybdenum cofactor from nitrogenase. Proc Natl Acad Sci U S A *74*, 3249-3253.
Shaul, O., Galili, S., Volpin, H., Ginzberg, I., Elad, Y., Chet, I., and Kapulnik, Y. (1999). Mycorrhiza-induced changes in disease severity and pr protein expression in tobacco leaves. Mol Plant-Microbe Interact *12*, 1000-1007.
Shaw, S.L., and Long, S.R. (2003). Nod factor elicits two separable calcium responses in Medicago truncatula root hair cells. Plant Physiol *131*, 976-984.
Shimamura, M. (2016). Marchantia polymorpha : Taxonomy, phylogeny and morphology of a model system. Plant and Cell Physiology *57*, 230-256.
Simpson, J.T., Wong, K., Jackman, S.D., Schein, J.E., Jones, S.J., and Birol, I. (2009). ABySS: a parallel assembler for short read sequence data. Genome Res *19*, 1117-1123.
Singh, S., Katzer, K., Lambert, J., Cerri, M., and Parniske, M. (2014). CYCLOPS, A DNA-binding transcriptional activator, orchestrates symbiotic root nodule development. Cell host & microbe *15*, 139-152.
Smit, P., Raedts, J., Portyanko, V., Debelle, F., Gough, C., Bisseling, T., and Geurts, R. (2005). NSP1 of the GRAS protein family is essential for rhizobial Nod factor-induced transcription. Science *308*, 1789-1791.
Smith, S.E., and Read, D.J. (2008). Mycorrhizal Symbiosis (New York: Academic).
Soltis, D.E., Soltis, P.S., Morgan, D.R., Swensen, S.M., Mullin, B.C., Dowd, J.M., and Martin, P.G. (1995). Chloroplast gene sequence data suggest a single origin of the predisposition for symbiotic nitrogen fixation in angiosperms. Proceedings of the National Academy of Sciences *92*, 2647-2651.
Sperschneider, J., Catanzariti, A.-M., DeBoer, K., Petre, B., Gardiner, D.M., Singh, K.B., Dodds, P.N., and Taylor, J.M. (2017). LOCALIZER: subcellular localization prediction of both plant and effector proteins in the plant cell. Scientific Reports *7*, 44598.
Steinhorst, L., and Kudla, J. (2013). Calcium - a central regulator of pollen germination and tube growth. Biochim Biophys Acta *1833*, 1573-1581.
Stracke, S., Kistner, C., Yoshida, S., Mulder, L., Sato, S., Kaneko, T., Tabata, S., Sandal, N., Stougaard, J., Szczyglowski, K*., et al.* (2002). A plant receptor-like kinase required for both bacterial and fungal symbiosis. Nature *417*, 959-962.
Strullu-Derrien, C., Kenrick, P., Pressel, S., Duckett, J.G., Rioult, J.P., and Strullu, D.G. (2014). Fungal associations in Horneophyton ligneri from the Rhynie Chert (c. 407 million year old) closely resemble those in extant lower land plants: novel insights into ancestral plant-fungus symbioses. New Phytol *203*, 964-979.

Sugano, S.S., Nishihama, R., Shirakawa, M., Takagi, J., Matsuda, Y., Ishida, S., Shimada, T., Hara-Nishimura, I., Osakabe, K., and Kohchi, T. (2018). Efficient CRISPR/Cas9-based genome editing and its application to conditional genetic analysis in Marchantia polymorpha. bioRxiv.
Sugano, S.S., Shirakawa, M., Takagi, J., Matsuda, Y., Shimada, T., Hara-Nishimura, I., and Kohchi, T. (2014). CRISPR/Cas9 mediated targeted mutagenesis in the liverwort Marchantia polymorpha L. Plant and Cell Physiology, pcu014.
Sun, J., Miller, J.B., Granqvist, E., Wiley-Kalil, A., Gobbato, E., Maillet, F., Cottaz, S., Samain, E., Venkateshwaran, M., Fort, S*., et al.* (2015). Activation of symbiosis signaling by arbuscular mycorrhizal fungi in legumes and rice. Plant Cell *27*, 823-838.
Sun, J., Miwa, H., Downie, J.A., and Oldroyd, G.E.D. (2007). Mastoparan activates calcium spiking analogous to nod factor-induced responses in Medicago truncatula root hair cells. Plant Physiol *144*, 695-702.
Sundell, D., Mannapperuma, C., Netotea, S., Delhomme, N., Lin, Y.-C., Sjödin, A., Van de Peer, Y., Jansson, S., Hvidsten, T.R., and Street, N.R. (2015). The plant genome integrative explorer resource: Plantgenie.org. New Phytol *208*, 1149-1156.
Sunkar, R., Kaplan, B., Bouché, N., Arazi, T., Dolev, D., Talke, I.N., Maathuis, F.J.M., Sanders, D., Bouchez, D., and Fromm, H. (2000). Expression of a truncated tobacco NtCBP4 channel in transgenic plants and disruption of the homologous Arabidopsis CNGC1 gene confer Pb2+ tolerance. The Plant Journal *24*, 533-542.
Swensen, S.M. (1996). The evolution of actinorhizal symbioses: evidence for multiple origins of the symbiotic association. Am J Bot, 1503-1512.
Takeda, N., Maekawa, T., and Hayashi, M. (2012). Nuclear-localized and deregulated calcium- and calmodulin-dependent protein kinase activates rhizobial and mycorrhizal responses in Lotus japonicus. Plant Cell *24*, 810-822.
Talke, I.N., Blaudez, D., Maathuis, F.J., and Sanders, D. (2003). CNGCs: prime targets of plant cyclic nucleotide signalling? Trends Plant Sci *8*, 286-293.
Tang, N., San Clemente, H., Roy, S., Bécard, G., Zhao, B., and Roux, C. (2016). A survey of the gene repertoire of Gigaspora rosea unravels conserved features among Glomeromycota for obligate biotrophy. Frontiers in microbiology *7*, 233.
Tatsukami, Y., and Ueda, M. (2016). Rhizobial gibberellin negatively regulates host nodule number. Scientific Reports *6*.
Taylor, W.A. (1995). Spores in earliest land plants. Nature *373*, 391.
Taylor, W.A. (1997). Ultrastructure of lower Paleozoic dyads from southern Ohio II: Dyadospora murusattenuata, functional and evolutionary considerations. Rev Palaeobot Palynol *97*, 1-8.
Tirichine, L., Imaizumi-Anraku, H., Yoshida, S., Murakami, Y., Madsen, L.H., Miwa, H., Nakagawa, T., Sandal, N., Albrektsen, A.S., and Kawaguchi, M. (2006). Deregulation of a Ca2+/calmodulin-dependent kinase leads to spontaneous nodule development. Nature *441*, 1153-1156.
Tisserant, E., Malbreil, M., Kuo, A., Kohler, A., Symeonidi, A., Balestrini, R., Charron, P., Duensing, N., dit Frey, N.F., and Gianinazzi-Pearson, V. (2013). Genome of an arbuscular mycorrhizal fungus provides insight into the oldest plant symbiosis. Proceedings of the National Academy of Sciences *110*, 20117-20122.
Trappe, J.M. (1987). Phylogenetic and ecologic aspects of mycotrophy in the angiosperms from an evolutionary standpoint. Ecophysiology of VA Mycorrhizal Plants, 2-25.
Tu, B., Hu, L., Chen, W., Li, T., Hu, B., Zheng, L., Lv, Z., You, S., Wang, Y., Ma, B*., et al.* (2015). Disruption of OsEXO70A1 causes irregular vascular bundles and perturbs mineral nutrient assimilation in rice. Scientific Reports *5*, 18609.

Tunc-Ozdemir, M., Rato, C., Brown, E., Rogers, S., Mooneyham, A., Frietsch, S., Myers, C.T., Poulsen, L.R., Malhó, R., and Harper, J.F. (2013). Cyclic Nucleotide Gated Channels 7 and 8 Are Essential for Male Reproductive Fertility. PloS one *8*, e55277.
Upson, R., Read, D., and Newsham, K. (2007). Widespread association between the ericoid mycorrhizal fungus Rhizoscyphus ericae and a leafy liverwort in the maritime and sub-Antarctic. New Phytol *176*, 460-471.
Urquhart, W., Chin, K., Ung, H., Moeder, W., and Yoshioka, K. (2011). The cyclic nucleotide-gated channels AtCNGC11 and 12 are involved in multiple Ca2+-dependent physiological responses and act in a synergistic manner. J Exp Bot *62*, 3671-3682.
Urquhart, W., Gunawardena, A.H.L.A.N., Moeder, W., Ali, R., Berkowitz, G.A., and Yoshioka, K. (2007). The chimeric cyclic nucleotide-gated ion channel ATCNGC11/12 constitutively induces programmed cell death in a Ca2+ dependent manner. Plant Mol Biol *65*, 747-761.
van Velzen, R., Holmer, R., Bu, F., Rutten, L., van Zeijl, A., Liu, W., Santuari, L., Cao, Q., Sharma, T., Shen, D*., et al.* (2018). Comparative genomics of the nonlegume Parasponia reveals insights into evolution of nitrogen-fixing rhizobium symbioses. Proceedings of the National Academy of Sciences.
Venkateshwaran, M., Cosme, A., Han, L., Banba, M., Satyshur, K.A., Schleiff, E., Parniske, M., Imaizumi-Anraku, H., and Ane, J.M. (2012). The recent evolution of a symbiotic ion channel in the legume family altered ion conductance and improved functionality in calcium signaling. Plant Cell *24*, 2528-2545.
Venkateshwaran, M., Jayaraman, D., Chabaud, M., Genre, A., Balloon, A.J., Maeda, J., Forshey, K., den Os, D., Kwiecien, N.W., Coon, J.J*., et al.* (2015). A role for the mevalonate pathway in early plant symbiotic signaling. P Natl Acad Sci USA *112*, 9781-9786.
Verret, F., Wheeler, G., Taylor, A.R., Farnham, G., and Brownlee, C. (2010). Calcium channels in photosynthetic eukaryotes: implications for evolution of calcium-based signalling. New Phytol *187*, 23-43.
Vierheilig, H., Coughlan, A.P., Wyss, U., and Piché, Y. (1998). Ink and vinegar, a simple staining technique for arbuscular-mycorrhizal fungi. Appl Environ Microbiol *64*, 5004-5007.
Villarreal, A.J.C., Crandall-Stotler, B.J., Hart, M.L., Long, D.G., and Forrest, L.L. (2016). Divergence times and the evolution of morphological complexity in an early land plant lineage (Marchantiopsida) with a slow molecular rate. New Phytol *209*, 1734-1746.
Volpe, V., Giovannetti, M., Sun, X.-G., Fiorilli, V., and Bonfante, P. (2016). The phosphate transporters LjPT4 and MtPT4 mediate early root responses to phosphate status in non mycorrhizal roots. Plant, Cell Environ *39*, 660-671.
Vujičić, M., Sabovljević, A., and Sabovljević, M. (2011). Axenically culturing the bryophytes: Establishment and propagation of the moss Hypnum cupressiforme Hedw.(Bryophyta, Hypnaceae) in in vitro conditions. Botanica Serbica *35*, 71-77.
Wais, R.J., Galera, C., Oldroyd, G., Catoira, R., Penmetsa, R.V., Cook, D., Gough, C., Dénarié, J., and Long, S.R. (2000). Genetic analysis of calcium spiking responses in nodulation mutants of Medicago truncatula. Proceedings of the National Academy of Sciences *97*, 13407-13412.
Walker, S.A., Viprey, V., and Downie, J.A. (2000). Dissection of nodulation signaling using pea mutants defective for calcium spiking induced by Nod factors and chitin oligomers. P Natl Acad Sci USA *97*, 13413-13418.
Wang, B., and Qiu, Y.L. (2006). Phylogenetic distribution and evolution of mycorrhizas in land plants. Mycorrhiza *16*.
Wang, B., Yeun, L.H., Xue, J.Y., Liu, Y., Ane, J.M., and Qiu, Y.L. (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*.

Wang, E., Yu, N., Bano, S.A., Liu, C., Miller, A.J., Cousins, D., Zhang, X., Ratet, P., Tadege, M., Mysore, K.S.*, et al.* (2014). A H+-ATPase that energizes nutrient uptake during mycorrhizal symbioses in rice and Medicago truncatula. Plant Cell *26*, 1818-1830.
Wang, M.B., Helliwell, C.A., Wu, L.M., Waterhouse, P.M., Peacock, W.J., and Dennis, E.S. (2008). Hairpin RNAs derived from RNA polymerase II and polymerase III promoter-directed transgenes are processed differently in plants. RNA *14*, 903-913.
Wang, Y., Kang, Y., Ma, C., Miao, R., Wu, C., Long, Y., Ge, T., Wu, Z., Hou, X., Zhang, J.*, et al.* (2017). CNGC2 is a Ca2+influx channel that prevents accumulation of apoplastic Ca2+ in the leaf. Plant Physiol *173*, 1342-1354.
Wang, Y., Wang, X., Tang, H., Tan, X., Ficklin, S.P., Feltus, F.A., and Paterson, A.H. (2011). Modes of gene duplication contribute differently to genetic novelty and redundancy, but show parallels across divergent angiosperms. PloS one *6*, e28150.
Wang, Y.F., Munemasa, S., Nishimura, N., Ren, H.M., Robert, N., Han, M., Puzorjova, I., Kollist, H., Lee, S., Mori, I.*, et al.* (2013). Identification of cyclic GMP-activated nonselective Ca2+-permeable cation channels and associated CNGC5 and CNGC6 genes in Arabidopsis guard cells. Plant Physiol *163*, 578-590.
Ward, J.M., Mäser, P., and Schroeder, J.I. (2009). Plant ion channels: Gene families, physiology, and functional genomics analyses. Annu Rev Physiol *71*, 59-82.
Weber, E., Engler, C., Gruetzner, R., Werner, S., and Marillonnet, S. (2011). A modular cloning system for standardized assembly of multigene constructs. PloS one *6*, e16765.
Wellman, C.H., Osterloff, P.L., and Mohiuddin, U. (2003). Fragments of the earliest land plants. Nature *425*, 282.
Werner, G., Cornwell, W., Sprent, J., Kattge, J., and Kiers, E. (2014). Data from: A single evolutionary innovation drives the deep evolution of symbiotic N2-fixation in angiosperms. Dryad Digital Repository, doi: 10.1038/ncomms5087.
Wewer, V., Brands, M., and Dörmann, P. (2014). Fatty acid synthesis and lipid metabolism in the obligate biotrophic fungus Rhizophagus irregularis during mycorrhization of Lotus japonicus. The Plant Journal *79*, 398-412.
Whipps, J.M. (2004). Prospects and limitations for mycorrhizas in biocontrol of root pathogens. Canadian Journal of Botany *82*, 1198-1227.
Wickett, N.J., Mirarab, S., Nguyen, N., Warnow, T., Carpenter, E., Matasci, N., Ayyampalayam, S., Barker, M.S., Burleigh, J.G., Gitzendanner, M.A.*, et al.* (2014). Phylotranscriptomic analysis of the origin and early diversification of land plants. Proceedings of the National Academy of Sciences *111*, E4859-E4868.
Xie, X., Huang, W., Liu, F., Tang, N., Liu, Y., Lin, H., and Zhao, B. (2013). Functional analysis of the novel mycorrhiza-specific phosphate transporter AsPT1 and PHT1 family from Astragalus sinicus during the arbuscular mycorrhizal symbiosis. New Phytol *198*, 836-852.
Xie, Y., Wu, G., Tang, J., Luo, R., Patterson, J., Liu, S., Huang, W., He, G., Gu, S., Li, S.*, et al.* (2014). SOAPdenovo-Trans: de novo transcriptome assembly with short RNA-Seq reads. Bioinformatics *30*, 1660-1666.
Xue, L., Cui, H., Buer, B., Vijayakumar, V., Delaux, P.M., Junkermann, S., and Bucher, M. (2015). Network of GRAS transcription factors involved in the control of arbuscule development in Lotus japonicus. Plant Physiol *167*, 854-871.
Yadav, V., Molina, I., Ranathunge, K., Castillo, I.Q., Rothstein, S.J., and Reed, J.W. (2014). ABCG transporters are required for suberin and pollen wall extracellular barriers in Arabidopsis. The Plant Cell *26*, 3569-3588.
Yang, S.Y., Gronlund, M., Jakobsen, I., Grotemeyer, M.S., Rentsch, D., Miyao, A., Hirochika, H., Kumar, C.S., Sundaresan, V., Salamin, N.*, et al.* (2012). Nonredundant regulation of rice arbuscular mycorrhizal symbiosis by two members of the phosphate transporter1 gene family. Plant Cell *24*, 4236-4251.

Yano, K., Yoshida, S., Muller, J., Singh, S., Banba, M., Vickers, K., Markmann, K., White, C., Schuller, B., Sato, S*., et al.* (2008). CYCLOPS, a mediator of symbiotic intracellular accommodation. Proc Natl Acad Sci U S A *105*, 20540-20545.
Young, N.D., and Bharti, A.K. (2012). Genome-enabled insights into legume biology. Annu Rev Plant Biol *63*, 283-305.
Young, N.D., Debelle, F., Oldroyd, G.E.D., Geurts, R., Cannon, S.B., Udvardi, M.K., Benedito, V.A., Mayer, K.F.X., Gouzy, J., Schoof, H*., et al.* (2011). The Medicago genome provides insight into the evolution of rhizobial symbioses. Nature *480*, 520-524.
Yu, N., Luo, D., Zhang, X., Liu, J., Wang, W., Jin, Y., Dong, W., Liu, J., Liu, H., Yang, W*., et al.* (2014). A DELLA protein complex controls the arbuscular mycorrhizal symbiosis in plants. Cell Res *24*, 130-133.
Zelman, A.K., Dawe, A., Gehring, C., and Berkowitz, G.A. (2012). Evolutionary and structural perspectives of plant cyclic nucleotide-gated cation channels. Front Plant Sci *3*, 95.
Zeng, W., Melotto, M., and He, S.Y. (2010). Plant stomata: a checkpoint of host immunity and pathogen virulence. Curr Opin Biotechnol *21*, 599-603.
Zhang, Q., Blaylock, L.A., and Harrison, M.J. (2010). Two Medicago truncatula half-ABC transporters are essential for arbuscule development in arbuscular mycorrhizal symbiosis. Plant Cell *22*, 1483-1497.
Zhang, S., Pan, Y., Tian, W., Dong, M., Zhu, H., Luan, S., and Li, L. (2017). Arabidopsis CNGC14 Mediates Calcium Influx Required for Tip Growth in Root Hairs. Molecular Plant *10*, 1004-1006.
Zhang, X., Pumplin, N., Ivanov, S., and Harrison, M.J. (2015). EXO70I is required for development of a sub-domain of the periarbuscular membrane during arbuscular mycorrhizal symbiosis. Curr Biol *25*, 2189-2195.
Zhang, Y., and Guo, L.D. (2007). Arbuscular mycorrhizal structure and fungi associated with mosses. Mycorrhiza *17*, 319-325.
Zhou, J., Shen, B., Zhang, W., Wang, J., Yang, J., Chen, L., Zhang, N., Zhu, K., Xu, J., Hu, B*., et al.* (2014a). One-step generation of different immunodeficient mice with multiple gene modifications by CRISPR/Cas9 mediated genome engineering. The International Journal of Biochemistry & Cell Biology *46*, 49-55.
Zhou, J., Wang, J., Shen, B., Chen, L., Su, Y., Yang, J., Zhang, W., Tian, X., and Huang, X. (2014b). Dual sgRNAs facilitate CRISPR/Cas9-mediated mouse genome targeting. The FEBS journal *281*, 1717-1725.
Zhu, Y., Wu, B., and Guo, W. (2017). The role of Exo70 in exocytosis and beyond. Small GTPases, 1-5.

# Acronyms and abbreviations

ABC ATP-binding cassette

AM Arbuscular Mycorrhiza/Mycorrhizal

AMT Ammonium Transporter

BLAST Basic Local Alignment Search Tool

BSA Bovine Serum Albumin

$Ca2^{+}$ Calcium

Cas 9 CRISPR-Associated Protein 9

CCaMK Calcium/Calmodulin-dependent Protein Kinase

CERK Ceramide Kinase

CNGC Cyclic Nucleotide Gated Channel

cNMP Cyclic Nucleotide Monophosphate

CO Chitin Oligomer

CRISPR Clustered Regularly Interspaced Short Palindromic Repeat

CSSP Common Symbiosis Signalling Pathway

DMI Does not Make Infections

DPI Days Post Inoculation

ER Endoplasmic Reticulum

FP Fahreus Plant

GlcNAc N-Acetylglucosamine

GPAT Glycerol-3-Phosphate Acyltransferase

GRAS Gibberellic Acid Insensitive (GAI), Repressor of Gai, and Scarecrow Family Protein

GUS β-Glucuronidase

HA Proton Pump ATPase

HR Homologous Recombination

HMGR1 3-Hydroxy-3-Methylglutaryl CoA Reductase 1

IPD3 Interacting Protein of DMI3

LB Lysogeny Broth

LCO Lipochitooligosaccharide

LRR Leucine-Rich Repeat

LYK LysM Domain-Containing Receptor-Like Kinases

LYS LysM Receptor Kinase

LysM Lysin Domain

β-MAG 16:0 β-monoacylglycerol

MAPK Mitogen Activated Protein Kinase

MCA M. truncatula Calcium ATPase

MSP Major Sperm Protein

MYA Million Years Ago

MYB Myeloblastosis Family Gene

NFN Nitrogen Fixing Root Nodule

NFP Nod Factor Perception

NFR Nod Factor Receptor

NSP Nodulation Signalling Pathway

Nup/NUP Nucleoporin

OD Optical Density/Absorbance

PAM Peri-Arbuscular Membrane

PCR Polymerase Chain Reaction

PHT1 Phosphate Transporter 1

PT Phosphate Transporters

RAD Required for Arbuscule Development

RAM Reduced Arbuscular Mycorrhization

RLK Receptor Like Kinase

RN Root Nodule

RNAi RNA interference

SERCA Sarco/Endoplasmic Reticulum Ca2+-ATPase

SIP2 SymRK-Interacting Protein 2

SNARE Soluble N-Ethylmaleimide-Factor Attachment Protein Receptor

SYMRK Symbiosis Receptor Kinase

VAMP Vesicle Associated Membrane Protein

STR Stunted Arbuscule

SYP Syntaxin of Plants

TUBa1 α-Tubulin

WGD Whole Genome Duplication

WPI Weeks Post Inoculation

WT Wild Type

YT Yeast Tryptone

# Appendix 1

| Name | Identity |
|---|---|
| AmtrPollux | XP_006847006.1_ion_channel_DMI1_isoform_X1_Amborella_trichopoda |
| AmtrCastor | XP_011628711.1_ion_channel_CASTOR_isoform_X2_Amborella_trichopoda |
| AmtrPlxL | XP_011621607.1_putative_ion_channel_POLLUX-like_2_Amborella_trichopoda |
| AqcoPollux | PIA47269.1_hypothetical_protein_AQUCO_01400149v1_Aquilegia_coerulea |
| AqcoCastor | PIA45162.1_hypothetical_protein_AQUCO_01700594v1_Aquilegia_coerulea |
| AcqoPlxL | PIA31537.1_hypothetical_protein_AQUCO_04900084v1_Aquilegia_coerulea |
| ArtaDMI1 | AT5G49960 |
| ArtaPlxL2 | AT5G02940 |
| ArtaPlxL1 | AT5G43745 |
| MapoDMI1 | Mapoly0173s0003.1 |
| MapoDMI1f | Mapoly0024s0089.1 |
| PhpaDMI1a | Pp3c24_1450V3.1 |
| PhpaDMI1b | Pp3c6_21060V3.2 |
| PhpaDMI1c | Pp3c17_14700V3.1 |
| PhpaDMI1d | Pp3c10_22510V3.2 |
| SpfaDMI1a | Sphfalx0154s0016.1 |
| SpfaDMI1c | Sphfalx0019s0067.2 |
| SpfaDMI1b | Sphfalx0154s0025.1 |
| SpfaDMI1d | Sphfalx0081s0011.1 |
| SemoDMI1a | 60947_Selaginella_moellendorffii_scaffold_7:1153295..1156194_reverse |
| SemoDMI1b | 84001_Selaginella_moellendorffii_scaffold_6:459418..462003_reverse |
| SemoDMI1c | 86132__Selaginella_moellendorffii_scaffold_7:1778439..1781145_reverse |
| SemoDMI1d | 169248_Selaginella_moellendorffii__scaffold_9:806507..809924_reverse |
| SemoDMI1e | 104898_Selaginella_moellendorffii_scaffold_30:1411261..1418385_reverse_e_gw1.30.299.1 |
| BrdiPollux | XP_010232620.2_probable_ion_channel_POLLUX_Brachypodium_distachyon |
| BrdiCastor | XP_010235421.1_probable_ion_channel_CASTOR_isoform_X2_Brachypodium_distachyon |
| BrdiPlxL | XP_024317391.1_putative_ion_channel_POLLUX-like_2_isoform_X2_Brachypodium_distachyon |
| BrraPollux | XP_009127363.1_PREDICTED:_probable_ion_channel_POLLUX_isoform_X1_Brassica_rapa |
| BrraPlxL | XP_009117335.1_PREDICTED:_putative_ion_channel_POLLUX-like_2_isoform_X3_Brassica_rapa |
| CosuDMI1a | 584-_Coccomyxa_subellipsoidea-_scaffold_8:640213..646710_reverse |
| CosuDMI1b | 55736Coccomyxa_subellipsoidea-_scaffold_1:3656427..3666755_reverse |
| MispDMI1a | 80696_-_Micromonas_sp._RCC299_Chr_04:480099..481727_reverse |
| MispDMI1b | 60543__Micromonas_sp._RCC299_Chr_08:246417..248939_reverse |
| DusaDMI1 | Dusal.0939s00002.1__Dunaliella_salina |
| OsluDMI1 | 94483_Ostreococcus_lucimarinus_Chr_6:261452..263344_reverse |

| | |
|---|---|
| CisiPollux | XP_006471154.1_ion_channel_DMI1_isoform_X2_Citrus_sinensis |
| CisiCastor | XP_006492774.1_ion_channel_CASTOR-like_isoform_X1_Citrus_sinensis |
| CisiPlxL | XP_006489497.1_putative_ion_channel_POLLUX-like_2_isoform_X1_Citrus_sinensis |
| CusaPollux | XP_011655542.1_PREDICTED:_probable_ion_channel_POLLUX_Cucumis_sativus |
| CusaCastor | XP_011655555.1_PREDICTED:_ion_channel_CASTOR-like_isoform_X2_Cucumis_sativus |
| CusaPlxL | XP_011650951.1_PREDICTED:_putative_ion_channel_POLLUX-like_2_isoform_X2_Cucumis_sativus |
| EugrPollux | XP_010030165.1_PREDICTED:_ion_channel_DMI1_isoform_X2_Eucalyptus_grandis |
| EugrCastor | XP_010069236.2_PREDICTED:_ion_channel_CASTOR_isoform_X2_Eucalyptus_grandis |
| EugrPlxL | XP_018717425.1_PREDICTED:_putative_ion_channel_POLLUX-like_2_isoform_X4_Eucalyptus_grandis |
| GlmaSym8 | XP_003539390.1_PREDICTED:_probable_ion_channel_SYM8_isoform_X1_Glycine_max |
| GlmaCastor | XP_003554802.1_PREDICTED:_ion_channel_CASTOR_isoform_X1_Glycine_max |
| GlmaPlxL | XP_006589264.1_PREDICTED:_putative_ion_channel_POLLUX-like_2_isoform_X6_Glycine_max |
| GoraPollux | XP_012480607.1_PREDICTED:_ion_channel_DMI1-like_Gossypium_raimondii |
| GoraCastor | XP_012481430.1_PREDICTED:_ion_channel_CASTOR-like_isoform_X2_Gossypium_raimondii |
| GoraPlxL | XP_012487233.1_PREDICTED:_putative_ion_channel_POLLUX-like_2_isoform_X4_Gossypium_raimondii |
| EnbaDMI1 | onekp:GNQG_scaffold_2015334_Encephalartos_barteri |
| MadoPollux | XP_008347915.1_PREDICTED:_ion_channel_DMI1-like_Malus_domestica |
| MadoCastor | XP_017183007.1_PREDICTED:_ion_channel_CASTOR-like_Malus_domestica |
| MadoPlxL | XP_008341043.1_PREDICTED:_putative_ion_channel_POLLUX-like_2_Malus_domestica |
| MetrDMI1 | XP_003592931.1_ion_channel_DMI1_isoform_X1_Medicago_truncatula |
| MetrCastor | XP_013450641.1_ion_channel_CASTOR_isoform_X1_Medicago_truncatula |
| MetrPlxL | XP_013468755.1_putative_ion_channel_POLLUX-like_2_isoform_X3_Medicago_truncatula |
| OrsaPollux | BAD81711.1_putative_DMI1_protein_Oryza_sativa_Japonica_Group |
| OrsaCastor | XP_015629685.1_PREDICTED:_probable_ion_channel_CASTOR_isoform_X1_Oryza_sativa_Japonica_Group |
| OrsaPlxL | XP_015632613.1_PREDICTED:_putative_ion_channel_POLLUX-like_2_Oryza_sativa_Japonica_Group |
| PotrPollux | XP_006389539.2_ion_channel_DMI1_isoform_X1_Populus_trichocarpa |
| PotrPollux2 | PNT42617.1_hypothetical_protein_POPTR_004G223400v3_Populus_trichocarpa |
| PotrCastor | XP_024438968.1_ion_channel_CASTOR_isoform_X4_Populus_trichocarpa |
| PotrDMI1a | PNT43017.1_hypothetical_protein_POPTR_003G008800v3_Populus_trichocarpa |
| PotrDMI1b | XP_024446986.1_ion_channel_CASTOR-like_partial_Populus_trichocarpa |
| PotrPlxL | XP_024466973.1_putative_ion_channel_POLLUX-like_2_isoform_X2_Populus_trichocarpa |
| RicoPollux | XP_002526461.2_ion_channel_DMI1_partial_Ricinus_communis |
| RicoCastor | XP_002517736.1_ion_channel_CASTOR_isoform_X1_Ricinus_communis |
| RicoPlxL | XP_025012750.1_putative_ion_channel_POLLUX-like_2_isoform_X1_Ricinus_communis |

| | |
|---|---|
| SolyPollux | XP_004248837.1_PREDICTED:_probable_ion_channel_POLLUX_isoform_X1_Solanum_lycopersicum |
| SolyCastor | XP_019071651.1_PREDICTED:_ion_channel_DMI1-like_isoform_X2_Solanum_lycopersicum |
| SolyPlxL | XP_004247688.1_PREDICTED:_putative_ion_channel_POLLUX-like_2_Solanum_lycopersicum |
| SobiPollux | XP_021310983.1_probable_ion_channel_POLLUX_Sorghum_bicolor |
| SobiPlxL | XP_021306564.1_putative_ion_channel_POLLUX-like_2_isoform_X1_Sorghum_bicolor |
| ZemaPollux | XP_008674775.1_probable_ion_channel_POLLUX_Zea_mays |
| ZemaCastor | AQK61629.1_Putative_ion_channel_protein_CASTOR_Zea_mays |
| ZemaPlxL | XP_008651290.1_putative_ion_channel_POLLUX-like_2_isoform_X1_Zea_mays |
| MapaDMI1 | MpaDMI1 |
| ArthPLXL1 | AT5G43745 |
| ArthPLXL2 | AT5G02940 |
| ArthPollux | AT5G49960 |
| AmarDMI1 | onekp:IAJW_scaffold_2011822_Amentotaxus_argotaenia |
| AtcuDMI1a | onekp:XIRK_scaffold_2003216_Athrotaxis_cupressoides |
| AtcuDMI1b | onekp:XIRK_scaffold_2003787_Athrotaxis_cupressoides |
| AtcuDMI1c | onekp:XIRK_scaffold_2003788_Athrotaxis_cupressoides |
| AuchDMI1a | onekp:YYPE_scaffold_2003933_Austrocedrus_chilensis |
| AuchDMI1b | onekp:YYPE_scaffold_2003934_Austrocedrus_chilensis |
| AuchDMI1c | onekp:YYPE_scaffold_2006322_Austrocedrus_chilensis |
| AuchDMI1d | onekp:YYPE_scaffold_2063613_Austrocedrus_chilensis |
| CehaDMI1a | onekp:WYAJ_scaffold_2005739_Cephalotaxus_harringtonia |
| CehaDMI1b | onekp:WYAJ_scaffold_2005740_Cephalotaxus_harringtonia |
| CehaDMI1c | onekp:WYAJ_scaffold_2005741_Cephalotaxus_harringtonia |
| CehaDMI1d | onekp:WYAJ_scaffold_2016128_Cephalotaxus_harringtonia |
| DacoDMI1a | onekp:FMWZ_scaffold_2005809_Dacrycarpus_compactus |
| DacoDMI1b | onekp:FMWZ_scaffold_2005810_Dacrycarpus_compactus |
| EnbaDMI1 | onekp:GNQG_scaffold_2004541_Encephalartos_barteri |
| LafrDMI1a | onekp:ZQWM_scaffold_2008468_Lagarostrobos_franklinii |
| LafrDMI1b | onekp:ZQWM_scaffold_2013136_Lagarostrobos_franklinii |
| PocoDMI1a | onekp:SCEB_scaffold_2009007_Podocarpus_coriaceus |
| PocoDMI1b | onekp:SCEB_scaffold_2010877_Podocarpus_coriaceus |
| TacrDMI1a | onekp:QSNJ_scaffold_2003354_Taiwania_cryptomerioides |
| TacrDMI1b | onekp:QSNJ_scaffold_2067490_Taiwania_cryptomerioides |
| WonoDMI1a | onekp:RSCE_scaffold_2008206_Wollemia_nobilis |
| WonoDMI1b | onekp:RSCE_scaffold_2009658_Wollemia_nobilis |
| WonoDMI1c | onekp:RSCE_scaffold_2009869_Wollemia_nobilis |
| ZyspDMI1 | onekp:MFZO_scaffold_2004223_Zygnemopsis_sp. |
| CycuDMI1 | onekp:JOJQ_scaffold_2040819_Cylindrocystis_cushleckae |
| CybrDMI1a | onekp:YOXI_scaffold_2054777_Cylindrocystis_brebissonii |
| MecaDMI1 | onekp:HKZW_scaffold_2002405_Mesotaenium_caldariorum |
| CybrDMI1b | onekp:RPGL_scaffold_2016947_Cylindrocystis_brebissonii |
| CyspDMI1 | onekp:VAZE_scaffold_2009657_Cylindrocystis_sp. |

| | |
|---|---|
| PeexDMI1 | onekp:YSQT_scaffold_2036866_Penium_exiguum |
| MeenDMI1 | onekp:WDCW_scaffold_2005938_Mesotaenium_endlicherianum |
| CoocDMI1 | onekp:HJVM_scaffold_2030069_Cosmarium_ochthodes |
| CoirDMI1 | onekp:QPDY_scaffold_2031354_Coleochaete_irregularis |
| CosubDMI1 | onekp:WDGV_scaffold_2010788_Cosmarium_subtumidum |
| DacoPLXL | onekp:FMWZ_scaffold_2012933_Dacrycarpus_compactus |
| CoocPLXL | onekp:HJVM_scaffold_2021464_Cosmarium_ochthodes |
| MecaPLXL | onekp:HKZW_scaffold_2003519_Mesotaenium_caldariorum |
| CycuPLXL | onekp:JOJQ_scaffold_2007839_Cylindrocystis_cushleckae |
| TacrPLXL | onekp:QSNJ_scaffold_2069164_Taiwania_cryptomerioides |
| WonoPLXL | onekp:RSCE_scaffold_2006858_Wollemia_nobilis |
| PocoPLXL | onekp:SCEB_scaffold_2055654_Podocarpus_coriaceus |
| CyspPLXL | onekp:VAZE_scaffold_2009975_Cylindrocystis_sp. |
| CosuPLXL | onekp:WDGV_scaffold_2010788_Cosmarium_subtumidum |
| CosbPLXL1 | onekp:WDGV_scaffold_2051323_Cosmarium_subtumidum |
| CosbPLXL2 | onekp:WDGV_scaffold_2055529_Cosmarium_subtumidum(2) |
| CosbPLXL3 | onekp:WDGV_scaffold_2056355_Cosmarium_subtumidum |
| CehaPLXL | onekp:WYAJ_scaffold_2016296_Cephalotaxus_harringtonia |
| AtcuPLXL | onekp:XIRK_scaffold_2009540_Athrotaxis_cupressoides |
| CybrPLXL | onekp:YOXI_scaffold_2010364_Cylindrocystis_brebissonii |
| PeenPLXL1 | onekp:YSQT_scaffold_2036866_Penium_exiguum |
| PeenPLXL2 | onekp:YSQT_scaffold_2036899_Penium_exiguum |
| AuchPLXL | onekp:YYPE_scaffold_2006564_Austrocedrus_chilensis |
| LafrPLXL1 | onekp:ZQWM_scaffold_2009865_Lagarostrobos_franklinii |
| LafrPLXL2 | onekp:ZQWM_scaffold_2087326_Lagarostrobos_franklinii |
| KlniDMI1a | kfl00177_0270_v1.1_Protein |
| KlniDMI1b | kfl01079_0020_v1.1_Protein |
| KlniDMI1c | kfl00806_0020_v1.1_Protein |
| KlniDMI1d | kfl00911_0030_v1.1_Protein |

**Sequences used in phylogenetic analysis: DMI1**

| Name | Identity |
|---|---|
| AmCNGC15 | evm_27_model_AmTr_v1_0_scaffold00148.48 |
| AmCNGCI | evm_27_model_AmTr_v1_0_scaffold00053.2 |
| AmCNGCII | evm_27_model_AmTr_v1_0_scaffold00210.4 |
| AmCNGCIIIa | gnl_Ambo_Trinity_comp11271_c3_seq2 |
| AmCNGCIIIb | evm_27_model_AmTr_v1_0_scaffold00019.141 |
| AmCNGCIVA | evm_27_model_AmTr_v1_0_scaffold00024.111 |
| AmCNGCIVB | evm_27_model_AmTr_v1_0_scaffold00003.425 |
| AtCNGC1 | AT5G53130.1 |
| AtCNGC10 | AT5G15410.1 |
| AtCNGC11 | AT2G46430.1 |
| AtCNGC12 | AT5G54250.1 |
| AtCNGC13 | AT5G57940.1 |
| AtCNGC14 | AT2G23980.1 |
| AtCNGC15 | AT1G15990.1 |
| AtCNGC16 | AT1G19780.1 |
| AtCNGC17 | AT4G30560.1 |
| AtCNGC18 | AT1G01340.2 |
| AtCNGC19 | AT2G46440.1 |
| AtCNGC2 | AT2G46450.1 |
| AtCNGC20 | AT4G01010.1 |
| AtCNGC3 | AT2G24610.1 |
| AtCNGC4 | AT2G28260.1 |
| AtCNGC5 | AT3G48010.1 |
| AtCNGC6 | AT4G30360.1 |
| AtCNGC7 | AT5G14870.1 |
| AtCNGC8 | AT3G17690.1 |
| AtCNGC9 | AT3G17700.1 |
| BbCNGC1 | onekp:QWFV_scaffold_2021887_Bambusina_borreri |
| BdCNGC15a | BRADI3G49290.1 |
| BdCNGC15b | BRADI3G47900.1 |
| BdCNGC2 | BRADI1G08060.1 |
| BdCNGC4a | BRADI2G51836.1 |
| BdCNGC4b | BRADI2G20660.1 |
| BdCNGCIa | BRADI1G38680.1 |
| BdCNGCIb | BRADI3G09550.1 |
| BdCNGCIIa | BRADI5G23700.1 |
| BdCNGCIIb | BRADI1G13740.1 |
| BdCNGCIIc | BRADI1G78010.1 |
| BdCNGCIIIa | BRADI2G57697.1 |
| BdCNGCIIIb | BRADI1G46970.1 |
| BdCNGCIIIc | BRADI3G53747.1 |
| BdCNGCIIId | BRADI4G37740.1 |
| BdCNGCIVAa | BRADI1G46260.1 |

| | |
|---|---|
| BdCNGCIVAb | BRADI3G57510.1 |
| CbCNGC1 | onekp:RPGL_scaffold_2019733_Cylindrocystis_brebissonii |
| CbrCNGC1 | onekp:YOXI_scaffold_2010374_Cylindrocystis_brebissonii |
| CbrCNGC2 | onekp:YOXI_scaffold_2010406_Cylindrocystis_brebissonii |
| CbrCNGC3 | onekp:YOXI_scaffold_2056905_Cylindrocystis_brebissonii |
| CcCNGC1 | onekp:JOJQ_scaffold_2007960_Cylindrocystis_cushleckae |
| CcCNGC2 | onekp:JOJQ_scaffold_2008465_Cylindrocystis_cushleckae |
| CcoCNGC1 | onekp:RQFE_scaffold_2002666_Cosmocladium_cf._constrictum |
| CcoCNGC2 | onekp:RQFE_scaffold_2002668_Cosmocladium_cf._constrictum |
| CcoCNGC3 | onekp:RQFE_scaffold_2002669_Cosmocladium_cf._constrictum |
| CcoCNGC4 | onekp:RQFE_scaffold_2004277_Cosmocladium_cf._constrictum |
| CcoCNGC5 | onekp:RQFE_scaffold_2009830_Cosmocladium_cf._constrictum |
| CcoCNGC6 | onekp:RQFE_scaffold_2038209_Cosmocladium_cf._constrictum |
| CgCNGC1 | onekp_DRGY_scaffold_2006531_Chaetosphaeridium_globosum |
| CgrCNGC1 | onekp:MNNM_scaffold_2044837_Cosmarium_granatum |
| CiCNGC1 | onekp:QPDY_scaffold_2006863_Coleochaete_irregularis |
| CiCNGC2 | onekp_QPDY_scaffold_2000889_Coleochaete_irregularis |
| CiCNGC3 | onekp_QPDY_scaffold_2004650_Coleochaete_irregularis |
| CoCNGC1 | onekp:HJVM_scaffold_2001370_Cosmarium_ochthodes |
| CoCNGC2 | onekp:HJVM_scaffold_2014616_Cosmarium_ochthodes |
| CoCNGC3 | onekp:HJVM_scaffold_2030295_Cosmarium_ochthodes |
| CrCNG1 | EDP02961 |
| CrCNG2 | EDP02965 |
| CrCNG3 | EDP04816 |
| CsCNGC1 | onekp_VQBJ_scaffold_2000954_Coleochaete_scutata |
| CsCNGC2 | onekp_VQBJ_scaffold_2005840_Coleochaete_scutata |
| CsuCNGC1 | onekp:WDGV_scaffold_2003289_Cosmarium_subtumidum |
| CsuCNGC2 | onekp:WDGV_scaffold_2013520_Cosmarium_subtumidum |
| CsuCNGC3 | onekp:WDGV_scaffold_2056132_Cosmarium_subtumidum |
| CsuCNGC4 | onekp:WDGV_scaffold_2056182_Cosmarium_subtumidum |
| CtCNGC1 | onekp:BHBK_scaffold_2012034_Cosmarium_tinctum |
| CylCNGC1 | onekp:VAZE_scaffold_2009847_Cylindrocystis_sp. |
| CylCNGC2 | onekp:VAZE_scaffold_2009953_Cylindrocystis_sp. |
| CylCNGC3 | onekp:VAZE_scaffold_2009996_Cylindrocystis_sp. |
| CylCNGC4 | onekp:VAZE_scaffold_2055681_Cylindrocystis_sp. |
| CylCNGC5 | onekp:VAZE_scaffold_2055731_Cylindrocystis_sp. |
| DaCNGC1 | onekp:DFDS_scaffold_2007972_Desmidium_aptogonum |
| EfCNGC1 | onekp_BFIK_scaffold_2030348_Entransia_fimbriat |
| EfCNGC2 | onekp:BFIK_scaffold_2028632_Entransia_fimbriat |
| GkCNGC1 | onekp:KEYW_scaffold_2004396_Gonatozygon_kinahanii |
| GkCNGC2 | onekp:KEYW_scaffold_2006513_Gonatozygon_kinahanii |
| GmCNGC14a | GLYMA06G08170.1 |
| GmCNGC14b | GLYMA04G08085.1 |
| GmCNGC15a | GLYMA13G20420.2 |

| | |
|---|---|
| GmCNGC15b | GLYMA10G06121.1 |
| GmCNGC15c | GLYMA12G08160.1 |
| GmCNGC15d | GLYMA12G29840.1 |
| GmCNGC15e | GLYMA13G39965.1 |
| GmCNGC16a | GLYMA04G35210.1 |
| GmCNGC16b | GLYMA06G19570.1 |
| GmCNGC17a | GLYMA04G24950.1 |
| GmCNGC17b | GLYMA06G30030.1 |
| GmCNGC18 | GLYMA06G08110.1 |
| GmCNGC1a | GLYMA14G31940.1 |
| GmCNGC1b | GLYMA06G13200.1 |
| GmCNGC1c | GLYMA04G41610.1 |
| GmCNGC2a | GLYMA18G49890.1 |
| GmCNGC2b | GLYMA08G26340.1 |
| GmCNGC4a | GLYMA12G34740.1 |
| GmCNGC4b | GLYMA06G42310.1 |
| GmCNGC4c | GLYMA12G16146.1 |
| GmCNGCIa | GLYMA07G06220.2 |
| GmCNGCIb | GLYMA16G02853.1 |
| GmCNGCIc | GLYMA19G44451.1 |
| GmCNGCId | GLYMA03G41790.2 |
| GmCNGCIe | GLYMA03G41780.1 |
| GmCNGCIf | GLYMA19G44430.1 |
| GmCNGCIIa | GLYMA07G02560.2 |
| GmCNGCIIb | GLYMA08G23460.1 |
| GmCNGCIIc | GLYMA12G23890.1 |
| GmCNGCIId | GLYMA02G36560.1 |
| GmCNGCIIe | GLYMA17G08120.3 |
| GmCNGCIVAa | GLYMA16G34420.1 |
| GmCNGCIVAb | GLYMA09G29880.1 |
| GmCNGCIVAc | GLYMA09G29860.1 |
| GmCNGCIVAd | GLYMA16G34370.3 |
| GmCNGCIVAe | GLYMA09G29850.2 |
| GmCNGCIVAf | GLYMA09G29850.5 |
| GmCNGCIVAg | GLYMA16G34381.1 |
| GmCNGCIVAh | GLYMA16G34390.1 |
| GmCNGCIVAi | GLYMA09G29870.2 |
| IpCNGC | onekp_FPCO_scaffold_2029041_Interfilum_paradoxum |
| IpCNGC1 | onekp_FPCO_scaffold_2029663_Interfilum_paradoxum |
| IpCNGC2 | onekp_FPCO_scaffold_2029041_Interfilum_paradoxum |
| IpCNGC3 | onekp:FPCO_scaffold_2029304_Interfilum_paradoxum |
| KnCNGC1 | kfl00018_0440_v1.1 |
| KnCNGC2 | kfl00007_0170_v1.1 |
| KnCNGC3 | kfl00007_0180_v1.1 |

| | |
|---|---|
| KnCNGC4 | kfl00007_0150_v1.1 |
| KnCNGC5 | kfl00280_0030_v1.1 |
| KnCNGC6 | kfl00044_0210_v1.1 |
| KsCNGC1 | onekp_FQLP_scaffold_2009700_Klebsormidium_subtile |
| KsCNGC2 | onekp_FQLP_scaffold_2006237_Klebsormidium_subtile |
| MbCNGC1 | onekp:WSJO_scaffold_2007784_Mesotaenium_braunii |
| MbCNGC2 | onekp:WSJO_scaffold_2047771_Mesotaenium_braunii |
| McCNGC1 | onekp:HKZW_scaffold_2010365_Mesotaenium_caldariorum |
| MeCNGC1 | onekp:WDCW_scaffold_2048497_Mesotaenium_endlicherianum |
| MkCNGC1 | onekp:NBYP_scaffold_2010046_Mesotaenium_kramstei |
| MkCNGC2 | onekp:NBYP_scaffold_2056215_Mesotaenium_kramstei |
| MpaCNGC1 | Unpublished assembly |
| MpaCNGC2 | Unpublished assembly |
| MpaCNGC3 | Unpublished assembly |
| MpaCNGC4 | Unpublished assembly |
| MpaCNGC5 | Unpublished assembly |
| MpoCNGC1 | Mapoly0044s0062.1 |
| MpoCNGC2 | Mapoly0004s0205.1 |
| MpoCNGC3 | Mapoly0052s0012.1 |
| MpoCNGC4 | Mapoly0127s0006.1 |
| MpoCNGC5 | Mapoly0011s0149.1 |
| MsCNGC1 | onekp:ZRMT_scaffold_2007883_Mougeotia_sp. |
| MsCNGC2 | onekp:ZRMT_scaffold_2032242_Mougeotia_sp. |
| MtCNGC14 | Medtr3g109030.1 |
| MtCNGC15a | Medtr3g088875.1 |
| MtCNGC15b | Medtr2g460590.1 |
| MtCNGC15c | Medtr7g034605.1 |
| MtCNGC16 | Medtr1g064240.1 |
| MtCNGC17 | Medtr4g058730.1 |
| MtCNGC18 | Medtr2g094860.1 |
| MtCNGC1a | Medtr1g028370.1 |
| MtCNGC1b | Medtr3g091080.1 |
| MtCNGC2 | Medtr7g012260.1 |
| MtCNGC4a | Medtr2g081110.1 |
| MtCNGC4b | Medtr8g464570.1 |
| MtCNGCIa | Medtr8g027755.1 |
| MtCNGCIb | Medtr7g117310.3 |
| MtCNGCII | Medtr4g130820.1 |
| MtCNGCIVAa | Medtr6g075290.3 |
| MtCNGCIVAb | Medtr5g007630.1 |
| MtCNGCIVAc | Medtr6g075440.1 |
| MtCNGCIVAd | Medtr6g075390.1 |
| MtCNGCIVAe | Medtr6g075410.1 |
| MtCNGCIVAf | Medtr6g477760.1 |

| | |
|---|---|
| NdCNGC1 | onekp:FFGR_scaffold_2004292_Netrium_digitus |
| OlCNGC1 | onekp:GGWH_scaffold_2011308_Onychonema_laeve |
| OlCNGC2 | onekp:GGWH_scaffold_2011309_Onychonema_laeve |
| OlCNGC3 | onekp:GGWH_scaffold_2011864_Onychonema_laeve |
| OlCNGC4 | onekp:GGWH_scaffold_2013994_Onychonema_laeve |
| OsCNGC15 | LOC_Os02g41710.1 |
| OsCNGC2 | LOC_Os03g55100.1 |
| OsCNGC4a | LOC_Os01g57370.1 |
| OsCNGC4b | LOC_Os05g42250.1 |
| OsCNGCIa | LOC_Os06g33570.1 |
| OsCNGCIb | LOC_Os06g33610.1 |
| OsCNGCIc | LOC_Os02g15580.1 |
| OsCNGCIIa | LOC_Os04g55080.1 |
| OsCNGCIIb | LOC_Os03g44440.1 |
| OsCNGCIIc | LOC_Os12g28260.1 |
| OsCNGCIIIa | LOC_Os12g06570.1 |
| OsCNGCIIIb | LOC_Os06g08850.1 |
| OsCNGCIIIc | LOC_Os02g54760.1 |
| OsCNGCIIId | LOC_Os09g38580.1 |
| OsCNGCIVAa | LOC_Os06g10580.1 |
| OsCNGCIVAb | LOC_Os02g53340.2 |
| PeCNGC1 | onekp:YSQT_scaffold_2005266_Penium_exiguum |
| PeCNGC2 | onekp:YSQT_scaffold_2007244_Penium_exiguum |
| PeCNGC3 | onekp:YSQT_scaffold_2009106_Penium_exiguum |
| PeCNGC4 | onekp:YSQT_scaffold_2036319_Penium_exiguum |
| PltCNGC1 | onekp:MOYY_scaffold_2005663_Pleurotaenium_trabecul |
| PnCNGC1 | onekp:RPQV_scaffold_2037163_Phymatodocis_nordstedtiana |
| PnCNGC2 | onekp:RPQV_scaffold_2037259_Phymatodocis_nordstedtiana |
| PnCNGC3 | onekp:RPQV_scaffold_2037266_Phymatodocis_nordstedtiana |
| PoCNGC1 | onekp:SNOX_scaffold_2002309_Planotaenium_ohtanii |
| PoCNGC2 | onekp:SNOX_scaffold_2024515_Planotaenium_ohtanii |
| PoCNGC3 | onekp:SNOX_scaffold_2024736_Planotaenium_ohtanii |
| PpCNGCa | Pp1s68_102V6.1 |
| PpCNGCb | Pp1s90.245F4.1 |
| PpCNGCc | Pp1s37_149V6.1 |
| PpCNGCd | Pp1s99_40E1.1 |
| PpCNGCe | Pp1s183_83E1.1 |
| PpCNGCf | Pp1s189_85V6.1 |
| PpCNGCg | Pp1s204_120E1.1 |
| PpCNGCh | Pp1s211_70E2.1 |
| PtCNGC1 | Potri.015G019100.1 |
| PtCNGC14a | Potri.009G010700.1 |
| PtCNGC14b | Potri.017G067000.1 |
| PtCNGC14c | Potri.018G097600.1 |

| | |
|---|---|
| PtCNGC15 | Potri.013G108200.1 |
| PtCNGC16 | Potri.018G009200.1 |
| PtCNGC17 | Potri.006G271300.1 |
| PtCNGC18 | Potri.018G009200.1 |
| PtCNGC2a | Potri.017G089900.1 |
| PtCNGC2b | Potri.017G089800.1 |
| PtCNGCIa | Potri.002G170000.1 |
| PtCNGCIb | Potri.014G097900.1 |
| PtCNGCIIa | Potri.001G043900.1 |
| PtCNGCIIb | Potri.003G183000.1 |
| PtCNGCIVAa | Potri.012G038700.1 |
| PtCNGCIVAb | Potri.015G034000.1 |
| PtCNGCIVAc | Potri.015G033900.1 |
| RoCNGC1 | onekp:XRTZ_scaffold_2003431_Roya_obtusa |
| RoCNGC2 | onekp:XRTZ_scaffold_2008265_Roya_obtusa |
| RoCNGC3 | onekp:XRTZ_scaffold_2027383_Roya_obtusa |
| ScCNGC1 | onekp:WCQU_scaffold_2012809_Staurodesmus_convergens |
| ScCNGC2 | onekp:WCQU_scaffold_2062616_Staurodesmus_convergens |
| SiCNGC15 | Si016531m |
| SiCNGC2 | Si034425m |
| SiCNGC4 | Si000574m |
| SiCNGCIa | Si008042m |
| SiCNGCIb | Si016638m |
| SiCNGCIc | Si020209m |
| SiCNGCIIa | Si011861m |
| SiCNGCIib | Si034469m |
| SiCNGCIIc | Si024468m |
| SiCNGCIIIa | Si028368m |
| SiCNGCIIIb | Si005993m |
| SiCNGCIIIc | Si016502m |
| SiCNGCIIId | Si029078m |
| SiCNGCIVA | Si016450m |
| SlCNGC14 | Solyc08g069140.2_1 |
| SlCNGC15a | Solyc03g114110.2_1 |
| SlCNGC15b | Solyc07g006510.2_1 |
| SlCNGC15c | Solyc02g086990.2_1 |
| SlCNGC16 | Solyc09g007840.2_1 |
| SlCNGC17 | Solyc06g010190.1_1 |
| SlCNGC18 | Solyc11g069580.1_1 |
| SlCNGC1a | Solyc06g051920.2_1 |
| SlCNGC1b | Solyc03g007260.2_1 |
| SlCNGC2 | Solyc02g088560.2_1 |
| SlCNGC4a | Solyc12g005400.1_1 |
| SlCNGC4b | Solyc10g006800.2_1 |

| | |
|---|---|
| SlCNGCIa | Solyc01g095770.2_1 |
| SlCNGCIb | Solyc05g050380.2_1 |
| SlCNGCIc | Solyc05g050360.2_1 |
| SlCNGCId | Solyc05g050350.1_1 |
| SlCNGCIIa | Solyc03g116850.2_1 |
| SlCNGCIIb | Solyc07g005590.2_1 |
| SlCNGCIIc | Solyc12g010010.1_1 |
| SlCNGCIVA | Solyc03g098210.2_1 |
| SmCNGCB1 | EFJ17870 |
| SmCNGCB2 | EFJ16402 |
| SmCNGCB3 | EFJ15042 |
| SmCNGCB4 | EFJ37650 |
| SmCNGCIVAa | EFJ30735 |
| SmCNGCIVAb | EFJ28778 |
| SmCNGCIVAc | EFJ34267 |
| SmCNGCIVAd | EFJ31333 |
| SmCNGCIVBa | EFJ22043 |
| SmCNGCIVBb | EFJ18327 |
| SmCNGCIVBc | EFJ09701 |
| SmCNGCIVBd | EFJ33072 |
| SoCNGC1 | onekp:RPRU_scaffold_2002627_Staurodesmus_omearii |
| SoCNGC2 | onekp:RPRU_scaffold_2012913_Staurodesmus_omearii |
| SoCNGC3 | onekp:RPRU_scaffold_2042802_Staurodesmus_omearii |
| SpfaCNGC1 | Sphfalx0079s0096.1 |
| SpfaCNGC2 | Sphfalx0169s0035.1 |
| SpfaCNGC3 | Sphfalx0087s0046.1 |
| SpfaCNGC4 | Sphfalx0006s0193.1 |
| SpfaCNGC5 | Sphfalx0030s0179.2 |
| SpfaCNGC6 | Sphfalx0066s0099.1 |
| SpfaCNGC7 | Sphfalx0087s0046.2 |
| SpfaCNGC8 | Sphfalx0059s0054.1 |
| SpfaCNGC9 | Sphfalx0139s0069.1 |
| SpspCNGC1 | onekp:TPHT_scaffold_2002714_Spirotaenia_sp. |
| SpspCNGC2 | onekp:TPHT_scaffold_2010622_Spirotaenia_sp. |
| SpspCNGC3 | onekp_TPHT_scaffold_2014339_Spirotaenia_sp. |
| SpspCNGC4 | onekp:TPHT_scaffold_2016273_Spirotaenia_sp. |
| SsCNGC1 | onekp:ISHC_scaffold_2005487_Staurastrum_sebaldi |
| SsCNGC2 | onekp:ISHC_scaffold_2009636_Staurastrum_sebaldi |
| SspCNGC1 | onekp:HAOX_scaffold_2001667_Spirogyra_sp. |
| VvCNGC1 | VIT_16s0050g02560.t01 |
| VvCNGC14 | VIT_04s0069g00790.t01 |
| VvCNGC15a | VIT_08s0040g01770.t01 |
| VvCNGC15b | VIT_06s0004g02670.t01 |
| VvCNGC16 | VIT_17s0000g08060.t01 |

| | |
|---|---|
| VvCNGC18 | VIT_14s0066g01560.t01 |
| VvCNGC2 | VIT_14s0108g01420.t01 |
| VvCNGC4 | VIT_19s0014g03700.t01 |
| VvCNGCIIa | VIT_09s0018g00410.t01 |
| VvCNGCIIb | VIT_11s0037g00230.t01 |
| XaCNGC1 | onekp:GBGT_scaffold_2011668_Xanthidium_antilopaeum |
| XaCNGC2 | onekp:GBGT_scaffold_2011670_Xanthidium_antilopaeum |
| XaCNGC3 | onekp:GBGT_scaffold_2024971_Xanthidium_antilopaeum |
| ZmCNGC15 | GRMZM2G068904_P01 |
| ZmCNGC2a | GRMZM5G858887_P01 |
| ZmCNGC2b | GRMZM2G074317_P01 |
| ZmCNGC4 | GRMZM2G078781_P02 |
| ZmCNGCIa | GRMZM2G148118_P02 |
| ZmCNGCIb | GRMZM2G090381_P01 |
| ZmCNGCIIa | GRMZM2G023037_P01 |
| ZmCNGCIIb | AC197150.3_FGP003 |
| ZmCNGCIIIa | GRMZM2G005791_P01 |
| ZmCNGCIIIb | GRMZM2G135651_P01 |
| ZsCNGC1 | onekp:MFZO_scaffold_2026840_Zygnemopsis_sp. |
| ZygCNGC1 | onekp:STKJ_scaffold_2019889_Zygnema_sp. |
| ZsCNGC2 | onekp:MFZO_scaffold_2027092_Zygnemopsis_sp. |
| ZygCNGC2 | onekp:WGMD_scaffold_3002778_Zygnema_sp. |

**Sequences used in phylogenetic analysis: CNGC**

| Name | Identity |
|---|---|
| AmtrExo70A | evm_27.model.AmTr_v1.0_scaffold00030.155 |
| AmtrExo70B | evm_27.model.AmTr_v1.0_scaffold00023.35 |
| AmtrExo70C | evm_27.model.AmTr_v1.0_scaffold00153.51 |
| AmtrExo70D | evm_27.model.AmTr_v1.0_scaffold00131.81 |
| AmtrExo70E | evm_27.model.AmTr_v1.0_scaffold00154.13 |
| AmtrExo70F | evm_27.model.AmTr_v1.0_scaffold00013.132 |
| AmtrExo70G | evm_27.model.AmTr_v1.0_scaffold00019.401 |
| AmtrExo70H | evm_27.model.AmTr_v1.0_scaffold00034.49 |
| AmtrExo70I | evm_27.model.AmTr_v1.0_scaffold00065.130 |
| AqcoEx70F1 | PIA33729.1 |
| AqcoEx70F2 | PIA51152.1 |
| AqcoEx70H1 | PIA57737.1 |
| AqcoEx70H2 | PIA38868.1 |
| AqcoExo70A | PIA54369.1 |
| AqcoExo70B | PIA59987.1 |
| AqcoExo70C | PIA41954.1 |
| AqcoExo70D | PIA27607.1 |
| AqcoExo70E | PIA57741.1 |
| AqcoExo70G | PIA40916.1 |
| AqcoExo70H | PIA52971.1 |
| AqcoExo70I | PIA50738.1 |
| ATE70A1 | ArathA1_Exo70_NP_195974.2_ATEXO70A1_(exocyst_subunit_EXO70_family_protein_A1)_protein_binding_Arabidopsis_thaliana |
| ATE70A2 | ArathA2_Exo70_NP_200047.3_ATEXO70A2_(exocyst_subunit_EXO70_family_protein_A2)_protein_binding_Arabidopsis_thaliana |
| ATE70A3 | ArathA3_Exo70_NP_200048.2_ATEXO70A3_(exocyst_subunit_EXO70_family_protein_A3)_protein_binding_Arabidopsis_thaliana |
| ATE70B1 | ArathB1_Exo70_NP_200651.1\|_ATEXO70B1_(exocyst_subunit_EXO70_family_protein_B1)_protein_binding_Arabidopsis_thaliana |
| ATE70B2 | ArathB2_Exo70_NP_172181.1\|_ATEXO70B2_(exocyst_subunit_EXO70_family_protein_B2)_protein_binding_Arabidopsis_thaliana |
| ATE70C1 | ArathC1_Exo70_NP_196819.1\|_ATEXO70C1_(exocyst_subunit_EXO70_family_protein_C1)_protein_binding_Arabidopsis_thaliana |
| ATE70C2 | ArathC2_Exo70_NP_196903.1\|_ATEXO70C2_(exocyst_subunit_EXO70_family_protein_C2)_protein_binding_Arabidopsis_thaliana |
| ATE70D1 | ArathD1_Exo84_NP_177391.1\|_ATEXO70D1_(exocyst_subunit_EXO70_family_protein_D1)_protein_binding_Arabidopsis_thaliana |
| ATE70D2 | ArathD2_Exo84_NP_175811.1\|_ATEXO70D2_(exocyst_subunit_EXO70_family_protein_D2)_protein_binding_Arabidopsis_thaliana |
| ATE70D3 | ArathD3_Exo84_NP_566477.2\|_ATEXO70D3_(exocyst_subunit_EXO70_family_protein_D3)_protein_binding_Arabidopsis_thaliana |
| ATE70E1 | ArathE1_Exo70_NP_189586.1\|_ATEXO70E1_(exocyst_subunit_EXO70_family_protein_E1)_protein_binding_Arabidopsis_thaliana |
| ATE70E2 | ArathE2_Exo70_NP_200909.1\|_ATEXO70E2_(EXOCYST_SUBUNIT_EXO70_FAMILY_PROTEIN_E2)_protein_binding_Arabidopsis_thaliana |
| ATE70F1 | ArathF1_Exo70_NP_199849.2\|_ATEXO70F1_(exocyst_subunit_EXO70_family_protein_F1)_protein_binding_Arabidopsis_thaliana |
| ATE70G1 | ArathG1_Exo70_NP_194882.2\|_ATEXO70G1_(exocyst_subunit_EXO70_family_protein_G1)_protein_binding_Arabidopsis_thaliana |

| | |
|---|---|
| ATE70G2 | ArathG2_Exo70_NP_175575.1\|_ATEXO70G2_(exocyst_subunit_EXO70_family_protein_G2)_protein_binding_Arabidopsis_thaliana |
| ATE70H1 | ArathH1_Exo70_NP_191075.2\|_ATEXO70H1_(exocyst_subunit_EXO70_family_protein_H1)_protein_binding_Arabidopsis_thaliana |
| ATE70H2 | ArathH2_Exo70_NP_181470.1\|_ATEXO70H2_(EXOCYST_SUBUNIT_EXO70_FAMILY_PROTEIN_H2)_protein_binding_Arabidopsis_thaliana |
| ATE70H3 | ArathH3_Exo70_NP_187564.1\|_ATEXO70H3_(exocyst_subunit_EXO70_family_protein_H3)_protein_binding_Arabidopsis_thaliana |
| ATE70H4 | ArathH4_Exo70_NP_187563.1\|_ATEXO70H4_(exocyst_subunit_EXO70_family_protein_H4)_protein_binding_Arabidopsis_thaliana |
| ATE70H5 | ArathH5_Exo70_NP_180432.2\|_ATEXO70H5_(exocyst_subunit_EXO70_family_protein_H5)_protein_binding_Arabidopsis_thaliana |
| ATE70H6 | ArathH6_Exo70_NP_683286.2\|_ATEXO70H6_(exocyst_subunit_EXO70_family_protein_H6)_protein_binding_Arabidopsis_thaliana |
| ATE70H7 | ArathH7_Exo70_NP_200781.1\|_ATEXO70H7_(EXOCYST_SUBUNIT_EXO70_FAMILY_PROTEIN_H7)_protein_binding_Arabidopsis_thaliana |
| ATE70H8 | ArathH8_Exo70_NP_180433.1\|_ATEXO70H8_(exocyst_subunit_EXO70_family_protein_H8)_protein_binding_Arabidopsis_thaliana |
| BaboExo70 | onekp:QWFV_scaffold_2008982_Bambusina_borreri |
| Bradi1 | Bradi1_Exo70_Bradi1g33940.1 |
| Bradi10 | Bradi10_Exo70_Bradi3g40510.1 |
| Bradi11 | Bradi11_Exo70_Bradi3g41160.1 |
| Bradi12 | Bradi12_Exo70_Bradi4g24960.1 |
| Bradi13 | Bradi13_Exo70_Bradi4g31270.1 |
| Bradi14 | Bradi14_Exo70_Bradi4g41750.1 |
| Bradi15 | Bradi15_Exo70_Bradi4g41980.1 |
| Bradi16 | Bradi16_exo70_Bradi5g01930.1 |
| Bradi17 | Bradi17_Exo70_Bd17FGENESH_Bradi5g06870.1 |
| Bradi18 | Bradi18_Exo70_Bradi5g25512.1 |
| Bradi19 | Bradi19_Exo70_Bradi5g25520.1 |
| Bradi2 | Bradi2_Exo70__Bradi2g22490.1 |
| Bradi20 | Bradi20_Exo70_Bradi5g25530.1 |
| Bradi21 | Bradi21_Exo70_Bradi5g26580.1 |
| Bradi22 | Bradi22_Exo70_FGENESH |
| Bradi23 | Bradi23_Exo70_Bradi4g24680.1 |
| Bradi24 | Bradi24_Exo70_Bradi5g07740.1 |
| Bradi25 | Bradi25_Exo70_15368.BRADI5G26590.2 |
| Bradi26 | Bradi26_Exo70_Bradi3g43960.1 |
| Bradi27 | Bradi27_Exo70_15368.BRADI2G59060.1 |
| Bradi3 | Bradi3_Exo70_Bradi2g50510.1_REVISED_by_fgenesh |
| Bradi4 | Bradi4_Exo70_Bradi2g50730.2 |
| Bradi5 | Bradi5_Exo70_Bradi2g53820.1 |
| Bradi6 | Bradi6_Exo70_Bradi2g59057.1 |
| Bradi7 | Bradi7_Exo70_Bradi3g03860.1 |
| Bradi8 | Bradi8_Exo70_Bradi3g32640.1 |
| Bradi9 | Bradi9_Exo70_Bradi3g37580.1 |
| ChatExo70 | onekp:AZZW_scaffold_2005384_Chlorokybus_atmophyticus |
| ChglExo701 | onekp:DRGY_scaffold_2003180_Chaetosphaeridium_globosum |

| | |
|---|---|
| ChglExo702 | onekp:DRGY_scaffold_2003710_Chaetosphaeridium_globosum |
| ChreExo70 | Cre10.g421250.t1.2 |
| CIluExo70 | onekp:DRFX_scaffold_2011657_Closterium_lunula |
| CocoExo701 | onekp:RQFE_scaffold_2004091_Cosmocladium_cf._constrictum |
| CocoExo702 | onekp:RQFE_scaffold_2037554_Cosmocladium_cf._constrictum |
| CoirExo70 | onekp:QPDY_scaffold_2030145_Coleochaete_irregularis |
| CoocExo701 | onekp:HJVM_scaffold_2002513_Cosmarium_ochthodes |
| CoocExo702 | onekp:HJVM_scaffold_2002514_Cosmarium_ochthodes |
| CoocExo703 | onekp:HJVM_scaffold_2002516_Cosmarium_ochthodes |
| CoocExo704 | onekp:HJVM_scaffold_2026990_Cosmarium_ochthodes |
| CosuExo70 | 48354_Coccomyxa_subellipsoidea_C-169__scaffold_14:194437..204287_forward |
| CycuExo701 | onekp:JOJQ_scaffold_2002773_Cylindrocystis_cushleckae |
| CycuExo702 | onekp:JOJQ_scaffold_2008616_Cylindrocystis_cushleckae |
| CycuExo703 | onekp:JOJQ_scaffold_2041354_Cylindrocystis_cushleckae |
| CycuExo704 | onekp:JOJQ_scaffold_2041897_Cylindrocystis_cushleckae |
| DeapEx070 | onekp:DFDS_scaffold_2039543_Desmidium_aptogonum |
| EnfiExo70 | onekp:BFIK_scaffold_2004040_Entransia_fimbriat |
| EuafExo701 | onekp:GYRP_scaffold_2003756_Euastrum_affine |
| EuafExo702 | onekp:GYRP_scaffold_2003757_Euastrum_affine |
| EuafExo703 | onekp:GYRP_scaffold_2003759_Euastrum_affine |
| EuafExo704 | onekp:GYRP_scaffold_2027489_Euastrum_affine |
| GokiExo701 | onekp:KEYW_scaffold_2005679_Gonatozygon_kinahanii |
| GokiExo702 | onekp:KEYW_scaffold_2026598_Gonatozygon_kinahanii |
| InpaExo70 | onekp:FPCO_scaffold_2028156_Interfilum_paradoxum |
| KlsuExo70 | onekp:FQLP_scaffold_2008246_Klebsormidium_subtile |
| LeobExo70 | onekp:ZNUM_scaffold_2026326_Leptosira_obovata |
| MapaEx70GI | Mapal_EXOG |
| MapaExo70A | Mapal_EXOA |
| MapaExo70X | Mapal_EXO70X |
| MapoEx70GI | Mapoly0113s0036.1 |
| MapoExo70A | Mapoly0134s0013.1 |
| MapoExo70X | Mapoly0153s0012.1 |
| MekrExo701 | onekp:NBYP_scaffold_2005586_Mesotaenium_kramstei |
| MekrExo702 | onekp:NBYP_scaffold_2005587_Mesotaenium_kramstei |
| MekrExo703 | onekp:NBYP_scaffold_2054749_Mesotaenium_kramstei |
| MekrExo704 | onekp:NBYP_scaffold_2055412_Mesotaenium_kramstei |
| MekrExo705 | onekp:NBYP_scaffold_2055686_Mesotaenium_kramstei |
| MetrE70A1 | Medtr1g090620_EXO70A1 |
| MetrE70A2 | Medtr1g057560_EXO70A2 |
| MetrE70B1 | Medtr5g073450_EXO70B1 |
| MetrE70B3 | Medtr5g073920_EXO70B3 |
| MetrE70B4 | Medtr2g096230 Exo70B4 |
| MetrE70B5 | Medtr5g093530_EXO70B5 |
| MetrE70B6 | Medtr3g464010 Exo70B6 |

| | |
|---|---|
| MetrE70B8 | Medtr3g031130_EXO70B7 |
| MetrE70C1 | Medtr7g026630_EXO70C1 |
| MetrE70C2 | Medtr6g033310_EXO70C2 |
| MetrE70D | Medtr8g021330 Exo70D |
| MetrE70E1 | Medtr3g025900_EXO70E1 |
| MetrE70E2 | Medtr7g013770_EXO70E2 |
| MetrE70F | Medtr4g103540_EXO70F |
| MetrE70G1 | Medtr2g041650_EXO70G1 |
| MetrE70G2 | Medtr4g008260_EXO70G2 |
| MetrE70G3 | Medtr1g009620_EXO7G3 |
| MetrE70H1 | Medtr4g062330_EXO70H1 |
| MetrE70H3 | Medtr1g067480_EXO70H3 |
| MetrE70H4 | Medtr1g067460_EXO70H4 |
| MetrE70I | Medtr1g017910_EXO70I |
| MifiExo701 | onekp:MCHJ_scaffold_2004757_Micrasterias_fimbriata |
| MifiExo702 | onekp:MCHJ_scaffold_2007485_Micrasterias_fimbriata |
| MipuExo70 | 182960 Micromonas pusilla CCMP1545 scaffold_1:94866..97787 reverse |
| MispExo70 | 98475 Micromonas sp. RCC299 Chr_01:95453..98515 reverse |
| MospExo701 | onekp:ZRMT_scaffold_2008964_Mougeotia_sp. |
| MospExo702 | onekp:ZRMT_scaffold_2009254_Mougeotia_sp. |
| MospExo703 | onekp:ZRMT_scaffold_2009524_Mougeotia_sp. |
| MospExo704 | onekp:ZRMT_scaffold_2010142_Mougeotia_sp. |
| MospExo705 | onekp:ZRMT_scaffold_2032192_Mougeotia_sp. |
| NediExo701 | onekp:FFGR_scaffold_2000085_Netrium_digitus |
| NediExo702 | onekp:FFGR_scaffold_2001998_Netrium_digitus |
| NediExo703 | onekp:FFGR_scaffold_2014204_Netrium_digitus |
| NueiExo70 | onekp:KMNX_scaffold_2037855_Nucleotaenium_eifelense |
| OnlaExo701 | onekp:GGWH_scaffold_2004702_Onychonema_laeve |
| OnlaExo702 | onekp:GGWH_scaffold_2010036_Onychonema_laeve |
| OrysaA1 | OrysaA1_Exo70_LOC_Os04g58880.1 |
| OrysaA2 | OrysaA2_Exo70_LOC_Os04g58870.2_revised |
| OrysaA3 | OrysaA3_Exo70_LOC_Os11g05880.1 |
| OrysaA4 | OrysaA4_Exo70_REVISED_LOC_Os12g06270.1 |
| OrysaB1 | OrysaB1_Exo70_OsB1_LOC_Os01g61180.1 |
| OrysaB2 | OrysaB2_Exo70_OsB2_LOC_Os01g61190.1 |
| OrysaB3 | OrysaB3_Exo70_OsB3_LOC_Os05g39610.1 |
| OrysaC1 | OrysaC1_Exo70_OsC1_LOC_Os12g06840.1 |
| OrysaC2 | OrysaC2_Exo70_LOC_Os11g06700.1 |
| OrysaD1 | OrysaD1_Exo70__OsD1_LOC_Os09g26820.1 |
| OrysaD2 | OrysaD2_Exo70_OsD2_ex_Fx13_LOC_Os08g35470.1 |
| OrysaE1 | OrysaE1_Exo70_tr\|C7JA30\|C7JA30_ORYSJ_Os12g0165500_protein__OS_Oryza_sativa_subsp._japonica_GN_Os12g0165500_fragment |
| OrysaF1 | OrysaF1_Exo70_OsF1_LOC_Os01g69230.1 |
| OrysaF2 | OrysaF2_Exo70_OsF2_LOC_Os02g30230.2 |
| OrysaF3 | OrysaF3_Exo70_OsF3_LOC_Os04g31330.1 |

| | |
|---|---|
| OrysaF4 | OrysaF4_Exo70_OsF4_LOC_Os08g41820.1 |
| OrysaF5 | OrysaF5_Exo70_OsF5_LOC_Os06g14450.1 |
| OrysaFX1 | OrysaFX1_Exo70_gi\|54287483\|gb\|AAV31227.1\|_unknown_protein_Oryza_sativa_Japonica_Group |
| OrysaFX10 | OrysaFX10_Exo70_LOC_Os07g10910.1 |
| OrysaFX11 | OrysaFX11_Exo70_LOC_Os07g10940.1 |
| OrysaFX12 | OrysaFX12_Exo70_LOC_Os07g10960.1 |
| OrysaFX13 | OrysaFX13_Exo70_LOC_Os04g02070.1 |
| OrysaFX14 | OrysaFX14_Exo70_LOC_Os01g67820.1 |
| OrysaFX15 | OrysaFX15_Exo70_LOC_Os01g67810.1 |
| OrysaFX16 | OrysaFX16_Exo70_6gi\|297728641\|ref\|NP_001176684.1\|_Os11g0650100_Oryza_sativa_Japonica_Group |
| OrysaFX2 | OrysaFX2_Exo70_LOC_Os05g30640.1 |
| OrysaFX3 | OrysaFX3_Exo70_LOC_Os05g30660.1 |
| OrysaFX4 | OrysaFX4_Exo70_LOC_Os08g13570.1 |
| OrysaFX5 | OrysaFX5_Exo70_LOC_Os09g17810.1 |
| OrysaFX6 | OrysaFX6_Exo70_LOC_Os02g36619.1 |
| OrysaFX7 | OrysaFX7_Exo70_LOC_Os01g28600.1 |
| OrysaFX8 | OrysaFX8_Exo70_LOC_Os11g36400.1 |
| OrysaFX9 | OrysaFX9_Exo70_LOC_Os06g08460.1 |
| OrysaG1 | OrysaG1_Exo70_OsG1_LOC_Os02g05620.1 |
| OrysaG2 | OrysaG2_Exo70_OsG2_LOC_Os06g48330.1 |
| OrysaH1 | OrysaH1_Exo70_LOC_Os11g42989.1 |
| OrysaH2 | OrysaH2_Exo70_OsH2_LOC_Os03g33520.1 |
| OrysaH3a | OrysaH3a_Exo70_OsH3a_LOC_Os12g01040.1 |
| OrysaH3b | OrysaH3b_Exo70_OsH3b_LOC_Os11g01050.1 |
| OrysaI | OrysaI1_Exo70_OsI1_LOC_Os08g40840.1 |
| OrysaX1 | OrysaX1_Exo70_OsLOC_Os01g55799.2 |
| OrysaX2 | OrysaX2_Exo70_OsLOC_Os10g33850.1 |
| OrysaX3 | OrysaX3_Exo70_OsLOC_Os07g10970.1 |
| OrysaX4 | OrysaX4_Exo70_OsLOC_Os07g10920.4 |
| OrysaX5 | OrysaX5_Exo70_OsLOC_Os01g56210.1 |
| OrysaX6 | OrysaX6_Exo70_OsLOC_Os05g30620.1 |
| OrysaX7 | OrysaX7_Exo70_OsLOC_Os05g30680.1 |
| PeenExo701 | onekp:YSQT_scaffold_2004439_Penium_exiguum |
| PeenExo702 | onekp:YSQT_scaffold_2035369_Penium_exiguum |
| PhnoExo701 | onekp:RPQV_scaffold_2005972_Phymatodocis_nordstedtiana |
| PhnoExo702 | onekp:RPQV_scaffold_2033226_Phymatodocis_nordstedtiana |
| PhnoExo703 | onekp:RPQV_scaffold_2036760_Phymatodocis_nordstedtiana |
| Phypa1 | Phypa1_Exo70_Phypa1a_Pp1s36_119V6.1 |
| Phypa10 | Phypa10_Exo70_Phypa3a_Pp1s38_352V6.1 |
| Phypa11 | Phypa11_Exo70_Phypa3b_Pp1s6_436V6.1 |
| Phypa12 | Phypa12_Exo70_Phypa3c_Pp1s80_89V6.1 |
| Phypa13 | Phypa13_Exo70_Phypa3d_Pp1s97_91V6.2 |
| Phypa2 | Phypa2_Exo70_Phypa1b_Pp1s39_74V6.1 |

| | |
|---|---|
| Phypa3 | Phypa3_Exo70_Phypa1c_Pp1s59_232V6.1 |
| Phypa4 | Phypa4_Exo70_Phypa2a_Pp1s70_34V6.1 |
| Phypa5 | Phypa5_Exo70_Phypa2b_Pp1s376_59V6.1 |
| Phypa6 | Phypa6_Exo70_Phypa2c_Pp1s227_61V6.1 |
| Phypa7 | Phypa7_Exo70_Phypa2d_Pp1s142_86V6.1 |
| Phypa8 | Phypa8_Exo70_Phypa2e_Pp1s169_51V6.1 |
| Phypa9 | Phypa9_Exo70_Phypa2f_Pp1s177_127V6.1 |
| PiabEx7010 | lcl\|MA_89549g0010_high_confidence |
| PiabExo701 | lcl\|MA_10430208g0010_high_confidence |
| PiabExo702 | lcl\|MA_106977g0010_high_confidence |
| PiabExo703 | lcl\|MA_121699g0010_high_confidence |
| PiabExo704 | lcl\|MA_130238g0010_high_confidence |
| PiabExo705 | lcl\|MA_138348g0010_high_confidence |
| PiabExo706 | lcl\|MA_281161g0010_high_confidence |
| PiabExo707 | lcl\|MA_418978g0010_high_confidence |
| PiabExo708 | lcl\|MA_482507g0010_high_confidence |
| PiabExo709 | lcl\|MA_605718g0020_high_confidence |
| PitaExo701 | lcl\|PITA_000013408_unnamed_protein_product |
| PitaExo702 | lcl\|PITA_000019001_unnamed_protein_product |
| PitaExo703 | lcl\|PITA_000022501_unnamed_protein_product |
| PitaExo704 | lcl\|PITA_000023019_unnamed_protein_product |
| PitaExo705 | lcl\|PITA_000023020_unnamed_protein_product |
| PitaExo706 | lcl\|PITA_000024493_unnamed_protein_product |
| PitaExo707 | lcl\|PITA_000026669_unnamed_protein_product |
| PitaExo708 | lcl\|PITA_000036336_unnamed_protein_product |
| PitaExo709 | lcl\|PITA_000047030_unnamed_protein_product |
| PitaExo710 | lcl\|PITA_000060555_unnamed_protein_product |
| PlohExo701 | onekp:SNOX_scaffold_2003556_Planotaenium_ohtanii |
| PlohExo702 | onekp:SNOX_scaffold_2004567_Planotaenium_ohtanii |
| PltrExo70 | onekp:MOYY_scaffold_2021916_Pleurotaenium_trabecul |
| Potr10 | Potri10_Exo70_PtExo70D1.1_POPTR27_gi\|224070782\|ref\|XP_002303234.1\|_predicted_protein_Populus_trichocarpa |
| Potr11 | Potri11_Exo70_PtExo70D1.2POPTR15_gi\|224119604\|ref\|XP_002331201.1\|_predicted_protein_Populus_trichocarpa |
| Potr12 | Potri12_Exo70_Ptrichocarpa\|POPTR_0017s12950\|POPTR_0017s12950.2 |
| Potr13 | Potri13_Exo70_PtExo70E1.2_POPTR7_gi\|224141507\|ref\|XP_002324113.1\|_predicted_protein_Populus_trichocarpa |
| Potr14 | Potri14_Exo70_PtExo70E1.4_POPTR1_gi\|224141509\|ref\|XP_002324114.1\|_predicted_protein_Populus_trichocarpa |
| Potr15 | Potri15_Exo70_PtExo70E2_POPTR10_gi\|224130150\|ref\|XP_002328666.1\|_predicted_protein_Populus_trichocarpa |
| Potr16 | Potri16_Exo70_PtExo70E1.5POPTR6_gi\|224141511\|ref\|XP_002324115.1\|_predicted_protein_Populus_trichocarpa |
| Potr17 | Potri17_Exo70_PtExo70F1.1_POPTR13_gi\|224121990\|ref\|XP_002318723.1\|_predicted_protein_Populus_trichocarpa |
| Potr18 | Potri18_Exo70_PtExo70F1.2_POPTR9_gi\|224136133\|ref\|XP_002322248.1\|_predicted_protein_Populus_trichocarpa |

| | |
|---|---|
| Potr19 | Potri19_Exo70_PtExo70G1.1_POPTR14_gi\|224121152\|ref\|XP_002330756.1\|_predicted_protein_Populus_trichocarpa |
| Potri1 | Potri1_Exo70_POPTR8_gi\|224138342\|ref\|XP_002322790.1\|_predicted_protein_Populus_trichocarpa |
| Potri2 | Potri2_Exo70_POPTR17_gi\|224110126\|ref\|XP_002315422.1\|_predicted_protein_Populus_trichocarpa |
| Potri20 | Potri20_Exo70_PtExo70G1.2_POPTR5_gi\|224142073\|ref\|XP_002324383.1\|_predicted_protein_Populus_trichocarpa |
| Potri21 | Potri21_Exo70_PtExo70G2.1_POPTR24_gi\|224097464\|ref\|XP_002310945.1\|_predicted_protein_Populus_trichocarpa |
| Potri22 | Potri22_Exo70_PtExo70G2.2_POPTR18_gi\|224110008\|ref\|XP_002315384.1\|_predicted_protein_Populus_trichocarpa |
| Potri23 | Potri23_Exo70_PtExo70H1/H2_POPTR30_gi\|224054869\|ref\|XP_002298379.1\|_predicted_protein_Populus_trichocarpa |
| Potri24 | Potri24_Exo70_PtExo70H1/H2_POPTR20_gi\|224104493\|ref\|XP_002313454.1\|_predicted_protein_Populus_trichocarpa |
| Potri25 | Potri25_Exo70_PtExo70H1/H2_POPTR22_gi\|224100905\|ref\|XP_002312061.1\|_predicted_protein_Populus_trichocarpa |
| Potri26 | Potri26_Exo70_PtExo70H1/H2_POPTR19_gi\|224109610\|ref\|XP_002315254.1\|_predicted_protein_Populus_trichocarpa |
| Potri27 | Potri27_Exo70_POPTR26_gi\|224088585\|ref\|XP_002308485.1\|_predicted_protein_Populus_trichocarpa |
| Potri28 | Potri28_Exo70_POPTR2_gi\|224141505\|ref\|XP_002324112.1\|_predicted_protein_Populus_trichocarpa |
| Potri29 | Potri29_Exo70_POPTR11_gi\|224127256\|ref\|XP_002320026.1\|_predicted_protein_Populus_trichocarpa |
| Potri3 | Potri3_Exo70_POPTR23_gi\|224100587\|ref\|XP_002311935.1\|_predicted_protein_Populus_trichocarpa |
| Potri4 | Potri4_Exo70_POPTR25_gi\|224091423\|ref\|XP_002309247.1\|_predicted_protein_Populus_trichocarpa |
| Potri5 | Potri5_Exo70_POPTR4_gi\|224145517\|ref\|XP_002325672.1\|_predicted_protein_Populus_trichocarpa |
| Potri6 | Potri6_Exo70_POPTR12_gi\|224124704\|ref\|XP_002319401.1\|_predicted_protein_Populus_trichocarpa |
| Potri7 | Potri7_Exo70_PtExo70C1_POPTR29_gi\|224057804\|ref\|XP_002299332.1\|_predicted_protein_Populus_trichocarpa |
| Potri8 | Potri8_Exo70_Ptrichocarpa\|POPTR_0001s33320\|POPTR_0001s33320.1 |
| Potri9 | Potri9_Exo70_PtExo70C2.2_POPTR16_gi\|224115404\|ref\|XP_002332164.1\|_predicted_protein_Populus_trichocarpa |
| RoobExo701 | onekp:XRTZ_scaffold_2002197_Roya_obtusa |
| RoobExo702 | onekp:XRTZ_scaffold_2002198_Roya_obtusa |
| RoobExo703 | onekp:XRTZ_scaffold_2007583_Roya_obtusa |
| Selmo1 | Selmo1_Exo70_Sel_gi\|302806856\|ref\|XP_002985159.1\|_hypothetical_protein_SELMODRAFT_121876_Selaginella_moellendorffii |
| Selmo2 | Selmo2_Exo70_Sel_gi\|302754910\|ref\|XP_002960879.1\|_hypothetical_protein_SELMODRAFT_74444_Selaginella_moellendorffii |
| Selmo3 | Selmo3_Exo70_Sel_gi\|302789430\|ref\|XP_002976483.1\|_hypothetical_protein_SELMODRAFT_232778_Selaginella_moellendorffii |
| Selmo4 | Selmo4_Exo70_sel_gi\|302815211\|ref\|XP_002989287.1\|_hypothetical_protein_SELMODRAFT_184479_Selaginella_moellendorffii |
| Selmo5 | Selmo5_Exo70_Sel_gi\|302798414\|ref\|XP_002980967.1\|_hypothetical_protein_SELMODRAFT_154195_Selaginella_moellendorffii |
| Selmo6 | Selmo6_Exo70_sel_gi\|302783302\|ref\|XP_002973424.1\|_hypothetical_protein_SELMODRAFT_98911_Selaginella_moellendorffii |

| | |
|---|---|
| Selmo7 | Selmo7_Exo70_Sel_gi\|302772811\|ref\|XP_002969823.1\|_hypothetical_protein_SELMODRAFT_171149_Selaginella_moellendorffii |
| Selmo8 | Selmo8_Exo70_Sel_gi\|302767426\|ref\|XP_002967133.1\|_hypothetical_protein_SELMODRAFT_87112_Selaginella_moellendorffii |
| Solyc1 | Solyc1_Exo70_SlExo70.1_Solyc01g009880.2.1 |
| Solyc10 | Solyc10_Exo70_SlExo70.10_Solyc05g054820.1.1 |
| Solyc11 | Solyc11_Exo70_SlExo70.11_Solyc06g005280.1.1 |
| Solyc12 | Solyc12_Exo70_SlExo70.12_Solyc06g062990.1.1 |
| Solyc13 | Solyc13_Exo70_SlExo70.13_Solyc06g075610.1.1 |
| Solyc14 | Solyc14_Exo70_SlExo70.14_Solyc08g066070.1.1 |
| Solyc15 | Solyc15_Exo70_SlExo70.15_Solyc08g069100.1.1 |
| Solyc16 | Solyc16_Exo70_SlExo70.16_Solyc09g005830.1.1 |
| Solyc17 | Solyc17_Exo70_SlExo70.17_Solyc09g005840.1.1 |
| Solyc18 | Solyc18_Exo70_SlExo70.18_Solyc10g081940.1.1 |
| Solyc19 | Solyc19_Exo70_SlExo70.19_Solyc11g006620.1.1 |
| Solyc2 | Solyc2_Exo70_SlExo70.2_Solyc02g088800.1.1 |
| Solyc20 | Solyc20_Exo70_SlExo70.20_Solyc11g073010.1.1 |
| Solyc21 | Solyc21_Exo70_SlExo70.21_Solyc12g055770.1.1 |
| Solyc22 | Solyc22_Exo70_SlExo70.22_Solyc12g096370.1.1 |
| Solyc3 | Solyc3_Exo70_SlExo70.3_Solyc02g094370.1.1 |
| Solyc4 | Solyc4_Exo70_SlExo70.4_Solyc03g111320.1.1 |
| Solyc5 | Solyc5_Exo70_SlExo70.5_Solyc03g115070.1.1 |
| Solyc6 | Solyc6_Exo70_SlExo70.6_Solyc03g122240.2.1 |
| Solyc7 | Solyc7_Exo70_SlExo70.7_Solyc04g009740.2.1 |
| Solyc8 | Solyc8_Exo70_SlExo70.8_Solyc04g077760.1.1 |
| Solyc9 | Solyc9_Exo70_SlExo70.9_Solyc05g024340.1.1 |
| Sorbi1 | Sorbi1_Exo70__Sb03g013570.1 |
| Sorbi10 | Sorbi10_Exo70_Sb07g022280.1 |
| Sorbi11 | Sorbi11_Exo70_Sb07g026300.1 |
| Sorbi12 | Sorbi12_Exo70_Sb07g027000.1 |
| Sorbi13 | Sorbi13_Exo70_Sb08g000230.1 |
| Sorbi14 | Sorbi14_Exo70_Sb08g003730.1_REVISED |
| Sorbi15 | Sorbi15_Exo70_Sb08g004090.1 |
| Sorbi16 | Sorbi16_Exo70_Sb09g023210.1 |
| Sorbi17 | Sorbi17_Exo70_Sb10g003910.1 |
| Sorbi18 | Sorbi18_Exo70_Sb10g005610.1 |
| Sorbi19 | Sorbi19_Exo70_Sb10g028680.1 |
| Sorbi2 | Sorbi2_Exo70_jgi |
| Sorbi20 | Sorbi20_Exo70_Sb02g025750.1 |
| Sorbi21 | Sorbi21_Exo70_Sb05g004020.1 |
| Sorbi22 | Sorbi22_Exo70_Sb03g043090.1_REVISED |
| Sorbi23 | Sorbi23_Exo70_Sb10g030700.1 |
| Sorbi24 | Sorbi24_Exo70_Sb10g030630.1 |
| Sorbi25 | Sorbi25_Exo70_Sb10g030710.1_REVISED |
| Sorbi26 | Sorbi26_Exo70_Sb10g030640.1 |

| | |
|---|---|
| Sorbi27 | Sorbi27_Exo70__Sb10g030620.1 |
| Sorbi28 | Sorbi28_Exo70_Sb10g030650.1 |
| Sorbi29 | Sorbi29_Exo70_Sb04g003570.1_REVISED |
| Sorbi3 | Sorbi3_Exo70_Sb3_Sb03g038570.1 |
| Sorbi30 | Sorbi30_Exo70_Sb03g038580.1 |
| Sorbi31 | Sorbi31_exo70_Sb10g030480.1 |
| Sorbi4 | Sorbi4_Exo70__Sb03g043100.1 |
| Sorbi5 | Sorbi5_Exo70_Sb03g044110.1 |
| Sorbi6 | Sorbi6_Exo70_Sb04g020520.1 |
| Sorbi7 | Sorbi7_Exo70__Sb05g003650.1 |
| Sorbi8 | Sorbi8_Exo70_Sb06g011830.1 |
| Sorbi9 | Sorbi9_exo70_Sb06g033600.1 |
| Sphfal1 | Sphfalx0009s0094.1 |
| Sphfal10 | Sphfalx0113s0008.1 |
| Sphfal11 | Sphfalx0134s0016.1 |
| Sphfal12 | Sphfalx0182s0003.1 |
| Sphfal13 | Sphfalx0206s0036.1 |
| Sphfal2 | Sphfalx0010s0089.1 |
| Sphfal3 | Sphfalx0011s0070.1 |
| Sphfal4 | Sphfalx0014s0208.1 |
| Sphfal5 | Sphfalx0015s0170.1 |
| Sphfal6 | Sphfalx0020s0192.2 |
| Sphfal7 | Sphfalx0022s0123.1 |
| Sphfal8 | Sphfalx0032s0143.1 |
| Sphfal9 | Sphfalx0052s0109.1 |
| SpmiExo70 | onekp:NNHQ_scaffold_2033888_Spirotaenia_minuta |
| SpspExo701 | onekp:HAOX_scaffold_2000161_Spirogyra_sp. |
| SpspExo702 | onekp:HAOX_scaffold_2023833_Spirogyra_sp. |
| StcoExo701 | onekp:WCQU_scaffold_2001849_Staurodesmus_convergens |
| StcoExo702 | onekp:WCQU_scaffold_2001851_Staurodesmus_convergens |
| StcoExo703 | onekp:WCQU_scaffold_2062930_Staurodesmus_convergens |
| StseExo701 | onekp:ISHC_scaffold_2003568_Staurastrum_sebaldi |
| StseExo702 | onekp:ISHC_scaffold_2008020_Staurastrum_sebaldi |
| TecoExo70 | onekp:DUMA_scaffold_2001587_Tetraselmis_cordiformis |
| TrarExo70 | onekp:NKXU_scaffold_2026304_Trebouxia_arboricola |
| Vivin1 | Vivin1_Exo70_VITIS6_gi\|225440706\|ref\|XP_002280545.1\|_PREDICTED:_hypothetical_protein_Vitis_vinifera |
| Vivin10 | Vivin10_Exo70_VITIS9_gi\|225443302\|ref\|XP_002274042.1\|_PREDICTED:_hypothetical_protein_Vitis_vinifera |
| Vivin11 | Vivin11_Exo70_VITIS2_gi\|225429756\|ref\|XP_002280486.1\|_PREDICTED:_hypothetical_protein_Vitis_vinifera |
| Vivin12 | Vivin12_Exo70_VITIS5_gi\|225439838\|ref\|XP_002274342.1\|_PREDICTED:_hypothetical_protein_Vitis_vinifera |
| Vivin13 | Vivin13_Exo70_VITIS10_gi\|225448817\|ref\|XP_002276070.1\|_PREDICTED:_hypothetical_protein_Vitis_vinifera |
| Vivin14 | Vivin14_Exo70_VITIS4_gi\|225435104\|ref\|XP_002281519.1\|_PREDICTED:_hypothetical_protein_Vitis_vinifera |

| | |
|---|---|
| Vivin15 | Vivin15_Exo70_VITIS14_gi\|225458854\|ref\|XP_002283343.1\|_PREDICTED:_hypothetical_protein_Vitis_vinifera |
| Vivin2 | Vivin2_Exo70_VITIS15_gi\|225464400\|ref\|XP_002268110.1\|_PREDICTED:_hypothetical_protein_Vitis_vinifera |
| Vivin3 | Vivin3_Exo70_VITIS3_gi\|225434439\|ref\|XP_002272867.1\|_PREDICTED:_hypothetical_protein_Vitis_vinifera |
| Vivin4 | Vivin4_Exo70_VITIS1_gi\|225425194\|ref\|XP_002264953.1\|_PREDICTED:_hypothetical_protein_Vitis_vinifera |
| Vivin5 | Vivin5_Exo70_VITIS7_gi\|225442065\|ref\|XP_002272396.1\|_PREDICTED:_hypothetical_protein_Vitis_vinifera |
| Vivin6 | Vivin6_Exo70_VITIS11_gi\|225451747\|ref\|XP_002280135.1\|_PREDICTED:_hypothetical_protein_Vitis_vinifera |
| Vivin7 | Vivin7_Exo70_VITIS13_gi\|225457209\|ref\|XP_002280826.1\|_PREDICTED:_hypothetical_protein_Vitis_vinifera |
| Vivin8 | Vivin8_Exo70_VITIS12_gi\|225456309\|ref\|XP_002279988.1\|_PREDICTED:_hypothetical_protein_Vitis_vinifera |
| Vivin9 | Vivin9_Exo70_VITIS8_gi\|225443300\|ref\|XP_002273932.1\|_PREDICTED:_hypothetical_protein_Vitis_vinifera |
| VocaExo70 | Vocar.0016s0034.1 |
| XaanExo701 | onekp:GBGT_scaffold_2008100_Xanthidium_antilopaeum |
| XaanExo702 | onekp:GBGT_scaffold_2008101_Xanthidium_antilopaeum |
| XaanExo703 | onekp:GBGT_scaffold_2010365_Xanthidium_antilopaeum |
| XaanExo704 | onekp:GBGT_scaffold_2025843_Xanthidium_antilopaeum |
| ZysoExo701 | onekp:WGMD_scaffold_3002357_Zygnema_sp. |
| ZysoExo702 | onekp:WGMD_scaffold_3002358_Zygnema_sp. |
| ZysoExo703 | onekp:WGMD_scaffold_3010468_Zygnema_sp. |
| ZysoExo704 | onekp:WGMD_scaffold_3014799_Zygnema_sp. |
| ZyspExo701 | onekp:MFZO_scaffold_2004191_Zygnemopsis_sp. |
| ZyspExo702 | onekp:MFZO_scaffold_2005064_Zygnemopsis_sp. |
| ZyspExo703 | onekp:MFZO_scaffold_2005065_Zygnemopsis_sp. |
| ZyspExo704 | onekp:MFZO_scaffold_2005066_Zygnemopsis_sp. |
| ZyspExo705 | onekp:MFZO_scaffold_2026859_Zygnemopsis_sp. |

**Sequences used in phylogenetic analysis: Exo70**

| Name | Identity |
|---|---|
| AcPHT1;1 | Aqcoe2G029400.1 |
| AcPHT1;10 | Aqcoe6G162300.1 |
| AcPHT1;11 | Aqcoe6G164500.1 |
| AcPHT1;2 | Aqcoe2G075000.1 |
| AcPHT1;3 | Aqcoe2G372300.1 |
| AcPHT1;4 | Aqcoe2G372300.2 |
| AcPHT1;5 | Aqcoe5G026500.1 |
| AcPHT1;6 | Aqcoe5G026500.2 |
| AcPHT1;7 | Aqcoe5G026600.1 |
| AcPHT1;8 | Aqcoe5G026700.1 |
| AcPHT1;9 | Aqcoe5G445500.1 |
| AmtrPHT1;1 | lcl\|evm_27.model.AmTr_v1.0_scaffold00003.306_unnamed_protein_product |
| AmtrPHT1;2 | lcl\|evm_27.model.AmTr_v1.0_scaffold00009.71_unnamed_protein_product |
| AmtrPHT1;3 | lcl\|evm_27.model.AmTr_v1.0_scaffold00009.73_unnamed_protein_product |
| AmtrPHT1;4 | lcl\|evm_27.model.AmTr_v1.0_scaffold00009.74_unnamed_protein_product |
| AmtrPHT1;5 | lcl\|evm_27.model.AmTr_v1.0_scaffold00009.76_unnamed_protein_product |
| AmtrPHT1;6 | lcl\|evm_27.model.AmTr_v1.0_scaffold00077.130_unnamed_protein_product |
| AmtrPHT1;7 | lcl\|evm_27.model.AmTr_v1.0_scaffold00097.45_unnamed_protein_product |
| AtcuPHT1;1 | onekp:XIRK_scaffold_2064071_Athrotaxis_cupressoides |
| AtcuPHT1;2 | onekp:XIRK_scaffold_2064545_Athrotaxis_cupressoides |
| AtPHT1;1 | AT5G43350.1species_A_thaliana |
| AtPHT1;10 | AT4G08878.1species_A_thaliana |
| AtPHT1;11 | AT4G08895.1species_A_thaliana |
| AtPHT1;2 | AT5G43370.1species_A_thaliana |
| AtPHT1;3 | AT5G43360.1species_A_thaliana |
| AtPHT1;4 | AT2G38940.1species_A_thaliana |
| AtPHT1;5 | AT2G32830.1species_A_thaliana |
| AtPHT1;6 | AT5G43340.1species_A_thaliana |
| AtPHT1;7 | AT3G54700.1species_A_thaliana |
| AtPHT1;8 | AT1G20860.1species_A_thaliana |
| AtPHT1;9 | AT1G76430.1species_A_thaliana |
| BaboPHT1;1 | onekp:QWFV_scaffold_2021765_Bambusina_borreri |
| BdPHT1;1 | Bradi1g00700.1species_B_distachyon |
| BdPHT1;10 | Bradi5g02730.1species_B_distachyon |
| BdPHT1;11 | Bradi5g02750.1species_B_distachyon |
| BdPHT1;12 | Bradi5g02760.1species_B_distachyon |
| BdPHT1;13 | Bradi5g02770.1species_B_distachyon |
| BdPHT1;14 | Bradi3g21015.1species_B_distachyon |
| BdPHT1;2 | Bradi1g42610.1species_B_distachyon |
| BdPHT1;3 | Bradi1g52590.1species_B_distachyon |
| BdPHT1;4 | Bradi1g75020.1species_B_distachyon |
| BdPHT1;5 | Bradi1g75030.1species_B_distachyon |
| BdPHT1;6 | Bradi1g76010.1species_B_distachyon |

| | |
|---|---|
| BdPHT1;7 | Bradi2g45520.1species_B_distachyon |
| BdPHT1;8 | Bradi3g12590.1species_B_distachyon |
| BdPHT1;9 | Bradi3g27680.1species_B_distachyon |
| CagrPHT1;1 | onekp:IFLI_scaffold_2002406_Callitris_gracilis |
| CagrPHT1;2 | onekp:IFLI_scaffold_2002407_Callitris_gracilis |
| CagrPHT1;3 | onekp:IFLI_scaffold_2145319_Callitris_gracilis |
| CamaPHT1;1 | onekp:RMMV_scaffold_2002116_Callitris_macleayana |
| CamaPHT1;2 | onekp:RMMV_scaffold_2005805_Callitris_macleayana |
| CamaPHT1;3 | onekp:RMMV_scaffold_2057120_Callitris_macleayana |
| CamaPHT1;4 | onekp:RMMV_scaffold_2057562_Callitris_macleayana |
| CehaPHT1;1 | onekp:GJTI_scaffold_2003088_Cephalotaxus_harringtonia |
| CehaPHT1;2 | onekp:GJTI_scaffold_2009436_Cephalotaxus_harringtonia |
| ChatPHT1;1 | onekp:AZZW_scaffold_2021223_Chlorokybus_atmophyticus |
| ClluPHT1;1 | onekp:DRFX_scaffold_2001384_Closterium_lunula |
| CocoPHT1;1 | onekp:RQFE_scaffold_2037607_Cosmocladium_cf._constrictum |
| CogrPHT1;1 | onekp:MNNM_scaffold_2003545_Cosmarium_granatum |
| CogrPHT1;2 | onekp:MNNM_scaffold_2043942_Cosmarium_granatum |
| CoocPHT1;1 | onekp:HJVM_scaffold_2006988_Cosmarium_ochthodes |
| CoocPHT1;2 | onekp:HJVM_scaffold_2006989_Cosmarium_ochthodes |
| CostPHT1;1 | onekp:WDGV_scaffold_2003053_Cosmarium_subtumidum |
| CostPHT1;2 | onekp:WDGV_scaffold_2003056_Cosmarium_subtumidum |
| CostPHT1;3 | onekp:WDGV_scaffold_2055959_Cosmarium_subtumidum |
| CotiPHT1;1 | onekp:BHBK_scaffold_2050531_Cosmarium_tinctum |
| CuguPHT1;1 | onekp:QNGJ_scaffold_2013258_Cupressus_dupreziana |
| CuguPHT1;2 | onekp:QNGJ_scaffold_2074368_Cupressus_dupreziana |
| CybrPHT1;1 | onekp:RPGL_scaffold_2008947_Cylindrocystis_brebissonii |
| CybrPHT1;2 | onekp:RPGL_scaffold_2010035_Cylindrocystis_brebissonii |
| CybrPHT1;3 | onekp:YOXI_scaffold_2008956_Cylindrocystis_brebissonii |
| CybrPHT1;4 | onekp:YOXI_scaffold_2056003_Cylindrocystis_brebissonii |
| CycuPHT1;1 | onekp:JOJQ_scaffold_2004279_Cylindrocystis_cushleckae |
| CycuPHT1;2 | onekp:JOJQ_scaffold_2038471_Cylindrocystis_cushleckae |
| CycuPHT1;3 | onekp:JOJQ_scaffold_2039067_Cylindrocystis_cushleckae |
| CyspPHT1;1 | onekp:VAZE_scaffold_2000655_Cylindrocystis_sp. |
| CyspPHT1;2 | onekp:VAZE_scaffold_2001763_Cylindrocystis_sp. |
| CyspPHT1;3 | onekp:VAZE_scaffold_2001765_Cylindrocystis_sp. |
| DacoPHT1;1 | onekp:FMWZ_scaffold_2005430_Dacrycarpus_compactus |
| DacoPHT1;2 | onekp:FMWZ_scaffold_2005431_Dacrycarpus_compactus |
| DacoPHT1;3 | onekp:FMWZ_scaffold_2011494_Dacrycarpus_compactus |
| EnfiPHT1;1 | onekp:BFIK_scaffold_2029084_Entransia_fimbriat |
| EuafPHT1;1 | onekp:GYRP_scaffold_2205464_Euastrum_affine |
| EuafPHT1;2 | onekp:GYRP_scaffold_2205546_Euastrum_affine |
| GokiPHT1;1 | onekp:KEYW_scaffold_2001487_Gonatozygon_kinahanii |
| KlsuPHT1;1 | onekp:FQLP_scaffold_2004782_Klebsormidium_subtile |
| KlsuPHT1;2 | onekp:FQLP_scaffold_2004783_Klebsormidium_subtile |

| | |
|---|---|
| LjPHT1;1 | LjSGA_099256.1species_L_japonicus |
| LjPHT1;10 | LjSGA_133701.1species_L_japonicus |
| LjPHT1;11 | chr6.CM1613.60.ncspecies_L_japonicus |
| LjPHT1;2 | LjSGA_025342.1species_L_japonicus |
| LjPHT1;3 | chr6.CM1613.70.ncspecies_L_japonicus |
| LjPHT1;4 | LjSGA_014433.2species_L_japonicus |
| LjPHT1;5 | LjSGA_018600.1species_L_japonicus |
| LjPHT1;6 | LjSGA_023515.1species_L_japonicus |
| LjPHT1;7 | LjSGA_031438.1species_L_japonicus |
| LjPHT1;8 | LjSGA_077849.1species_L_japonicus |
| LjPHT1;9 | LjSGA_096922.1species_L_japonicus |
| MacoPHT1;1 | onekp:CDFR_scaffold_2005665_Manoao_colensoi |
| MacoPHT1;2 | onekp:CDFR_scaffold_2009297_Manoao_colensoi |
| MacoPHT1;3 | onekp:CDFR_scaffold_2009299_Manoao_colensoi |
| MapaPHT1A | PTA_LFVP-2086236__contig13142Node |
| MapaPHT1B | PTB_LFVP-2015963_marpal_contig3131_1525Node3220 |
| MapaPHT1C | NODE2296PTC_marpal_contig_14345_LFVP7378 |
| MapaPHT1D | PTD_LFVP-2011355_contig_15483Node2759 |
| MapaPHT1E | Node901 |
| MapoPHT1;1 | Mapoly0080s0076.1 |
| MapoPHT1;2 | Mapoly0129s0007.1 |
| MapoPHT1;3 | Mapoly0195s0005.1 |
| MapoPHT1;4 | Mapoly0195s0006.1 |
| MapoPHT1;5 | Mapoly0195s0006.1 |
| MapoPHT1;6 | Mapoly0195s0007.1 |
| MapoPHT1;7 | Mapoly0195s0008.1 |
| MapoPHT1;8 | Mapoly0773s0001.1 |
| MapoPHT1;9 | Mapoly0773s0001.1 |
| MebrPHT1;1 | onekp:WSJO_scaffold_2000514_Mesotaenium_braunii |
| MebrPHT1;2 | onekp:WSJO_scaffold_2000515_Mesotaenium_braunii |
| MebrPHT1;3 | onekp:WSJO_scaffold_2004621_Mesotaenium_braunii |
| MecaPHT1;1 | onekp:HKZW_scaffold_2000804_Mesotaenium_caldariorum |
| MecaPHT1;2 | onekp:HKZW_scaffold_2013879_Mesotaenium_caldariorum |
| MeenPHT1;1 | onekp:WDCW_scaffold_2009956_Mesotaenium_endlicherianum |
| MeglPHT1;1 | onekp:NRXL_scaffold_2008330_Metasequoia_glyptostroboides |
| MeglPHT1;2 | onekp:NRXL_scaffold_2009108_Metasequoia_glyptostroboides |
| MeglPHT1;3 | onekp:NRXL_scaffold_2065292_Metasequoia_glyptostroboides |
| MifiPHT1;1 | onekp:MCHJ_scaffold_2008723_Micrasterias_fimbriata |
| MtPHT1;1 | Medtr1g043220.1species_M_truncatula |
| MtPHT1;10 | Medtr1g069930.1species_M_truncatula |
| MtPHT1;11 | Medtr1g069935.1species_M_truncatula |
| MtPHT1;12 | Medtr7g096870.1species_M_truncatula |
| MtPHT1;13 | Medtr7g096880.1species_M_truncatula |
| MtPHT1;2 | Medtr1g043290.1species_M_truncatula |

| | |
|---|---|
| MtPHT1;3 | Medtr1g043200.1species_M_truncatula |
| MtPHT1;4 | Medtr1g028600.1species_M_truncatula |
| MtPHT1;5 | Medtr1g074930.1species_M_truncatula |
| MtPHT1;6 | Medtr3g082700.1species_M_truncatula |
| MtPHT1;7 | Medtr1g074940.1species_M_truncatula |
| MtPHT1;8 | Medtr5g068140.1species_M_truncatula |
| MtPHT1;9 | Medtr4g083960.1species_M_truncatula |
| NediPHT1;1 | onekp:FFGR_scaffold_2007606_Netrium_digitus |
| NueiPHT1;1 | onekp:KMNX_scaffold_2005384_Nucleotaenium_eifelense |
| NueiPHT1;2 | onekp:KMNX_scaffold_2037961_Nucleotaenium_eifelense |
| NueiPHT1;3 | onekp:KMNX_scaffold_2038103_Nucleotaenium_eifelense |
| OnlaPHT1;1 | onekp:GGWH_scaffold_2008969_Onychonema_laeve |
| OnlaPHT1;2 | onekp:GGWH_scaffold_2047940_Onychonema_laeve |
| OsPHT1;1 | LOC_Os03g05620.1species_O_sativa |
| OsPHT1;10 | LOC_Os06g21950.1species_O_sativa |
| OsPHT1;11 | LOC_Os01g46860.1species_O_sativa |
| OsPHT1;12 | LOC_Os03g05610.1species_O_sativa |
| OsPHT1;13 | LOC_Os04g10800.1species_O_sativa |
| OsPHT1;14 | LOC_Os06g21930.1species_O_sativa |
| OsPHT1;2 | LOC_Os03g05640.1species_O_sativa |
| OsPHT1;3 | LOC_Os10g30770.1species_O_sativa |
| OsPHT1;4 | LOC_Os04g10750.1species_O_sativa |
| OsPHT1;5 | LOC_Os04g10690.1species_O_sativa |
| OsPHT1;6 | LOC_Os08g45000.1species_O_sativa |
| OsPHT1;7 | LOC_Os03g04360.1species_O_sativa |
| OsPHT1;8 | LOC_Os10g30790.1species_O_sativa |
| OsPHT1;9 | LOC_Os06g21920.1species_O_sativa |
| PeenPHT1;1 | onekp:YSQT_scaffold_2036504_Penium_exiguum |
| PemaPHT1;1 | onekp:AEKF_scaffold_2008198_Penium_margaritaceum |
| PemaPHT1;2 | onekp:AEKF_scaffold_2012309_Penium_margaritaceum |
| PhnoPHT1;1 | onekp:RPQV_scaffold_2036343_Phymatodocis_nordstedtiana |
| PhpaPHT1;1 | Pp3c10_24280V3.1 |
| PhpaPHT1;2 | Pp3c16_15980V3.2 |
| PhpaPHT1;3 | Pp3c2_31940V3.3 |
| PhpaPHT1;4 | Pp3c2_35170V3.4 |
| PhpaPHT1;5 | Pp3c6_26510V3.4 |
| PhpaPHT1;6 | Pp3c9_15090V3.1 |
| PiabPHT1;1 | lcl\|MA_10434434g0010_high_confidence |
| PiabPHT1;10 | lcl\|MA_8087333g0010_high_confidence |
| PiabPHT1;2 | lcl\|MA_10434434g0020_high_confidence |
| PiabPHT1;3 | lcl\|MA_10434456g0010_high_confidence |
| PiabPHT1;4 | lcl\|MA_15374g0010_high_confidence |
| PiabPHT1;5 | lcl\|MA_228607g0010_high_confidence |
| PiabPHT1;6 | lcl\|MA_366260g0010_high_confidence |

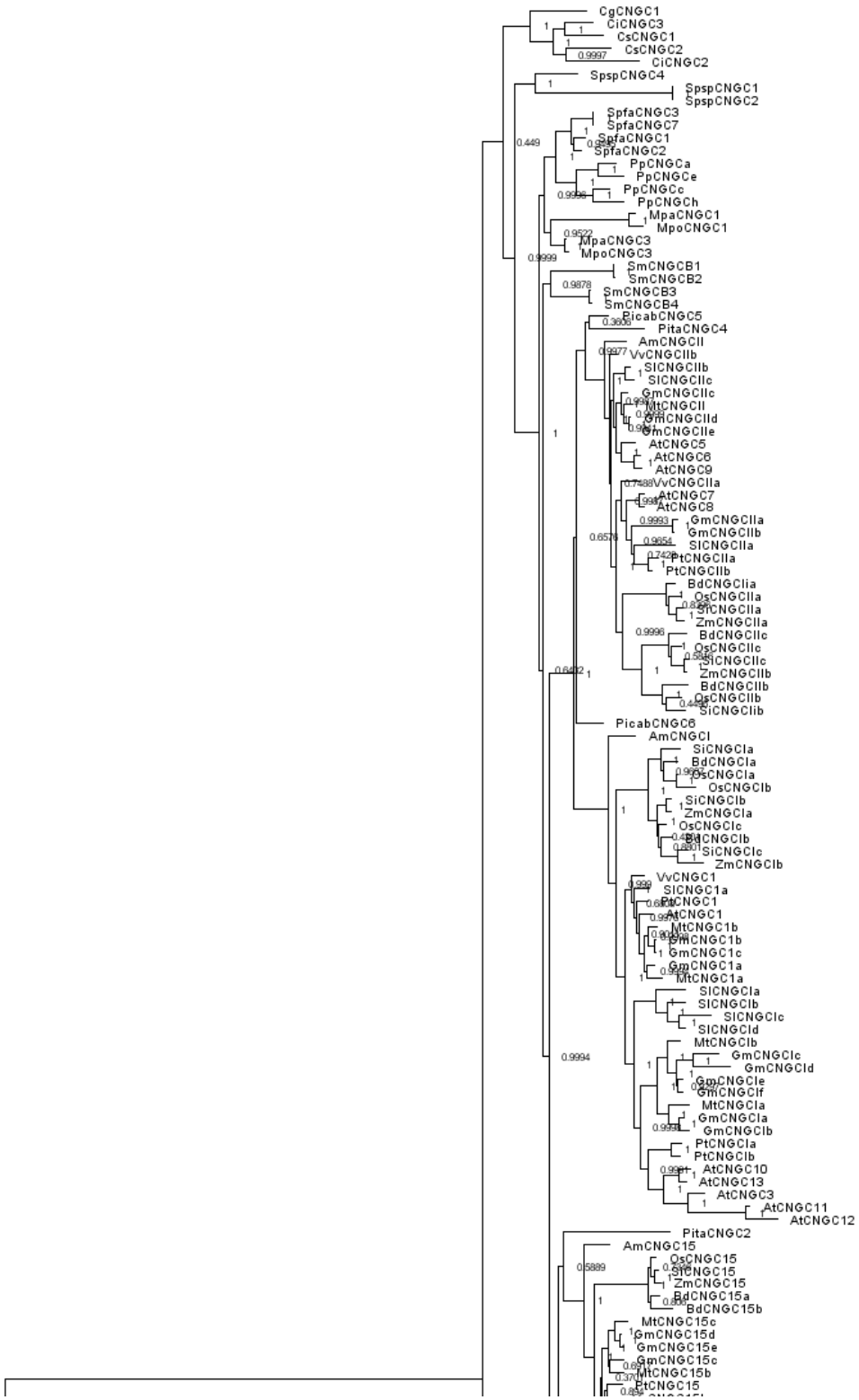
| | |
|---|---|
| PiabPHT1;7 | lcl\|MA_3730215g0010_high_confidence |
| PiabPHT1;8 | lcl\|MA_6448663g0010_high_confidence |
| PiabPHT1;9 | lcl\|MA_774908g0010_high_confidence |
| PitaPHT1;1 | lcl\|PITA_000015134_unnamed_protein_product |
| PitaPHT1;10 | lcl\|PITA_000096157_unnamed_protein_product |
| PitaPHT1;2 | lcl\|PITA_000020370_unnamed_protein_product |
| PitaPHT1;3 | lcl\|PITA_000027780_unnamed_protein_product |
| PitaPHT1;4 | lcl\|PITA_000045954_unnamed_protein_product |
| PitaPHT1;5 | lcl\|PITA_000057365_unnamed_protein_product |
| PitaPHT1;6 | lcl\|PITA_000060429_unnamed_protein_product |
| PitaPHT1;7 | lcl\|PITA_000067857_unnamed_protein_product |
| PitaPHT1;8 | lcl\|PITA_000074631_unnamed_protein_product |
| PitaPHT1;9 | lcl\|PITA_000093210_unnamed_protein_product |
| PlohPHT1;1 | onekp:SNOX_scaffold_2003804_Planotaenium_ohtanii |
| PltrPHT1;1 | onekp:MOYY_scaffold_2010768_Pleurotaenium_trabecul |
| PocoPHT1;1 | onekp:SCEB_scaffold_2054251_Podocarpus_coriaceus |
| PocoPHT1;2 | onekp:SCEB_scaffold_2054678_Podocarpus_coriaceus |
| PoruPHT1;1 | onekp:XLGK_scaffold_2006322_Podocarpus_rubens |
| PoruPHT1;2 | onekp:XLGK_scaffold_2006323_Podocarpus_rubens |
| PoruPHT1;3 | onekp:XLGK_scaffold_2058585_Podocarpus_rubens |
| PsmePHT1;1 | onekp:IOVS_scaffold_2005887_Pseudotsuga_menziesii |
| PsmePHT1;2 | onekp:IOVS_scaffold_2009133_Pseudotsuga_menziesii |
| PsmePHT1;3 | onekp:IOVS_scaffold_2009134_Pseudotsuga_menziesii |
| RoobPHT1 | onekp:XRTZ_scaffold_2002578_Roya_obtusa |
| SbPHT1;1 | Sb03g029970.1species_S_bicolor |
| SbPHT1;2 | Sb06g002540.1species_S_bicolor |
| SbPHT1;3 | Sb06g002560.1species_S_bicolor |
| SbPHT1;4 | Sb06g002800.1species_S_bicolor |
| SbPHT1;A | Sb01g046900.1species_S_bicolor |
| SbPHT1;B | Sb01g020570.1species_S_bicolor |
| SbPHT1;C | Sb01g020580.1species_S_bicolor |
| SbPHT1;D | Sb01g046890.1species_S_bicolor |
| SbPHT1;E | Sb07g023780.1species_S_bicolor |
| SbPHT1;F | Sb02g009880.1species_S_bicolor |
| SbPHT1;G | Sb01g047910.1species_S_bicolor |
| SbPHT1;H | Sb10g012710.1species_S_bicolor |
| SemoPHT1;1 | SELMODRAFT_126850 |
| SemoPHT1;2 | SELMODRAFT_135913 |
| SemoPHT1;3 | SELMODRAFT_165572 |
| SemoPHT1;4 | SELMODRAFT_438081 |
| SemoPHT1;5 | SELMODRAFT_78488 |
| SemoPHT1;6 | SELMODRAFT_78514 |
| SemoPHT1;7 | SELMODRAFT_83417 |
| SemoPHT1;8 | SELMODRAFT_90341 |

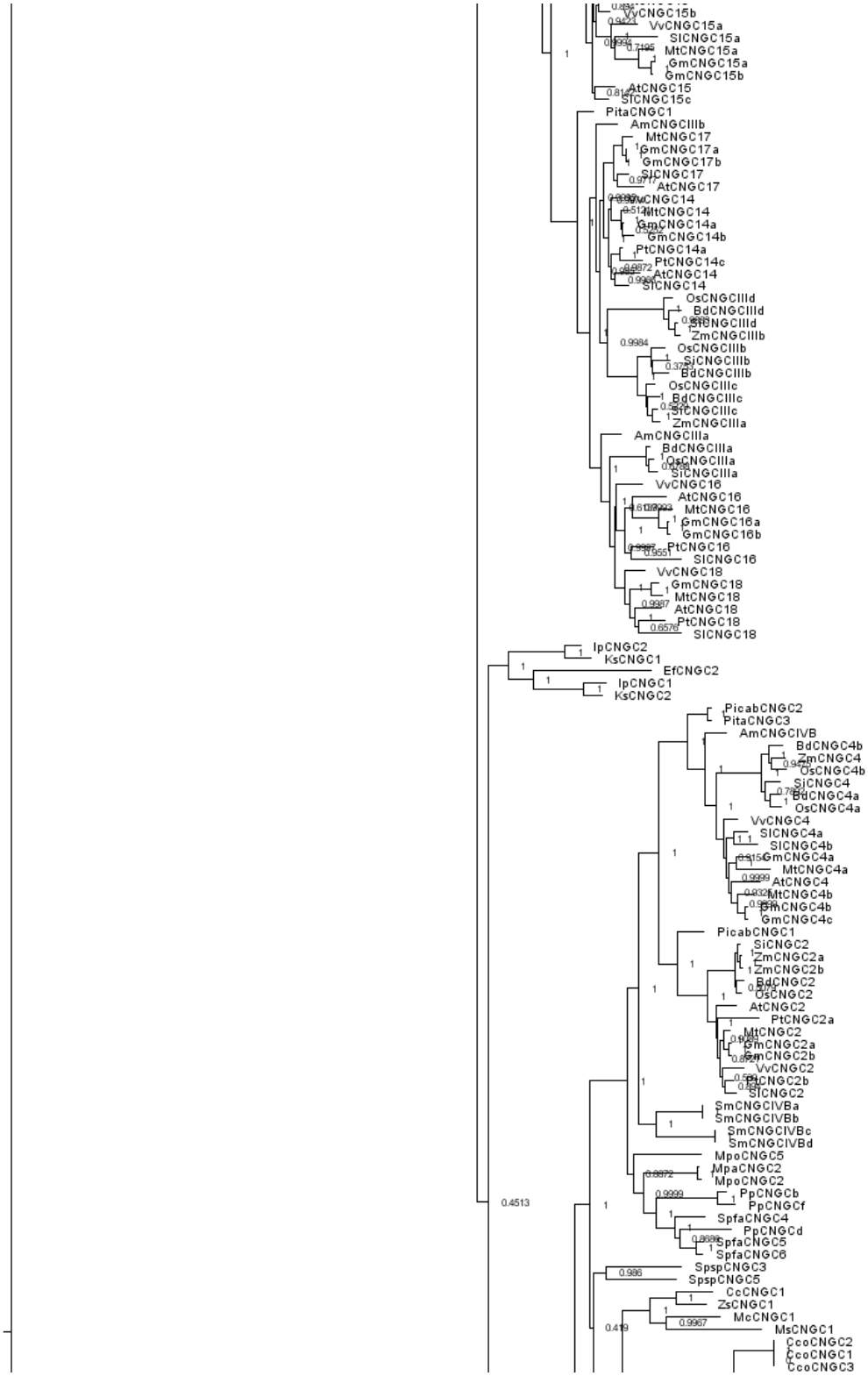
| | |
|---|---|
| SlPHT1;1 | Solyc09g090070.1.1species_S_lycopersicum |
| SlPHT1;2 | Solyc03g005530.1.1species_S_lycopersicum |
| SlPHT1;3 | Solyc09g090080.1.1species_S_lycopersicum |
| SlPHT1;4 | Solyc06g051850.1.1species_S_lycopersicum |
| SlPHT1;5 | Solyc06g051860.1.1species_S_lycopersicum |
| SlPHT1;6 | Solyc03g005560.1.1species_S_lycopersicum |
| SlPHT1;7 | Solyc09g073010.1.1species_S_lycopersicum |
| SlPHT1;A | Solyc06g034200.1.1species_S_lycopersicum |
| SlPHT1;B | Solyc09g066410.1.1species_S_lycopersicum |
| SpfaPHT1;1 | Sphfalx0003s0381.1 |
| SpfaPHT1;2 | Sphfalx0012s0181.2 |
| SpfaPHT1;3 | Sphfalx0026s0004.1 |
| SpfaPHT1;4 | Sphfalx0037s0008.1 |
| SpfaPHT1;5 | Sphfalx0043s0036.1 |
| SpfaPHT1;6 | Sphfalx0083s0064.1 |
| SpfaPHT1;7 | Sphfalx0146s0049.1 |
| SpfaPHT1;8 | Sphfalx0205s0009.1 |
| SpspPHT1;1 | onekp:TPHT_scaffold_2126991_Spirotaenia_sp. |
| StcoPHT1;1 | onekp:WCQU_scaffold_2009413_Staurodesmus_convergens |
| StcoPHT1;2 | onekp:WCQU_scaffold_2014933_Staurodesmus_convergens |
| StomPHT1;1 | onekp:RPRU_scaffold_2042823_Staurodesmus_omearii |
| StsePHT1;1 | onekp:ISHC_scaffold_2007068_Staurastrum_sebaldi |
| StsePHT1;2 | onekp:ISHC_scaffold_2009773_Staurastrum_sebaldi |
| TshePHT1;1 | onekp:GAMH_scaffold_2000146_Tsuga_heterophylla |
| TshePHT1;2 | onekp:GAMH_scaffold_2003690_Tsuga_heterophylla |
| TshePHT1;3 | onekp:GAMH_scaffold_2010861_Tsuga_heterophylla |
| ViviPHT1;1 | gi\|147778722\|emb\|CAN76105.1VITISV_020362 |
| ViviPHT1;10 | gi\|296083647\|emb\|CBI23636.3\|Vitisvinifera |
| ViviPHT1;11 | gi\|297736464\|emb\|CBI25335.3\|Vitisvinifera |
| ViviPHT1;12 | gi\|297744139\|emb\|CBI37109.3\|Vitisvinifera |
| ViviPHT1;13 | gi\|302143699\|emb\|CBI22560.3\|Vitisvinifera |
| ViviPHT1;14 | gi\|45356803\|gb\|AAS58441.1\|Vitisvinifera |
| ViviPHT1;15 | gi\|731389702\|ref\|XP_010650086.1\|Vitisvinifera |
| ViviPHT1;2 | gi\|147783645\|emb\|CAN70378.1\|VITISV_002165 |
| ViviPHT1;3 | gi\|147794581\|emb\|CAN78030.1\|VITISV_017530 |
| ViviPHT1;4 | gi\|225433120\|ref\|XP_002281264.1\|Vitisvinifera |
| ViviPHT1;5 | gi\|225433122\|ref\|XP_002285153.1\|Vitisvinifera |
| ViviPHT1;6 | gi\|225438157\|ref\|XP_002279070.1\|Vitisvinifera |
| ViviPHT1;7 | gi\|225457638\|ref\|XP_002274930.1Vitisvinifera |
| ViviPHT1;8 | gi\|225462697\|ref\|XP_002267327.1\|Vitisvinifera |
| ViviPHT1;9 | gi\|225462699\|ref\|XP_002267369.1\|Vitisvinifera |
| WonoPHT1;1 | onekp:RSCE_scaffold_2005978_Wollemia_nobilis |
| WonoPHT1;2 | onekp:RSCE_scaffold_2008014_Wollemia_nobilis |
| WonoPHT1;3 | onekp:RSCE_scaffold_2009850_Wollemia_nobilis |

| | |
|---|---|
| XaanPHT1;1 | onekp:GBGT_scaffold_2000929_Xanthidium_antilopaeum |
| ZynmPHT1;1 | onekp:STKJ_scaffold_2024366_Zygnema_sp. |
| ZyspPHT1;1 | onekp:MFZO_scaffold_2026548_Zygnemopsis_sp. |

**Sequences used in phylogenetic analysis: PHT1**

# Appendix 2


CgCNGC1
CiCNGC3
CsCNGC1
CsCNGC2
CiCNGC2
SpspCNGC4
SpspCNGC1
SpspCNGC2
SpfaCNGC3
SpfaCNGC7
SpfaCNGC1
SpfaCNGC2
PpCNGCa
PpCNGCe
PpCNGCc
PpCNGCh
MpaCNGC1
MpoCNGC1
MpaCNGC3
MpoCNGC3
SmCNGCB1
SmCNGCB2
SmCNGCB3
SmCNGCB4
PicabCNGC5
PitaCNGC4
AmCNGCII
VvCNGCIIb
SlCNGCIIb
SlCNGCIIc
GmCNGCIIc
MtCNGCII
GmCNGCIId
GmCNGCIIe
AtCNGC5
AtCNGC6
AtCNGC9
VvCNGCIIa
AtCNGC7
AtCNGC8
GmCNGCIIa
GmCNGCIIb
SlCNGCIIa
PtCNGCIIa
PtCNGCIIb
BdCNGCIIa
OsCNGCIIa
SlCNGCIIa
ZmCNGCIIa
BdCNGCIIc
OsCNGCIIc
SlCNGCIIc
ZmCNGCIIb
BdCNGCIIb
OsCNGCIIb
SlCNGCIIb
PicabCNGC6
AmCNGCI
SlCNGCIa
BdCNGCIa
OsCNGCIa
OsCNGCIb
SlCNGCIb
ZmCNGCIa
OsCNGCIc
MtCNGCIb
SlCNGCIc
ZmCNGCIb
VvCNGC1
SlCNGC1a
PtCNGC1
AtCNGC1
MtCNGC1b
GmCNGC1b
GmCNGC1c
GmCNGC1a
MtCNGC1a
SlCNGCIa
SlCNGCIb
SlCNGCIc
SlCNGCId
MtCNGCIb
GmCNGCIc
GmCNGCId
GmCNGCIe
GmCNGCIf
MtCNGCIa
GmCNGCIa
GmCNGCIb
PtCNGCIa
PtCNGCIb
AtCNGC10
AtCNGC13
AtCNGC3
AtCNGC11
AtCNGC12
PitaCNGC2
AmCNGC15
OsCNGC15
SlCNGC15
ZmCNGC15
BdCNGC15a
BdCNGC15b
MtCNGC15c
GmCNGC15d
GmCNGC15e
GmCNGC15c
MtCNGC15b
PtCNGC15


VvCNGC15b
VvCNGC15a
SlCNGC15a
MtCNGC15a
GmCNGC15a
GmCNGC15b
AtCNGC15
SlCNGC15c
PitaCNGC1
AmCNGCIIIb
MtCNGC17
GmCNGC17a
GmCNGC17b
SlCNGC17
AtCNGC17
VvCNGC14
MtCNGC14
GmCNGC14a
GmCNGC14b
PtCNGC14a
PtCNGC14c
AtCNGC14
SlCNGC14
OsCNGCIIId
BdCNGCIIId
SlCNGCIIId
ZmCNGCIIIb
OsCNGCIIIb
SiCNGCIIIb
BdCNGCIIIb
OsCNGCIIIc
BdCNGCIIIc
SlCNGCIIIc
ZmCNGCIIIa
AmCNGCIIIa
BdCNGCIIIa
OsCNGCIIIa
SlCNGCIIIa
VvCNGC16
AtCNGC16
MtCNGC16
GmCNGC16a
GmCNGC16b
PtCNGC16
SlCNGC16
VvCNGC18
GmCNGC18
MtCNGC18
AtCNGC18
PtCNGC18
SlCNGC18
IpCNGC2
KsCNGC1
EfCNGC2
IpCNGC1
KsCNGC2
PicabCNGC2
PitaCNGC3
AmCNGCIVB
BdCNGC4b
ZmCNGC4
OsCNGC4b
SiCNGC4
BdCNGC4a
OsCNGC4a
VvCNGC4
SlCNGC4a
SlCNGC4b
GmCNGC4a
MtCNGC4a
AtCNGC4
MtCNGC4b
SlCNGC4b
GmCNGC4c
PicabCNGC1
SiCNGC2
ZmCNGC2a
ZmCNGC2b
BdCNGC2
OsCNGC2
AtCNGC2
PtCNGC2a
MtCNGC2
GmCNGC2a
GmCNGC2b
VvCNGC2
SlCNGC2b
SlCNGC2
SmCNGCIVBa
SmCNGCIVBb
SmCNGCIVBc
SmCNGCIVBd
MpoCNGC5
MpaCNGC2
MpoCNGC2
PpCNGCb
PpCNGCf
SpfaCNGC4
PpCNGCd
SpfaCNGC5
SpfaCNGC6
SpspCNGC3
SpspCNGC5
CcCNGC1
ZsCNGC1
McCNGC1
MsCNGC1
CcoCNGC2
CcoCNGC1
CcoCNGC3
0.4513
0.9967
0.986

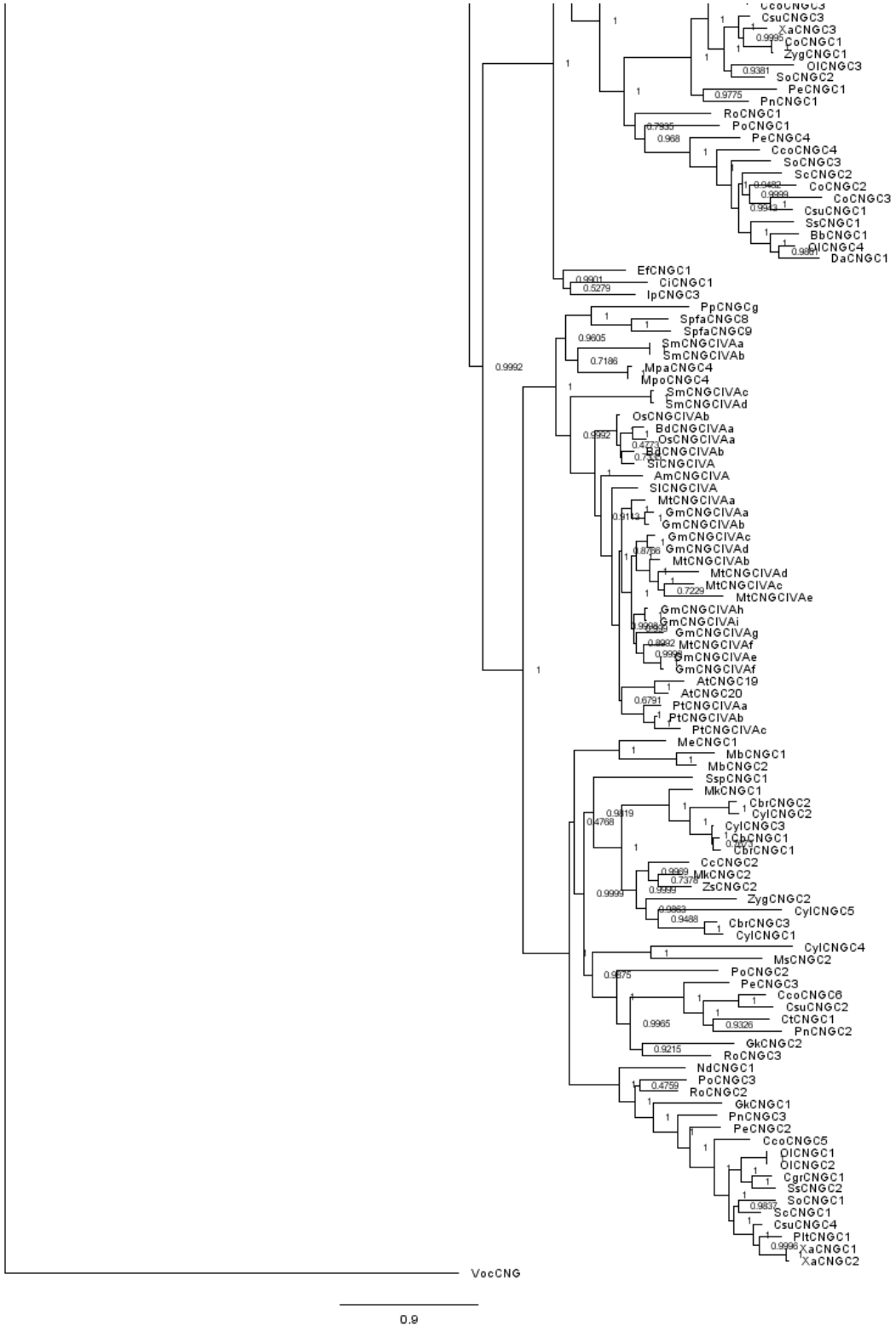
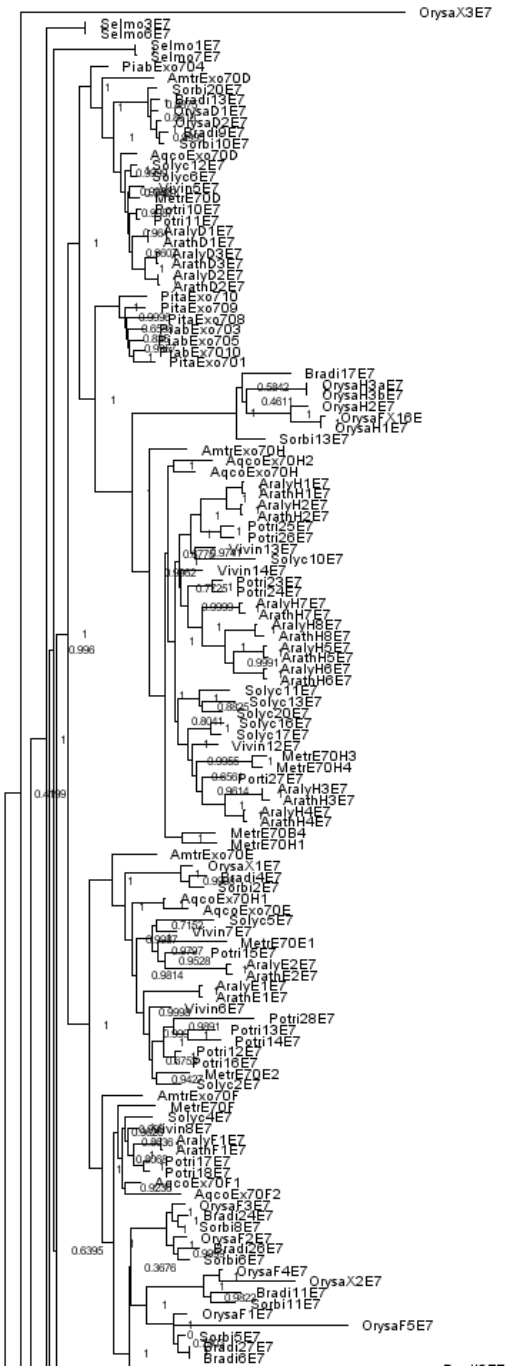
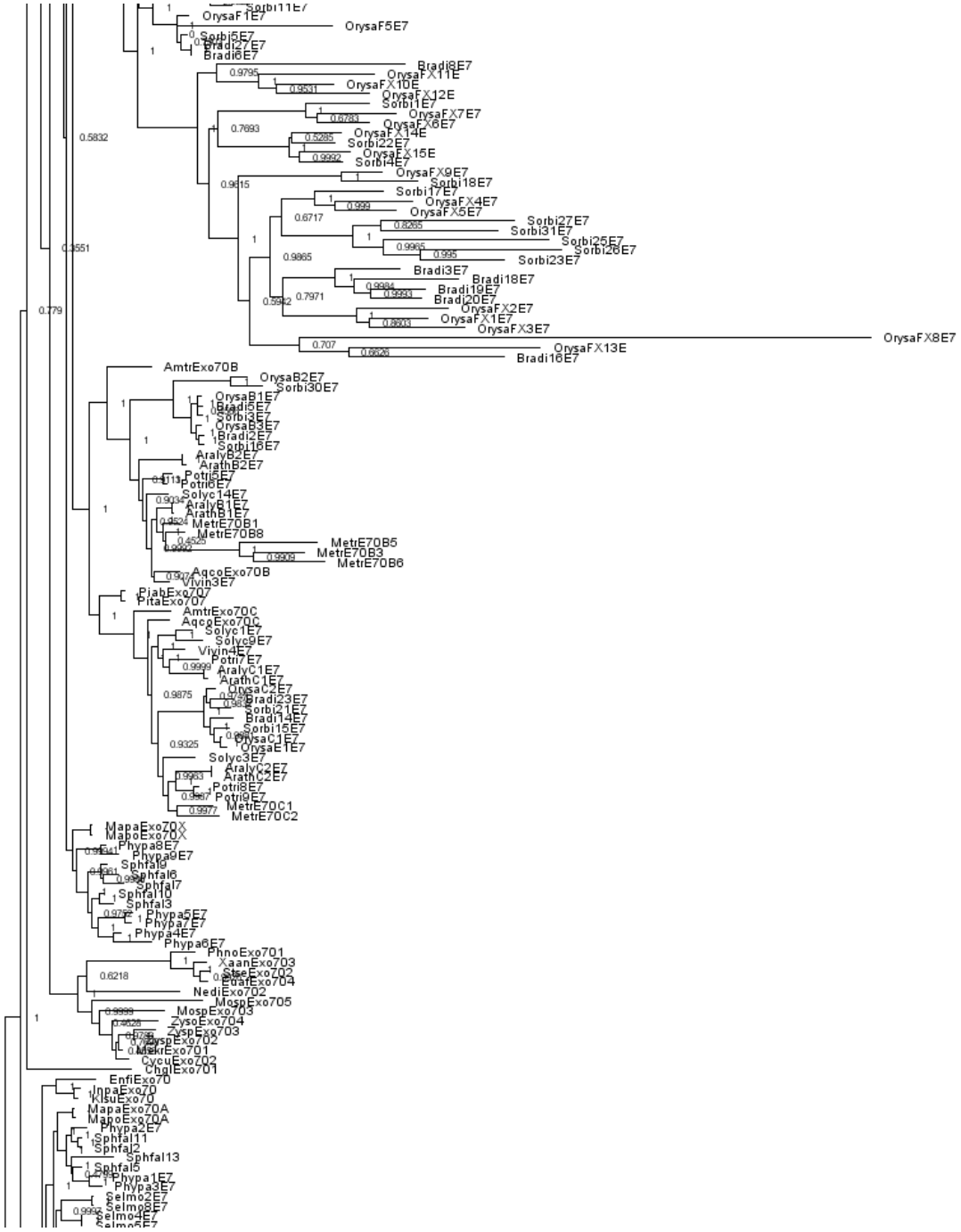



**Figure A.1** Maximum likelihood CNGC phylogeny in full.


OrysaX3E7
Selmo3E7
Selmo8E7
Selmo1E7
Selmo7E7
PiabExo70D4
AmtrExo70D
Sorbi20E7
Bradi13E7
OrysaD1E7
OrysaD2E7
Bradi9E7
Sorbi10E7
AqcoExo70D
Solyc12E7
Solyc8E7
Vivin5E7
MetrE70D
Potri10E7
Potri11E7
AralyD1E7
ArathD1E7
AralyD3E7
ArathD3E7
AralyD2E7
ArathD2E7
PitaExo710
PitaExo709
PitaExo708
PiabExo703
PiabExo705
PiabExo701D
PitaExo701
Bradi17E7
OrysaH3aE7
OrysaH3bE7
OrysaH2E7
OrysaFX16E
OrysaH1E7
Sorbi13E7
AmtrExo70H
AqcoEx70H2
AqcoExo70H
AralyH1E7
ArathH1E7
AralyH2E7
ArathH2E7
Potri25E7
Potri26E7
Vivin13E7
Solyc10E7
Vivin14E7
Potri23E7
Potri24E7
AralyH7E7
ArathH7E7
AralyH8E7
ArathH8E7
AralyH5E7
ArathH5E7
AralyH6E7
ArathH6E7
Solyc11E7
Solyc13E7
Solyc20E7
Solyc16E7
Solyc17E7
Vivin12E7
MetrE70H3
MetrE70H4
Potri27E7
AralyH3E7
ArathH3E7
AralyH4E7
ArathH4E7
MetrE70B4
MetrE70H1
AmtrExo70E
OrysaX1E7
Bradi4E7
Sorbi2E7
AqcoEx70H1
AqcoExo70E
Solyc5E7
Vivin7E7
MetrE70E1
Potri15E7
AralyE2E7
ArathE2E7
AralyE1E7
ArathE1E7
Vivin6E7
Potri28E7
Potri13E7
Potri14E7
Potri12E7
Potri16E7
MetrE70E2
Solyc2E7
AmtrExo70F
MetrE70F
Solyc4E7
Vivin8E7
AralyF1E7
ArathF1E7
Potri17E7
Potri18E7
AqcoEx70F1
AqcoExo70F2
OrysaF3E7
Bradi24E7
Sorbi8E7
OrysaF2E7
Bradi26E7
Sorbi6E7
OrysaF4E7
OrysaX2E7
Bradi11E7
Sorbi11E7
OrysaF1E7
OrysaF5E7
Sorbi5E7
Bradi27E7
Bradi6E7


Sorbi11E7
OrysaF1E7
OrysaF5E7
Sorbi5E7
Bradi27E7
Bradi6E7
Bradi8E7
OrysaFX11E
OrysaFX10E
OrysaFX12E
Sorbi1E7
OrysaFX7E7
OrysaFX6E7
OrysaFX14E
Sorbi22E7
OrysaFX15E
Sorbi4E7
OrysaFX9E7
Sorbi18E7
Sorbi17E7
OrysaFX4E7
OrysaFX5E7
Sorbi27E7
Sorbi31E7
Sorbi25E7
Sorbi26E7
Sorbi23E7
Bradi3E7
Bradi18E7
Bradi19E7
Bradi20E7
OrysaFX2E7
OrysaFX1E7
OrysaFX3E7
OrysaFX8E7
OrysaFX13E
Bradi16E7
AmtrExo70B
OrysaB2E7
Sorbi30E7
OrysaB1E7
Bradi5E7
Sorbi3E7
OrysaB3E7
Bradi2E7
Sorbi16E7
AralyB2E7
ArathB2E7
Potri5E7
Potri6E7
Solyc14E7
AralyB1E7
ArathB1E7
MetrE70B1
MetrE70B8
MetrE70B5
MetrE70B3
MetrE70B6
AqcoExo70B
Vivin3E7
PiabExo707
PitaExo707
AmtrExo70C
AqcoExo70C
Solyc1E7
Solyc9E7
Vivin4E7
Potri7E7
AralyC1E7
ArathC1E7
Bradi23E7
Sorbi21E7
Bradi14E7
Sorbi15E7
OrysaC1E7
OrysaE1E7
Solyc3E7
AralyC2E7
ArathC2E7
Potri8E7
Potri9E7
MetrE70C1
MetrE70C2
MapaExo70X
MapoExo70X
Phypa8E7
Phypa9E7
Sphfal4
Sphfal6
Sphfal7
Sphfal10
Sphfal3
Phypa5E7
Phypa7E7
Phypa4E7
Phypa6E7
PhnoExo701
XaanExo703
EdafExo704
NediExo702
MospExo705
MospExo703
ZysoExo704
ZyspExo703
ZyspExo702
MakrExo701
CycuExo702
ChglExo701
EnfiExo70
JnpaExo70
KlsuExo70
MapaExo70A
MapoExo70A
Phypa2E7
Sphfal11
Sphfal2
Sphfal13
Sphfal5
Phypa1E7
Phypa3E7
Selmo2E7
Selmo8E7
Selmo4E7

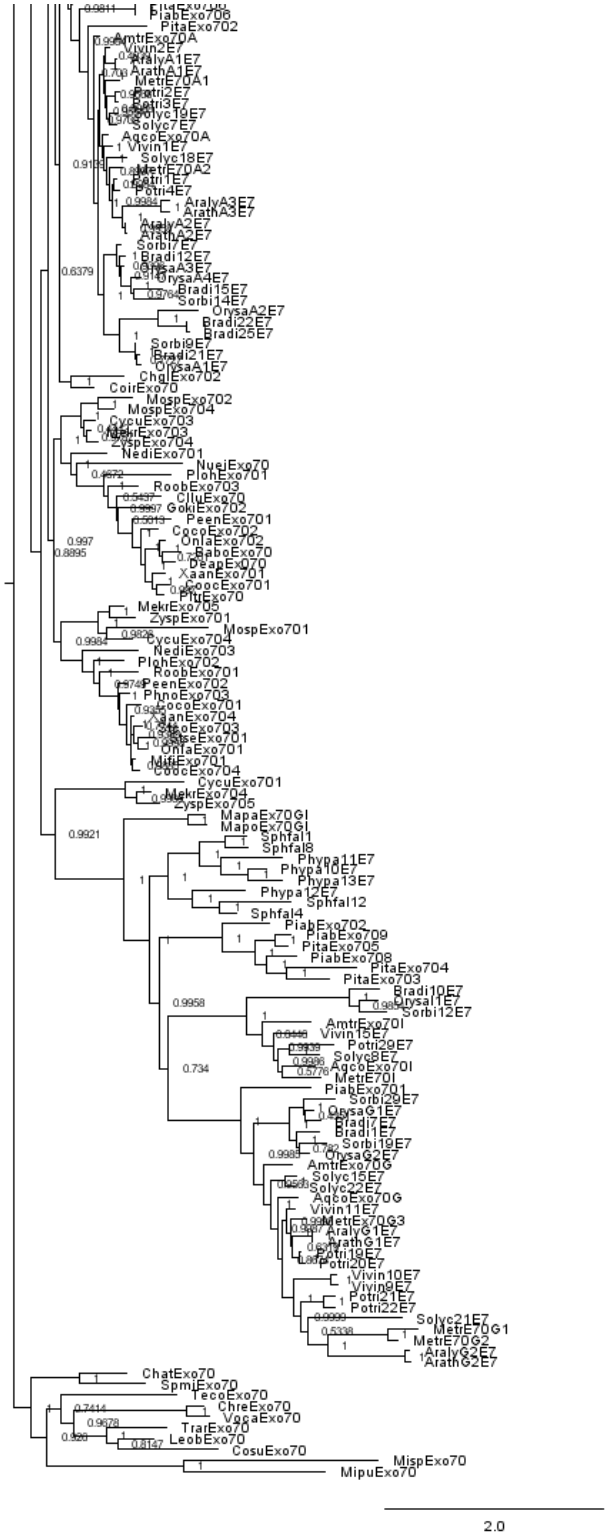
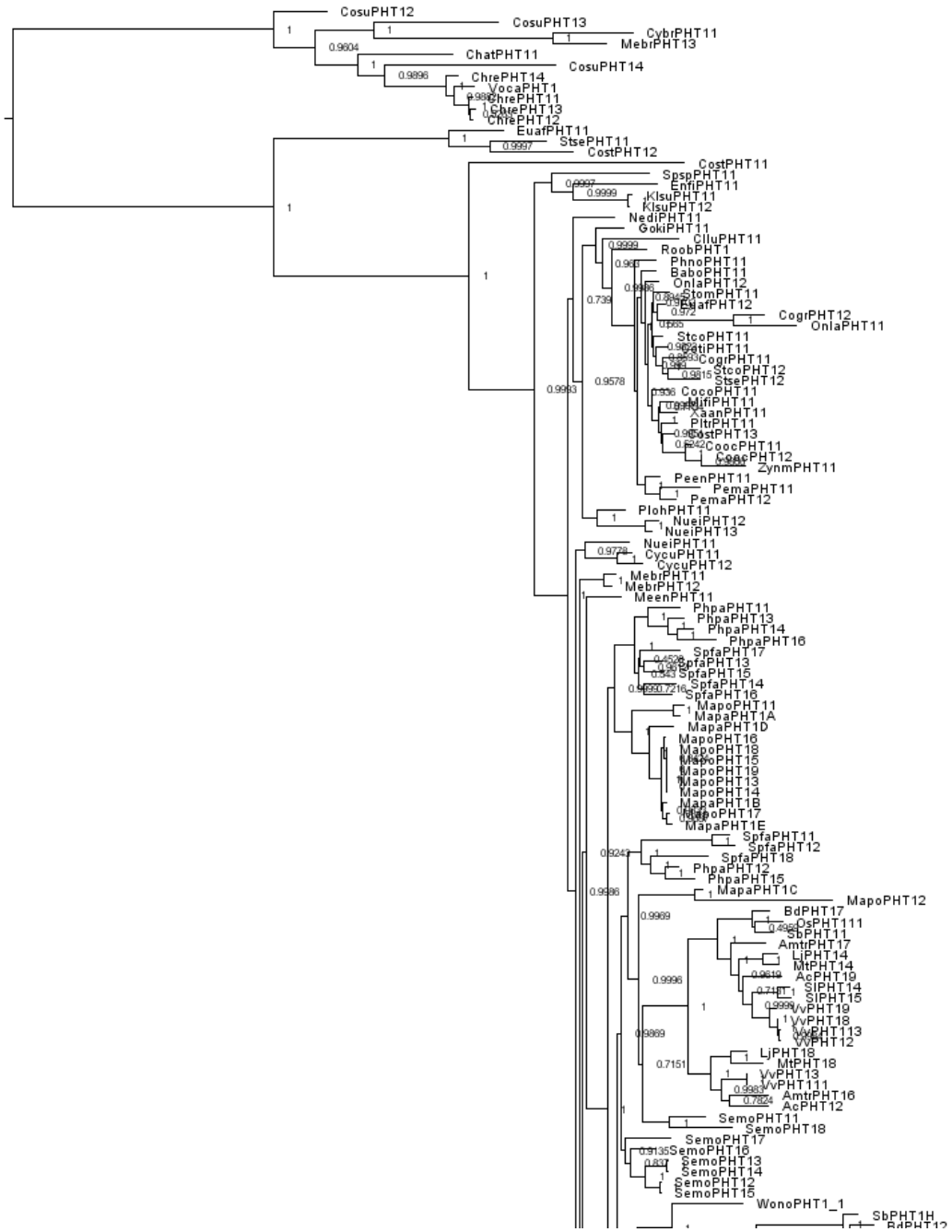
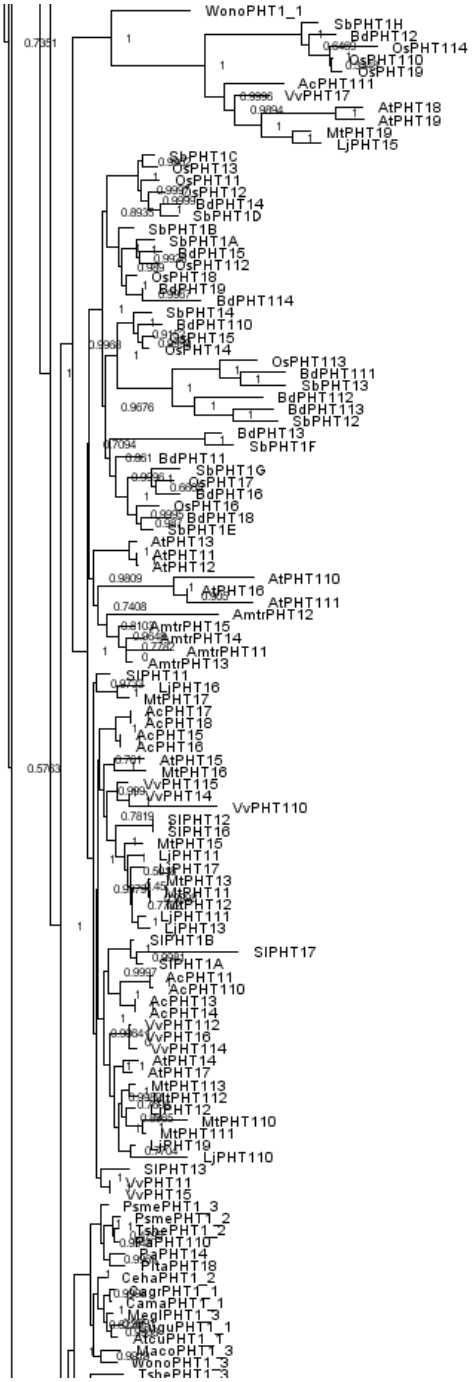



**Figure A.2** Maximum likelihood Exo70 phylogeny in full.


CosuPHT12
CosuPHT13
CybrPHT11
MebrPHT13
ChatPHT11
CosuPHT14
ChrePHT14
VocaPHT1
ChrePHT11
ChrePHT13
ChrePHT12
EuafPHT11
StsePHT11
CostPHT12
CostPHT11
SpspPHT11
EnfiPHT11
KlsuPHT11
KlsuPHT12
NediPHT11
GokiPHT11
CllaPHT11
RoobPHT1
PhnoPHT11
BaboPHT11
OnlaPHT12
StomPHT11
EuafPHT12
CogrPHT12
OnlaPHT11
StcoPHT11
CogrPHT11
StcoPHT12
StsePHT12
CocoPHT11
XaanPHT11
PltrPHT11
CoocPHT11
CoocPHT12
ZynmPHT11
PeenPHT11
PemaPHT11
PemaPHT12
PlohPHT11
NueiPHT12
NueiPHT13
NueiPHT11
CycuPHT11
CycuPHT12
MebrPHT11
MebrPHT12
MeenPHT11
PhpaPHT11
PhpaPHT13
PhpaPHT14
PhpaPHT16
SpfaPHT17
SpfaPHT13
SpfaPHT15
SpfaPHT14
SpfaPHT16
MapoPHT11
MapaPHT1A
MapaPHT1D
MapoPHT16
MapoPHT18
MapoPHT15
MapoPHT19
MapoPHT13
MapoPHT14
MapaPHT1B
MapoPHT17
MapaPHT1E
SpfaPHT11
SpfaPHT12
SpfaPHT18
PhpaPHT12
PhpaPHT15
MapaPHT1C
MapoPHT12
BdPHT17
OsPHT111
SbPHT11
AmtrPHT17
LjPHT14
MtPHT14
AcPHT19
SlPHT14
SlPHT15
VvPHT19
VvPHT18
VvPHT113
VvPHT12
LjPHT18
MtPHT18
VvPHT13
VvPHT111
AmtrPHT16
AcPHT12
SemoPHT11
SemoPHT18
SemoPHT17
SemoPHT16
SemoPHT13
SemoPHT14
SemoPHT12
SemoPHT15
WonoPHT1_1
SbPHT1H


WonoPHT1_1
SbPHT1H
BdPHT12
OsPHT114
OsPHT110
OsPHT19
AcPHT111
VvPHT17
AtPHT18
AtPHT19
MtPHT19
LjPHT15
SbPHT1C
OsPHT13
OsPHT11
OsPHT12
BdPHT14
SbPHT1D
SbPHT1B
SbPHT1A
BdPHT15
OsPHT112
OsPHT18
BdPHT19
BdPHT114
SbPHT14
BdPHT110
OsPHT15
OsPHT14
OsPHT113
BdPHT111
SbPHT13
BdPHT112
BdPHT113
SbPHT12
BdPHT13
SbPHT1F
BdPHT11
SbPHT1G
OsPHT17
BdPHT16
OsPHT16
BdPHT18
SbPHT1E
AtPHT13
AtPHT11
AtPHT12
AtPHT110
AtPHT16
AtPHT111
AmtrPHT12
AmtrPHT15
AmtrPHT14
AmtrPHT11
AmtrPHT13
SlPHT11
LjPHT16
MtPHT17
AcPHT17
AcPHT18
AcPHT15
AcPHT16
AtPHT15
MtPHT16
VvPHT115
VvPHT14
VvPHT110
SlPHT12
SlPHT16
MtPHT15
LjPHT11
LjPHT17
MtPHT13
LjPHT111
LjPHT13
SlPHT1B
SlPHT17
SlPHT1A
AcPHT11
AcPHT110
AcPHT13
AcPHT14
VvPHT112
VvPHT16
VvPHT114
AtPHT14
AtPHT17
MtPHT113
MtPHT112
LjPHT12
MtPHT110
MtPHT111
LjPHT19
LjPHT110
SlPHT13
VvPHT11
VvPHT15
PsmePHT1_3
PsmePHT1_2
TshePHT1_2
PsaPHT110
BpPHT14
PitaPHT18
CehaPHT1_2
CagrPHT1_1
CamaPHT1_1
MeglPHT1_3
AtcuPHT1_1
MacoPHT1_3
WonoPHT1_3
TshePHT1_3

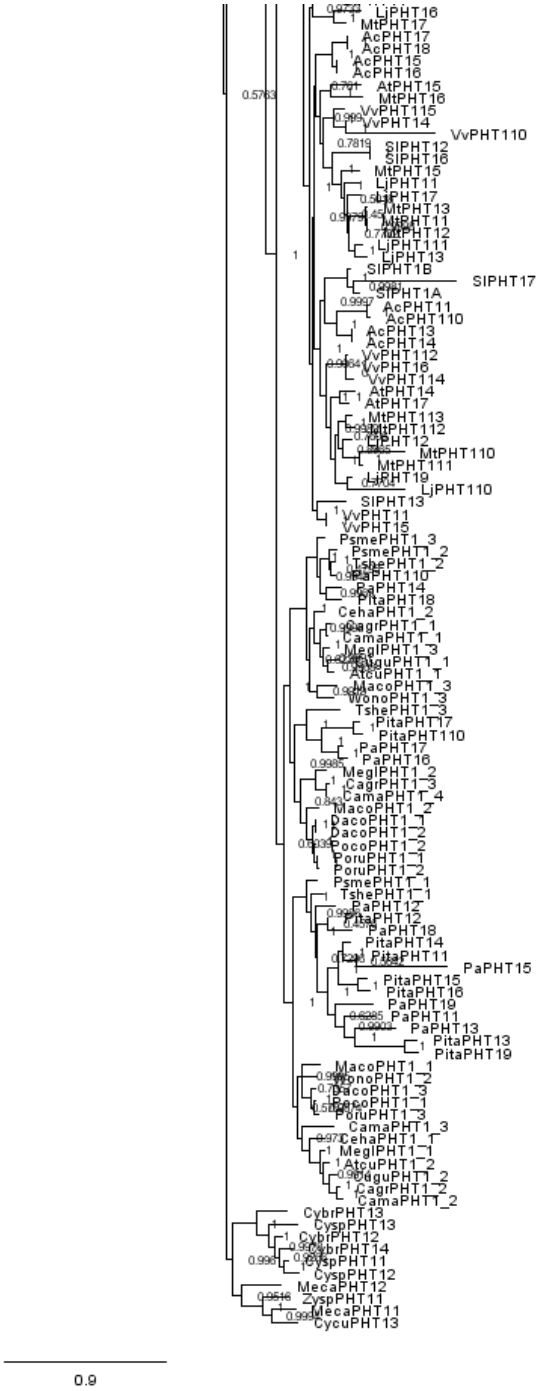



**Figure A.3** Maximum likelihood PHT1 phylogeny in full.

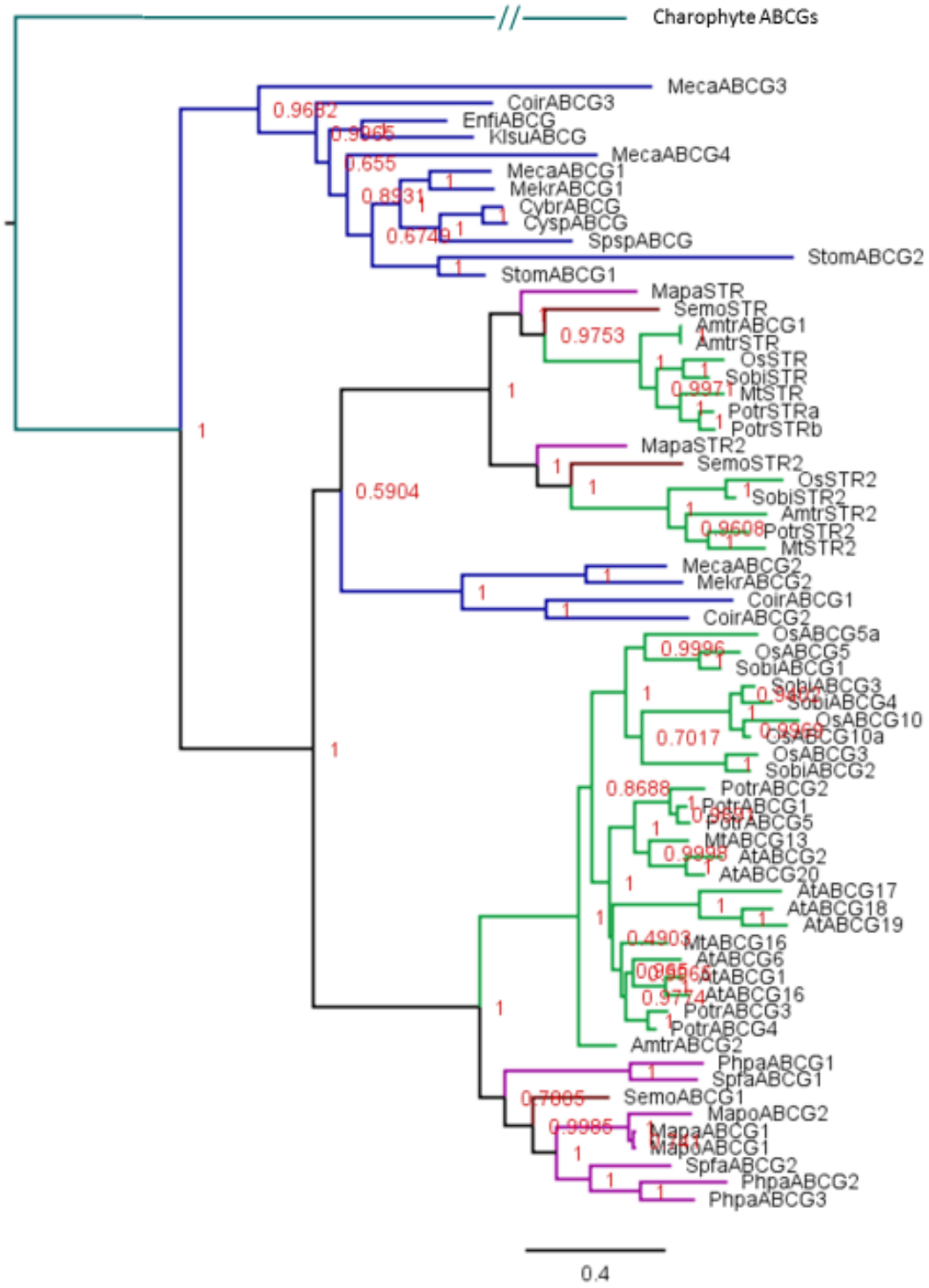



**Figure A.4** Maximum likelihood phylogeny of STR and the ABCGs.