The genetic control of rhizoid development in the liverwort *Marchantia polymorpha*

Victor Arnold Shivas Jones

Lincoln College
Department of Plant Sciences
University of Oxford
Supervised by Liam Dolan

Thesis submitted for the degree of Doctor of Philosophy
Michaelmas 2015

For Whamsie

The genetic control of rhizoid development in the liverwort *Marchantia polymorpha*

Victor Arnold Shivas Jones

Lincoln College
Department of Plant Sciences
University of Oxford

Thesis submitted for the degree of Doctor of Philosophy
Michaelmas 2015

Abstract

The first land plants faced a harsh terrestrial environment when they emerged from the water over 470 million years ago, and one of the key adaptations that allowed them radiate across the land was the development of a rooting system. To investigate the genetic mechanism that controlled the differentiation of rooting cells in ancient land plants, I carried out a mutant screen to identify genes that regulate rhizoid development in the liverwort *Marchantia polymorpha*, a member of the earliest-diverging lineage of land plants.

I used insertional mutagenesis to generate a population of 105,000 lines from which I selected 61 mutants with defects in rhizoid development, and identified 10 genes that are part of the network of genes that influence the differentiation and growth of rhizoids. Eight of these are late-acting genes that are required for the elongation of rhizoids by tip growth, while two are transcription factors that direct early events in the adoption of rhizoid fate. I identified the bHLH transcription factor Mp*ROOT HAIR DEFECTIVE 6-LIKE1* (Mp*RSL1*) as a key regulator of rhizoid differentiation, as gain-of-function mutations in Mp*RSL1* cause rhizoids to develop in ectopic locations. The homologues of Mp*RSL1* in the angiosperm *Arabidopsis* and the moss *Physcomitrella* control the differentiation of their root hairs and rhizoids, respectively, which suggests that a gene regulatory network that included *RSL* genes controlled the development of filamentous rooting cells in the last common ancestor of all land plants. I also identified Mp*WIP*, which encodes a member of a plant-specific family of zinc finger proteins, as a putative regulator of the development of both rhizoids and the cells of the air pore complex, a second specialized epidermal cell type. *WIP* genes have not been implicated in the control of rooting cell development in other species, so this role in *Marchantia* may be either inherited from the earliest land plants or a derived character.

This work demonstrates the suitability of *M. polymorpha* as a subject for large-scale mutageneses and screens for gene discovery. The genes I have found to be involved in rhizoid development indicate that the last common ancestor of all land plants already possessed a gene regulatory network that controlled the development of rooting cells, and that at least some of its components, such as *RSL* genes, have been conserved in its descendents since the divergence of the liverworts and other land plants.

Abbreviations

aLRT	Approximate likelihood ratio test
amiR	Artificial microRNA
CaMV	Cauliflower Mosaic Virus
CDS	Coding DNA sequence
CLSM	Confocal laser scanning microscope
EAR motif	Ethylene-responsive element binding factor-associated amphiphilic repression motif
gDNA	Genomic DNA
LB	Left border of T-DNA
miRNA	microRNA
NLS	Nuclear localisation signal
ORF	Open reading frame
PCR	Polymerase chain reaction
PI	Propidium iodide
PI4P	Phosphatidylinositol 4-phosphate
QC	Quiescent centre
RB	Right border of T-DNA
ROS	Reactive oxygen species
RT-qPCR	Reverse transcription-quantitative PCR
SEM	Scanning electron micrograph
SRDX	EAR-motif repression domain
T-DNA	Transfer DNA of *Agrobacterium tumefaciens*
TAIL-PCR	Thermal asymmetric interlaced PCR
Tak-1	Male wild type *Marchantia polymorpha* accession Takaragaike-1
Tak-2	Female wild type *Marchantia polymorpha* accession Takaragaike-2
UTR	Untranslated region
YFP	Yellow fluorescent protein

My thanks

I would like to express my profound gratitude to Professor Liam Dolan. Without his advice, guidance and support none of this work would have been possible. I also thank my second supervisor Professor Nick Harberd and my assessors, who have helped keep me going in the right direction throughout my studies.

I am grateful too to all the past and current members of the Dolan lab, who made my years in Oxford some of the best of my life. To my fellow screeners Giulia and Suvi, who have been there too. To Helen and Lida, for running such a tight ship. To Sourav, Chulmin, Hélène, Clémence, and Holger, for their help, advice, and words of wisdom. To the students of the Dolan lab, Tom, Bruno, Anna, Sandy, and Clément, for their insights, solidarity and camaraderie. To my adoptive lab family, Charlotte, Niloufer, Mara, and Eftychios. To the rest of the Department for their support, and especially Gem, for shepherding me through.

I am deeply obliged to the Newton Abraham Studentship and to the Vice Chancellor’s Fund, who generously funded this work.

Finally, I would not have been able to get this far without my family, whom I love.

Thanks to you all.

Victor

Chapter 1: General introduction

1.1 Introduction

The colonization of the land by plants over 470 million years ago was an event of outstanding importance in the history of the Earth, which made possible subsequent invasions of the land by animals and the establishment of complex terrestrial ecosystems (Berner 1997; Bateman et al. 1998) and had far-reaching effects on geochemical cycles, the atmosphere, and the climate. Through the physical action of roots and the secretion of acids into the rhizosphere, plants greatly accelerated the weathering of silicate minerals. This increase in the rate of reaction of atmospheric CO_2 with calcium and magnesium silicates caused a shift in the long-term carbon cycles and drastically reduced atmospheric CO_2 levels, causing global climate cooling (Berner 1997; Raven and Edwards 2001; Lenton et al. 2012).

The land plants are a monophyletic group that evolved from multicellular freshwater algae related to the extant charophytes, and a number of innovations were involved in their adaptation to the terrestrial environment. Important steps included the elaboration of meristems and tissue systems leading to the formation of complex body plans; the evolution of desiccation tolerance, which allowed plants to survive surrounded by air rather than water; and eventually the evolution of transport systems that allowed plants to move water from soil stores to the tops of tall trees. Among the key innovations that allowed early plants to survive in the harsh terrestrial environment, and eventually spread across dry land, was the evolution of rooting systems for anchorage, water uptake and nutrient acquisition (Bateman et al. 1998; Raven and Edwards 2001).

1.2 The evolutionary history of land plants

Land plants (or embryophytes) and the streptophyte algae (Charales, Coleochaetales, Zygnematales, Klebsormidiales, Chlorokybales and *Mesostigma*) together constitute a monophyletic group called the streptophytes. The streptophyte algae are a paraphyletic group, in that they do not include all descendents of a single common ancestor, while the land plants are monophyletic (Figure 1.1). It is unclear which algal group is most closely related to the land plants, though it is likely that one of three groups, Coleochaetales, Charales or Zygnematales, is sister to the land plants. Further phylogenetic analyses are needed to unequivocally define the closest algal relative of the land plants (Karol et al. 2001; Qiu et al. 2006; Finet et al. 2010; Wodniok et al. 2011; Leliaert et al. 2012; Ruhfel et al. 2014).

Three early diverging clades of extant land plants, liverworts, mosses and hornworts, are generally held to constitute a paraphyletic grade known as the bryophytes (Qiu et al. 2006), though this point is not entirely settled as some molecular phylogenies resolve the bryophytes as a monophyletic group that is sister to the vascular plants (Finet et al. 2010). Most modern analyses, however, find that the liverworts are sister to all other land plants (Kenrick and Crane 1997a; Karol et al. 2001; Qiu et al. 2006, 2007; Ruhfel et al. 2014). The relationships of the other two bryophyte groups are also not settled, but recent analyses have placed the hornworts sister to the vascular plants (Qiu et al. 2006, 2007; Ruhfel et al. 2014) (Fig. 1.1). However, this conclusion has been disputed by Cox and coworkers (2014), who suggest that the apparent phylogenetic closeness of the hornworts and vascular plants arises from similar nucleotide composition biases in the two groups driven by synonymous substitutions, and that in fact the bryophytes do form a monophyletic group. In either case, the liverworts likely belong to the earliest-diverging land plant lineage, either alone or with the other bryophytes.

The vascular plants are so named because they develop water-conducting tissues made up of cells with thickened, lignified cell walls. They are a monophyletic group that includes all land plants other than the bryophytes, namely the lycophytes, monilophytes (ferns and horsetails) and seed plants (gymnosperms and angiosperms) (Kenrick and Crane 1997b; Qiu et al. 2006).

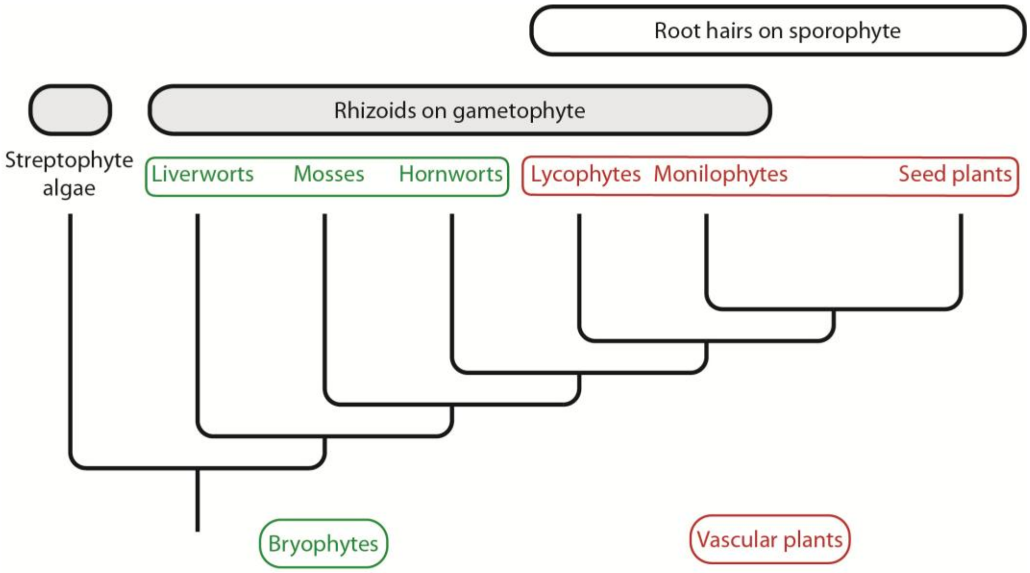


Figure 1.1 A phylogeny of land plants, including the distribution of root hairs and rhizoids among lineages. The freshwater stretophyte algae are sister to all landplants, and some develop filamentous rooting cells. The bryophytes (liverworts, mosses and hornworts) constitute a grade at the base of the land plant phylogeny, with the liverworts sister to all other land plants. In bryophytes the haploid gametophyte phase of the life cycle is dominant and develops rhizoids. In the vascular plants (lycophytes, monilophytes and seed plants) the diploid sporophyte phase is dominant, and develops root hairs. In the lycophytes and monilophytes both the gametophyte and sporophyte are free living, and develop rhizoids and root hairs, respectively. Tree after Ruhfel et al. (2014).

The fossil record of the early colonization of the land by plants is very patchy. The earliest clear evidence for plants on land is fossil cryptospores produced by embryophytes dating from the mid Ordovician, some 470 million years ago (Rubinstein et al. 2010). Analysis of the ultrastructure of younger cryptospores from the late Ordovician, approximately 450 million years ago, suggests that these early land plants had affinities to extant liverworts (Wellman et al. 2003), although macrofossils of liverwort bodies have not been found earlier than the mid Devonian (approximately 390 million years ago). Vascular plants appear in the mid-late Silurian, around 420 million years ago (Kenrick and Crane 1997a; Gensel 2008).

1.3 Rhizoids and root hairs develop at the plant-substrate interface in streptophytes

One of the key challenges facing land plants is the acquisition of water and nutrients. In order to explore the environment and gain access to these resources, almost all land plants develop tip-growing filamentous cells at the interface between the plant and substrate (the soil). Tip-growing cells elongate by growth in a small area at the apex of the cell, in contrast to the diffuse growth typical of most plant cells. These rooting cells are known as rhizoids or root hairs, depending on whether they are found in the haploid or diploid phase of the life cycle, respectively.

Although they live in the water, some of the streptophyte algal relatives of land plants, such as *Chara* (Charophytales) and *Spirogyra* (Zygnematales), develop rhizoids during the haploid phase of their life cycle that are morphologically similar to the rhizoids and root hairs of land plants (Braun 1997; Graham et al. 2009) while others (such as the Coleochaetales) do not (Lewis and McCourt 2004). The diploid phase of the life cycle of the streptophyte algae is unicellular, consisting only of a zygote that undergoes meiosis, and does not develop tip-growing cells.

In contrast to that of the streptophyte algae, the land plant life cycle consists of two distinct multicellular phases, comprising the diploid sporophyte and the haploid gametophyte. The gametophyte produces gametes that fuse to form a zygote; this undergoes mitosis to form the multicellular diploid sporophyte. In turn, cells of the sporophyte undergo meiosis to form haploid spores; these divide to form multicellular haploid gametophytes. Life cycles with multicellular haploid and diploid phases are said to display alternation of generations (Hofmeister 1851; Strasburger 1894; Kenrick and Crane 1997a).

In the early diverging bryophyte lineages (the liverworts, mosses and hornworts) the gametophyte is the dominant, and only free-living, stage of the life cycle. The gametophyte is in direct contact with the substrate, and develops a rooting system consisting of rhizoids to explore it. In contrast the relatively simple sporophyte is either entirely (liverworts and mosses) or mostly (hornworts) nutritionally dependent upon the haploid phase, and does not make contact with the substrate or develop rooting cells (Campbell 1934; McManus and Qiu 2008). The rhizoids of liverworts and hornworts are unicellular and unbranched, whereas those of mosses are multicellular and commonly branched (Crandall-Stotler and Stotler 2008; Goffinet et al. 2008; Renzaglia et al. 2008). Within a species there can be distinct populations of rhizoids; for instance, the basal and midstem rhizoids of the moss *Physcomitrella* follow separate developmental patterns (Sakakibara et al. 2003), while the liverwort *Marchantia* develops smooth and pegged rhizoids, which differ morphologically (Cao et al. 2014). The rooting system of most bryophytes consists solely of rhizoids that grow down into the soil, but exceptions are known. In some mosses, such as the Polytrichaceae, the shoot-like gametophore grows diageotropically as a rhizoid-bearing rhizome just below the soil before emerging as a leafy aerial shoot (Raven and Edwards 2001), while the rhizoidless genera *Takakia* and *Haplomitrium*, which are basal mosses and liverworts respectively, possess cylindrical branches that grow down into the soil (Grubb 1970; Raven and Edwards 2001; Qiu et al. 2007).

Unlike most bryophytes, the rooting system of vascular plants consists of more than filamentous rooting cells. These lineages evolved roots, axial organs that anchor the sporophyte and are involved in the absorption of nutrients and water. Roots possess unique defining characteristics including the formation of a protective root cap at the distal, growing end of the axes, in contrast with shoots where there is no cap (Raven and Edwards, 2001). Almost all roots develop filamentous cells, in this case known as root hairs, along their surface at the plant-soil interface. Root hairs are found in most vascular plant species and occur in all three major vascular plant lineages, the lycophytes, monilophytes, and seed plants (Dittmer 1949; Pearson 1969; Banks 2009). In some vascular plants, namely monilophytes and lycophytes, the gametophyte is still present as a free-living but ephemeral organism and develops rhizoids (Banks 1999, 2009) (Figure 1.1). In contrast the gametophyte is retained by and parasitic upon the sporophyte in all seed plants and reduced to a few cells (the pollen grains and embryo sac) in the angiosperms, where no rhizoids develop.

In short, among extant plants rhizoids are filamentous, tip-growing cells that develop on the gametophyte phase of the life cycle, while root hairs are such cells that grow on the root of the sporophyte. These cells form a major part of the interface between the plant and soil, and, as will be explored in the following sections, play vital roles in the uptake of water and nutrients.

1.4 Root hairs are important for nutrient uptake

Most angiosperms form root hairs at some stage during the development of the root system. They may be very short in some species such as onion (*Alium cepa*) or much longer in other species such as members of the Brassicales (reviewed in Jungk, 2001). Root hairs play a crucial role in the uptake of essential inorganic nutrients from the soil (Nye, 1966; Föhse *et al.*, 1991; Gahoonia and Nielsen, 1997, 1998). These essential nutrients are taken up in ionic form from the soil water at the root surface (Marschner, 2012). As the nutrient is transported into the plant it is replaced at the root surface by diffusion if it is present in sufficient concentrations in the soil water. Nitrate and ammonium are soluble and diffuse through the soil water, thereby replenishing the supply of these ions at the root surface. Phosphate on the other hand is not mobile in the soil water because of its tendency to bind to clay particles and form insoluble precipitates in the soil (Brady and Weil, 2008). As a result there is little diffusion of phosphate through the soil water to the root surface, where its concentration remains low after uptake into the root. Consequently the concentration of phosphate in soil water in the vicinity of the root remains low. This region surrounding the root where nutrients are present in very low concentrations is known as a depletion zone. The length of root hairs determines the volume of this zone. Plants with long root hairs develop large depletion zones (the zone has a relatively large diameter) while plants with short root hairs develop smaller depletion zones. All else being equal, long root hairs enable the plant to extract nutrients from a greater volume of soil compared to plants with short root hairs (Nye, 1966; Gahoonia and Nielsen, 1996, 1998; Gahoonia *et al.*, 1997). This explains why root hair length is positively correlated with the absorption of phosphate from the soil. Cultivars of barley with short root hairs take up less phosphate than cultivars with long root hairs when grown in field conditions with low available phosphate (Gahoonia and Nielsen, 1998, 2004). Furthermore root hairless barley mutants take up less phosphate from the soil and yield much less grain than wild type (long hair) cultivars in low phosphate conditions (Gahoonia *et al.*, 2001). The role of root hairs in the uptake of other nutrients with limited mobility in the soil has also been demonstrated. For example, the rate of K^+ uptake from the soil is also positively correlated with root hair length (Jungk, 2001). Together these data indicate that root hairs are important for the uptake of relatively immobile ions such as phosphate from the soil.

For many species root hairs are not the only means of mining a large volume of soil for nutrients, as they can form mycorrhizal associations with fungi to achieve this. Mycorrhizae are symbioses between fungi and plants in which fungi provide inorganic nutrients to the plant in exchange for reduced carbon compounds (Parniske, 2008). Approximately 80% of land plants develop symbioses with glomeromycotan fungal partners called vesicular arbuscular mycorrhizae (AM) (Van Der Heijden et al. 1998; Wang and Qiu 2006). Glomeromycota are obligate symbionts, have been found in 411 million year old fossil plants from the Rhynie Chert and are likely to have coevolved with the land plants (Remy et al. 1994; Wang et al. 2010). They form extensive hyphal networks in the soil and interface with the plant at branched intracellular structures called arbuscules where nutrients and carbon compounds are exchanged. There is a general trend among angiosperms that those species that develop mycorrhizae tend to form relatively short root hairs, while those that do not form mycorrhizae develop relatively long root hairs (Baylis, 1975; St John, 1980). Given that the ability to form mycorrhizae is an ancestral state among land plants, and that many unrelated groups of plants do not form mycorrhizae, it can be inferred that the ability to form mycorrhizae was lost independently among different lineages of plants (such as the Brassicales) (reviewed in Parniske, 2008). These non-mycorrhizal plants have evolved long root hairs that facilitate nutrient uptake in the absence of the fungal symbiont. Extreme examples of this are found among the Cyperaceae, where some species that lack mycorrhizae growing in nutrient poor soils develop “dauciform roots” where very long root hairs develop in patches along the root system. The formation of cluster roots with very long root hairs among species of the Proteaceae that lack mycorrhizae is another example where long root hairs provide the ability to extract limiting nutrients from the soil in the absence of mycorrhizae (reviewed in Lambers *et al.*, 2010).

1.5 Rhizoids have roles in anchorage and the uptake of water and nutrients

It is well established that root hairs are important in the uptake of nutrients. In contrast, there is little evidence about the function of rhizoids. It has been asserted in the literature that the primary role of rhizoids is in attachment to the substrate (Duckett *et al.*, 1998; Goffinet *et al.*, 2008; Crandall-Stotler *et al.*, 2009). The rhizoids of many liverworts form discs or ramify at their tips when they contact solid particles and adhere strongly (Haberlandt, 1914; Odu and Richards, 1976; Pocock and Duckett, 1985; Duckett *et al.*, 1991). Similar branching has also been observed at the tips of moss rhizoids in contact with hard substrates (Duckett, 1994a; Pressel and Duckett, 2009), as well as in the rhizoids of filmy fern gametophytes (Hymenophyllaceae) (Duckett *et al.*, 1996), while moss rhizoids can also display thigmotropic responses, coiling around objects in the substrate (Duckett, 1994b; Duckett and Matcham, 1995). A role in attachment is also suggested by the observation that, at least among the highly branched pleurocarpous mosses, rhizoids are more abundant and highly branched in plants growing on bare, hard substrates like rocks than ones growing on soil (Odu, 1978). Attachment to the substrate may be facilitated by the production of adhesive sulphated non-cellulose polysaccharides by rhizoid tips (Odu, 1989). These observations suggest that one of the important roles played by rhizoids in many species is mechanical attachment to the substrate, aside from any potential physiological roles in uptake.

Beyond their importance in attachment, rhizoids are also involved in the uptake and transport of water. Many bryophytes lack thick cuticles and absorb water over their whole surface (Proctor, 2000). Though rhizoids are not required for direct uptake of water in these species, some mosses produce a tomentum, a thick covering of rhizoids growing from the stem, and the spaces that form between the hairs aid water transport by capillary action (Proctor, 1984). Internal water transport occurs in some bryophytes, such as the moss *Polytrichum*, the rhizoids of which have been shown to take up water from the substrate, though the importance of this route of water uptake is probably minor compared to uptake across aerial surfaces of the plant (Mägdefrau, 1938; Trachtenberg and Zamski, 1979). Similarly, rhizoids are involved in the uptake and transport of water from the substrate in the complex thalloid liverworts of the order Marchantiales. These liverworts possess two kinds of rhizoids: smooth-walled rhizoids and tuberculate rhizoids, the latter having peg-shaped thickenings that project into the lumen of the rhizoid. These thickened rhizoids form bundles (like the moss tomenta) that run along the ventral surface of the thallus. Aqueous dyes are rapidly transported by capillarity between these rhizoids (Bowen, 1935; Czaja, 1936; McConaha, 1939; McConaha, 1941; Kobiyama and Crandall-Stotler, 2011). In addition to this external movement along bundles of rhizoids, water can travel inside both smooth-walled and tuberculate rhizoids, and into the cells of the thallus surrounding the rhizoid base (Kamerling, 1897; Clee, 1943). In *Conocephalum conicum* and *C. japonicum* the movement of water from the rhizoids into the thallus is promoted by specialized pitted cells on the ventral surface (Kobiyama and Crandall-

Stotler, 2011). Together these observations indicate that rhizoids are important for water transport in the Marchantiales. Rhizoids are thus involved in water uptake in at least some species of liverwort and moss.

Along with the uptake of water, rhizoids can also act in the uptake of inorganic nutrients in both algae and bryophytes. Amongst the algae, rhizoids of *Chara* species grow into the substrate where they play an important role in anchoring the plant. In addition, these rhizoids contain a higher concentration of mineral nutrients than the open water (Barko *et al.*, 1991) and take up nitrate, ammonium and phosphates from sediments (Box, 1986; Vermeer *et al.*, 2003). Nutrient acquisition by mosses is not well understood, but it is generally thought that most mosses get most of their nutrients from precipitation and the deposition of dust (Bates, 1992). Soil-growing mosses have been shown to be able to obtain nutrients from the substrate (Chapin *et al.*, 1987; Bates and Farmer, 1990; Van Tooren *et al.*, 1990), though it has not been shown whether this is a result of direct absorption by the rhizoids or external transport of nutrient-bearing soil water over the plant surface to leafy parts. There are to my knowledge no reports demonstrating that liverwort rhizoids have a direct role in nutrient uptake. However, many liverworts are colonized through their rhizoids by fungi to form mycorrhiza-like associations (Read *et al.*, 2000; Russell and Bulman, 2005), which can substantially increase the uptake of nutrients from the soil (Humphreys *et al.*, 2010). In general, studies into the uptake of nutrients by rhizoids have been limited by the difficulty of separating the absorptive functions of the rhizoids from those of the rest of the plant body, but there are indications that rhizoids contribute at least in part to this process.

Various lines of evidence from diverse species thus indicate that the rhizoids of algae and bryophytes can contribute significantly to attachment and to the uptake of water and nutrients. Similarly to the better-studied root hairs of vascular plants, they are key contributors to the fulfilment of the rooting function in these land plant lineages.

1.6 Genetic control of root hair and rhizoid differentiation and growth

Root hairs and rhizoids elongate by a distinctive mode of cell expansion called tip growth, in which new cell surface (membrane and wall) is deposited at the growing tip of the cell and there is no deposition along the sides. The root hairs of the angiosperm *Arabidopsis thaliana* are a model of cell differentiation and morphogenesis, and the cytological organization and cellular processes involved in their growth are understood in some detail. The knowledge of rhizoid development is less extensive; however, in some respects it is known to parallel the development of root hairs. Therefore an understanding of the findings concerning root hair development may be instructive when studying the development of rhizoids.

Root hair development in *Arabidopsis* can be divided into three major phases (Wymer et al. 1997). In the initiation phase the site of root hair outgrowth from a trichoblast (hair cell) is selected, the cell wall loosens at this position and a bulge forms. Tip growth from the bulge is then established, and subsequently a phase of rapid elongation begins. During this phase, the cytoplasm of root hairs shows a characteristic polarized organization. At the extreme apex of the root hair is a region of cytoplasm that is dense with secretory vesicles, required for the targeted delivery of cell wall material to the growth point, and endocytotic vesicles that recycle excess membrane delivered to the tip (Ketelaar et al. 2008; Rounds and Bezanilla 2013). The sub-apical region is organelle-rich, with many mitochondria and Golgi bodies, as well as compacted endoplasmic reticulum (Galway et al. 1997; Ridge et al. 1999). The cytoplasmic region extends about 70 μm behind the tip, and the more basal region is highly vacuolated, with the cytoplasm restricted to the periphery of the cells (Galway et al. 1997). The rhizoids of both mosses and ferns display a similar polarized cytological organization (Duckett et al. 1998; Parton et al. 2000; Pressel et al. 2008). After a time the root hair reaches its mature length and growth ceases (Carol and Dolan 2002; Datta et al. 2011).

Numerous physiological processes and genes are involved in each of these phases of root hair development. Below I present some general features of this process, without exhaustively reviewing the literature on root hair growth.

1.6.1 The patterning of root hair fate in *Arabidopsis* relies on positional information and lateral inhibition with feedback

One of the key features of complex multicellular organisms is the arrangement of different cell types in space, which relies on the generation of patterns of differing cell fates. Well-studied examples include the distribution of sensory bristles in the fruit fly *Drosophila*, which arises from *Delta*-*Notch*-mediated lateral inhibition (Simpson 1990), and, in plants, the spacing of stomata and trichomes (Ishida et al. 2008; Pillitteri and Torii 2012). In angiosperms, cells of the root epidermis differentiate into root hair cells according to three main patterns (Dolan and Roberts 1995). Most species have Type I patterning, in which all cells are competent to grow a root hair, though not all actually do (Clowes 2000; Pemberton 2001). Type II patterning is found in the grasses and some magnoliids; root hair (trichoblast) and non root hair (atrichoblast) cells alternate in the same file, usually as the result of an asymmetric cell division, the smaller daughter becoming a hair cell (Clowes 2000). In plants with Type III patterning, files of trichoblasts alternate with one or more files of atrichoblasts. *Arabidopsis* falls into this class, which is found only in the Brassicales, Caryophyllales and Boraginaceae (Pemberton 2001).

In *Arabidopsis* root epidermal cells that overlie an anticlinal cell wall in the layer below become trichoblasts, while those that do not will remain hairless. The genetic network that determines this patterning has been characterized; it also controls epidermal cell fate in the hypocotyl (Berger et al. 1998) and utilizes components similar to those that determine trichome patterning in the leaf (reviewed in Ishida et al. 2008). The pattern is generated by a combination of positional information relative to the underlying cell layer, mediated by the receptor-like kinase SCRAMBLED (SCM), and lateral inhibition with feedback (See Figure 1.2) (Kwak and Schiefelbein 2008; Savage et al. 2008; Schiefelbein et al. 2009). In the non-hair position, a complex of WEREWOLF (WER), GLABRA3/ENHANCER OF GLABRA3 (GL3/EGL3) and TRANSPARENT TESTA GLABRA (TTG) is preferentially formed and promotes the expression of GLABRA2 (GL2), which represses the expression of trichoblast-specific genes and induces the expression of atrichoblast-specific genes. In the hair position, the expression of WER is repressed by SCM, and so WER-complex cannot be formed. Lateral inhibition with feedback is accomplished through the induction by the WER-complex of the expression of the mobile protein CAPRICE (CPC) in non-hair cells, which moves into the neighbouring cell and prevents the accumulation of WER-complex there; thus no GL2 is expressed in the trichoblast, and root hair development can occur. In turn, the trichoblast produces GL3, which moves into the neighbouring atrichoblast and allows the formation of WER-complexes. In this way the neighbouring cells reinforce each other's fate. However, the phylogenetic distribution of the Type III root epidermal patterning found in *Arabidopsis* is limited, and it is not a general feature of root hair and rhizoid development. It is unknown whether the patterning mechanism outlined above is active in other species.

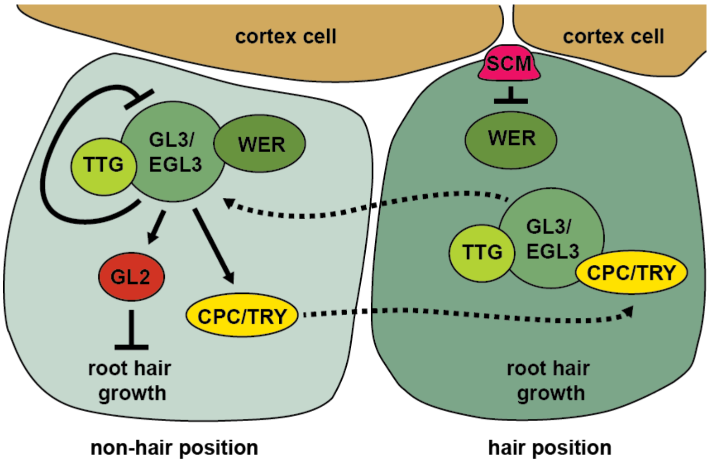


Figure 1.2 A model of cell-type patterning determination in *Arabidopsis*. Solid arrows represent positive gene regulation, barred lines represent negative gene regulation, dashed arrows represent protein movement.

Very little is known about the patterning of rhizoids in bryophytes, at either the morphological or molecular levels. Rhizoid patterning is at least partly position-dependent in the model moss *Physcomitrella patens*, where mid-stem rhizoids develop from epidermal cells adjacent to a leaf trace (Sakakibara et al. 2003). In *Marchantia polymorpha,* smooth rhizoids and tuberculate rhizoids develop in different regions of the thallus, smooth rhizoids in distinct patches adjacent to the midrib, and tuberculate rhizoids from epidermal cells beneath ventral scales or from cells of the marginal scales themselves (McConaha 1941; Cao et al. 2014). Within these regions, the rhizoids rarely develop adjacent to each other, each normally being surrounded by a rosette of non-rhizoid cells (Cao et al. 2014). The molecular bases of the formation of these patterns has not been described, but clearly rhizoid patterning mechanisms exist in at least some bryophytes.

1.6.2 An ancient conserved mechanism regulates the differentiation of root hairs and moss rhizoids

While little is known about the means by which rooting cell patterning is determined outside of *Arabidopsis*, a common genetic mechanism regulates the differentiation of both root hairs and moss rhizoids. Two basic helix-loop-helix transcription factors, encoded by the class I *RSL* genes *ROOT HAIR DEFECTIVE 6* (*AtRHD6*) and *ROOT HAIR DEFECTIVE 6-LIKE 1* (*AtRSL1*) control root hair differentiation in *Arabidopsis*. The proteins are expressed in cells that will go on to develop root hairs, where they promote the transcription of genes necessary for tip growth such as the Class II *RSL ROOT HAIR DEFECTIVE 6-LIKE 4* (*RSL4*) (Menand et al. 2007; Yi et al. 2010). Mutants that lack AtRHD6 and AtRSL1 function do not develop root hairs (Menand et al. 2007; Yi et al. 2010). Two similar genes were identified in the genome of the model moss *Physcomitrella patens* and named *PpRSL1* and *PpRSL2*. Few rhizoids develop in double mutants that lack both PpRSL1 and PpRSL2 function (Menand et al. 2007). Furthermore, constitutive co-expression of *PpRSL1* and *PpRSL2* transforms the gametophore into a mass of rhizoids (Jang et al. 2011). These data indicate that *PpRSL1* and *PpRSL2* together are necessary and sufficient to direct rhizoid development, and so are key regulators of rhizoid development in mosses.

Experiments have shown that *Arabidopsis* mutants lacking RHD6 function develop root hairs if made to express a functional *PpRSL1* gene from *Physcomitrella* (Menand et al. 2007). This indicates that the function of these Class I RSL proteins has been conserved for more than 420 million years since mosses and angiosperms diverged from a common ancestor. As they are transcription factors that regulate the transcription of target genes, this finding suggests that it is likely that the regulatory network downstream of the two regulatory genes has been largely conserved during the intervening period.

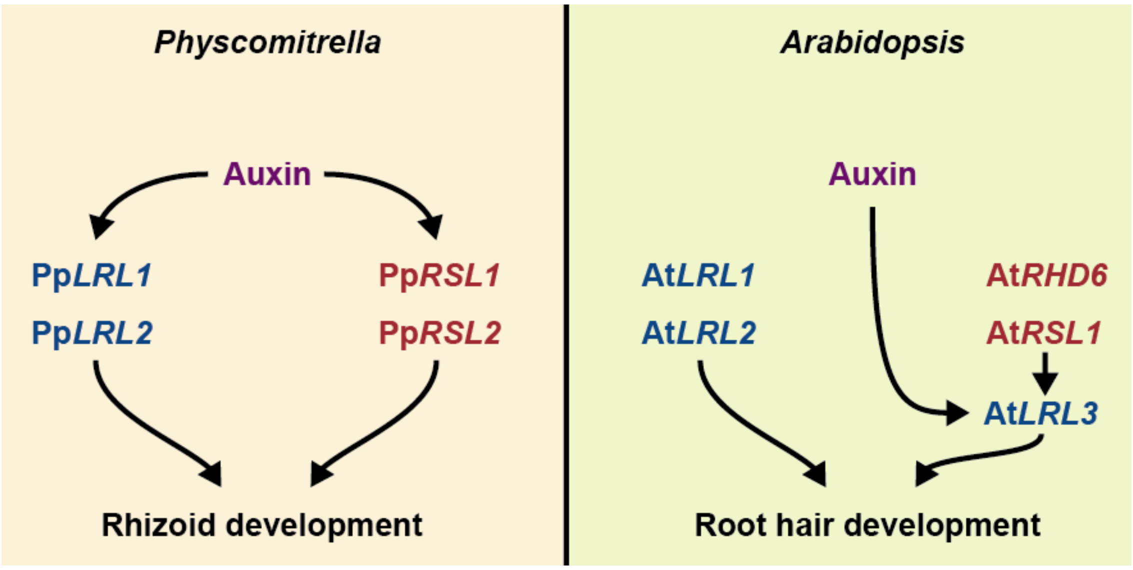


Figure 1.3 Conserved regulators of rooting cell development in a moss and angiosperm. Models of the transcriptional interactions between auxin, *LRLs* and class I *RSL* genes in the control of rhizoid development in *Physcomitrella* and root hair development in *Arabidopsis*.

A second group of bHLH transcription factors and the plant hormone auxin are also part of this conserved genetic network that regulates the development of tip-growing rooting cells. *LOTUS ROOTHAIRLESS-LIKE1* (*LRL*) genes encode group XI bHLH transcription factors that are redundantly required for the development of root hairs in *Arabidopsis* and rhizoids in *Physcomitrella* (Tam et al. 2015). The expression of both class I Pp*RSL* genes and Pp*LRL* genes is promoted by auxin. In *Physcomitrella* the expression of Pp*RSL* and Pp*LRL* genes is promoted by auxin, but neither group affects the expression of the other (Jang et al. 2011; Tam et al. 2015). In contrast, in *Arabidopsis* the expression of At*LRL3* is promoted by both class I RSLs and auxin, while the expression of At*RHD6* and At*RSL1* is affected by neither AtLRLs nor auxin, and the expression of At*LRL1* and At*LRL2* is affected by neither RSLs nor auxin (Bruex et al. 2012; Tam et al. 2015). The components of this gene regulatory network have been conserved between *Physcomitrella* and *Arabidopsis*, even though the precise connections between the components differ (Figure 1.3).

1.6.3 Auxin and ethylene promote both the differentiation and elongation of rooting cells

The plant hormone auxin promotes the differentiation of root hairs and rhizoids as part of a network that includes *RSL* and *LRL* transcription factors in *Arabidopsis* and *Physcomitrella*. Another plant hormone, the gas ethylene, also promotes root hair differentiation, and both auxin and ethylene promote elongation later in root hair development. Pharmacological studies indicate that auxin and ethylene positively regulate root hair initiation in *Arabidopsis* (Tanimoto et al. 1995; Rahman et al. 2002). The targets of auxin and ethylene lie downstream of At*RHD6*; At*rhd6* mutants usually develop few, short root hairs, but wild type root hair growth is restored by treatment with exogenous auxin or the ethylene precursor ACC (Masucci and Schiefelbein 1994). A transcriptomic study has found a set of genes whose expression is regulated by both auxin and ethylene, among them At*LRL3* (Bruex et al. 2012).

As well as their roles in the early stages of fate determination, auxin and ethylene also influence the downstream development of root hairs. In *Arabidopsis* auxin gradients generated by auxin influx carriers are responsible for limiting the position of root hair initiation to the region of the cell closest to the root tip, since root hairs develop in more shootward parts of the cell in mutants with defective auxin signalling (Pitts et al. 1998; Grebe et al. 2002; Ikeda et al. 2009). Downstream of this, auxin and ethylene positively regulate root hair elongation. The root hairs of auxin and ethylene signalling mutants are shorter than wild type (Pitts et al. 1998; Knox et al. 2003), while treatment with auxin or ACC increases root hair length (Pitts et al. 1998; Rahman et al. 2002). Auxin exerts its effect at least partly by influencing the expression of the gene encoding the Class II bHLH transcription factor AtRSL4, which determines cell size by controlling the expression of genes whose products are involved in tip growth (Yi et al. 2010; Datta et al. 2015).

Among the liverworts, auxin promotes the development of rhizoids in *Marchantia*. Plants that are treated with auxins initiate more rhizoids, including in ectopic locations, and rhizoid length is increased (Tarén 1958; Maravolo and Voth 1966; Maravolo 1980; Ishizaki et al. 2012; Eklund et al. 2015; Flores-Sandoval, Eklund, et al. 2015). The role of auxin in promoting rooting cell differentiation and elongation is thus conserved between flowering plants and the earliest diverging land plant lineages.

1.6.4 The selection of the site of outgrowth and its maintenance requires Rop GTPases, reactive oxygen species production and ion influx

Once a cell has become committed to initiate root hair outgrowth, the site where this outgrowth will occur must be determined, as tip growth requires the focused delivery of new cell material to a restricted area. An important part in delineating the region of growth is played by Rho of plant (Rop) GTPases, which act as molecular switches that define localized and polarized regions of the plasma membrane, cycling between an active, GTP-bound form and an inactive GDP-bound form. During the initiation of the *Arabidopsis* root hair the site of future outgrowth is marked by the accumulation in the membrane of AtROP2, AtROP4 and AtROP6, which remain localized at the tip during elongation (Molendijk et al. 2001; Jones et al. 2002). Plants that overexpress At*ROP2* form extra and misplaced root hairs, while plants that overexpress At*ROP4* and At*ROP6* possess swollen, non-elongating root hair cells, suggesting a role for these genes in site selection, and later in the maintenance of the site of growth (Molendijk et al. 2001; Jones et al. 2002).

The activity of Rop GTPases is regulated by guanine nucleotide exchange factors (RhoGEFs) and GTPase-activating proteins (RhoGAPs). RhoGAPs increase the GTPase activity of Rops, causing them to hydrolyze the bound GTP and become inactive, while RhoGEFs promote the exchange of GDP for GTP, returning Rops to their active form. The localization and activity of AtROP2 is restricted and controlled by the Rho GTPase GDP dissociation inhibitor encoded by At*SUPERCENTIPEDE1* (At*SCN1*); in At*scn1* loss-of-function mutants AtROP2 accumulates in ectopic sites, where extra outgrowths initiate from the same cell (Carol et al. 2005). In turn, AtROP2 regulates the localization and activity of the NADPH oxidase encoded by At*ROOT HAIR DEFECTIVE 2* (At*RHD2*) (Jones et al. 2007; Takeda et al. 2008). Reactive oxygen species (ROS) produced by AtRHD2 activate hyperpolerization-activated Ca^{2+} channels, causing an influx of Ca^{2+} (Foreman et al. 2003). Growing root hairs, and other tip-growing cells, form a Ca^{2+} gradient focused at the tip (Reiss and Herth 1979; Schiefelbein et al. 1992), and the influx of Ca^{2+} is essential for tip growth, since pharmacologically blocking the channels stops the influx and arrests tip growth (Schiefelbein et al. 1992). Through calcium and other effectors such as Rop-interactive CRIB motif-containing proteins (RICs), Rop GTPases regulate the cytoskeleton and other downstream factors (Craddock et al. 2012). The activity of Rop GTPases is thus vital both for defining the initial site of root hair outgrowth and the maintenance of a focused region of growth.

1.6.5 Cell wall loosening causes a bulge to form at the site of outgrowth

During the initiation phase a bulge begins to form in a small area of the cell surface. Localized acidification of the cell wall accompanies the formation of this bulge, and is necessary for root hair initiation (Bibikova et al. 1998). It is thought that this drop in pH may loosen the cell wall by activating expansins, which transiently disturb non-convalent interactions between components of the cell wall. In *Arabidopsis* the expansin genes *AtEXP7* and *AtEXP18* are expressed specifically in root hair cells starting just before the appearance of bulges (Cho and Cosgrove 2002). Patches of high xyloglucan endotransglycosylase activity, which is dependent on ethylene and local acidification, appear at the site of future outgrowth before and during the formation of the bulge, and by cleaving and joining xyloglucans these enzymes also modulate cell wall loosening (Vissenberg et al. 2001). Taken together these data indicate that a loosening of the cell wall accompanies the formation of a bulge on the outer face of the root hair bearing cell.

1.6.6 Actin microfilaments are required for normal tip growth

The cytoskeleton is integral to most growth processes in plants. For instance, microtubules influence cell growth anisotropy by directing cellulose synthase complexes and so the orientation of cellulose bundles in the cell wall (Wasteneys and Ambrose 2009), while the actin cytoskeleton is required for cytoplasmic streaming and guiding vesicles during exocytosis (Hussey et al. 2006).The actin cytoskeleton in particular is very important in root hair development. A dense mesh of F-actin forms in the bulge before and during outgrowth, which is maintained in the apex of the root hair throughout elongation (Baluška et al. 2000; Kiefer et al. 2015). The correct positioning of this F-actin patch, and so the site of root hair outgrowth, depends on an interaction between AtACTIN7 and AtACTIN-INTERACTING PROTEIN 1-2 (Kiefer et al. 2015). The importance of actin microfilaments at the site selection stage is reinforced by the phenotype of At*actin2* (At*act2*) mutants, which initiate root hairs closer to the shootwards end of the cell than the wild type (Ringli et al. 2002). The actin cap is required for tip growth, as treatment with latrunculin B, which inhibits actin polymerization, halts root hair growth, while At*act2* mutants develop very short root hairs or none at all (Baluška et al. 2000; Ringli et al. 2002). The actin cytoskeleton is also vital for the development of rhizoids and other tip-growing cells in *Physcomitrella.* PpAIP1 controls actin turnover by mediating the severing of actin bundles, and Pp*aip1* mutants completely lack rhizoids and caulonemata, another rapidly tip-growing cell type (Augustine et al. 2011). Similarly, silencing or deleting subunits of the Arp2/3 complex, which nucleates new actin filaments and regulates branching in the actin network, abolishes the development of rhizoids and caulonemata (Harries et al. 2005; Perroud and Quatrano 2006; Finka et al. 2008). In *Marchantia*, as in *Arabidopsis* root hairs and *Physcomitrella* rhizoids, there is a dense actin network near the apex of the growing rhizoid, but its importance has not been experimentally tested (Era et al. 2009). Taken together these data demonstrate that tip growth in root hairs and moss rhizoids crucially depends on the functions of the actin cytoskeleton.

1.6.7 Rapid elongation requires the delivery of material to the growing tip

The high rate of tip growth during the elongation phase requires the rapid deposition of a large amount of new cell surface at the tip, which is delivered by secretory vesicles. The trafficking of vesicles containing the material required for growth to the correct site depends on the coordinated activity of numerous genes.

New cell wall matrix must be deposited at the tip to maintain the integrity of the cell as it grows. Cellulose microfibrils are vital to the mechanical strength of plant cell walls, and the primary cell wall of the growing tip contains short, randomly oriented microfibrils (Akkerman et al. 2012). The cellulose synthases AtCESA3 and AtCESA6 are required for cellulose microfibril deposition in the primary cell wall in diffuse growth, yet tip-growing root hairs are able to develop in At*cesa1*, At*cesa3* and At*cesa6* mutants (Park et al. 2011), and the proteins encoded by these genes do not localize to the plasma membrane of the tip. This indicates that one or more other proteins are responsible for the deposition of cellulose in the tip. Unlike the cellulose synthase mutants, the root hairs of mutants in *AtCELLULOSE SYNTHASE-LIKE D 3*, also known as At*kojak*, burst at the tip soon after initiation, suggesting that the mechanical integrity of the tip is compromised (Favery et al. 2001). AtCSLD3 is localized to root hair tips, and a chimaeric protein in which the catalytic domain of AtCSLD3 was replaced with that of AtCESA6 was able to complement the At*csld3* tip-bursting defect (Park et al. 2011). This indicates that the deposition of cellulose at the growing tip by AtCSLD3 is crucial for maintaining the integrity of the wall. Cellulose microfibrils are embedded in a matrix of other polysaccharides, such as pectin and hemicellulose. The production of these components is also vital for building the growing wall; for instance, mutants in genes required for the biosynthesis of the hemicellulose xyloglucan develop short, bulging root hairs (Cavalier et al. 2008; Pena et al. 2012). Unlike cellulose, the other polysaccharide components of the cell wall matrix are produced in the Golgi, and so their delivery to the surface of the growing tip relies on the trafficking of vesicles (Scheller and Ulvskov 2010; Driouich et al. 2012).

The exocytosis of secretory vesicles at the apex delivers cell wall material, membranes and proteins that facilitate root hair elongation. Since the fusion of these vesicles delivers much more membrane than is required, the excess membrane is recycled through endocytosis (Ketelaar et al. 2008). In order for vesicles to be delivered to the appropriate location, their identity, and that of the target, needs to be marked so that they can be trafficked accordingly. Numerous factors are involved in conferring identity to membrane compartments and directing fusion events, such as members of the Rab GTPase and SNARE protein families, and the lipid composition of the membrane of the compartment (Lipka et al. 2007; Nielsen et al. 2008; Ischebeck et al. 2010; Ovecka et al. 2010; Kim and Brandizzi 2014). These play important roles in root hair elongation; for instance, AtRabA4b and AtPHOSPHATIDYLINOSITOL 4-KINASE III BETA1 (AtPI4Kβ1), which produces the membrane lipid phosphotidylinositol 4 phosphate, localize to the membrane at the tip of growing root hairs, and At*pi4kβ1/2* double mutants develop short, misshapen root hairs (Preuss et al. 2006). Once vesicles carrying cell surface material reach the tip, they must fuse with it to deliver their cargo. This requires the activity of the exocyst, a protein complex that tethers the vesicle to the plasma membrane. The importance of the exocyst to root hair elongation is demonstrated by the exocyst subunit mutants At*exo70A1-1* and At*exo70A1-2,* which both develop short root hairs (Synek et al. 2006).

The trafficking of vesicles takes place along the actin cytoskeleton, and relies mainly on the activity of myosin motor proteins (Prokhnevsky et al. 2008). In *Arabidopsis*, the class XI myosin mutants At*xi-k* and At*xi-2* both develop short root hairs that have defects in organelle trafficking (Peremyslov et al. 2008). AtXI-K associates with vesicles that traffic along F-actin and accumulate at the tip of growing hairs, via the activity of specific AtXI-K receptors, AtMYOSIN BINDING 1/2 (Peremyslov et al. 2012, 2013). In *Physcomitrella*, silencing both Pp*XI* genes abolishes tip growth in the protonema, which is transformed into small spherical cells (Vidali et al. 2010). Class VIII myosins are not required for tip growth in *Physcomitrella*, but are involved in processes such as protonemal patterning and branching (Wu et al. 2011). These results indicate that tip growth relies on the delivery of vesicles to the tip by class XI myosins.

1.7 The genetic control of rhizoid development in early-diverging land plants

The finding that *RSL* genes control the development of moss rhizoids and angiosperm root hairs suggests that an ancient gene network controls the development of tip-growing filamentous cells at the interface between the plant and the soil, the *RSL* component of which originated before the last common ancestor of *Physcomitrella* and *Arabidopsis*. Other conserved members of this network include the *LRL* genes and auxin signalling (Figure 1.3) (Menand et al. 2007; Tam et al. 2015), as well as several later-acting genes, such as AIP and myosin XI (Vidali et al. 2010; Augustine et al. 2011). We know that this mechanism is ancient, but how ancient is it? Was the development of filamentous rooting cells at the plant–soil interface an evolutionary novelty acquired by the ancestor of all land plants during the conquest of the land? Or was the mechanism already present in plants' algal ancestors, a pre-existing developmental network that facilitated the transition to the terrestrial environment? Characterizing the wider regulatory network controlling rhizoid development in an early-diverging lineage would provide insight into how this ancient gene network evolved during the radiation of plants on land. Addressing this question will require studies in species that diverged from the other land plants earlier than *Physcomitrella*, such as liverworts and the algal relatives of land plants.

1.8 *Marchantia polymorpha* as a genetic model for early-diverging land plants

1.8.1 *Marchantia* is amenable for forwards genetics

Physcomitrella can be a useful model system for reverse genetics because of the ability to perform targeted gene replacement and the published genome, but it presents some difficulties for forward genetics approaches that would allow the unbiased identification of genes that are important in rhizoid development. *Agrobacterium*-mediated insertional mutagenesis has been a successful strategy in *Arabidopsis* for generating mutations whose genomic location can be rapidly identified by sequencing approaches (for instance, Alonso et al. (2003)), but efficient transformation using *Agrobacterium* is not available in *Physcomitrella*.

Marchantia polymorpha is an attractive alternative model system in which to carry out investigations that allow inferences to be drawn about events early in land plant evolution. Of foremost importance is its position at the base of the land plant phylogeny; according to most phylogenies the last common ancestor of *Marchantia* and other model species such as *Arabidopsis* was the last common ancestor of all land plants, and so findings in *Marchantia* can be used to draw inferences about the earliest stages of land plant evolution. *Marchantia* is emerging as a model liverwort and may offer a better prospect than *Physcomitrella* for forward genetic approaches. It has several features that make it especially suitable as the subject of a genetic screen. The key advantage is the possibility of efficient *Agrobacterium*-mediated transformation of young plants derived from spores, allowing large numbers of transformants to be produced, each with T-DNA insertions at more-or-less random positions in the genome (Ishizaki et al. 2008). This makes insertional mutagenesis to produce mutations that are tagged with a known DNA sequence feasible, which allows rapid identification of the site of the mutation in the mutants identified. The *Marchantia* genome is relatively small, at about 280 Mbp (Okada et al. 2000). It also seems that its genome is relatively streamlined; for instance *Arabidopsis* possesses 600 receptor-like kinases, while only 29 have been found in

Marchantia (Sasaki et al. 2007). This means that redundancy may be less of a problem than in other species. Furthermore, the fact that the main phase of the life cycle is haploid (Figure 1.4) means that the effects of mutations are visible directly in the transformed generation. Compared to *Physcomitrella* and *Arabidopsis,* crossing is relatively simple in *Marchantia*, which facilitates genetic analyses. The species is dioecious, with sex determined by the possession of an X or Y chromosome (Bischler 1986). Far-red light induces the development of the reproductive structures, and crossing is achieved by transferring sperm-containing water from the antheridiophore of one plant to the archegoniophore of another (Chiyoda et al. 2008) (Figure 1.4). Isogenic populations of the same age can also be rapidly produced from gemmae, asexual propagules that are produced in large numbers by mature plants (Figure 1.5)

Other genetic tools recently developed in *Marchantia* include gene targeting by homologous recombination (Ishizaki, Johzuka-Hisatomi, et al. 2013), CRISPR/Cas9-mediated genome editing (Sugano et al. 2014), and suppression of target transcripts by artificial microRNAs (Flores-Sandoval, Dierschke, et al. 2015). Because of this growing molecular toolkit, *Marchantia* provides a feasible genetic model in which to investigate how the development of rooting cells is controlled in the earliest-diverging lineage of land plants.

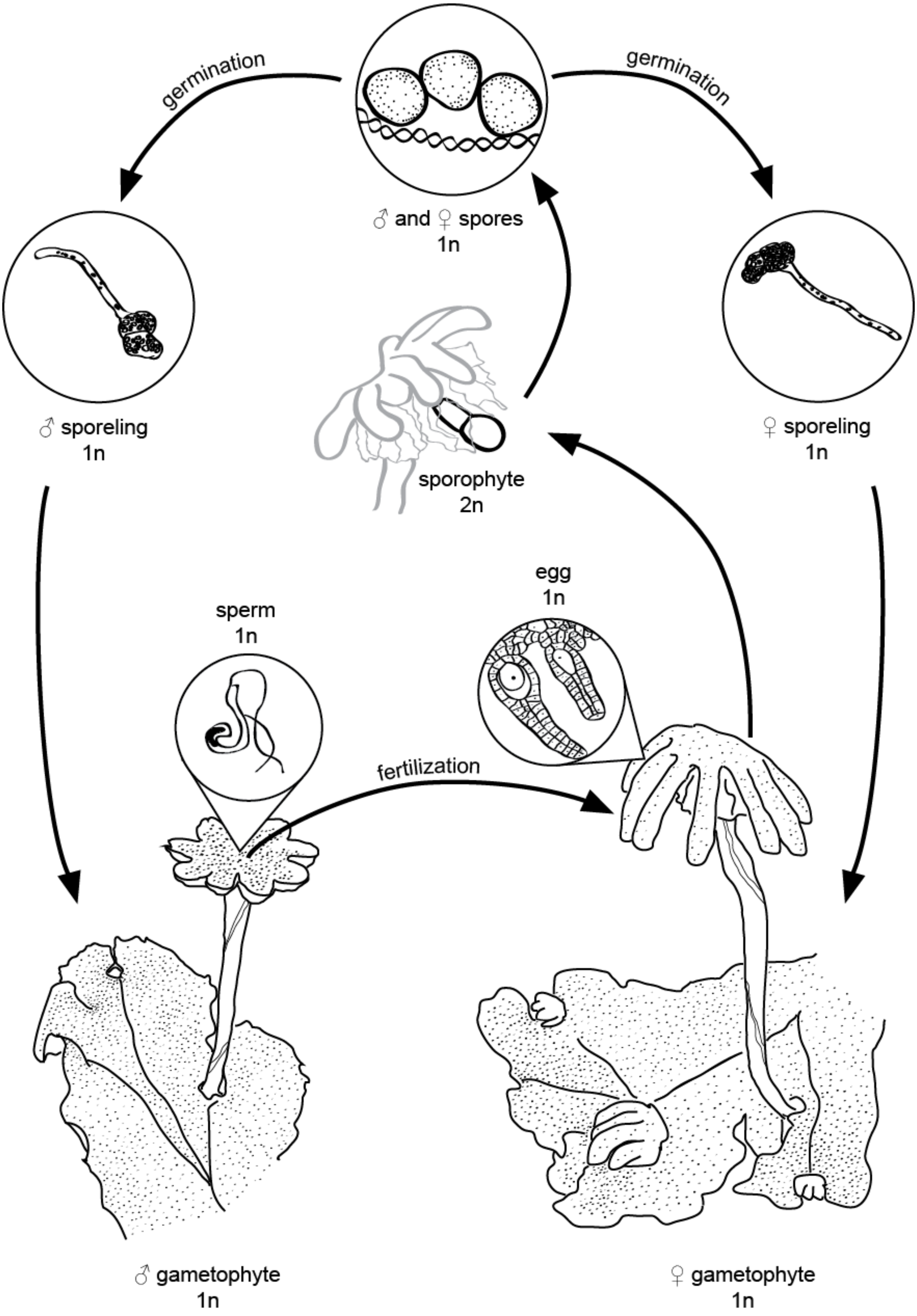


Figure 1.4 Sexual reproduction in *Marchantia polymorpha*.

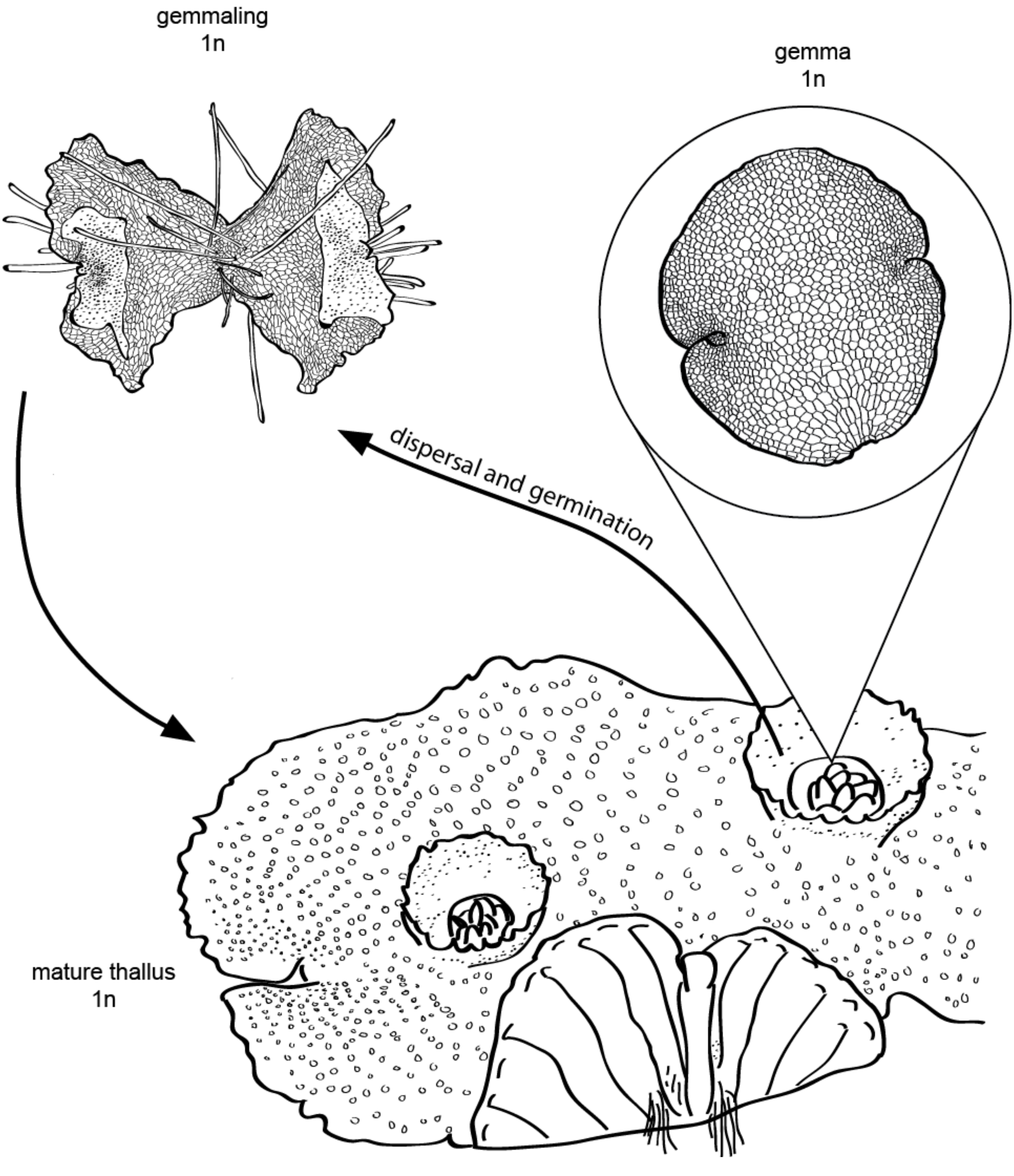


Figure 1.5 Asexual reproduction through gemmae in *Marchantia polymorpha.*

1.8.2 The anatomy and life cycle of *Marchantia*

Marchantia polymorpha is a complex thalloid liverwort. Growth occurs from the apical notch, and the thallus branches by bifurcation of the notch (Flores-Sandoval, Eklund, et al. 2015). The thallus is differentiated dorsiventrally, with rhizoids and membranous scales produced on the ventral (lower) epidermis (Figure 1.6 a) and photosynthetic air chambers and gemma cups on the dorsal (upper) surface (Figure 1.6 b). The rhizoids produced on the ventral surface are of two kinds (Mirbel 1835; Crandall-Stotler and Stotler 2008; Cao et al. 2014). Smooth rhizoids are initiated in distinct patches next to the midrib of the thallus and grow perpendicularly from the thallus into the soil. Tuberculate, or pegged, rhizoids are arise from all across the ventral epidermis, lie appressed to the epidermis and run together along the midrib (Figure 1.2 a). Rhizoids rarely develop adjacent to each other, normally being surrounded by a rosette of non-rhizoid cells (Figure 1.2 c) (Cao et al. 2014). The air chambers of the dorsal surface each consist of a schizogenous intercellular space filled with filaments of photosynthetic cells, which communicates with the atmosphere outside by means of an air pore (Mirbel 1835; Apostolakos and Galatis 1985). The air pore usually consists of 16 cells arranged in a stack of 4 layers, each made up of 4 cells (Figure 1.2 d). (Mirbel 1835; Kny 1890). No rhizoids develop on the dorsal epidermis (Figure 1.2 b).

The dominant phase of the life cycle of *Marchantia* is haploid. Both sexual (Figure 1.4) and asexual reproduction (Figure 1.5) are found in *Marchantia,* which, being dioecious, has separate male and female individuals. Mature plants produce aerial reproductive structures called gametangiophores, which are modified thallus branches known as antheridiophores in males and archegoniophores in females. Antheridia develop on the antheridiophore, and produce biflagellate motile sperm, while the archegoniophores produce egg-cell-containing archegonia. Sperm are transferred by water from the antherdiophore of a male plant to the archegoniophore of a female, where they fertilize the eggs within the archegonia. The diploid

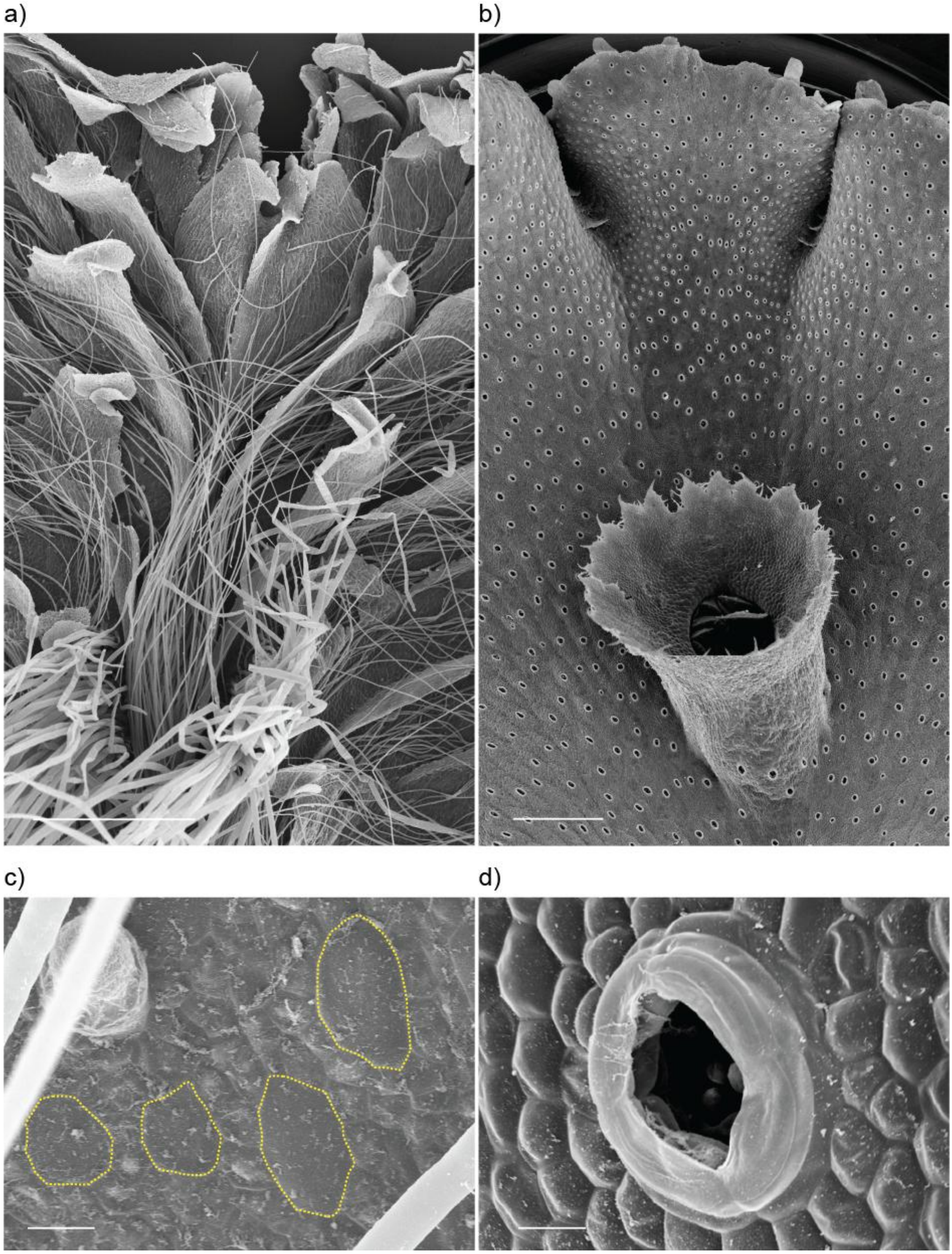


Figure 1.6 Epidermal features of *Marchantia polymorpha*. a) Mature thallus, ventral view. Scales, tuberculate rhizoids (narrow, darker) and smooth rhizoids (wider, lighter) are produced on the ventral side. Scale 1 mm, apex at top. b) Mature thallus, dorsal view. Air pores and gemma cups are produced on the dorsal side. Scale 1 mm, apex at top. c) Detail of rhizoid patch. Cells that will grow out into rhizoids (outlined) are separated by non-rhizoid cells. Scale 20 µm. d) Detail of air pore. Scale 20 µm.

phase is ephemeral; the zygote develops into a small and simple sporophyte, the cells of which undergo meiosis to produce haploid spores carrying either an X or Y chromosome. These spores germinate and go on to develop into mature thalli.

Asexual reproduction is achieved by the formation of propagules known as gemmae in structures called gemma cups on the dorsal surface of the thallus. A gemma is a lenticular disc of cells that originates from a single progenitor cell (Cavers 1903). The gemmae remain dormant inside the gemma cup, and this dormancy is promoted by auxin (Tarén 1958; Eklund et al. 2015). The gemma possesses two meristematic regions, from which growth is initiated after dispersal (Tarén 1958). In the inner region of each face of the gemma are large, pale cells that grow out into rhizoids once the gemma begins to grow (Tarén 1958; Völker and Halbsguth 1966; Eklund et al. 2015). Initially, rhizoids develop from both surfaces of the gemma, but within 24 hours dorsiventrality is established under the influence of light and gravity in an auxin-dependent manner (Fitting 1936, 1939; Flores-Sandoval, Eklund, et al. 2015), and subsequently rhizoids only develop from the ventral surface.

1.9 Approach and objectives of thesis

Forward genetic screens are an indispensible tool for identifying genes that underpin biological processes. At the most basic level all that is required for a mutant screen is a method of introducing genetic variation into a population of the organism of interest, and a way of detecting perturbations in the biological process under study. After mutants have been identified genetic techniques can be used to classify them into allelic groups and uncover epistatic relationships, while the location of the mutation can be identified by various means. In this way an understanding of the genes that are involved in a biological process can be built up, without any initial assumptions about what genes may be important. Thus it is a particularly powerful approach for uncovering novel genes involved in a previously-unstudied process, or in species that are distantly related to well-studied model organisms, in a manner that is not biased by focussing on candidate genes chosen because of knowledge of their functions in other processes or species.

Mutant screens have been carried out to investigate a number of processes in *Marchantia*. An early study used X-rays to mutagenize female *Marchantia* plants, resulting in the isolation of a monoecious plant and one that was phenotypically male, but which produced immobile spermatozoa (Lorbeer 1938). X-ray mutagenesis was later used to identify nutritionally-deficient mutants with defects in the biosynthesis of various amino acids (Miller et al. 1962). With the advent of molecular tools for *Marchantia*, insertional mutagenesis has been used in mutant screens. Particle bombardment was the first method of transformation to be developed, and in a screen of 2,100 transformants produced by this method a mutant was identified that constitutively produces sexual organs even in the absence of far-red light (Yamaoka et al. 2004), though the identity of the mutated gene has yet to be reported. More recently, T-DNA transformation mediated by *Agrobacterium tumefaciens* has permitted the generation of larger mutagenised populations, and a screen of 10,000 transformants identified the E3 ubiquitin ligase Mp*NOPPERABO1* as necessary for air chamber development (Ishizaki, Mizutani, et al. 2013).

In the research programme presented in this thesis, my goal was to investigate the network of genes that control the development of rhizoids in *Marchantia*, and so to enable inferences to be made about the evolution of the mechanism controlling the rooting cell developmental programme of land plants. Nothing was known of the genes that control the development of *Marchantia* rhizoids, and a forward genetic screen represented a powerful tool for identifying genes that are important in this process. Before the establishment of *Marchantia* as a model species with molecular tools, a large forward genetic screen was not practical. However, the development of efficient *Agrobacterium*-mediated transformation has provided the opportunity to mutagenize substantial populations of *Marchantia* using T-DNA, producing insertional mutations that can be rapidly mapped. My approach was to use this technique as the basis of a large mutagenesis and screen to select *Marchantia* mutants with defects in rhizoid development, and from this resource to identify genes that influence rhizoid differentiation and growth. This was to be followed by the molecular characterization of potential regulators among the identified genes, in particular those likely to regulate the early stages of rhizoid fate determination, with the aim of beginning to determine the gene network that controls the development of *Marchantia* rhizoids.

Chapter 2: Materials and methods

2.1 Plant and bacterial materials

The wild-type *Marchantia polymorpha* used were male accession Takaragaike-1 (Tak-1) and female accession Takaragaike-2 (Tak-2), described in Okada et al. (2000) and Ishizaki et al. (2008). These were a gift of Kimitsune Ishizaki (Kobe University, Japan). *Agrobacterium tumefaciens* GV3101 and *Escherichia coli* XL-1 Blue were used in all cloning and transformations.

2.2 Plant growth conditions

M. polymorpha was grown in axenic culture in Petri dishes containing modified Johnson's medium (Johnson et al. 1957) (6 mM KNO_3 (Sigma); 0.5mM $MgSO_4$ (Sigma); 4 mM $Ca(NO_3)_2$ (Sigma); 25 μM KCl (Sigma); 10 μM H_3BO_3 (Fisher Scientific); 1 μM $MnSO_4$ (Fisher Scientific); 1 μM $ZnSO_4$ (Sigma); 0.25 μM $CuSO_4$ (Sigma); 0.25 μM $(NH_4)_6Mo_7O_{24}$ (Sigma); 25 μM $FeSO_4$ (BDH Laboratory Supplies, UK); 25 μM Na2-EDTA (Sigma); 0.6 mM $NH_4H_2PO_4$ (Sigma); 0.4 mM $(NH_4)_2SO_4$ (Sigma); 2.7 mM 2-(*N*-morpholino)ethanesulfonic acid (Sigma); 29 mM sucrose (Sigma); pH adjusted to 5.8 with KOH; solidified with 1.4% Agar (Sigma, Cat. No. A9799) or 0.8% phytagel (Sigma). The dishes were placed horizontal in a Sanyo growth cabinet at 18°C under continuous white light (56 μmol photons m^{-2} s^{-1}).

Plants to be used for sexual reproduction were grown on a 1:3 mix of vermiculite and John Innes No. 2 compost (Arthur Bower's) at 18°C under a 16h light: 8h dark photoperiod with a white light intensity of 50 μmol photons m^{-2} s^{-1}, supplemented with far red light (Phillips GreenPower LED HF module, 740 nm peak emission, 6 μmol photons m^{-2} s^{-1}).

2.3 RNA extraction and cDNA synthesis

Three entire gemmalings, or approximately 1 cm^2 mature thallus tissue, were frozen and ground to powder in liquid nitrogen and 700 µl Trizol ® was added. The samples were mixed thoroughly and incubated at 56°C for 10 minutes, then centrifuged at 13,000 rpm for 1 minute. The supernatant was transferred to a new microcentrifuge tube and 700 µl ethanol was added and mixed. From this point, RNA isolation proceeded using the Direct-zol RNA MINIprep kit (Zymo Research) according to the manufacturer's instructions. RNA was eluted in 40 µl nuclease-free water.

RNA concentration was measured using a NanoDrop ND1000 spectrophotometer (Thermo Fisher Scientific). The concentration of RNA in each sample was equalized, and contaminating genomic DNA was digested using the Turbo DNA-free kit (Ambion, Life Technologies) according to the manufacturer's instructions. cDNA was synthesised using ProtoScript II reverse transcriptase (New England Biolabs) according to the manufacturer's instructions, with oligo(dT) primer.

2.4 RT-qPCR

RT-qPCR analysis was carried out using the SensiFAST SYBR Hi-ROX Kit (Bioline) and the Applied Biosystems 7300 Real Time PCR system. Three technical replicates were used for each sample. Mp*ACT*, Mp*CUL*, Mp*EF1*, Mp*ELF5* and Mp*APT* were chosen as reference genes, on the basis of an analysis of *M. polymorpha* reference genes suitable for RT-qPCR (Saint-Marcoux et al. 2015). cDNA of genes of interest was detected using the following primer pairs: Mp*WIP* - VJ_66, VJ_67; Mp*RSL1*- MpRSL1 F, MpRSL1 R; Mp*ACT* – VJ_213, VJ_214; Mp*APT* – VJ_215, VJ_216; Mp*CUL* – MpCUL F, MpCUL R.; Mp*EF1α* – MpEF1α F, MpEF1α R; Mp*ELF5* – MpELF5 F, MpELF5 R. See Appendix 1 for primer sequences. The cycling conditions were 95°C for 2 min, then 40 cycles of 95°C for 15s, 60°C for 15s, 72°C for 1 min, with detection during the 72°C step. At the end of the cycles, a dissociation curve analysis was performed to confirm that only a single amplicon had been formed. The bundled 7000 System SDS Software (Applied Biosystems) was used to export the ΔRn for analysis using LinRegPCR v2012.0 (Ruijter et al. 2009). In LinRegPCR, the window of linearity was determined for each amplicon and the mean reaction efficiency calculated for each reaction. From this an N_0 was computed, which reflects the initial amount of template included in the reaction. The geometrical mean of the N_0 of the technical replicates of each biological repeat was calculated.

2.5 Microscopy

2.5.1 Stereomicroscopy

Images were obtained of plants growing in axenic culture using a Leica M165FC stereomicroscope with a Leica M series Plan APO 1.0x objective (Leica, Gemany). Images were captured using a Leica DFC310 FX camera. Epifluorescent imaging of YFP was carried out with UV excitation and the Leica YFP ET filter set.

2.5.2 Confocal laser scanning microscopy

Plants were grown from gemmae on modified Johnson’s medium in 50 mm deep petri dishes (Fisher Scientific). Cell walls were stained by immersing plants in 15 µM propidium iodide solution for 15 minutes, which was then removed and the plants submerged in water. Images were acquired with a Leica SP5 confocal microscope and the Leica Application Suite (LAS) software, using either a Leica HCX APO L20x/0.50 W U-V-I dipping lens, for low-magnification images, or a Leica HCX APO 40x/0.80 W U-V-I dipping lens, for higher-magnification images. Images were collected with sequential scans. YFP and chlorophyll fluorescence was excited at 514 nm using an argon laser and emission was measured between 524 and 568 nm (YFP) and between 679 and 798 nm (chlorophyll) using an Acousto Optical

Tunable Filter. PI was excited at 543 nm using a helium-neon laser and emission was measured between 568 and 659 nm. The images were processed using FIJI software to create brightest-point 3d projections (Schindelin et al. 2012).

2.5.3 Scanning electron microscopy

Samples were fixed by immersion in 10 ml methanol (Sigma) dried over 3 Å molecular sieves (Sigma) for 1 hour, then the methanol was replaced with 10 ml ethanol (Sigma), similarly dried. This was repeated two more times, leaving 1 hour between each change. The samples were dried at the critical point using a Tousimis Autosamdri-815, then mounted on aluminium stubs (Agar Scientific) using double-sided carbon adhesive discs (Agar Scientific), and coated with a mixture of gold and palladium using an SC7640 sputter coater (Quorum Technologies). The samples were imaged immediately with a JSM-5510 scanning electron microscope (JEOL).

2.6 Transformation of *M. polymorpha* sporelings

The method of transformation was based on that presented by Ishizaki and coworkers (2008). Intact wild-type sporangia obtained from crosses between Tak-1 and Tak-2 plants were surfaced sterilized with 45 µM sodium dichloroisocyanurate (Sigma) for 5 minutes, then rinsed 3 times with sterile water. Two sterilized sporangia were broken open in 1ml sterile water, added to a 125ml vented tissue culture flask containing 25 ml M51C medium (19.8 µM KNO_3 (Sigma); 5 µM NH_4NO_3 (Sigma); 1.5 µM $MgSO_4$ (Fisher Scientific); 2 µM $CaCl_2$ (Sigma); 2 µM KH_2PO_4 (Sigma); 110 nM EDTA-Fe(III) (Sigma); 1 µM Na_2MoO_4 (Sigma) ; 100 nM $CuSO_4$ (Sigma); 100 nM $CoCl_2$ (Sigma); 7 µM $ZnSO_4$ (Sigma); 40 µM $MnSO_4$ (Fisher Scientific); 49 µM H_3BO_3 (Fisher Scientific); 555 µM *myo*-inositol (Sigma); 8 µM nicotinic acid (Sigma); 5 µM pyridoxine hydrochloride (Sigma); 30 µM thiamine hydrochloride (Sigma); 58 mM sucrose (Sigma); 2 mM L-glutamine (Sigma); 0.1% w/v casein hydrolysate (Fluka Analytical); pH 5.8), and cultured without shaking at 18°C under continuous white light (56 µmol photons m^{-2} s^{-1}) for 7 days.

A single colony of *Agrobacterium tumefaciens* strain GV3101 carrying the vector of interest was inoculated into 5 ml LB medium containing 25 µg/ml gentamycin, 100 µg/ml rifampicin, and 50 µg/ml kanamycin, and cultured with shaking at 27 °C for 1 d. The bacteria were sedimented by centrifugation at 2000 rpm for 15 m, the supernatant discarded, the bacteria resuspended in 10 ml M51C liquid medium containing 100 µM acetosyringone (Sigma Aldrich, 97%), and cultured for a further 6 h with shaking at 27 °C. The 7 d old cultures of *Marchantia* sporelings were then supplemented with 100 µM acetosyringone, and 1 ml of the bacterial culture was added. The sporelings and bacteria were co-cultivated with shaking at 120 rpm under continuous white light (56 µmol photons m^{-2} s^{-1}) at 18°C for 2 d. Sporelings were then collected on a sterile 40 µm cell strainer (Fisher Scientific) and rinsed with 100 ml sterile distilled water. The sporelings were then spread on antibiotic selection plates of Johnson’s agar with 100 µg/ml cefotaxime to kill *Agrobacterium* and 10 µg/ml hygromycin to select for transformed plants. After 14 d transformants could be distinguished from antibiotic-sensitive non-transformed plants, and were transferred, at regular spacing, onto 12 cm square petri dishes of Johnson’s agar without antibiotics for further growth. Other than in the mutant screen, where primary transformants were screened, chimaerism was avoided by establishing transformant lines from a single gemma of a primary transformant.

2.7 Identification of mutants with defective rhizoid development

Sporelings resulting from a cross between Tak-1 and Tak-2 were transformed by co-cultivation with *Agrobacterium* carrying the binary vector pCambia1300, which contains $_{pro}$*35s:HygR:Term*. After approximately 14 d selection on hygromycin-containing medium transformants were transferred to 12 cm square plates of antibiotic-free Johnson’s agar, evenly spaced 56 plants to a plate. Plants were screened for aberrant rhizoid development using a Leica M165FC stereomicroscope at approximately 7 d, 14 d and 21 d after transfer. The aberrant phenotypes screened for were few or no rhizoids, short rhizoids, misshapen rhizoids and development of rhizoids in ectopic locations. Plants thought to display defective rhizoid growth were moved to new plates, spaced more widely at 16 per plate to allow further growth and observation. Gemmae of plants considered to have defects in rhizoid development were grown alongside wild type gemmae to better assess the development of their rhizoids. Plants still considered to have defective rhizoid development were accepted to be taken forward to the next stages of investigation, *viz.* sexual reproduction and identification of the T-DNA insertion site.

2.8 Analysis of T-DNA copy number and linkage to mutant phenotype

Mutant plants were propagated on soil and crossed with the wild type of the opposite sex (Tak-1 or Tak-2). The sporangia resulting from the cross were sterilized with 45 μM sodium dichloroisocyanurate (Sigma) for 5 minutes, then rinsed 3 times with sterile water and broken to release the spores. The spores were spread on Johnson's agar and grown for approximately 14 d. Approximately 216 F1 plants were transferred to 12 cm square plates of Johnson's agar, evenly spaced 36 to a plate. After 14 d further growth individuals were scored for inheritance of the mutant or wild type phenotype.

To test for the number of T-DNA insertions and linkage between them and the mutant phenotype, hygromycin resistance was used as a marker for the presence of an insertion. If the mutant phenotype was inherited, a fragment of each F1 individual was transferred to Johnson's agar containing 10 μg/ml hygromycin and grown for 14 d. The number of plants in each of the following categories was recorded: mutant phenotype and hygromycin sensitive; mutant phenotype and hygromycin resistant; wild type phenotype and hygromycin sensitive;

wild type and hygromycin resistant. If the ratio of hygromycin resistant: hygromycin sensitive plants was not significantly greater than 1:1 (χ^2 test), it was inferred that there was a T-DNA insertion at only one locus in the genome. In lines with only one insertion, the T-DNA was inferred to be closely linked to the locus of the mutation responsible for the defect in rhizoid development if all F1 plants with the mutant phenotype were hygromycin resistant and all plants with the wild type phenotype were hygromycin sensitive.

2.9 DNA extraction

An approximately 1 cm^2 piece of thallus was frozen and ground to powder with liquid nitrogen and transferred to a microcentrifuge tube, to which was added 500 μl DNA extraction buffer (100 mM TRIS pH8; 1.4 M $NaCl_2$; 20 mM EDTA; 2% w/v CTAB; 1% w/v polyvinylpyrrolidone, 0.3% v/v β-mercaptoethanol), and incubated at 60°C for 30 minutes. 250 μl chloroform:isoamyl alcohol 24:1 was added, the sample was mixed by vortexing and centrifuged at 13,000 rpm for 1 minute, and the upper phase was transferred to a new microcentrifuge tube. This was repeated one more time, then 500 μl isopropanol was added and the sample centrifuged at 13,000 rpm for 15 minutes. The supernatant was removed, the pellet washed with 200 μl 70% ethanol, and centrifuged at 13,000 rpm for 2 minutes. The supernatant was removed and the pellet air-dried for 15 minutes, then dissolved in 70 μl water.

2.10 Location of T-DNA insertion sites by TAIL-PCR

Thermal asymmetric interlaced polymerase chain reaction (TAIL-PCR) was used to amplify genomic sequence flanking the right or left border of T-DNA insertions, using a set of nested primers specific to either the left (primers LB1, LB2, LB3) or right (primers RB1, RB2, RB3) borders, in combination with six different arbitrary degenerate primers (AD1-6). See

Appendix 1 for primer sequences. For each genotype 12 primer combinations were used, each of the two nested primer sets with each of the six AD primers. Three rounds of TAIL-PCR were carried out, the first using primer LB1 or RB1, the second using LB2 or RB2, and the third using LB3 or RB3. In each round the concentration of each primer was as follows: border-specific primer 0.5 µM; AD1-3 3µM; AD4 and AD6 4 µM; AD5 2 µM.

The first round of TAIL-PCR was carried out using Ultra Polymerase (PCR Biosystems) in a 20 µl reaction according to the manufacturer's instructions, with 25 ng genomic DNA as the template. The thermal cycling conditions were as follows: 1 m at 93 °C; 30 s at 95 °C, 1 m at 62 °C, 2 m 30 s at 72 °C, repeated 3 more times; 30 s at 94 °C; 3 m at 25 °C; 2 m 30 s at 72 °C with a 50% ramp rate; 30 s at 94 °C, 1 m at 68 °C, 2 m 30 s at 72 °C, 10 s at 94 °C, 1 m at 68 °C, 2 m 30 s at 72 °C, 10 s at 94 °C, 1 m at 44 °C, 2 m 30 s at 72 °C, repeated 13 more times; 5 m at 72 °C.

The second and third rounds were carried out using MangoTaq DNA polymerase (Bioline) in a 20 µl or 50 µl reaction, respectively, according to the manufacturer's instructions. A 96-pin replicator was used to transfer a small amount of reaction from the previous round as the template. The thermal cycling conditions for the second round were as follows: 2 m at 95 °C; 15 s at 95 °C, 1 m at 64 °C, 2 m 30 at 72 °C, 15 s at 95 °C, 1 m at 64 °C, 2 m 30 s at 72 °C, 15 s at 95 °C, 1 m at 44 °C, 2 m 30 at 72 °C, repeated 11 more times; 5 m at 72 °C. The thermal cycling conditions for the third round were as follows: 2 m at 95 °C; 15 s at 95 °C, 1 m at 44 °C, 2 m 30 s at 72°C, repeated 34 more times; 5 m at 72 °C.

The products of the second and third rounds were separated by electrophoresis through 1.2% agarose gel. Products of the third round that showed a reduction in size between the second and third rounds were cut from the gel and extracted using the QIAquick gel extraction kit (Qiagen) and sequenced by Sanger sequencing with either primer LB3 or RB3 by Source Biosciences, Oxford.

T-DNA flanking sequence obtained by TAIL-PCR was used to query a draft *M. polymorpha* genome (prepared by Giulia Morieri, Hélène Proust, Clément Champion, Sandy Hetherington and Steve Kelly, unpublished) using the blastn algorithm, to identify the site at which the T-DNA border sequence meets *M. polymorpha* genomic sequence. See Appendix 3 for the sequences used to locate the insertions.

2.11 Plasmid construction

Sequences of all primers used can be found in Appendix 1, and maps of plasmids in Appendix 2.

2.11.1 Constitutive expression of Mp*WIP*

Mp*WIP*, including the UTRs, was amplified from wild type gDNA using Phusion High-Fidelity DNA Polymerase (New England Biolabs) with primers VJ_81 and VJ_82 and recombined into the pCR8/GW/TOPO Gateway entry vector (Invitrogen) to make plasmid Vp24. To create a binary plasmid for plant transformation, an LR reaction was carried out between this plasmid and plasmid HB047 using Gateway LR Clonase II (Invitrogen), creating plasmid Vp26, which contains $_{pro}$*35S:*Mp*WIP:Term* and confers plant hygromycin resistance.

2.11.2 $_{pro}$*MpWIP* reporter contruct

4.7 kb of sequence 5’ to the start of the CDS of Mp*WIP* was amplified in 4 overlapping segments using Phusion High-Fidelity DNA Polymerase (New England Biolabs) with the primer pairs VJ_46 and VJ_141; VJ_186 and VJ_187; VJ_143 and VJ_144; and VJ_45 and VJ_188. These fragments were joined by overlap PCR using Phusion High-Fidelity DNA Polymerase (New England Biolabs) and primers VJ_45 and VJ_186, and subcloned into pGEM-T.

The In-Fusion HD Cloning Kit (Clontech Laboratories) was used to introduce this promoter into the binary vector Vp57, which confers hygromycin resistance for plant selection. The promoter was amplified with the primers VJ_190 and VJ_191, which add 16 bp of sequence homologous to the desired insertion site on either side, as well as a SacI site at the 3' terminus of the promoter, and the In-Fusion reaction was performed with Vp57 linearized with SmaI (New England Biolabs). The resulting plasmid was digested with SacI (New England Biolabs), dephosphorylated with Antarctic Phosphatase (New England Biolabs), and the Gateway Vector Conversion System (Thermo Fisher) was used to ligate GW Cassette C.1 in between the promoter and terminator, to generate plasmid Vp65. An LR reaction was carried out between plasmid Vp65 and HBp039 to create Vp73, containing $_{pro}$Mp*WIP:3xYFP-NLS:Term*.

2.11.3 Mp*WIP-SRDX* fusion construct

To generate a fusion between the EAR-motif repression domain (SRDX) and the C-terminus of Mp*WIP*, the Mp*WIP* CDS was amplified from plasmid Vp24 with Phusion High-Fidelity DNA Polymerase and the primers VJ_83 and VJ_200. VJ_200 replaces the STOP codon of *MpWIP* with sequence encoding the SRDX domain (LDLDLELRLGFA*). This product was recombined into the pCR8/GW/TOPO Gateway entry vector to create plasmid Vp66. An LR reaction was carried out between Vp66 and HBp047 to generate Vp80, which contains $_{pro}$*35S:*Mp*WIP-SRDX :Term* and confers plant hygromycin resistance.

2.11.4 Mp*WIP* artificial microRNA

The *MpmiR160* pre-miR backbone was used as the basis of amiR design (Flores-Sandoval, Dierschke, et al. 2015), with the endogenous *miR160* sequence replaced with 21 nt targeting the Mp*WIP* transcript. miRs were designed using the WMD3 software (http://wmd3.weigelworld.org/), with the minimum number of included targets set to 1, the accepted off-site targets set to 3, the full-length Mp*WIP* transcript as the target, and the transcript library set to Marchantia_polymorpha_v183.mRNA.PUT.fasta. The highest-ranked *amiR,* targeting the 3' UTR of *MpWIP,* was chosen. The amiR* was designed to have mismatches with the amiR sequence at positions 7, 13 and 18. These were then used to replace the native *miR160* miR and miR* sequences in the backbone. This backbone was bracketed by attB1 and attB2 sites and synthesized by Life Technologies to create the plasmid Vp40. This was recombined with pDONR221 using BP Clonase II (Invitrogen) to create an entry clone, which was recombined with plasmid HBp350 using LR Clonase II (Thermo Fisher) to generate the expression clone Vp46 (${}_{pro}$*OsACT:amiR*-Mp*WIP*$^{\mathrm{Mp}\mathit{miR160}}$*:Term*).

2.12 Phylogenetic analysis

M. polymorpha transcript sequences were obtained from a gametophyte transcriptome and genomic sequences were obtained from a *M. polymorpha* genome, both prepared from Tak-1 and Tak-2 accessions by Hélène Proust, Giulia Morieri, Sandy Hetherington, Clément Champion and Steve Kelly. The likely amino acid sequence of each gene identified in the screen as putatively involved in rhizoid development was found by using the transcript to query the *Arabidopsis thaliana* genome v. *TAIR10* using the tblastx algorithm to identify *Arabidopsis* sequences with conserved domains. The longest ORF on the transcript that encoded a protein containing these domains was designated as the CDS.

OrthoFinder (Emms and Kelly 2015) was used to identify homologues of the protein encoded by the CDS of each gene from *Chlamydomonas reinhardtii*, *Volvox carteri*, *M. polymorpha*, *Physcomitrella patens*, *Selaginella moellendorfii* and *A. thaliana*. The tblastn algorithm was used to search a transcriptome of *Coleochaete nitellarum* prepared by Holger Breuninger, Sandy Hetherington, Clément Champion and Steve Kelly. MAFFT v. 7 (Katoh and Standley 2013) was used to align these sequences, implementing the L-INS-i strategy. The alignments were manually trimmed to remove poorly-aligned regions. Maximum likelihood phylogenies were estimated with PhyML 3.0 (Guindon et al. 2010), using the LG substitution model and NNI tree improvement. Branch support was estimated using the aLR-T SH-like method.

Chapter 3: A mutant screen to identify genes involved in *Marchantia* rhizoid development

3.1 Abstract

To identify components of the gene regulatory network that controls rhizoid development in *Marchantia polymorpha*, I carried out a mutagenesis and screen to find mutants with defective rhizoid development. I created a population of 105,000 insertionally-mutagenised plants by transformation with T-DNA, screened these for defects in rhizoid development, and identified 61 lines with aberrant rhizoid development. In 14 of these lines, the mutant phenotype was inherited through a sexual generation and found to be closely linked to the T-DNA insertion. I located the sites of these insertions using TAIL-PCR, and so identified the genes putatively affected by the mutations. Some lines were unable to reproduce sexually, but by using TAIL-PCR to identify the site of the insertions in these lines, further alleles of already-identified genes were found. In total ten genes were identified that cause defects in rhizoid development when mutated. Among these were two putative regulators of rhizoid development, and eight genes that are likely to function further downstream, in tip growth. In this chapter I present the results of the mutant screen and the genes that are likely to be involved in rhizoid elongation. The homologues of several of these genes have previously been implicated in rooting-cell development in other species, while the role of others in rhizoid development is here presented for the first time.

3.2 Introduction

In order to identify components of the rhizoid developmental network in *Marchantia* I carried out an insertional mutagenesis screen for mutants with defects in rhizoid development. In this chapter I present the results of the screen and the genes I identified that likely act late in rhizoid development. In the subsequent chapters I focus on two putative early regulators of rhizoid development that I identified in the screen.

3.2.1 Lessons from *Arabidopsis* root hair development mutant screens

Arabidopsis root hairs are attractive as a model system for studying how distinct plant cell types are specified and subsequently develop, since they are easily accessible for observation, and because they develop in predictable positions on the root, deviations from normal fate specification can be readily detected (Schiefelbein 2000). Mutant screens in *Arabidopsis* have identified many genes that play a role in various stages of root hair development, from the patterning of root hair cells, through the determination of root hair fate, to the cell biology of root hair elongation. The genes operating during the different stages of root hair development may not be homologous to those that function in *Marchantia* rhizoid development. Nevertheless, the findings of mutant screens in *Arabidopsis* can be useful in forming hypotheses about what kinds of genes might be important in each step of *Marchantia* rhizoid development, and what sort of mutant phenotypes might be expected of plants with defects in each stage.

3.2.2 Patterning genes can mutate to give hairless or ectopic-root hair phenotypes

Marchantia rhizoids develop in specific parts of the plant, and rhizoid cells are usually surrounded by a rosette of non-rhizoid cells (Cao et al. 2014). Thus the earliest-acting components of the gene network that regulates the development of rhizoids will be involved in patterning which cells will adopt the rhizoid fate, and which will not.

In *Arabidopsis* root hair specification is position dependent, with the epidermis consisting of alternating files of root-hair-producing trichoblasts and hairless atrichoblasts. Epidermal cells that overlie a junction between cortical cells become trichoblasts, while ones that overlie a single cortical cell become atrichoblasts (Dolan et al. 1994). Loss-of-function mutants of the MYB transcription factor *CAPRICE* (At*CPC*) produce few root hairs, even in positions where they would develop in the wild type, a phenotype that is enhanced by mutations in the related genes *TRYPTICHON* (At*TRY*) and *ENHANCER OF TRY AND CPC1* (At*ETC1*) (Wada et al. 1997; Schellmann et al. 2002; Kirik et al. 2004; Simon et al. 2007). Conversely, mutants in the WD-40 gene *TRANSPARENT TESTA GLABRA 1* (At*TTG1*), the R2R3 MYB *WEREWOLF* (At*WER*) or the bHLH transcription factors *GLABRA3* (At*GL3*) and *ENHANCER OF GLABRA3* (At*EGL3*) develop root hairs in the files that would be atrichoblasts in the wild type (Galway et al. 1994; Masucci et al. 1996; Lee and Schiefelbein 1999; Bernhardt et al. 2003). These findings have been integrated into a lateral-inhibition model for the patterning of trichoblast and atrichoblast cells (see Figure 1.2) (Ueda et al. 2005; Savage et al. 2008; Schiefelbein et al. 2009).

Whether or not this specific mechanism operates in liverworts, these findings from *Arabidopsis* suggest that mutations in genes that control the patterning of rhizoid and non-rhizoid cells in *Marchantia* may be found amongst mutants that develop few or no rhizoids, or that develop rhizoids in ectopic locations.

3.2.3 Mutations in positive regulators of root hair fate reduce root hair number

Cells that are to develop into rhizoids will need to coordinate the induction, repression and modulation of the expression of numerous genes in order to differentiate and initiate tip growth. In *Arabidopsis* the bHLH transcription factor At*ROOT HAIR DEFECTIVE 6* (At*RHD6*) is a master regulator of root hair development that is expressed in trichoblasts and is under the control of the patterning genes described above (Menand et al. 2007). Mutants lacking functional At*RHD6* and its close relative At*ROOT HAIR DEFECTIVE SIX-LIKE 1* (At*RSL1*) develop no root hairs, even in the trichoblast position (Menand et al. 2007). The original At*rhd6* mutant was identified in a screen for abnormal root hairs, and develops few root hairs (Masucci and Schiefelbein 1994).

From this it can be hypothesized that, in *Marchantia*, positive regulators of rhizoid development that act downstream of patterning genes might mutate to give a phenotype in which few or no rhizoids develop in positions where they would in the wild type. Alternatively, there may be negative regulators of rhizoid fate that are active in cells that have been patterned as non-rhizoid; it would be expected that loss-of-function mutations in such genes would lead to the development of rhizoids in ectopic positions. Gain-of-function mutations could be expected to give the opposite phenotypes. A gain-of-function mutation in a positive regulator of rhizoid fate might cause the ectopic development of rhizoids, and in a negative regulator, to cause fewer or no rhizoids to develop. Therefore, mutants in which genes that control rhizoid or non-rhizoid differentiation are defective are likely to develop few or no rhizoids, or rhizoids in ectopic positions. Mutants in early-acting genes that determine rhizoid patterning or regulate rhizoid differentiation are thus likely to present similar phenotypes.

3.2.4 Mutations in genes that act in root hair initiation cause root hair development to arrest very early

Once a cell has adopted the rhizoid developmental fate, regulators of rhizoid development will begin to promote its differentiation and outgrowth. Three phases have been described in

Arabidopsis root hair development: the formation of a bulge at the site of outgrowth, an initial slow phase of tip growth, and then a transition to rapid elongation by tip growth (Dolan et al. 1994). Naturally, mutations in genes required for the earliest phases of root hair development cause root hair development to arrest very early. For example, the NADPH oxidase AtROOT HAIR DEFECTIVE 2 (AtRHD2) produces reactive oxygen species at the forming bulge and at the elongating root hair tip, and is necessary for the establishment of a Ca^{2+} gradient at the tip that is required for tip growth (Foreman et al. 2003). At*rhd2* mutants develop very short root hairs that arrest very soon after bulge formation (Foreman et al. 2003). It is to be expected that, in *Marchantia*, genes required for the early events of rhizoid outgrowth and transition to elongation when mutated will give plants in which rhizoid growth stops very soon after initiation, causing the development of very short rhizoids, or even small bulges where rhizoids should be.

3.2.5 Many kinds of genes mutate to a short root hair phenotype

In tip growing cells, a large number of cellular processes are involved in cell elongation. Accordingly, in *Arabidopsis* a very wide variety of genes, when mutated, condition a short root hair phenotype. These include any number of the many structural genes that have been found to have a role in the cell biology of tip growth, such as those involved in vesicle trafficking, signalling or cell wall biosynthesis, as reviewed in Chapter 1 of this thesis. Similarly, mutations in effector genes that act downstream of the high-level regulatory genes to carry out tip growth can also be expected to cause the development of short rhizoids in *Marchantia*.

A subgroup of short root hair mutants in *Arabidopsis* possess a distinctive phenotype in which the tips of the root hairs burst, and consists of mutants in genes involved in cell wall biosynthesis and the maintenance of cell wall integrity. For example, KOJAK (AtKJK) is a cellulose synthase-like protein that produces constituents of the cell wall, and the root hairs of *kjk* mutants burst soon after initiation because of the weakness of the cell wall (Favery et al. 2001; Wang et al. 2001). Bursting root hairs are also seen in At*rhd2* mutants, which are proposed to arise from defects in balancing the loosening and strengthening of the cell wall (Monshausen et al. 2007), and in At*feronia*, which encodes a receptor-like kinase that has been implicated in mechanosensing and the maintenance of cell wall integrity (Kessler et al. 2010; Rounds and Bezanilla 2013; Shih et al. 2014). *Marchantia* mutants that suffer from short, bursting rhizoids might therefore have mutations in genes that are important for the biosynthesis of the cell wall or the maintenance of its integrity.

Aside from these effector genes, mutants with short root hairs could also have mutations in regulatory genes that control the process of tip growth. The cell growth-regulating bHLH transcription factor At*RSL4* was identified by reverse genetics approaches, but the At*rsl4* mutant develops root hairs that are significantly shorter than those of the wild type, and could in principle have been identified in a forward genetic screen (Yi et al. 2010). Therefore, *Marchantia* mutants that develop short rhizoids might also have mutations in genes that are part of a network of regulators that control rhizoid development.

3.2.6 Likely phenotypes of mutants in regulators of rhizoid development

The aim of this work is to investigate the genetic regulation of rhizoid development in *Marchantia*. This includes the mechanisms that determine whether or not a cell adopts rhizoid identity, and that promote its differentiation into a rhizoid and eventual outgrowth. Genes whose expression is controlled by these regulators to bring about tip growth are also part of the rhizoid development gene network. From the considerations in the preceding sections, I can hypothesize what phenotypic classes are likely to result from mutations in genes that control the early events of rhizoid fate determination and differentiation. These are mutants that produce few or no rhizoids, or ones that develop rhizoids in ectopic locations. Amongst these mutants are likely to be plants with loss- or gain-of-function mutations in genes that control the patterning of rhizoid fate or regulate differentiation into a rhizoid cell. Mutants that develop short rhizoids, or have defects in the directionality of rhizoid growth, are likely to have mutations in various genes that encode proteins that are important for the cell biology of tip growth, but regulatory proteins that control growth may also be represented.

In this chapter I report the results of a mutagenesis and screen for mutants with defects in rhizoid development, and present eight genes in which mutations cause phenotypic defects that suggest that the gene product acts relatively late in rhizoid development. In the following chapters I focus on two transcription factors, the mutant phenotypes of which suggest roles in early in rhizoid fate determination and differentiation.

3.3 Results

3.3.1 61 individuals with defective rhizoid development were identified in a population of 105,000 insertion lines

To identify genes that are part of the network that regulates the development of rhizoids in *M. polymorpha*, I carried out a screen to identify mutants in which rhizoid patterning or growth is defective. I transformed wild-type sporelings with the T-DNA binary vector pCambia1300, which carries a hygromycin resistance gene driven by a the constitutive CaMV 35S promoter, to generate insertional mutants. I screened approximately 105,000 transformants and identified 61 plants with defective rhizoid development, classified as having either no rhizoids, rhizoids developing in ectopic locations, short rhizoids, very short rhizoids, or misshapen rhizoids (Table 3.1). Approximately 0.06% of the 105,000 screened transformants were selected as putative mutants with defects in rhizoid development. To estimate the number of T-DNA insertions per transformant, my colleagues Hélène Proust and Steven Kelly sequenced the genomes of 15 transformants, and identified 25 T-DNA insertions. This indicates that there were approximately 1 $^{2}/_{3}$ T-DNA insertions per transformed line, and so approximately 175,000 T-DNA insertions in the population of 105,000 transformants that I screened. Therefore approximately 0.035% of T-DNA insertions gave rise to mutants with defective rhizoid development.

Phenotype	Number identified	Inherited	Linked to insertion	Not linked to insertion	Not crossed	Inherited but linkage unclear
No rhizoids	1	-	-	-	1	-
Ectopic rhizoids	6	6	5	0	0	1
Very short rhizoids	15	3	1	2	12	0
Short rhizoids	32	16	8	7	16	1
Misshapen rhizoids	7	2	0	2	5	0
Total	**61**	**28**	**14**	**12**	**33**	**2**

Table 3.1 **Summary table** of the number of mutants identified in each phenotypic class, and the number of each in which the T-DNA insertion was found to be linked to the mutant phenotype in the F1 generation.

3.3.2 In 14 mutants the causative mutation is linked to a T-DNA insertion

The insertional mutagenesis used to generate the population of mutants that I screened can disrupt a gene by the integration of a single intact copy of the insertion into the genome within or near the gene. However, T-DNA transformation can also give rise to other kinds of mutation, such as partial insertions, deletions or chromosomal rearrangements (Gheysen et al. 1987; Takano et al. 1997; Nacry et al. 1998). The plants that were screened contain at least one intact or partial T-DNA insertion that conferred on them hygromycin resistance, but could also contain further T-DNA insertions or other kinds of mutation at other loci. The location of a T-DNA insertion in a mutant can be readily identified, but the defective phenotype of the mutant could be caused by a mutation at another locus. In order to infer that a T-DNA insertion is the causative mutation it is necessary to demonstrate genetic linkage between the insertion and the mutant phenotype. I therefore crossed the mutants I isolated to a wild-type of the opposite sex and scored the F1 progeny for the presence of the defective rhizoid phenotype and for resistance to hygromycin.

If a T-DNA insertion is the mutation that causes the defective phenotype then all F1 progeny that present the mutant phenotype should be hygromycin resistant, because recombination cannot separate the T-DNA insertion that confers the hygromycin resistance from the locus of the causative mutation. Therefore the presence of hygromycin-sensitive F1 progeny with the mutant phenotype is strong evidence that the causative mutation is not a T-DNA insertion. For all mutants from which mutant-phenotype, hygromycin-sensitive F1 progeny were identified I inferred that an intact T-DNA insertion was not the causative mutation. The one exception is mutant *vj69*; for an explanation see section 3.3.5, below. I inferred that mutants from which all mutant-phenotype F1 progeny were resistant to hygromycin possess a T-DNA insertion at, or closely linked to, the locus of the mutation that gives rise to the mutant phenotype.

The presence of hygromycin-resistant plants with the wild type phenotype in the F1 generation could also be problematic, as it is indicative of a recombination event between the mutation causing the defective phenotype and a T-DNA insertion. However, it is not as serious a problem as hygromycin-sensitive mutant-phenotype F1s, as it could be due to the presence of multiple T-DNA insertions in the mutant parent, which then segregate independently in the next generation.

Not all mutants were crossed to wild type, as some failed to produce viable gametangiophores or mature sporangia. In total I successfully crossed 27 mutants to the wild type, and in 14 of these the hygromycin resistance conferred by the T-DNA insertion cosegregated with the mutant phenotype in the next generation; these are presented in Table 3.2 and summarized in Table 3.1. In two cases the mutant phenotype was inherited but I do not have clear evidence that the T-DNA insertion is closely linked to the mutant phenotype. One of these was a mutant with short rhizoids from which only a small number of F1 progeny were obtained, 14 with the mutant phenotype and 19 with the wild type phenotype. All the mutant F1 progeny were hygromycin resistant and all the wild type progeny were hygromycin sensitive, but the numbers are too small to be confident of linkage; these data are therefore not included here. The second case was mutant *vj8*. The cross between this mutant and the wild type was unsuccessful, but the phenotype was inherited in a cross to another mutant (*vj7*; see Chapter 5).

3.3.3 Mutations in 10 genes were found to cause defects in rhizoid development

In the mutants in which the T-DNA insertion cosegregates with the mutant phenotype the insertion is likely to be the mutation that gives rise to the defect in rhizoid development.To find out which genes are affected by the insertions I used TAIL-PCR to identify the genomic sequence flanking the T-DNA insertion (see Appendix 4 for sequences).

		Mut		WT		total		HygS: HygR		χ^2 test
line name	**phenotype**	**HygR**	**HygS**	**HygR**	**HygS**	**HygR**	**HygS**	**ratio**	**Insert. no.**	**p**
vj3	ectopic rhizoids	85	0	0	130	85	130	1: 0.7	1	0.00
vj4	ectopic rhizoids	104	0	45	67	149	67	1: 2.2	2	0.04
vj5	ectopic rhizoids	76	0	0	97	76	97	1: 0.8	1	0.11
vj6	ectopic rhizoids	116	0	0	102	116	102	1: 1.1	1	0.34
vj7	ectopic rhizoids	137	0	101	61	238	61	1: 3.9	2	0.07
vj10	short rhizoids	111	0	8	85	119	85	1: 1.4	1	0.02
vj11	short rhizoids	78	0	0	102	78	102	1: 0.8	1	0.07
vj12	short rhizoids	103	0	0	74	103	74	1: 1.4	1	0.03
vj13	very short rhizoids	86	0	0	89	86	89	1: 1.0	1	0.82
vj14	short rhizoids	87	0	0	93	87	93	1: 0.9	1	0.65
vj21	short rhizoids	113	0	0	103	113	103	1: 1.1	1	0.50
vj24	short rhizoids	123	0	5	132	128	132	1: 1.0	1	0.80
vj25	short rhizoids	112	0	69	57	181	57	1: 3.2	2	0.71
vj69	short rhizoids	70	2	20	20	90	22	1: 4.1	2	0.19

Table 3.2 Cosegregation data for mutants in which the T-DNA insertion is linked to the mutant phenotype, showing numbers of F1 progeny of a cross between the mutant and wild-type displaying mutant (Mut) or WT phenotype that are hygromycin resistant (HygR) and hygromycin sensitive (HygS). Insert. no. indicates the inferred number of T-DNA insertions present in the genome. χ2 test gives the probability that the observed ratio of HygR:HygS is not different from the ratio expected from the inferred number of T-DNA insertions.

A draft *Marchantia* genome (prepared by Giulia Morieri, Hélène Proust, Clément Champion, Sandy Hetherington and Steve Kelly) allowed me to use the flanking sequence to identify the genomic locus of each insertion. I found the genomic location of the insertions in all lines except for mutant *vj10*; I was unable to retrieve flanking sequence from this line using TAIL-PCR. I then used a draft gametophyte transcriptome (prepared by Giulia Morieri, Hélène Proust, Clément Champion and Sandy Hetherington) to identify and create gene models of putative genes at or near the insertion sites.

I also used TAIL-PCR to identify the insertion sites in mutants that I could not cross, in order to identify further alleles of the genes that had been identified in the mutants where the insertion is linked to the mutant phenotype. In some cases this identified putative mutations in new genes that are not mutated in any lines for which I have linkage data, but for which other members of my group (Suvi Honkanen, Giulia Morieri) have at least one mutant in which they demonstrated linkage between the insertion and the mutant phenotype. The identification of multiple mutants with similar phenotypes and insertions in the same gene, at least one of which has been shown to be closely linked to the mutant phenotype, constitutes strong evidence that mutations in that gene cause defects in rhizoid development.

In total I identified mutations in 10 genes that cause defects in rhizoid development, 8 with multiple alleles and 2 with only a single allele (Table 3.3). In all cases, the insertion in at least one line is linked to the mutant phenotype. Genes were named after the closest homologue in *Arabidopsis thaliana*; in cases where the *Arabidopsis* gene has not been named, the genes were given the name Mp*SRI*, for *SHORT RHIZOIDS*, followed by a numeral. The identified genes are summarised in Table 3.3, and described further in the following sections.

TRANSCRIPTIONAL REGULATION		***Arabidopsis* homologue**	**mutant phenotype**	**number of mutant alleles**	**further alleles from SH, GM**
Mp*RSL1*	bHLH transcription factor	AT5G37800	ectopic rhizoids	5	9
Mp*WIP*	zinc finger transcription factor		ectopic rhizoids	1	0
VESICLE TRAFFICKING					
Mp*PI4Ka*	1-phosphatidylinositol 4-kinase alpha	AT1G49340	very short rhizoids	3	3
Mp*MYO-XI*	class XI myosin	AT3G12130	short rhizoids	2	3
Mp*SPI*	WD-40 repeat protein	AT1G03060	short rhizoids	1	2
Mp*SRI1*	guanyl nucleotide exchange factor, similar to *ScRIC1*	AT3G61480	short rhizoids	2	1
OTHERS					
Mp*EMB2756*	domain of unknown function 616 (DUF616)	AT1G34550	short rhizoids	2	0
Mp*SRI3*	domain of unknown function 616 (DUF616)	AT5G42660	short rhizoids	1	1
Mp*EXL1*	Unknown function	AT4G08950	short rhizoids	1	1
Mp*HSP70-1*	Heat shock protein 70	AT1G79930	short rhizoids	1	0

Table 3.3 Genes identified in the screen for mutants with defects in rhizoid development, grouped according to their putative functions. The closest *Arabidopsis* homologue was identified by reciprocal best BLAST. The number of mutant alleles indicates the number of alleles identified in my screen. Further alleles from SH, GM indicates the number of mutations in these genes identified by Suvi Honkanen and Giulia Morieri. Grey shading indicates genes for which more than 1 allele was identified.

3.3.4 Mutations affecting two genes, Mp*RSL1* and Mp*WIP*, cause rhizoids to develop in ectopic positions

I identified six mutants in the screen (*vj3* to *vj8*) that develop rhizoids on the dorsal surface of the thallus, where rhizoids do not develop in the wild type. Five of these mutants have insertions 5’ to the coding sequence of the bHLH transcription factor Mp*RSL1*. These mutants are discussed in more detail in Chapter 4.

The ectopic-rhizoids phenotype of the remaining mutant is linked to an insertion 5’ to the transcriptional start site of a gene that encodes the zinc finger protein Mp*WIP*. This gene is the subject of Chapter 5.

All of the remaining mutants for which I was able to locate the mutation that causes the mutant phenotype display defects in rhizoid growth. In the following sections of this chapter I report the essential information about each of the genes I have identified as being required for rhizoid growth on the basis of these mutants. The morphology of the rhizoids of the mutants is compared to that of wild type rhizoids, to illustrate the developmental defect associated with the mutations. I present simple gene trees containing the gene of interest and related genes from *Arabidopsis* and outgroups, where necessary. My goal is not to use these trees to draw strong conclusions about the evolution of these gene families, but simply to identify which *Arabidopsis* genes are likely to be closely related to the genes of interest. Finally I include a schematic of the predicted structure of the gene, for use in forming hypotheses about the effects of the mutations on the products of the genes.

3.3.5 Mutations in four genes with possible roles in vesicle trafficking cause defects in rhizoid elongation

Eight mutants that develop short or very short rhizoids have an insertion in one of four genes that may have roles in vesicle trafficking.

Three mutants were identified with insertions in *PHOSPHATIDYLINOSITOL 4-KINASE III ALPHA1* (Mp*PI4Kα1*), *vj39, vj32,* and *vj13*. In *vj13* the mutant phenotype is linked to the T-DNA insertion; I was unable to cross the other two mutants to the wild type. Compared to wild-type plants, these mutants develop very short rhizoids (Figure 3.1 a, b), which are often thicker than those of wild-type plants, and club-shaped rather than tubular (Figure 3.1 b). Phylogenetic analysis (Figure 3.1 c) indicates that Mp*PI4Kα1* belongs to a clade that includes two *Arabidopsis PI4Kα* genes and a second *PI4Kα* from *Marchantia* (Mp*PI4Kα2). Marchantia* also possesses a single *PHOSPHATIDYLINOSITOL 4-KINASE BETA* (Mp*PI4Kβ*) gene, which belongs to a clade containing the two At*PI4Kβ* genes. The insertions in all three alleles of Mp*PI4Kα1* are in exons of the CDS (Figure 3.1 d), and so are likely to be loss-of-function mutations. A further three Mp*pi4k α1* alleles, which display the same phenotype, have been identified by my colleague Suvi Honkanen. In all three the insertion is linked to the mutant phenotype.

The shortness of the rhizoids in these mutants indicates that Mp*PI4Kα1* is important for elongation by tip growth. The thickened, club-shaped tips suggest that the mutants may be impaired in the ability to restrict growth to a limited region of the tip.

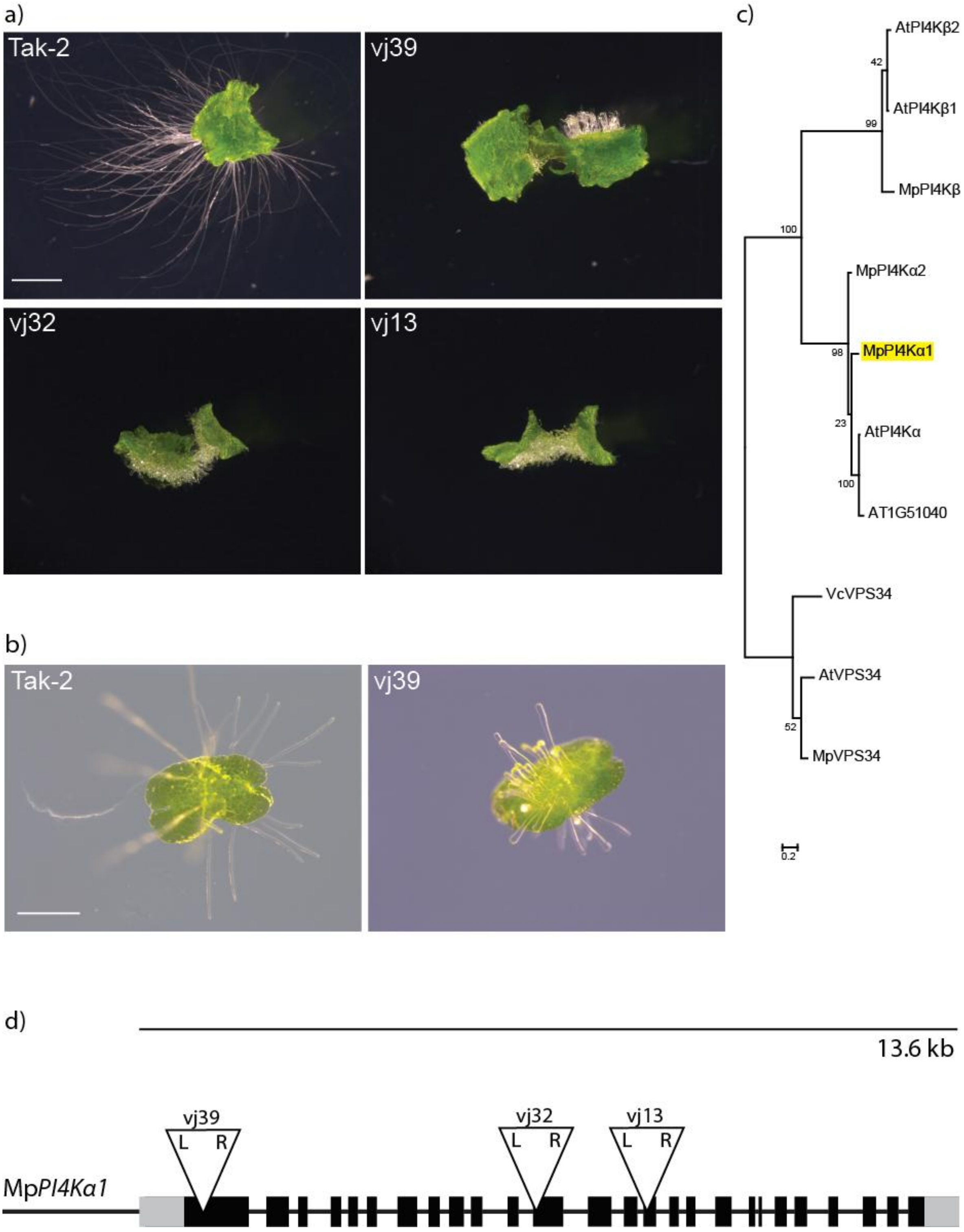


Figure 3.1 Mutations in Mp*PHOSPHATIDYLINOSITOL 4-KINASE ALPHA1* cause the development of very short rhizoids. a) 3 Mp*pi4kα* alleles (*vj39*, *vj32*, *vj13*) cause the development of very short rhizoids in comparison with the wild type Tak-2. 10 d old gemmalings, scale 2 mm. b) The very short rhizoids of Mp*pi4kα* mutants are thickened and club-shaped in comparison with those of Tak-2. 2 d gemmae, scale 500 µm. c) Maximum-likelihood phylogeny of phosphatidylinositol 4-kinase proteins from Arabidopsis thaliana (At) and Marchantia polymorpha (Mp), rooted with *VPS34* from At, Mp and *Volvox carteri* (Vc). Nodes are marked with aLRT values. d) Gene structure of Mp*PI4Kα*, with the insertion site in each allele marked. L and R indicate the orientation of the left and right borders of the T-DNA, respectively. Boxes represent exons; grey are untranslated, black are CDS.

Two alleles of Mp*MYOSIN-XI* (Mp*XI;* mutants *vj21* and *vj11)* cause the development of short rhizoids (Figure 3.2 a). In both mutants the T-DNA insertion is linked to the mutant phenotype (Table 3.2), and is located in the final intron, within the 3' UTR (Figure 3.2 b). I identified two myosin genes in the genome of *Marchantia*, one in class XI and one in class VIII (Figure 3.2 c). Suvi Honkanen has identified a further three alleles of Mp*xi*, in all of which the insertion is linked to the mutant phenotype and is located in the final intron. In all five Mp*xi* alleles the T-DNA insertion is oriented with the RB towards the transcriptional start site.

The short rhizoids of these mutants indicate that wild type Mp*XI* function is required for rhizoids to elongate to the extent seen in the wild type.

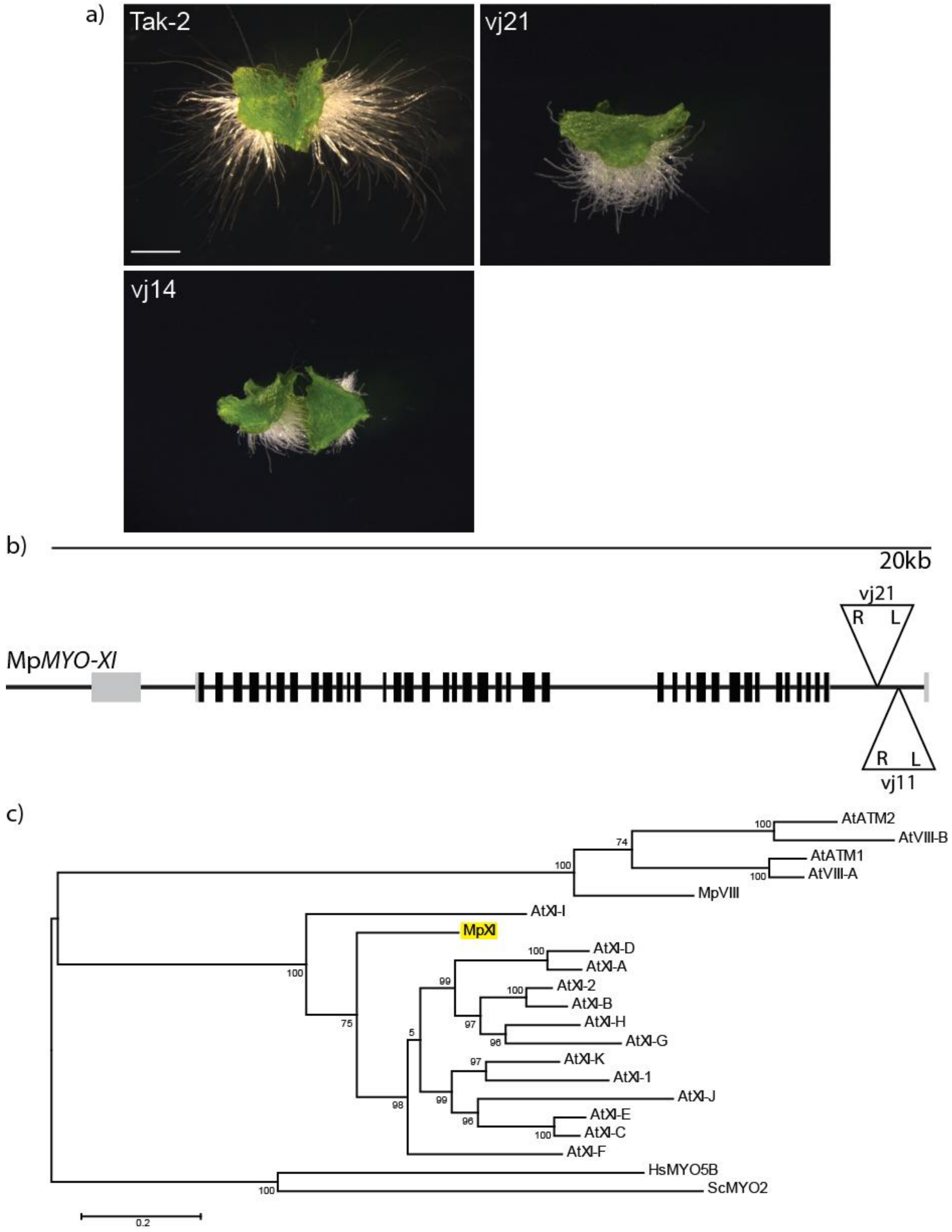


Figure 3.2 Mutations in *MpMYOSIN-XI* cause the development of short rhizoids. a) Rhizoids of *Mpmyo-xi* mutants *vj21* and *vj11* are shorter than those of Tak-2. 10 d old gemmalings, scale = 2 mm. b) Gene structure of Mp*XI*, with the insertion site in each allele marked. L and R indicate the orientation of the left and right borders of the T-DNA, respectively. Boxes represent exons; grey are untranslated, black are CDS. c) Maximum-likelihood phylogeny of myosin proteins from *Arabidopsis thaliana* (At) and *Marchantia polymorpha* (Mp), rooted with *Homo sapiens MYO5B* and *Saccharomyces cerevisiae MYO2*. Nodes are marked with aLRT values.

One mutant with short rhizoids (*vj69*; Figure 3.3 a) possesses an insertion in a WD40/BEACH domain protein (Figure 3.3 b) that is sister to a clade consisting of At*SPIRRIG* (AtSPI) and At*BEACH-DOMAIN HOMOLOG A2* (At*BCHA2*) (Figure 3.3 c). The insertion in *vj69* lies within the CDS of Mp*SPI*, albeit within an intron, and so this mutant may produce truncated MpSPI and so be impaired in MpSPI function.

From the cosegregation data, this insertion does not appear to be closely linked to the defective rhizoid elongation phenotype (Table 3.2). A limited number of F1 progeny were produced and analysed (112 in total), and 2 of the progeny scored as possessing the mutant phenotype were sensitive to hygromycin, suggesting that they lacked the T-DNA insertion. Furthermore, the number of progeny with the wild type phenotype that were resistant to hygromycin indicates that there are multiple T-DNA insertions segregating in the population. Nevertheless, the hypothesis that the insertion in Mp*SPI* is responsible for the mutant phenotype is supported by the identification of a second allele of Mp*SPI* with the same mutant phenotype by Suvi Honkanen, in which the insertion is clearly linked to the mutant phenotype. The presence in Table 3.2 of two hygromycin-sensitive F1 progeny with the short-rhizoid mutant phenotype might be explained by an error in scoring the phenotypes of the progeny, that is, these two plants may in fact have been phenotypically wild type. The phenotypes of this mutant and the allele identified by Suvi Honkanen, in which the rhizoids are shorter than the wild type, indicate that Mp*SPI* contributes to rhizoid elongation.

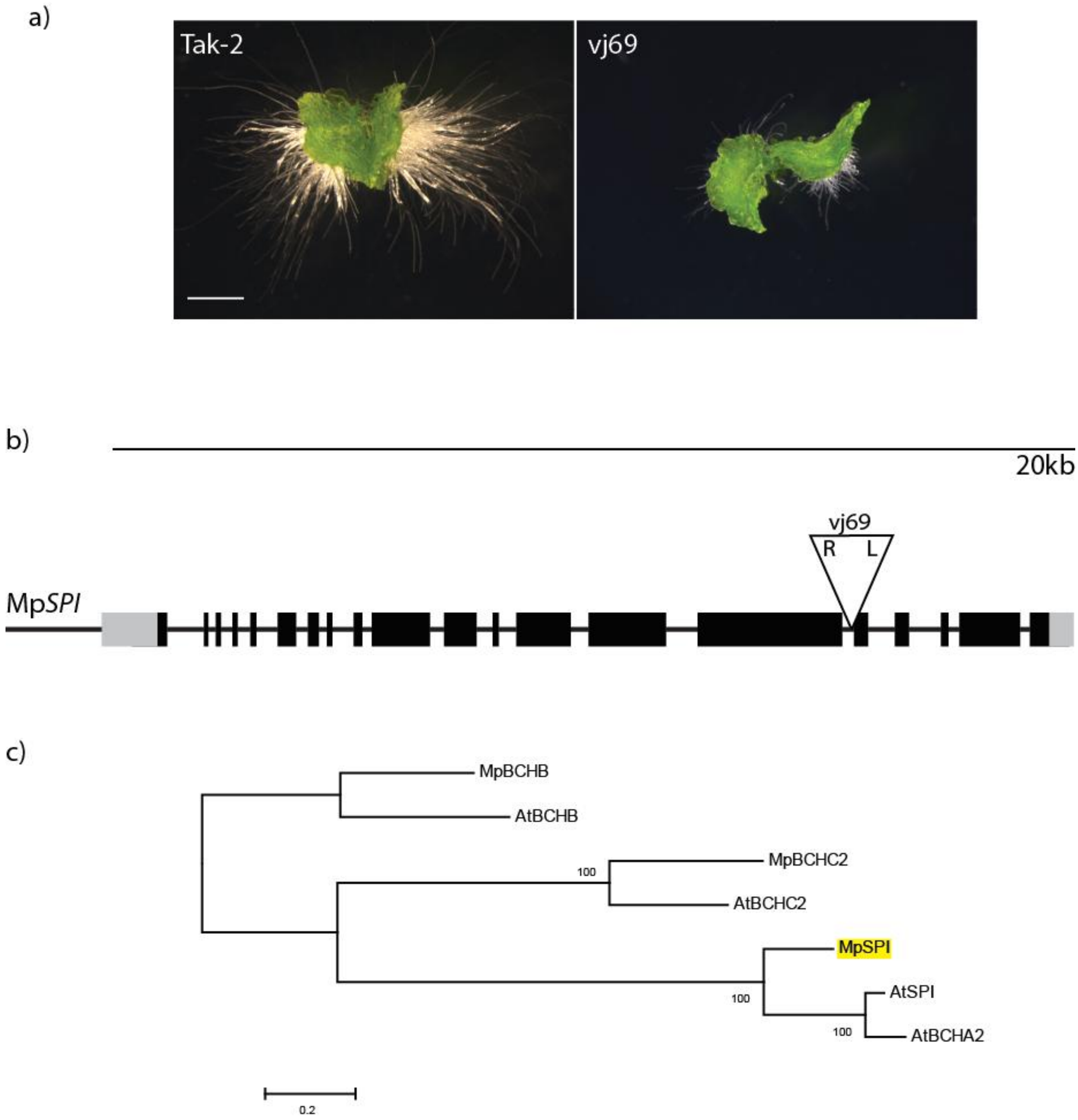


Figure 3.3 A mutation in *MpSPIRRIG* causes the development of short rhizoids. a) Mutant *vj69* develops short rhizoids in comparison with the wild type Tak-2. b) Gene structure of *MpSPI*, with the insertion site in *vj69* marked. L and R indicate the orientation of the left and right borders of the T-DNA, respectively. Boxes represent exons; grey are untranslated, black are CDS. c) Maximum-likelihood phylogeny of group A and B WD40/BEACH domain proteins from *Arabidopsis thaliana* (At) and *Marchantia polymorpha* (Mp). Nodes are marked with aLRT values.

Two mutants with short rhizoids (*vj25* and *vj14*, Figure 3.4 a) have insertions linked to the mutant phenotype (Table 3.2) in the same gene (Figure 3.4 b). The insertion in *vj25* lies within an exon of the CDS, and is likely to result in a truncation to the encoded protein. This gene is sister to a clade containing two *Arabidopsis* genes without names or characterized functions (Figure 3.4 c). It was therefore named Mp*SHORTRHIZOIDS1* (*MpSRI1*). A search for conserved protein domains using the phhmer algorithm (http://www.ebi.ac.uk/Tools/hmmer/) (Finn et al. 2011) identified a RIC1 domain (Pfam domain PF07064) in these genes. As I discuss in section 3.4.2, the RIC1 domain may be involved in the control of vesicle fusion events, which is why I include Mp*SRI1* among the genes involved in vesicle trafficking. No signal peptides were found in the amino acid sequence of the *Marchantia* or *Arabidopsis* proteins using SignalP 4.1 (http://www.cbs.dtu.dk/services/SignalP/) (Petersen et al. 2011). The short rhizoid phenotypes of these mutants suggests that the product of the gene is required for normal rhizoid elongation.

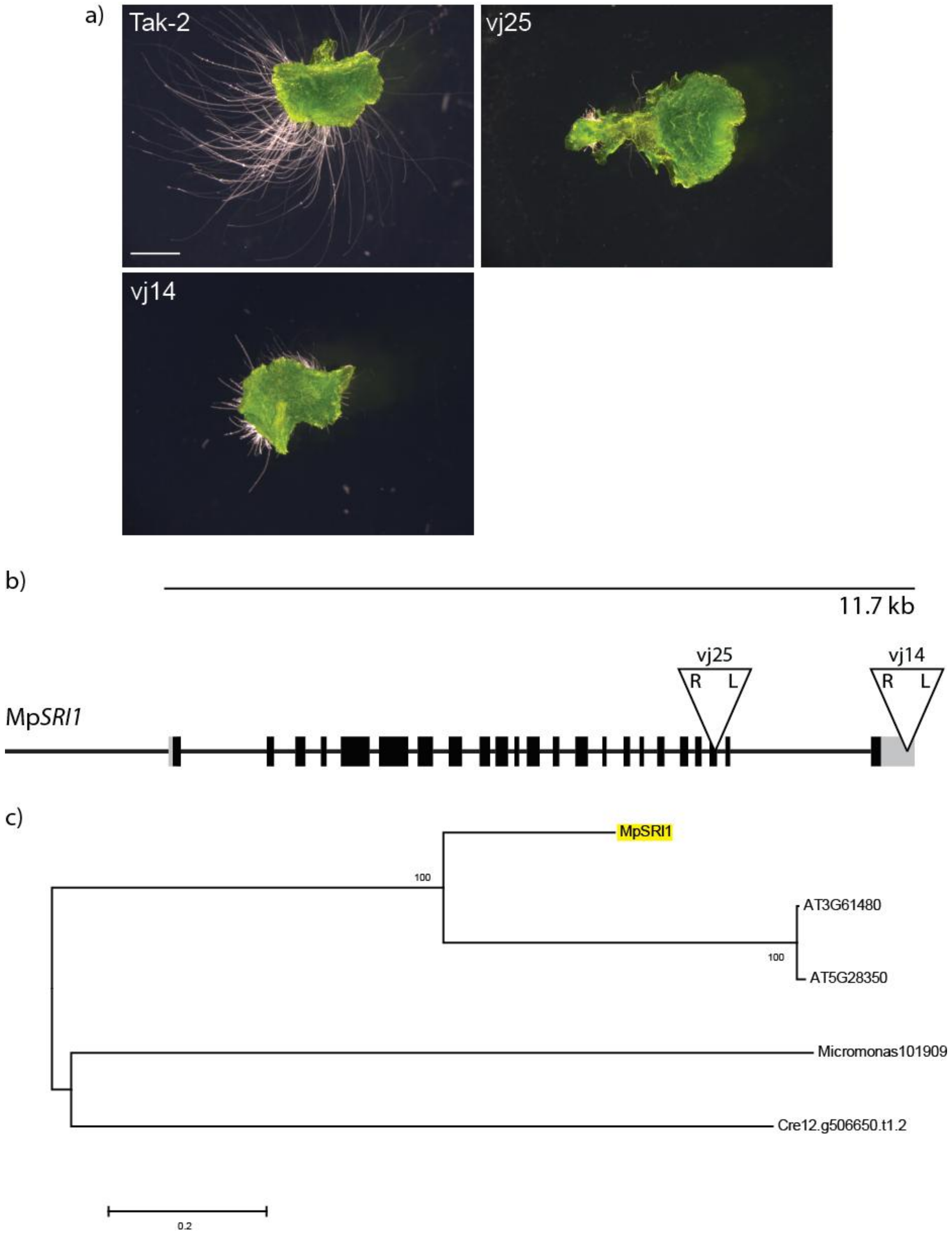


Figure 3.4 Mutations in *MpSHORTRHIZOIDS1* cause the development of short rhizoids. a) Mutants *vj25* and *vj14* develop short rhizoids in comparison with the wild type Tak-2. 10 d old gemmalings, scale = 2 mm. b) Gene structure of *MpSRI1*, with the insertion sites in *vj25* and *vj14* marked. L and R indicate the orientation of the left and right borders of the T-DNA, respectively. Boxes represent exons; grey are untranslated, black are CDS. c) Maximum-likelihood phylogeny of *SRI1* from *Marchantia polymorpha* (Mp) and related proteins from *Arabidopsis thaliana*, *Chlamydomonas reinhardtii* (Cre) and *Micromonas* sp. RCC299. Nodes are marked with aLRT values.

3.3.6 Mutations in four genes with other functions cause defects in rhizoid elongation

In the screen I identified four more genes that mutate to a short-rhizoid phenotype, the functions of which are unknown or do not fall into the classes of transcriptional regulation or vesicle trafficking.

Mutations in two closely related genes, Mp*EMB2756* and Mp*SRI3*, lead to the development of short rhizoids. The rhizoids of mutant *vj67* are short, and appear to be fewer in number compared to the wild type, although I was not able to quantify this (Figure 3.5 a). *vj67* has an insertion in an intron in the CDS of a gene of unknown function, dubbed Mp*SRI3* (Figure 3.5 b), which may result in a loss of functional MpSRI3. I have not demonstrated that this insertion is linked to the mutant phenotype; however, Giulia Morieri has identified a mutant with the same phenotype and an insertion in the same gene, which she has shown to be linked to the mutant phenotype. I identified two alleles of Mp*EMB2756*, mutants *vj24* and *vj68*, whose rhizoids are moderately shorter than those of the wild type (Figure 3.5 a). As the insertions in these lines are within the CDS of Mp*EMB2756*, they are likely to be loss-of-function mutants, especially *vj68*, which has an insertion into an exon. The insertion in *vj24* appears closely linked to the mutant phenotype, since all 123 F1 progeny with the wild type phenotype possessed the T-DNA insertion. However, there may be a second insertion nearby (3.6 cM) because 5 of the 137 F1 progeny with the wild type phenotype were hygromycin resistant (Table 3.2). Alternatively, because the mutant phenotype is rather moderate, I may have mis-scored mutant plants as wild type. I was unable to cross *vj68*.

Searches for conserved domains using the phmmer algorithm (Finn et al. 2011) (www.ebi.ac.uk/Tools/hmmer/) indicate that both MpEMB2756 and MpSRI3 contain a domain of unknown function 616 (DUF616). The phylogeny in Figure 3.5 (c) includes all *Marchantia* and *Arabidopsis* proteins that contain DUF616, and indicates that Mp*SRI3* and Mp*EMB2756* belong to sister clades. The short rhizoids of mutants in these two genes indicates that the products of the genes are involved in some process that contributes to rhizoid elongation.

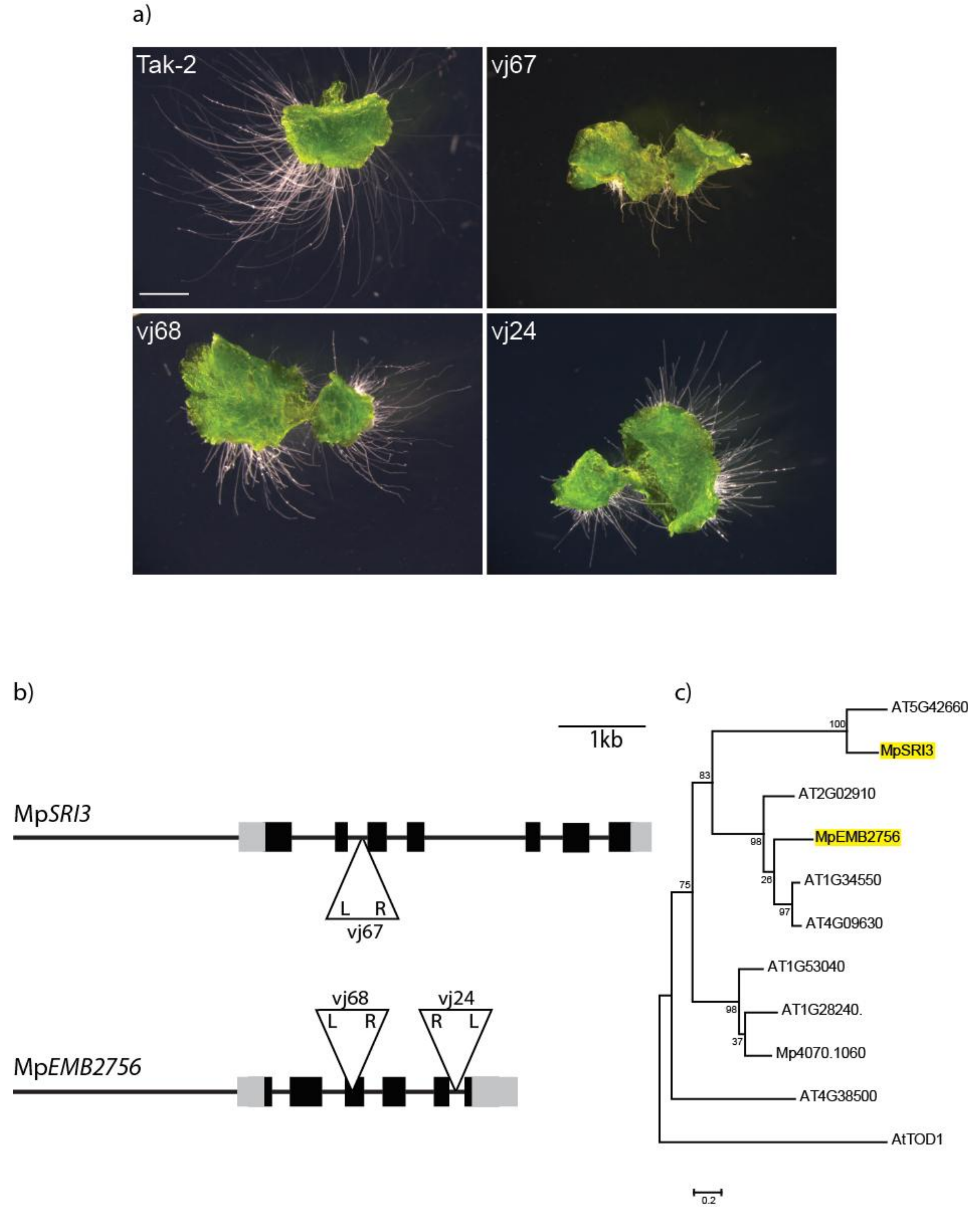


Figure 3.5 Mutations in *MpSHORTRHIZOIDS3* and *MpEMB2756* cause the development of short rhizoids. a) Mutant *vj67* develops few short rhizoids, and *vj24* and *vj68* develop short rhizoids in comparison with the wild type Tak-2. 10 d old gemmalings, scale = 2 mm. b) Gene structures of *MpSRI3* and *MpEMB2756*, with the insertion sites in mutants *vj67*, *vj24*, and *vj68* marked. L and R indicate the orientation of the left and right borders of the T-DNA, respectively. Boxes represent exons; grey are untranslated, black are CDS. c) Maximum-likelihood phylogeny of *MpSRI3*, *MpEMB2756* and related genes from *Marchantia polymorpha* (Mp) and *Arabidopsis thaliana* (At). Nodes are marked with aLRT values.

One mutant (*vj19*) develops short rhizoids (Figure 3.6 a), whose shape is distorted in comparison with the smooth, straight rhizoids of the wild type (Figure 3.6 a, insets). The mutant has an insertion 15,379 nt upstream of the CDS of a gene that is similar to the *Arabidopsis* gene *EXORDIUM* (Figure 3.6 b), and so was named Mp*EXORDIUM-LIKE 1* (Mp*EXL1*). I was unable to cross this mutant, but Giulia Morieri identified an allele of Mp*exl1* with an insertion approximately 500 nt upstream of the CDS, which is linked to the mutant phenotype.

The phylogeny in Figure 3.6 (c) contains all EXL proteins from the land plants *Arabidopsis thaliana*, *Selaginella moellendorffii*, *Physcomitrella patens*, and *Marchantia polymorpha*, the charaphycean alga *Coleochaete nitellarum*, and the chlorophyte alga *Coccomyxa subellipsoidea*. The number of EXL proteins in each species is as follows: *Coccomyxa*, 1; *Coleochaete*, 2; *Marchantia*, 18; *Physcomitrella*, 4; *Selaginella*, 5; *Arabidopsis*, 7. Lineage-specific expansion in *Marchantia* appears to have given rise to two well-supported clades containing eight and nine *Marchantia* genes, with one further gene (Mp*EXL9*) that falls outside these two clades.

Because AtEXO has been reported to be secreted and present in the cell wall (Bayer et al. 2006), I used SignalP 4.1 (http://www.cbs.dtu.dk/services/SignalP/) (Petersen et al. 2011) to search all EXL proteins included in Figure 3.6 (c) for predicted signal peptides. All *Arabidopsis*, *Physcomitrella*, and *Coleochaete* EXL proteins were predicted to include signal peptides. No signal peptide was predicted for Cs*EXL*, Sm*EXL2* or Sm*EXL3*, while the remaining *Selaginella* EXL proteins contain predicted signal peptides. Strikingly, the majority of *Marchantia* EXL proteins lack a signal peptide: MpEXL1, 2, 3, 6, 7, 8, 9, 10, 15, 16, and 17 all lack signal peptides, while MpEXL4, 5, 11, 12, 13, 14, and 18 contain one. It should be noted that Mp*EXL1*, the gene mutated in mutant *vj19*, is one of those without a signal peptide, which suggests that its function may differ from those of *Arabidopsis EXL*s There is no apparent phylogenetic clustering of *Marchantia* proteins with or without signal peptides, as both classes of protein are distributed across the two major clades. The shortness and misshapenness of the rhizoids of this mutant suggest that its defects in elongation might perhaps arise from deficiencies in regulating the deposition of cell surface material.

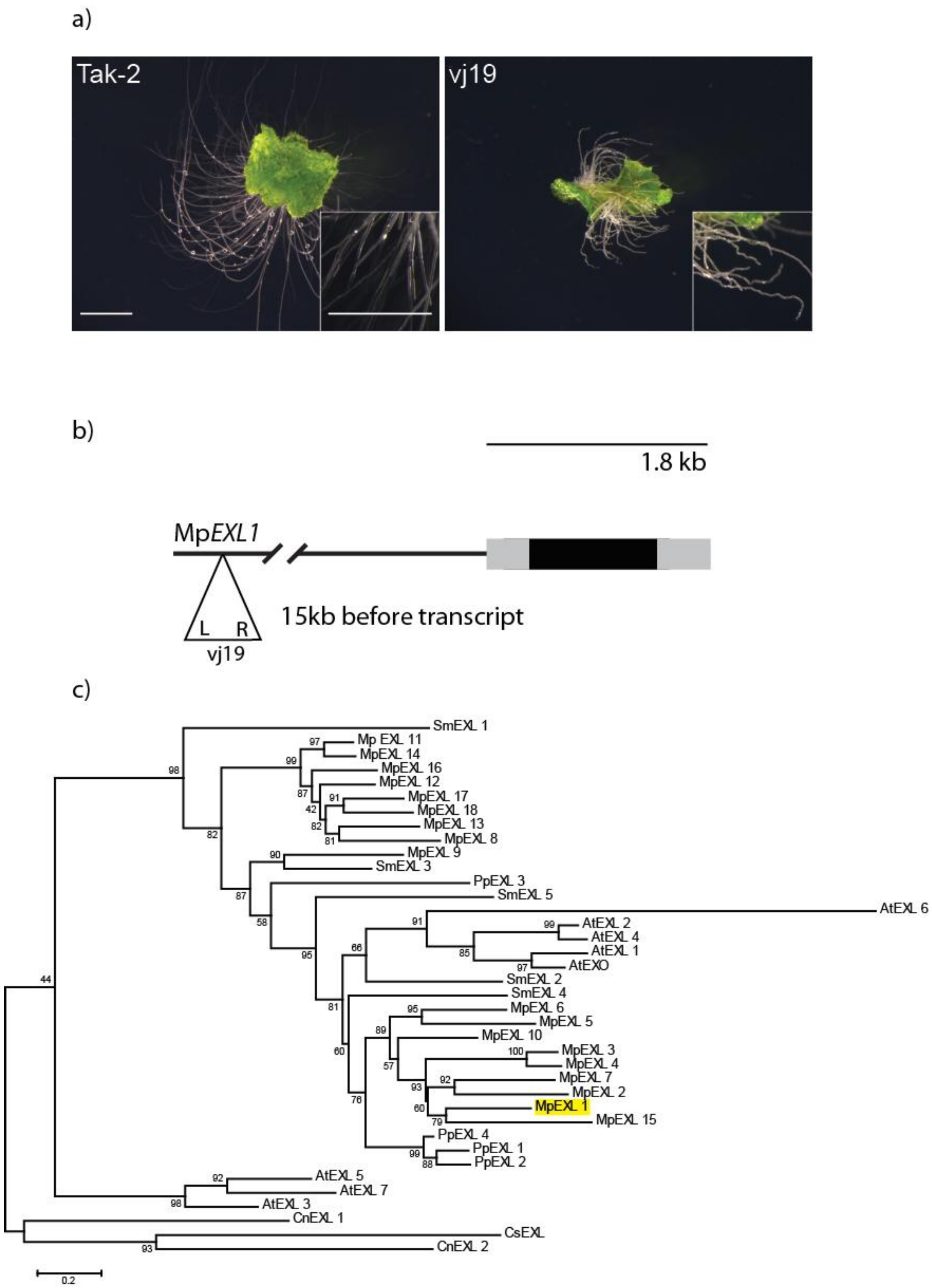


Figure 3.6 A mutation in Mp*EXORDIUM-LIKE 1* causes the development of short rhizoids. a) Mutant *vj19* develops short, crumpled rhizoids compared to the wild type Tak-2. 10 d old gemmalings, scale = 2 mm. b) Gene structure of Mp*EXL1*. Boxes represent exons; grey are untranslated, black are CDS. Mutant *vj19* possesses an insertion 15,379 nt upstream of the CDS. c) Maximum-likelihood phylogeny of EXL proteins from *Marchantia polymorpha* (Mp), *Arabidopsis thaliana* (At), *Selaginella moellendorffii* (Sm), *Physcomitrella patens* (Pp), *Coleochaete nitellarum* (Cn), and *Coccomyxa subellipsoidea* (Cs). Nodes are marked with aLRT values.

Mutant *vj12* initiates rhizoids, but their elongation arrests almost immediately (Figure 3.7 a). There is a single T-DNA insertion in this mutant, which is linked to the mutant phenotype in the F1 progeny (Table 3.2). The insertion is located 610 nt upstream of a gene, with the LB of the T-DNA oriented towards the CDS of the gene (Figure 3.7 b). The protein this gene encodes belongs to the Hsp110/SSE subfamily of the HEAT SHOCK PROTEIN 70 protein family (Figure 3.7 c; subfamilies as described in Lin et al. 2001), and so was named Mp*HSP70-1*. Despite the fact that the first identified HSP70 proteins were so called because of their weight of 70 kDa, the protein encoded by Mp*HSP70-1* has a mass of 93 kDa, similar to the 91 kDa of AtHSP91, which is in a clade sister to Mp*HSP70-1*.

The fact that the rhizoids of this mutant barely progress beyond the formation of bulges suggests that the protein encoded by Mp*HSP70-1* plays a role in the transition from the bulge stage to the establishment of tip growth and elongation.

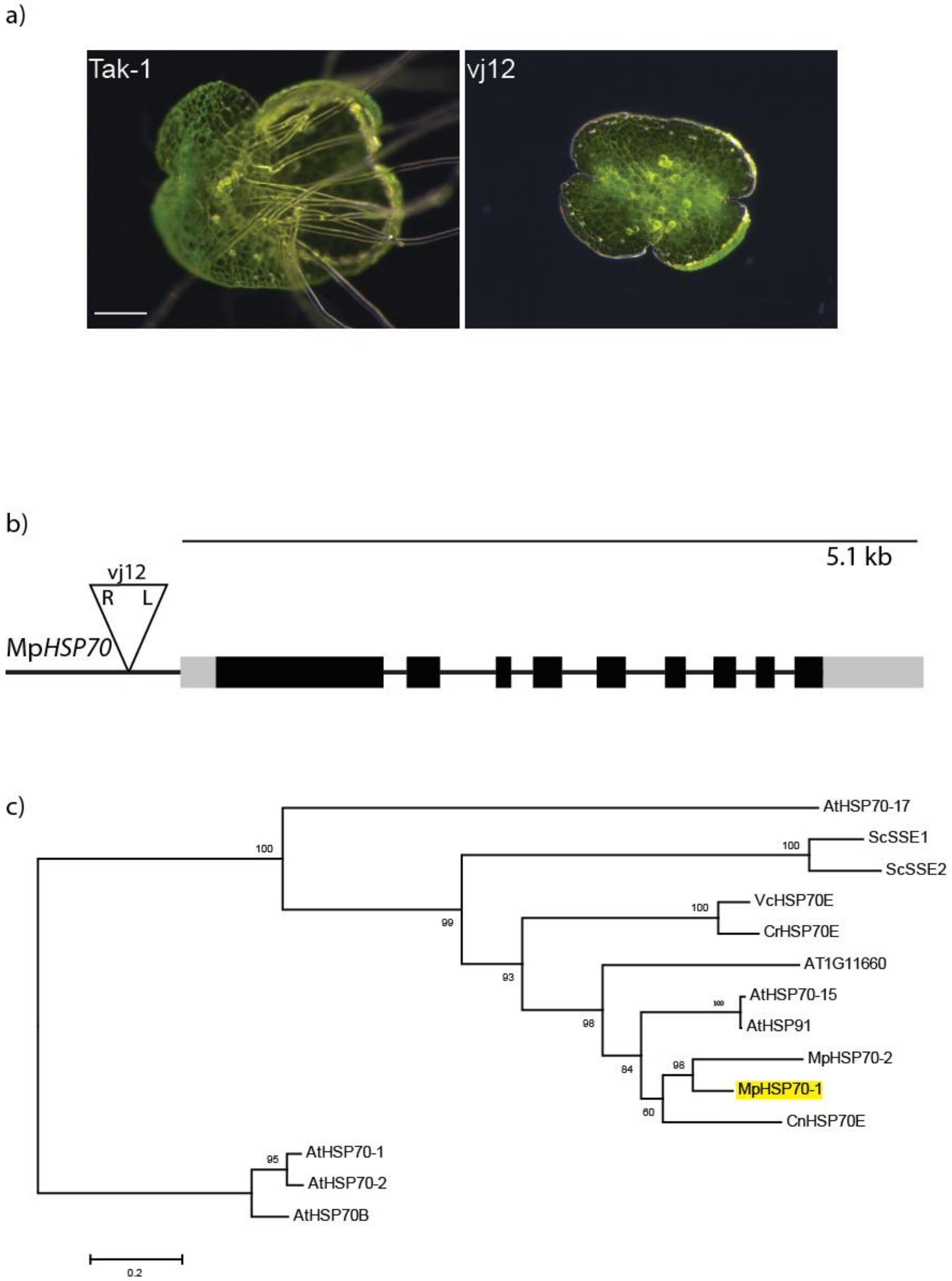


Figure 3.7 A mutation in *MpHSP70-1* causes the development of extremely short rhizoids. a) Mutant *vj12* possesses rhizoids that barely elongate after initiation. 2 d gemmalings, scale 200 µm. **b)** Gene structure of *MpHSP70-1*, with the position of the insertion in mutant *vj12* marked. Boxes represent exons; grey are untranslated, black are CDS. **c)** Maximum-likelihood phylogeny of HSP110/SSE1 subfamily of HSP70 proteins from *Marchantia polymorpha* (Mp), *Arabidopsis thaliana* (At), *Coleochaete nitellarum* (Cn), *Volvox carteri* (Vc), *Chlamydomonas reinhardtii* (Cr), and *Saccharomyces cerevisiae* (Sc), rooted with *Arabidopsis* cytoplasmic HSP70 proteins (after Lin et al. 2001).

3.4 Discussion

3.4.1 Genes with both known and novel roles in rooting cell growth identified

Three of the genes identified in this screen (Mp*RSL1*, Mp*XI*, and Mp*SPI*) are homologues of genes known to function in the development of root hairs in *Arabidopsis*. Two *Arabidopsis* class XI myosins have been shown to be necessary for wild type root hair elongation, as At*xi-2* and At*xi-k* mutant root hairs are only 28-40% the length of wild type root hairs (Peremyslov et al. 2008). The phenotype of the Mp*xi* mutants (Figure 3.2) is similar, as rhizoid length is reduced in these mutants. The root hairs of At*spi* mutants are less than half the length of those of the wild type (Steffens et al. 2015), and likewise the Mp*spi* mutant develops short rhizoids (Figure 3.3). The role of *RSL* genes in the development of root hairs and moss rhizoids has been well studied. Like the class I *RSL* genes of other species, Mp*RSL1* has been identified in this work as a likely positive regulator of rhizoid differentiation; this is discussed further in Chapter 4.

Several genes encoding proteins that mediate phosphoinositide signalling have been implicated in root hair development in *Arabidopsis*, such as *1-PHOSPHATIDYLINOSITOL-4-PHOSPHATE 5-KINASE 3* (At*PIP5K3*) and *PHOSPHATIDYLINOSITOL 4-OH KINASE III BETA 1* and *2* (At*PI4Kβ1* and At*PI4Kβ2*), mutants in which develop short, often deformed rhizoids (Preuss et al. 2006; Kusano et al. 2008; Stenzel et al. 2008). In this screen I identified mutants in the phosphoinositide-modifying enzyme MpPI4Kα that cause the development of short rhizoids. However, though PI4Kα proteins are involved in the same biochemical pathways as PIP5K3 and PI4Kβ, to my knowledge these are the first *pi4kα* mutants to be isolated from any plant species.

A second group of genes identified in this screen consists of ones that are homologues of genes in *Arabidopsis* that have not been reported to have a role in root hair development, or for which no function has been reported. I identified Mp*WIP* as a possible regulator of rhizoid development, while no role in root hair development has been reported for any member of the *WIP* family in *Arabidopsis*. I explore this more fully in Chapter 5. Of the remaining genes, which all cause the development of short rhizoids when mutated, the *Arabidopsis* homologues of Mp*EXL1* or Mp*HSP70-1* have not been implicated in root hair development. At*EMB2756* is essential for seed development, but otherwise nothing is known of its function (Tzafrir et al. 2004; Lloyd and Meinke 2012). No functions have been reported for the *Arabidopsis* homologues of Mp*SRI1* or Mp*SRI3*.

Thus in this screen I have identified a conserved positive regulator of rooting cell differentiation (Mp*RSL1*) and *Marchantia* relatives of genes known to be important in root hair growth in *Arabidopsis* (Mp*XI,* and Mp*SPI*). This indicates that the roles of genes that act both early and late in rooting cell differentiation are conserved between *Marchantia* and *Arabidopsis*. In addition I have identified a potential novel positive regulator of *Marchantia* rhizoid fate and report for the first time developmental roles for several genes, which may provide new insights into the mechanisms of tip growth.

3.4.2 Delivery of material to the growing tip is vital for rhizoid development

During tip growth, new material for the construction of the cell surface must be continuously delivered to the growing tip. Several mutants identified in this screen are likely to be necessary for the delivery of this material to the apex.

In plants, the actomyosin motility system is largely responsible for the intracellular movement of organelles and endomembrane compartments (Prokhnevsky et al. 2008). Plant myosins belong to one of two classes, VIII and XI. In *Arabidopsis* these have diversified into families with 4 and 13 members, respectively (Avisar et al. 2008) whereas *Marchantia* possesses a single member of each class. Three *Arabidopsis* class XI myosins have been found to be important for root hair growth, as the At*xi-2* At*xi-k* At*xi-b* triple mutant has root hairs 1/10th the length of the wild type (Peremyslov et al. 2010). This is associated with a more than two-fold reduction in the speed of trafficking of peroxisomes, Golgi stacks and mitochondria (Peremyslov et al. 2008). Receptors of AtXI-K have been identified that define a subset of XI-K trafficked vesicles localized in the growing root hair tip (Peremyslov et al. 2013). The short rhizoid phenotype of the Mp*myo-xi* is thus consistent with a role for Mp*MYO-XI* in delivering material to the growing tips of rhizoids. These mutants both have insertions after the CDS (Figure 3.2 a), as do the three alleles identified by Suvi Honkanen. I hypothesize that in these mutants the expression of Mp*XI* might be reduced, rather than abolished. In the discussion of Chapter 4 I speculate that the CaMV 35S promoter present in the T-DNA used in this screen may activate the expression of the gene downstream of the RB. In all five Mp*xi* alleles the insertion is oriented with the RB towards the beginning of the gene. Perhaps the RB drives transcription of the non-coding strand of the gene, leading to the formation of dsRNA, cleavage, and so a reduction of Mp*XI* transcript. Rhizoids may be more sensitive than other cell types to a reduction in the speed of trafficking because their rapid elongation requires a high rate of vesicle delivery to the tip. No mutants were found with insertions in the CDS that would likely cause null mutations, possibly because such mutations are lethal.

To achieve polarized cell growth, it is necessary to specifically deliver vesicles and macromolecules to particular parts of the cell. This requires that membrane compartments possess markers of identity that allow them to be transported correctly. One way in which identity is conferred on vesicles is through interactions with Rab GTPases, a large family of proteins that associate with distinct membrane compartments and organize membrane trafficking across the eukaryotes (Stenmark 2009). Another class of regulators of trafficking are the phosphoinositides, membrane phospholipids that are enriched in various membrane compartments. Phosphatidylinositol 4-kinases (PI4Ks), the family to which MpPI4Kα1 belongs, catalyze the phosphorylation of phosphatidylinositol to produce phosphatidylinositol 4-phosphate (PI4P). PI4P is the most abundant phosphoinositol (Krinke et al. 2007) and has diverse functions in the cell; it is important to intracellular signalling as a precursor of phosphatidylinositol 4,5-bisphosphate, the substrate of phospholipase C (Berridge and Irvine 1989), and contributes substantially to the negative charge of the inner leaflet of the plasma membrane (Hammond et al. 2012). It also plays a key role in vesicle trafficking. From studies on the yeast PI4K Pik1p it appears that PI4K activity is needed for trafficking from the Golgi to the plasma membrane (Ischebeck et al. 2010). It has also been suggested that an accumulation of PI4P in the plasma membrane of the sub-apical zone of the root hair tip might define a region into which proteins containing PI4P-binding domains are recruited, including AGD1, which is required for the maintenance of growth polarity (Yoo et al. 2012). Very little is known about PI4Kα in *Arabidopsis* as homozygous mutants are not viable (Delage et al. 2012), but PI4Kα and PI4Kβ possess similar catalytic domains and catalyze the same reaction, though PI4Kα is larger and possesses different domains outside the catalytic region, including pleckstrin-homology and ankyrin domains (Wong and Cantley 1994). AtPI4Kβ1 localizes to *trans*-Golgi vesicles and membranes of the root hair tip, and At*pi4kβ1/2* double mutants develop short, swollen root hairs (Preuss et al. 2006), similar to the rhizoid phenotype of the Mp*pi4kα* mutants identified in this screen (Figure 3.1). Because of the very wide range of processes influenced by PI4P and the other members of the phosphoinositide pool, along with the fact that the subcellular localization of PI4Kα is unknown in plants, it is difficult to infer exactly how the rhizoid phenotype of the Mp*pi4kα* identified in this screen comes about. The new availability of a plant *pi4kα* mutant will provide a chance to investigate the function of PI4Kα in plants.

The WD40/BEACH domain protein AtSPIRRIG has also been implicated in membrane trafficking in *Arabidopsis*. The trichomes of At*spi* mutants are twisted and malformed, their pavement cells are less highly lobed, and their root hairs are shorter than the wild type (Saedler et al. 2009). The vacuoles of the mutant root hairs are highly fragmented, and a similar defect is seen in the contractile vacuoles of mutants in the *Dictyolstelium* homologue

(Gerald et al. 2002; Du et al. 2008; Saedler et al. 2009). It has been suggested that AtSPIRRIG may be involved in membrane fusion events in localized regions (Saedler et al. 2009), and so the short rhizoid phenotype of Mp*spi* mutants (Figure 3.3) may result from an inability of vesicles to fuse and deliver material for the construction of the cell surface at the tip of the rhizoid.

Nothing is known of the function of the *Arabidopsis* homologue of MpSRI1, which means there is little that can be inferred about the function of this gene and the reason for the short rhizoids of the Mp*sri1* mutant (Figure 3.4). However, I speculate that the protein may play a role in vesicle trafficking. This is based on the presence in MpSRI1 of a predicted RIC1 domain. In yeast and humans, Ric1 and Rgp1 form a complex that acts as a guanine exchange factor for the Rab GTPase Rab6 (human) or Ypt6p (yeast) and is vital for the fusion of endosome-derived vesicles with the Golgi body (Siniossoglou et al. 2000; Pusapati et al. 2012).The *Arabidopsis* homologues AT3G61480 and AT5G28350 are ubiquitously expressed and annotated as Golgi-localized (www.arabidopsis.org), which would be consistent with a role in vesicle trafficking. Thus MpSRI1 may play a similar role in vesicle trafficking in *Marchantia*.

Overall, these mutants indicate that, as in the tip-growing cells of other species, the delivery of material to build new cell surface is vital for elongation by tip growth in *Marchantia* rhizoids.

3.4.3 Novel roles in rooting cell development for four genes

In the screen I identified four genes that mutate to a short rhizoid phenotype whose relatives in *Arabidopsis* have not been implicated in root hair development.

Mutations in either of two closely-related genes, encoding proteins that contain DUF616 (Mp*SRI3* and Mp*EMB2756*), cause the development of short rhizoids (Figure 3.5). I identified three *Marchantia* genes that encode DUF616 proteins, each of which falls into a clade with one to three *Arabidopsis* genes. No function has been described for any of the *Arabidopsis* genes in these clades, apart from At*EMB2756*; in this case, all that is known is that the At*emb2756* mutant is embryo lethal (Tzafrir et al. 2004; Lloyd and Meinke 2012). The *Arabidopsis* homologues are expressed in all tissue types, which would be consistent with a broad role in cell biology (except AT2G02910, whose expression is unknown because it is not present on the Affymetrix ATH1 GeneChip) (http://bbc.botany.utoronto.ca/efp). A recent *Arabidopsis* study has provided some information on a possible function for DUF616 proteins. The DUF616 protein AtTURGOR REGULATION DEFECT 1 (AtTOD1) is an alkaline ceramidase that cleaves ceramides to produce sphingosine and a fatty acid, whose DUF616 is predicted to be the phytoceramidase domain (Chen et al. 2015). Sphingolipids, including ceramides, are important structural components of membranes and have diverse roles in cellular signalling (Hannun and Obeid 2008). At*tod1* mutants have reduced pollen tube growth *in vivo*, and defects in the regulation of turgor pressure (Chen et al. 2015). No defect was reported in root hair growth. Mp*SRI3* and Mp*EMB2756* may therefore be ceramidases and possibly act as regulators of turgor pressure, which is of great importance in tip growth (Peterson and Farquhar 1996; Hepler and Vidali 2001). Alternatively, their putative ceramidase activity may be involved in some other cellular signalling pathway, or be required for their contribution to the sphingolipid fraction of cell membranes.

The Mp*exl1* mutant develops short, misshapen rhizoids (Figure 3.6). I have not determined whether they are gain- or loss-of-function mutants, but the insertions in the alleles Giulia Morieri and I isolated are both located in the promoter, and in both the T-DNA is inserted in the orientation that causes overexpression of the downstream gene in the Mp*rsl1* gain-of-function mutants, and so it is possible that Mp*EXL1* expression is elevated in these mutants (see discussion in Chapter 4). Some studies have been carried out on *Arabidopsis EXL*s. AtEXO was identified in a promoter trap screen for genes expressed in meristematic cells, and because the loss-of-function mutation is able to suppress the meristem patterning and root growth defects of the *hydra2* mutant, a role in meristem function has been suggested (Farrar et al. 2003). A later study found that At*EXO* expression is regulated by brassinosteroid signalling and that growth is diminished throughout the plant in the At*exo* mutant (Schröder et al. 2009). None of the At*EXL* genes are expressed in a root-hair-specific manner, except possibly At*EXL4*, the expression of which is elevated in the At*rhd6* mutant, indicating that its expression might be negatively regulated in root hairs (Birnbaum et al. 2003). AtEXO is secreted and found in the cell wall, as are the other AtEXLs (Bayer et al. 2006). However, no specific molecular function has been identified for EXL proteins, which presents difficulties for interpreting the phenotype of the Mp*exl1* mutants. A further complication is the fact that all *Arabidopsis* EXL proteins are secreted, while MpEXL1 is one of the majority of *Marchantia* EXL proteins that lack a predicted signal peptide for secretion, so its molecular function is likely to differ from those of the *Arabidopsis* proteins. It is notable that this is the only family of genes I identified that has undergone an expansion in *Marchantia* relative to the other land plants, which raises the possibility that the functions of genes in this family have diversified in *Marchantia.* Unfortunately with the information currently available I cannot speculate as to the exact molecular function of Mp*EXL1*, but the shortness and misshapenness of the rhizoids suggests that it may be involved in directing the deposition of cell wall material or maintaining the direction of growth.

I found a single mutant of Mp*hsp70-1* in which rhizoid development arrests very soon after initiation (Figure 3.7), which suggests that the mutant is defective in the early stages of bulge formation or the transition to rhizoid elongation. Although the data indicates that the insertion in this line is linked to the defective phenotype (Table 3.2), the absence of further alleles means that conclusions made on the basis of this mutant must remain relatively tentative.

MpHSP70-1 belongs to the HSP110/SSE1subfamily of HSP70 proteins. HSP70 proteins possess diverse functions but generally act as molecular chaperones, assisting in the refolding of proteins (reviewed in Wang et al. 2004). Some *HSP70* genes are expressed in response to heat stress, while others are constitutively expressed; the closest *Arabidopsis* homologues of Mp*HSP70-1* are expressed in all tissues. In addition to their role as chaperones, some AtHSP70 proteins interact with heat stress transcription factors (Hsfs) and negatively regulate their activity, autoregulating the heat shock response (Lee and Schöffl 1996; Morimoto 1998). Furthermore, members of the HSP110/SSE1subfamily, to which MpHSP70-1 belongs, act as nucleotide exchange factors for members of the DnaK HSP70 subfamily in yeast and possibly in *Arabidopsis* (Raviol et al. 2006; Jungkunz et al. 2011). Perhaps the rhizoid growth defect in Mp*hsp70-1* is a result of improper folding of proteins required for the early stages of rhizoid development, or of misregulation of heat stress transcription factors leading to a constitutive heat stress response that interferes with rhizoid growth.

Chapter 4: Mp*RSL1* is a positive regulator of rhizoid differentiation in *Marchantia*

4.1 Abstract

In *Arabidopsis* and *Physcomitrella* the Class I RSL clade of bHLH transcription factors have been characterised as master positive regulators of the acquisition of root hair and rhizoid fates, respectively. In the screen for *Marchantia* mutants with abnormal rhizoid development, I identified five mutants with ectopic rhizoid development that possess mutations affecting Mp*RSL1*, and Mp*RSL1* transcript levels are increased in these mutants. The phenotype of these gain-of-function mutants suggests that, as in other species, in *Marchantia* Mp*RSL1* is a positive regulator of rooting cell fate.

4.2 Introduction

In the screen for mutants with defects in rhizoid development, I found a class of mutants in which rhizoids develop on the dorsal surface of the thallus, where they do not develop in the wild type. To identify genes that control the specification of rhizoid cell fate, I focussed on these mutants that produce ectopic rhizoids, because they mis-specify cells that would not normally become rhizoids to develop into rhizoids. Five of these mutants possess gain-of-function alleles of Mp*rsl1*, a transcription factor whose homologues control rhizoid differentiation in other land plant lineages.

4.2.1 Class I *RSL* genes are key regulators of root hair development in *Arabidopsis*

In an early mutant screen using T-DNA insertional mutagenesis in *Arabidopsis*, (Masucci and Schiefelbein 1994) identified the mutant *root hair defective 6* (*rhd6*). *rhd6* mutants are defective in root hair initiation; root hair number is reduced, and site of root hair outgrowth is shifted to a more basal position on the root epidermal cell in comparison with the wild type. Genetic analysis using double mutants between *rhd6* and other mutants known to affect root hair development indicated that *RHD6* acts downstream of the root hair patterning genes *TRANSPARENT TESTA GLABROUS* (*TTG*) and *GLABRA2* (*GL2*), but upstream of genes involved in bulge formation, the transition to tip growth, and tip growth (Masucci and Schiefelbein 1996; Parker et al. 2000).

At*RHD6* was cloned using a different allele, *rhd6-2*, generated with the *DsE* element in an enhancer trap screen (Menand et al. 2007). It was found to encode a bHLH protein, a member of a large family of transcription factors that contain a highly conserved amino acid sequence, the bHLH domain (Pires and Dolan 2010). The bHLH domain is around 60 residues long, consisting of an N-terminal basic region that mediates DNA binding, and a C-terminal HLH region that is involved in dimerization. There are 158 bHLHs in *Arabidopsis*, falling into 26 subfamilies (Pires and Dolan 2010) with a wide range of characterized functions, from regulation of cell fate determination (Nadeau 2009) to light signalling (Castillon et al. 2007). At*RHD6* belongs to bHLH subfamily VIIIc(1), members of which possess an extra conserved “RSL1” motif at the carboxy end of the bHLH domain (Pires and Dolan 2010).

There is a second member of bHLH subfamily VIIIc(1) in *Arabidopsis*, *ROOT HAIR DEFECTIVE SIX-LIKE 1* (At*RSL1*). In the wild type these two genes are expressed in the cells that will go on to develop root hairs, where they promote the expression of other genes required for tip growth, such as the closely-related subfamily VIIIc(2) bHLH *ROOT HAIR DEFECTIVE SIX-LIKE 4* (At*RSL4*) (Menand et al. 2007; Yi et al. 2010). A double mutant lacking the function of both At*RHD6* and At*RSL1* has a more severe phenotype than the At*rhd6* single mutant, and completely lacks root hairs (Menand et al. 2007). These data indicate that At*RHD6* and At*RSL1* act partially redundantly and are necessary for the initiation of root hair development. This function is specific to root hair development, rather than being necessary for tip growth in general, as the development of pollen tubes, the other tip-growing cell type in *Arabidopsis*, is normal in the At*rhd6/*At*rsl1* double mutant (Menand et al. 2007).

4.2.2 Class I *RSL* genes are key regulators of rhizoid differentiation in *Physcomitrella*

The genome of the moss *Physcomitrella patens* encodes two members of bHLH subfamily VIIIc(1), Pp*RSL1* and Pp*RSL2*. Wild type *P. patens* develops rhizoids on the base and stem of the leafy gametophore, and Pp*RSL1* and Pp*RSL2* are expressed in the progenitor cells from which these rhizoids develop (Jang et al. 2011), and mutants lacking both Pp*RSL1* and Pp*RSL2* function develop very few, short rhizoids (Menand et al. 2007; Jang et al. 2011). Conversely, the gametophores of *P. patens* that constitutively express both Pp*RSL1* and Pp*RSL2* are entirely transformed into a mass of rhizoid cells (Jang et al. 2011). Taken together, these findings indicate that the co-expression of Pp*RSL1* and Pp*RSL2* is both necessary and sufficient to induce the development of rhizoids, and so these genes are key positive regulators of moss rhizoid fate.

4.2.3 Class I *RSL* genes are part of a conserved ancient gene network that regulates rhizoid development

These experiments in *Arabidopsis* and *Physcomitrella* demonstrate that *RSL* Class I genes are key regulators of the differentiation of cells with a rooting function in both species. It has also been found that there are several other conserved components in the genetic networks that regulate rooting cell development in the two species. The *LOTUS JAPONICUS ROOTHAIRLESS1-LIKE* transcription factors belong to bHLH subfamily XI, and are positive regulators of *Arabidopsis* root hair and *Physcomitrella* rhizoid development (Pires and Dolan 2010; Tam et al. 2015). In *Arabidopsis* At*LRL3* is downstream of At*RHD6* and At*RSL1*; in contrast, in *Physcomitrella* the two *LRL* genes are transcriptionally independent of the two *RSL* genes (Tam et al. 2015). Auxin also promotes root hair and rhizoid development in these species; however, it induces the expression of different members of the network in each species. In *Physcomitrella* auxin induces the expression of both *RSL* and *LRL* genes, while in *Arabidopsis* it only induces the expression of At*LRL3* (Tam et al. 2015). Therefore several components are common in the rooting cell development regulatory networks of both species, but the connections between them have diverged.

From these findings it can be inferred that the last common ancestor of *Physcomitrella* and *Arabidopsis* possessed a regulatory network including *RSL* genes that controlled the development of tip-growing cells with a rooting function, although it remains unknown at what point in evolution this network appeared. In this chapter I present Mp*rsl1* mutants I identified in the mutant screen that provide evidence that *RSL* genes control the differentiation of rhizoids in *Marchantia*, which in turn suggests that they controlled the differentiation of tip-growing rooting cells in the last common ancestor of all land plants.

4.3 Results

4.3.1 Five insertions at the same locus cause ectopic rhizoid development

Five mutants that I selected in the screen (Chapter 3) develop a dense growth of rhizoids on their dorsal surface (Fig 4.1 a). In four of these mutants, the mutant phenotype cosegregates with the hygromycin resistance marker in the T-DNA insertion (Table 4.1). Mutant *v8* was not successfully crossed with the WT.

		Mut		**WT**		**total**		**HygS:HygR**		
line name	**phenotype**	**HygR**	**HygS**	**HygR**	**HygS**	**HygR**	**HygS**	**ratio**	**Insert. no.**	**χ2 p**
vj3	ectopic rhizoids	85	0	0	130	85	130	1: 0.7	1	0.00
vj4	ectopic rhizoids	104	0	45	67	149	67	1: 2.2	2	0.04
vj5	ectopic rhizoids	76	0	0	97	76	97	1: 0.8	1	0.11
vj6	ectopic rhizoids	116	0	0	102	116	102	1: 1.1	1	0.34

Table 4.1 Cosegregation data for ectopic rhizoid mutants, showing numbers of F1 progeny of a cross between the mutant and wild-type displaying mutant (Mut) or WT phenotype that are hygromycin resistant (HygR) and hygromycin sensitive (HygS). Insert. no. indicates the inferred number of T-DNA insertions present in the genome. χ^2 p gives the probability that the observed ratio of HygR:HygS is not different from the ratio expected from the inferred number of T-DNA insertions, as inferred from a χ^2 test.

TAIL-PCR indicated that all five of these mutants possessed T-DNA insertions in the same region, upstream of a predicted gene (Fig. 4.1 b). The insertion in *vj5* is the furthest from the gene, over 31 kb before the CDS of the gene, and there are no predicted genes in the intervening region. A cross between mutants *vj5* and *vj6* indicates that these mutants are allelic, as they fail to complement the defective mutant phenotype (100% of 203 F1 progeny develop rhizoids on the dorsal surface).

There are a further 9 mutants with ectopic rhizoid development identified by Suvi Honkanen and Giulia Morieri that have insertions in the promoter of this gene, for a total of 14 alleles. All of the insertions identified 5’ to the gene are in the same orientation, with the RB of the T-DNA closest to the CDS of Mp*RSL1*.

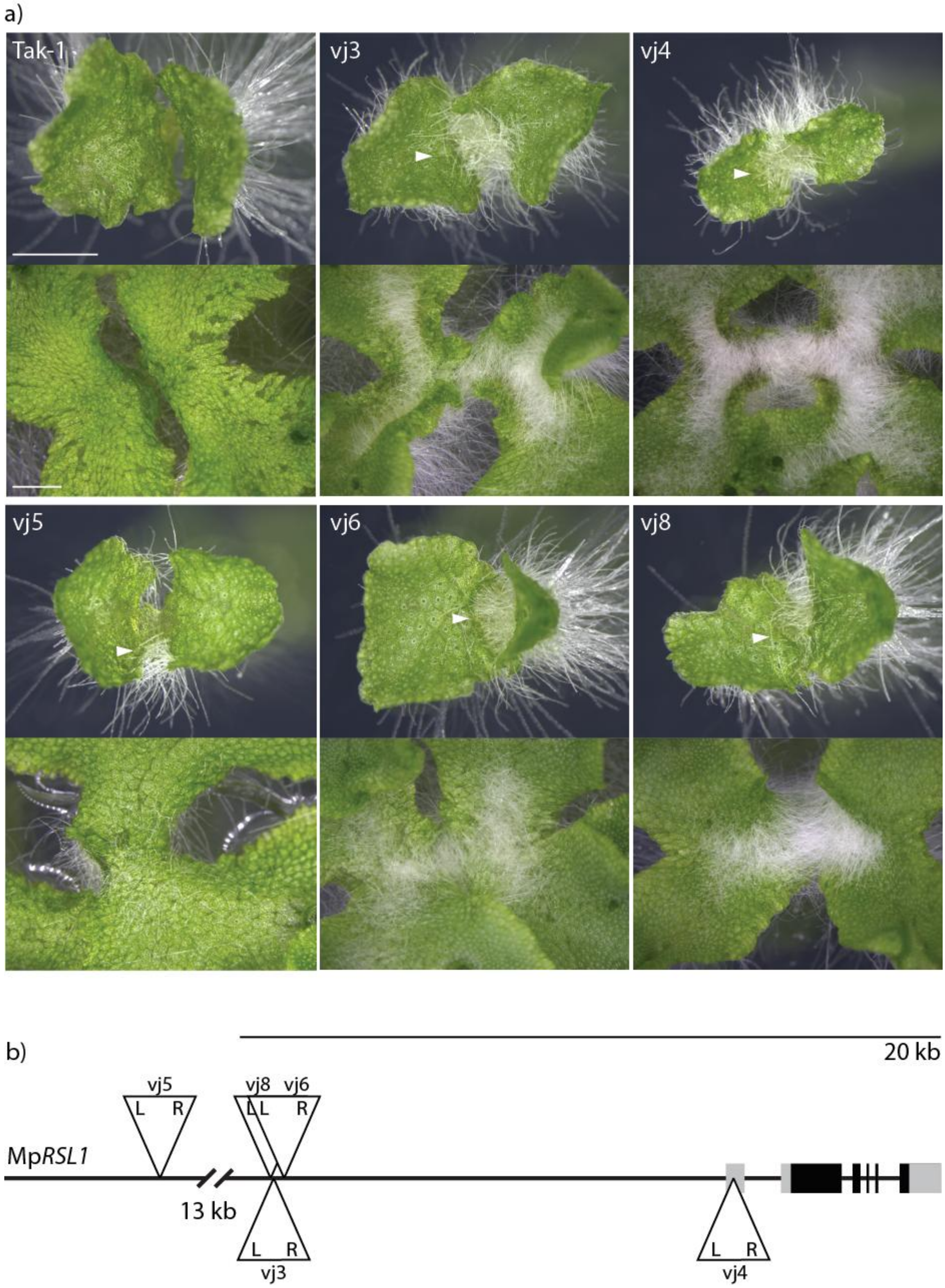


Figure 4.1 Five mutants that develop ectopic rhizoids have mutations in Mp*RSL1*. a) Mp*rsl1* mutants develop ectopic rhizoids on the dorsal surface of the thallus, unlike the wild type. Arrows indicate ectopic rhizoids. Upper row 11 d old gemmalings, lower row 25 d old gemmalings, scale 2 mm. b) Gene structure of Mp*RSL1*, with the insertion site in each allele marked. L and R indicate the orientation of the left and right borders of the T-DNA, respectively. Boxes represent exons; grey are untranslated, black are CDS.

4.3.2 The mutated gene encodes Mp*RSL1*

The protein encoded by this gene contains a bHLH domain, with an “RSL” motif at the carboxy end, which indicates that it is a member of the VII(c) subfamily of bHLHs (Figure 4.2 a). To test this hypothesis I constructed a phylogenetic tree, which demonstrates that this gene falls into a clade that contains At*ROOT HAIR DEFECTIVE6* (*AtRHD6*) and At*ROOT HAIR DEFECTIVE SIX-LIKE1* (*AtRSL1*) (Figure 4.2 b). For these reasons I designated this gene Mp*RSL1*. The *Marchantia* genome encodes one class I *RSL* gene (Mp*RSL1*) and one class II *RSL* gene (Mp*RSL2*).

4.3.3 Mp*RSL1* expression is elevated in these mutants

Because the insertions in these lines are outside of the CDS, I hypothesized that they might increase or decrease the expression of Mp*RSL1*, rather than being complete loss-of-function mutations. To ascertain the effect of these mutations on the levels of transcript, I carried out qRT-PCR using primers specific to Mp*RSL1*. This indicated that the level of Mp*RSL1* transcript is elevated in the mutant lines compared to wild type (Figure 4.2 c). I performed only one biological repeat, but these results have been confirmed in further repeats performed by Hélène Proust. Therefore these mutations are gain-of-function Mp*rsl1* alleles (Mp$rsl1^{GOF}$).

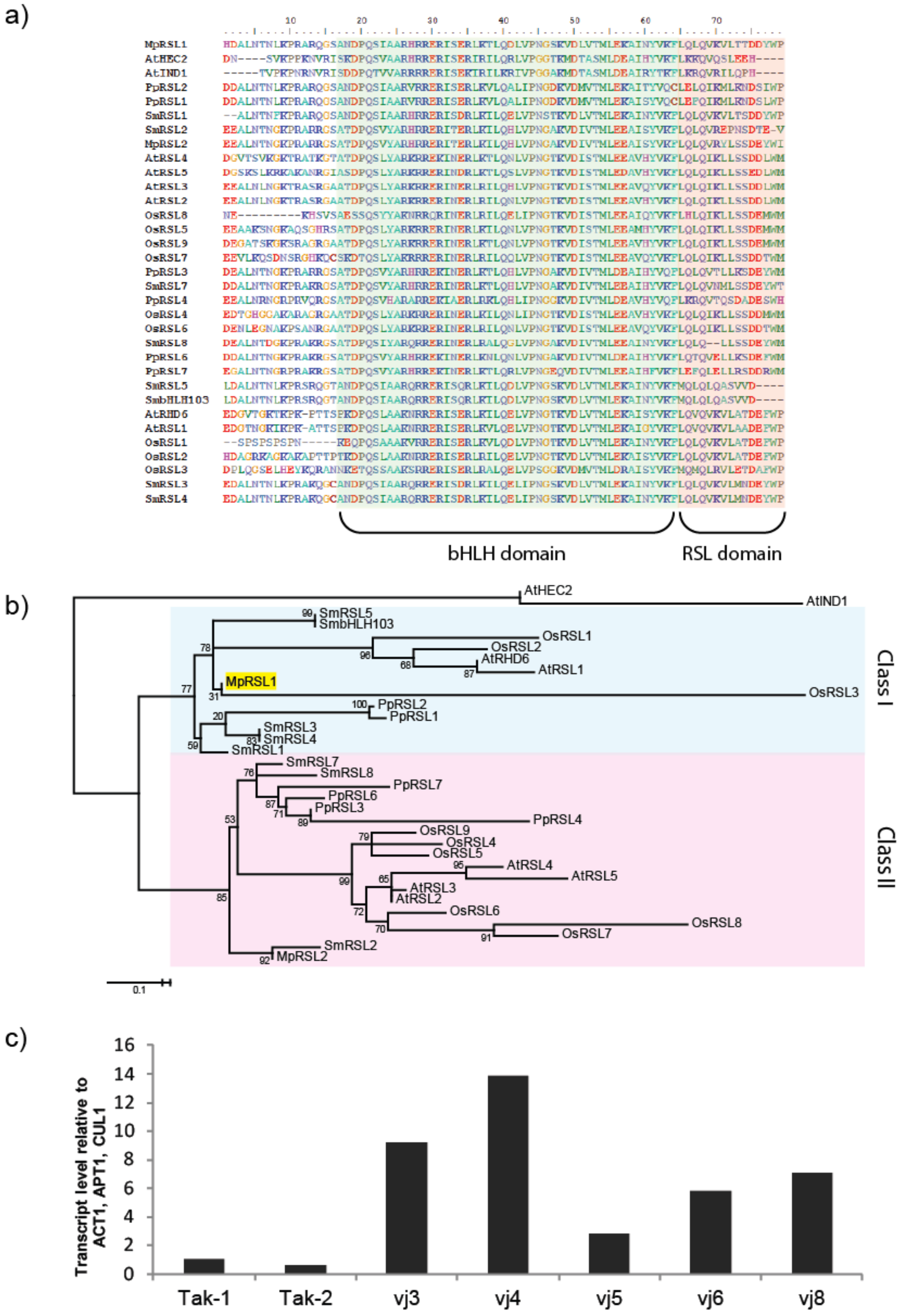


Figure 4.2 **Ectopic rhizoid mutants carry gain-of-function mutations in *MpRSL1*.** a) Alignment of RSL proteins from *Marchantia polymorpha* (Mp), *Physcomitrella* patens (Pp), *Selaginella moellendorfii* (Sm), *Oryza sativa* (Os) and *Arabidopsis thaliana* (At), along with *Arabidopsis* HECATE2 and INDEHISCENT1. b) Maximum-likelihood phylogeny of RSL proteins. Nodes are marked with aLRT values. c) Transcript levels of Mp*RSL1* in wild type and Mp*rsl1* mutants, normalized to Tak-1, using Mp*ACT*, Mp*CUL* and Mp*APT* as reference genes. 11 d old gemmalings.

4.4 Discussion

4.4.1 Mp*RSL1* is a positive regulator of rhizoid development

In order to identify regulators of rhizoid specification and differentiation, out of the mutants I selected in the screen I focussed on ones in which the wild type patterning of rhizoids was disrupted. Amongst these were five gain-of-function alleles of Mp*rsl1*, in which rhizoids develop in ectopic locations. In all of them the insertion lies in the promoter or 5' UTR of the gene (Figure 4.1 b), and in all of them the level of Mp*RSL1* transcript is elevated compared to the wild type (Figure 4.2 c). I therefore designate them gain-of-function mutants of Mp*RSL1* (Mp$rsl1^{GOF}$). The elevated Mp*RSL1* transcript levels may be caused by elevated expression in the cells that express Mp*RSL1* in the wild type, a broader expression domain than in the wild type, or both. There is a correlation between the level of Mp*RSL1* expression and the degree of ectopic rhizoid development; *vj5* has the lowest level of Mp*RSL1* expression among the mutants, and also the least pronounced ectopic rhizoid development. Conversely, *vj4* expresses Mp*RSL1* to the highest level and produces the most luxuriant growth of rhizoids on the dorsal surface (Figure 4.1 a, Figure 4.2 c). Assuming that the elevated levels of Mp*RSL1* transcript I observe in the mutants are due at least in part to ectopic expression, this indicates that ectopic expression of Mp*RSL1* is sufficient to drive a cell to adopt the rhizoid developmental fate.

The Mp$rsl1^{GOF}$ mutants identified in this screen therefore suggest that Mp*RSL1* may be a positive regulator of rhizoid development. This hypothesis is supported by loss-of-function Mp*rsl1* mutants identified by Suvi Honkanen and characterised by Hélène Proust, which do not develop any rhizoids (unpublished). Because they indicate that Mp*RSL1* is both necessary and sufficient to direct cells to differentiate as rhizoids, these two classes of mutant identify Mp*RSL1* as a key early-acting positive regulator of the adoption rhizoid fate.

4.4.2 T-DNA insertions can cause gain-of-function mutations in *Marchantia*

The large number of gain-of-function alleles of Mp*rsl1* identified in the screen was unexpected. It is also notable that in the Mp*rsl1*^{GOF} alleles in which the T-DNA was located by TAIL-PCR, all insertions were in the same orientation relative to the Mp*RSL1* gene, *viz.*, with the RB of the T-DNA insertion closest to the gene. I can only speculate on a mechanism that could cause increased expression of genes 3' to the RB of the insertion. In the construct used in the mutagenesis, between the LB and RB lies a hygromycin resistance gene under the control of the CaMV 35S promoter, which drives strong constitutive expression, oriented as follows: RB – $_{pro}$*35S:HygR:Terminator* – LB. It may be the case that enhancer elements in the 35S promoter increase the expression of *Marchantia* genes, even at a distance of more than 30 kb, but it is not clear why this should only happen when the T-DNA is inserted in one particular orientation relative to the affected gene.

I only have evidence that this might be the case for insertions upstream of Mp*RSL1*. However, it may be fruitful to hypothesize that alleles of other genes in which the T-DNA insertion is orientated in the same fashion as discussed here may also be causing elevated expression of the affected gene, that is, that they may be gain-of-function mutations.

4.4.3 The *RSL* gene regulatory network likely promoted rooting cell development in the earliest land plants

A gene regulatory network that includes Class I *RSL* genes as a key component controls the development of root hairs in the flowering plant *Arabidopsis* and rhizoids in the moss *Physcomitrella* (Menand et al. 2007; Jang et al. 2011; Tam et al. 2015). From this it has been inferred that the last common ancestor of mosses and flowering plants, sometime before 420 mya, possessed cells with a rooting function whose development was controlled by a gene regulatory network that included Class I *RSL* genes (Tam et al. 2015).

The finding that the Class I *RSL* gene of *Marchantia*, Mp*RSL1*, is also a key positive regulator of rhizoid development suggests that this mechanism of controlling the development of rooting cells was already present in the last common ancestor of all land plants, around 470 mya. It also supports the hypothesis that tip-growing rooting cells originated only once in the ancestor of all land plants, rather than arising separately in the lineages that lead to liverworts and to the other land plants. Further work in the algal relatives of land plants will be necessary to constrain the earliest point at which *RSL* genes began to control rooting cell development. It may have happened already in the common ancestor of land plants and some of their algal relatives, or it may have arisen after the divergence of the land plant lineage from the other streptophytes. The latter possibility would hint that tip-growing cells with a rooting function, and the elaboration of a gene regulatory network to control their development, was an innovation that occurred during, and may have been vital for, the conquest of the land by plants.

It should be noted, however, that homologues of the other known components of the *RSL* gene regulatory networks of *Arabidopsis* and *Physcomitrella* were not identified as important for rhizoid development in this screen, and that as yet no network connections are known between Mp*RSL1* and other genes. These include auxin response genes, *LRL* transcription factors and Class II *RSLs* (Pires et al. 2013; Tam et al. 2015), all of which are present in the *Marchantia* genome (Holger Breuninger and Helen Proust, personal communication; Flores-Sandoval, Eklund, et al. 2015) . There are several possible reasons for this. Functional redundancy between closely-related members of gene families is unlikely, as only a single *LRL*, a single class II *RSL*, and a single member of each ARF class are present in the genome. Mutations in these genes may cause a lethal phenotype, or a general growth defect in which any specific effect on rhizoid development is not apparent, which may be especially likely because of the probable lack of functional redundancy. Alternatively, it may simply be the case that these genes do not mutate to a defective rhizoid phenotype because they are not members of the gene regulatory network that controls their development. Reverse-genetics approaches to target these genes would thus shed light on the wider regulatory network that controls rhizoid development in *Marchantia*, and so trace how the components and connections in the rooting-cell developmental network have changed over the course of evolution.

Chapter 5: Mp*WIP* regulates specialized epidermal cell fate in *Marchantia*

5.1 Abstract

In the screen I identified a mutation in a potential regulator of rhizoid initiation. I isolated a single mutant (*vj7*) that develops ectopic rhizoids and is not a mutant allele of Mp*RSL1*. The T-DNA is inserted 5' of the Mp*WIP* gene. WIP proteins are a family of zinc finger transcription factors that have been implicated in diverse developmental processes in flowering plants. There is a single WIP–coding gene in the *Marchantia* genome. The level of Mp*WIP* transcript is elevated in mutant *vj7*, and expression of Mp*WIP* under the constitutive 35S promoter caused the development of ectopic rhizoids and elevated Mp*RSL1* expression. This suggests that increased expression of Mp*WIP* causes the ectopic rhizoid phenotype. The Mp*WIP* promoter is expressed in every cell type in the ventral thallus. In the young dorsal thallus it is preferentially active in the developing air pores. Plants expressing the CDS of Mp*WIP* fused to the SRDX domain to create a dominant transcriptional repressor develop abundant ectopic rhizoids on the dorsal surface of the thallus. This suggests that MpWIP promotes the adoption of rhizoid fate through the transcriptional repression of target genes. Plants transformed with a constitutively-expressed amiRNA targeting Mp*WIP* have defects in the development of the thallus, including the differentiation of air pores, which suggests that MpWIP is required for air pore development. From these data I tentatively conclude that Mp*WIP* may be a positive regulator of rhizoid fate determination, and may also be required for the differentiation of the air pore complex.

5.2 Introduction

I identified six mutants that develop ectopic root hairs on the dorsal surface of the thallus. Five of the six mutants are gain-of-function mutants of Mp*RSL1*, a homologue of known positive regulators of rhizoid differentiation in other species. The sixth mutation is unlinked to the MpRSL1 gene. The T-DNA is inserted in the 5’ region of Mp*WIP*. In this chapter I present evidence that Mp*WIP* is another component of the gene regulatory network that controls rhizoid differentiation, and a member of a gene family whose members have not previously been implicated in rooting cell development.

5.2.1 *WIP* genes encode a small family of zinc finger proteins

The *WIP* genes encode a small, plant-specific subfamily of C2H2 zinc finger proteins that belong to subclass A1d (Sagasser et al. 2002; Petricka et al. 2008; Appelhagen et al. 2010). Many C2H2 zinc finger proteins bind DNA and act as transcription factors, though some recognise and bind RNA or other proteins (Wolfe et al. 2000; Lee et al. 2006; Gamsjaeger et al. 2007). In *Arabidopsis* the WIP subfamily consists of six proteins that share a highly conserved domain, called the WIP domain after its first three strictly conserved residues (Sagasser et al. 2002). The WIP domain consists of four zinc finger motifs, the second and fourth of which contain nuclear localization signals. All six *Arabidopsis* WIP proteins localize to the nucleus, and at least in the case of AtTRANSPARENT TESTA 1 (AtTT1) this relies on the NLS in the WIP domain (Appelhagen et al. 2010). Outside of the WIP domain are a number of variable motifs that are not conserved in all family members (Appelhagen et al. 2010).

5.2.2 *WIP* genes function in diverse developmental processes

Several *WIP* genes are regulators of different developmental processes in *Arabidopsis*. The first *WIP* to be identified was At*TT1*, also known as At*WIP1*. The seeds of At*tt1* mutants have morphological defects in the endothelium, and they are pale because of a lack of condensed tannins in the seed coat (Sagasser et al. 2002). At*TT1* is specifically expressed in the endothelium and controls the expression of enzymes of the flavonoid biosynthesis pathway (Sagasser et al. 2002; Appelhagen et al. 2011). Ectopic expression of At*TT1* under the CaMV 35S promoter causes pleiotropic developmental defects, including aberrant, narrow cotyledons; lanceolate, serrated leaves; delayed flowering; abnormal floral morphology and sterility (Sagasser et al. 2002)

At*WIP6* is also known as At*DISORGANIZED TRIBUTARIES 5* (At*DOT5*). The At*dot5* mutant was identified in a screen for defects in vein patterning; At*dot5* cotyledons and leaves lack higher order venation, and the veins that do develop are not tightly aligned (Petricka et al. 2008). The developmental defects of At*dot5* mutants extend beyond vein patterning: phyllotaxy is altered and leaf initiation is delayed, apical dominance is reduced and the roots of the mutant are short.

At*WIP2*, also known as At*NO TRANSMITTING TRACT*, is essential for the development of the transmitting tract, a specialized tissue of the carpel through which the pollen tube grows to reach the ovule (Crawford et al. 2007). It was initially reported that the gene is expressed specifically in the transmitting tract, but further work found that it is also expressed in the root meristem. The roots of the At*ntt* single mutant develop normally, but a triple mutant that additionally lacks the function of the two closest relatives of At*ntt*, At*wip4* andAt*wip5*, lacks roots entirely (Crawford et al. 2015). In the wild type embryo the hypophysis divides asymmetrically at the 32-cell stage to form a basal daughter cell and a lens-shaped cell that gives rise to the quiescent centre (QC). In the triple mutant this asymmetric division does not take place, the lens-shaped cell is not produced, the QC is not established, and so the root meristem does not form (Crawford et al. 2015). Consistent with a role in promoting this asymmetric division, in the embryo At*NTT* is expressed in the hypophysis and suspensor, while At*WIP4* and At*WIP5* are expressed only in the hypophysis. All three *WIP* genes are direct targets of the auxin response factor AtMONOPTEROS (AtMP/AtARF5), and *WIP* expression is not seen in the hypophysis of At*mp* mutant embryos (Crawford et al. 2015). In the mature root all three genes are expressed in the QC and columella initials of the meristem, suggesting that they may play a role in the maintenance or functioning of the root meristem. The development of roots in the triple mutant can be rescued by the application of auxin, but these rescued roots fail to pattern distal stem cell fate, and the expression of several QC marker genes is abnormal (Crawford et al. 2015). These data indicate that At*NTT*, At*WIP4* and At*WIP5* are redundantly necessary for the proper development of distal stem cells and the QC in the root meristem. Furthermore, ectopic expression of At*NTT* is sufficient to cause other populations of stem cells, such as cotyledon primordia, to adopt distal stem cell fate (Crawford et al. 2015). All together these findings establish At*NTT* as a key regulator of basal and distal stem cell identity in the embryo and root, respectively, as well as being vital for the differentiation of the transmitting tract during floral development.

All *Arabidopsis WIP* genes have been shown to regulate developmental processes, except for At*WIP3*, about which no functional data are available. The only other species in which a *WIP* gene has been studied is the melon, *Cucumis melo*, in which it controls sex determination. Expression of Cm*WIP1* in carpel primordia causes the carpel to abort, and indirectly represses the expression of an inhibitor of stamen development (Martin et al. 2009). In melon, too, a *WIP* gene is a key developmental regulator.

5.2.3 WIP proteins may act as transcriptional repressors

The molecular mechanism by which WIP proteins act has not been thoroughly investigated, although there is some evidence that they act as transcriptional repressors of target genes. At*TT1* promotes the expression of enzymes of the flavonoid biosynthetic pathway, but AtTT1 was not able to activate the promoters of members of the pathway directly (Appelhagen et al.

2011). It could have been the case that this failure was due to the absence of interacting factors that are required for activation, as AtTT1 is known to interact with some proteins. Yeast two-hybrid and bimolecular fluorescence complementation experiments found that AtTT1 can homodimerize and interact with the R2-R3 MYB transcription factors AtTT2 and AtPAP1, which form complexes with basic helix-loop-helix (bHLH) and WD40 repeat-containing proteins to regulate the expression of target genes (Appelhagen et al. 2011). A familiar example of such a complex is the AtWER/AtCPC-AtGL3/AtEGL3-AtTTG complex that controls root hair patterning in *Arabidopsis*. However, a fusion of AtTT1 and the VP16 activation domain also could not activate the promoters of the enzymes, suggesting that the absence of an activating partner was not the cause of the failure (Appelhagen et al. 2011). It is also possible that AtTT1 alone is not able to bind DNA, but can only do so as part of a complex.

This of course does not rule out the possibility that AtTT1 promotes the expression of these genes indirectly. In order to explain the fact that in At*tt1* mutants the promoter of the flavonoid biosynthesis enzyme At*BANYULS* (At*BAN*) is active throughout most of the endothelium, but the proanthocyanidin it produces does not accumulate there, Debeaujon et al. (2003) hypothesized that At*TT1* might be a cadastral gene that represses the expression of a negative regulator of At*BAN* that acts post-transcriptionally, such as a microRNA. The hypothesis that AtTT1 acts as a transcriptional repressor is supported by the fact that it contains an EAR motif, an active transcriptional repressor domain (Kagale et al. 2010). EAR motifs were also identified in AtWIP4 and AtWIP5 (Kagale et al. 2010). The fact that any of the *Arabidopsis* WIP proteins, expressed under the At*TT1* promoter, was able to at least partially rescue the At*tt1* pale-seeds phenotype suggests that all AtWIPs might function in a similar way (Appelhagen et al. 2010). Despite the lack of direct evidence of the repression of target genes by WIP proteins, published data are consistent with the hypothesis that *WIP* genes encode transcriptional repressors.

5.3 Results

5.3.1 Identification of a second regulator of rhizoid development

One of the mutants I identified in the mutant screen, *vj7*, develops sparse rhizoids from the dorsal epidermis on the older parts of the thallus (Figure 5.1 a), which suggested that it has a mutation in a gene that controls the adoption of rhizoid fate. The mutant phenotype is linked to a T-DNA insertion, but the cosegregation data indicated that there was more than one T-DNA segregating in the F1 generation (Chapter 3, Table 3.2). Because I had identified five mutants with ectopic rhizoids that possessed gain-of-function alleles of Mp*rsl1*, I tested whether the mutation in *vj7* was allelic to Mp*rsl1* by crossing mutant *vj7* to the Mp*rsl1*^{GOF} mutants *vj4* and *vj5*. If the causative mutations in these mutants were allelic, they should fail to complement each other and all F1 progeny would be expected to develop ectopic rhizoids. If they were not allelic and were unlinked, a quarter of the F1 progeny should inherit the wild type allele of both loci and display the wild type phenotype. In the *vj4 x vj7* F1 generation 64 of 297 plants (22 %) displayed the wild type phenotype, and in the *vj5 x vj7* F1 generation 52 of 186 plants (28%) displayed the wild type phenotype. Since the mutations in *vj7* and the known Mp*rsl1*$^{\mathrm{GOF}}$ mutants successfully complemented each other in the F1 generation I concluded that the mutation causing the defective phenotype in mutant *vj7* is not allelic to Mp*rsl1*.

To locate the T-DNA insertion that is linked to the mutant phenotype I used TAIL-PCR and identified insertions at two locations in the genome, the first 764 bp upstream of the transcriptional start site of a predicted protein coding gene (Figure 5.1 b), the second with no predicted genes within 4 kb in either direction. From F1 progeny of a cross between *vj7* and Tak-1 I selected 72 hygromycin-resistant individuals with the mutant phenotype and 34 with the wild type phenotype. I genotyped these plants for both insertions, and found that all 72 mutant-phenotype plants possessed the first insertion, while none of the 34 wild type-phenotype plants possessed it. The second insertion was present in all 34 wild-type phenotype plants, but only in 29 of the 72 mutant-phenotype plants. I thus concluded that the T-DNA insertion linked to the mutant phenotype was the one 764 bp upstream of the predicted protein coding gene illustrated in Figure 5.1 (b), while the second insertion was not linked to the mutant phenotype.

In all the Mp*rsl1*^{GOF} mutants I isolated, the T-DNA is located upstream of the CDS, with the right border facing towards the CDS, suggesting that insertion in this orientation may cause increased expression of the downstream gene (see Chapter 4). In *vj7* the T-DNA is inserted in the same orientation, and so I hypothesized that expression of the downstream gene might be elevated. I used RT-qPCR to measure the level of the transcript of the gene in mutant *vj7* and in the wild type. In *vj7* the level of transcript was 165% of that in the wild type (Figure 5.1 c). Since an elevated level of the transcript is associated with the development of ectopic rhizoids, this suggested that the gene might encode a positive regulator of rhizoid differentiation.

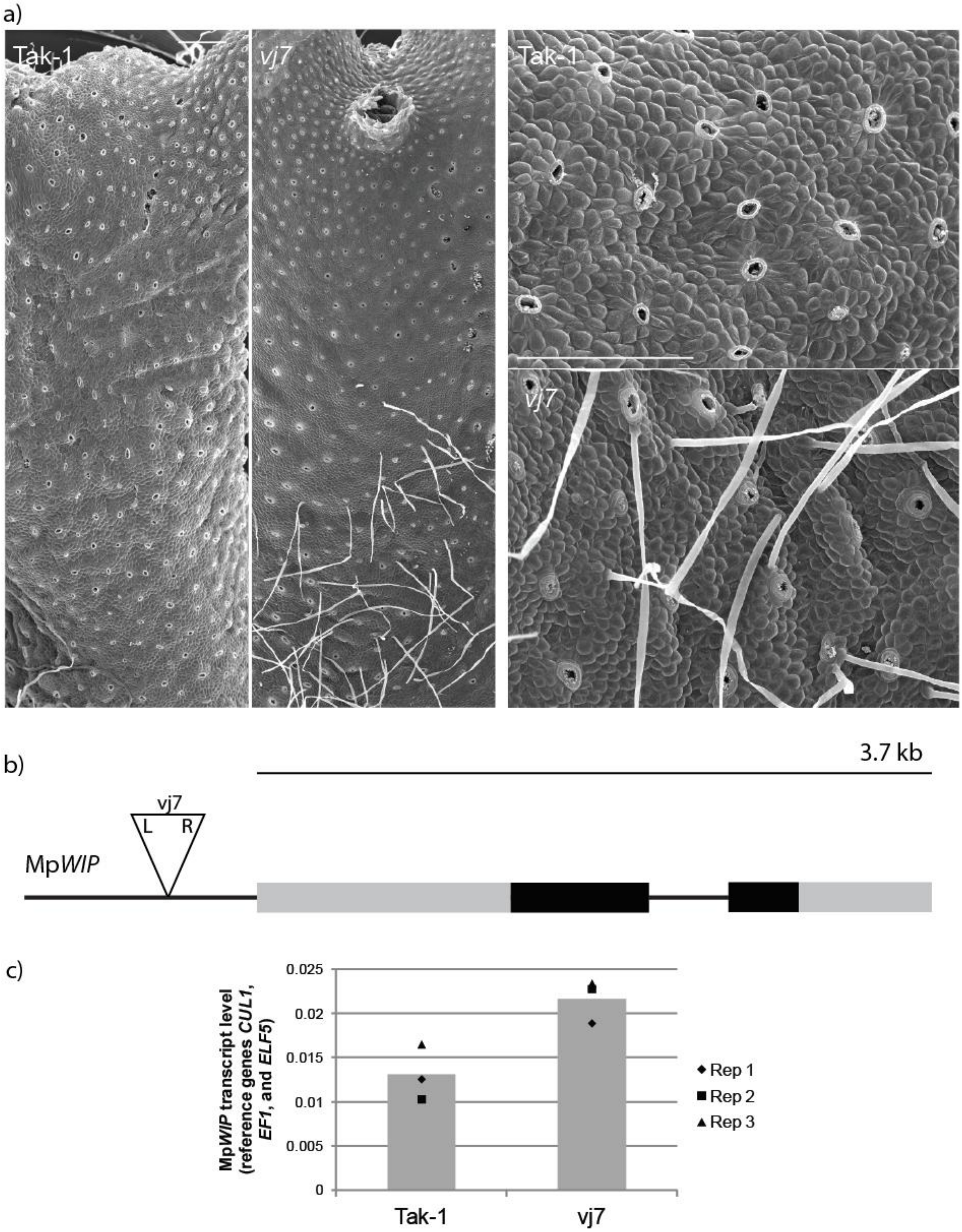


Figure 5.1 Mutant *vj7* develops ectopic rhizoids on the dorsal surface and possesses a gain-of-function mutation in Mp*WIP*. a) Rhizoids develop on the dorsal surface of old parts of the thallus of *vj7*, but not of Tak-1. SEM, 43d old gemmalings, scale 500 µm. b) Gene structure of Mp*WIP*, with the insertion site in mutant *vj7* marked. L and R indicate the orientation of the left and right borders of the T-DNA, respectively. Boxes represent exons; grey are untranslated, black are CDS. c) Mp*WIP* transcript level is elevated in mutant *vj7* compared to Tak-1. 21 d old gemmalings.

5.3.2 The gene downstream of the linked insertion in *vj7* encodes the zinc finger transcription factor MpWIP

To identify homologues of the gene downstream of the insertion linked to the mutant phenotype in mutant *vj7* I used the translation of the longest open reading frame to query the *Arabidopsis* genome using the tblastn algorithm (see Appendix 4 for sequences). The most similar match (71% identity) was At*NO TRANSMITTING TRACT* (At*NTT*), a member of the WIP family of zinc finger transcription factors. I created an alignment of the WIP domain of MpWIP with those of all WIP proteins from *Coleochaete nitellarum, Physcomitrella patens, Selaginella moellendorffii, Oryza sativa, Aquilegia caerulea* and *Arabidopsis thaliana,* and the homologous region of the related non-WIP C2H2 zinc finger transcription factors from *Arabidopsis,* AtSENSITIVE TO PROTON RHIZOTOXICITY 1 (AtSTOP1), AtJACKDAW (AtJKD) and AtINDETERMINATE(ID)-DOMAIN 7 (AtIDD7). The WIP domain defined by Appelhagen et al. (2010) is highly conserved among the WIP proteins (Figure 5.2 a), including the *Marchantia* protein (94% identity to AtNTT), though conservation is poor outside this domain.

I used this WIP domain alignment to construct a maximum-likelihood gene tree (Figure 5.2 b). The *Marchantia* protein falls into a well supported clade containing all the WIP proteins (aLRT value = 92). I was only able to identify one WIP protein encoded in the *Marchantia* genome. Resolution within the WIP clade is lacking in parts of the tree, likely because of the high degree of sequence similarity in the conserved WIP domain. Because it encodes a protein that contains the conserved WIP domain and falls within the WIP clade, I gave the name Mp*WIP* to the gene downstream of the insertion linked to the mutant phenotype in *vj7.* Since the level of Mp*WIP* transcript is increased in mutant *vj7*, I designated it Mp*wip*$^{\mathrm{GOF}}$.

AtTT1, AtWIP4 and AtWIP5 contain the transcriptional repressor EAR motif (Kagale et al. 2010), but I found neither the LxLxL nor DLNxxP EAR motif conserved patterns in MpWIP.

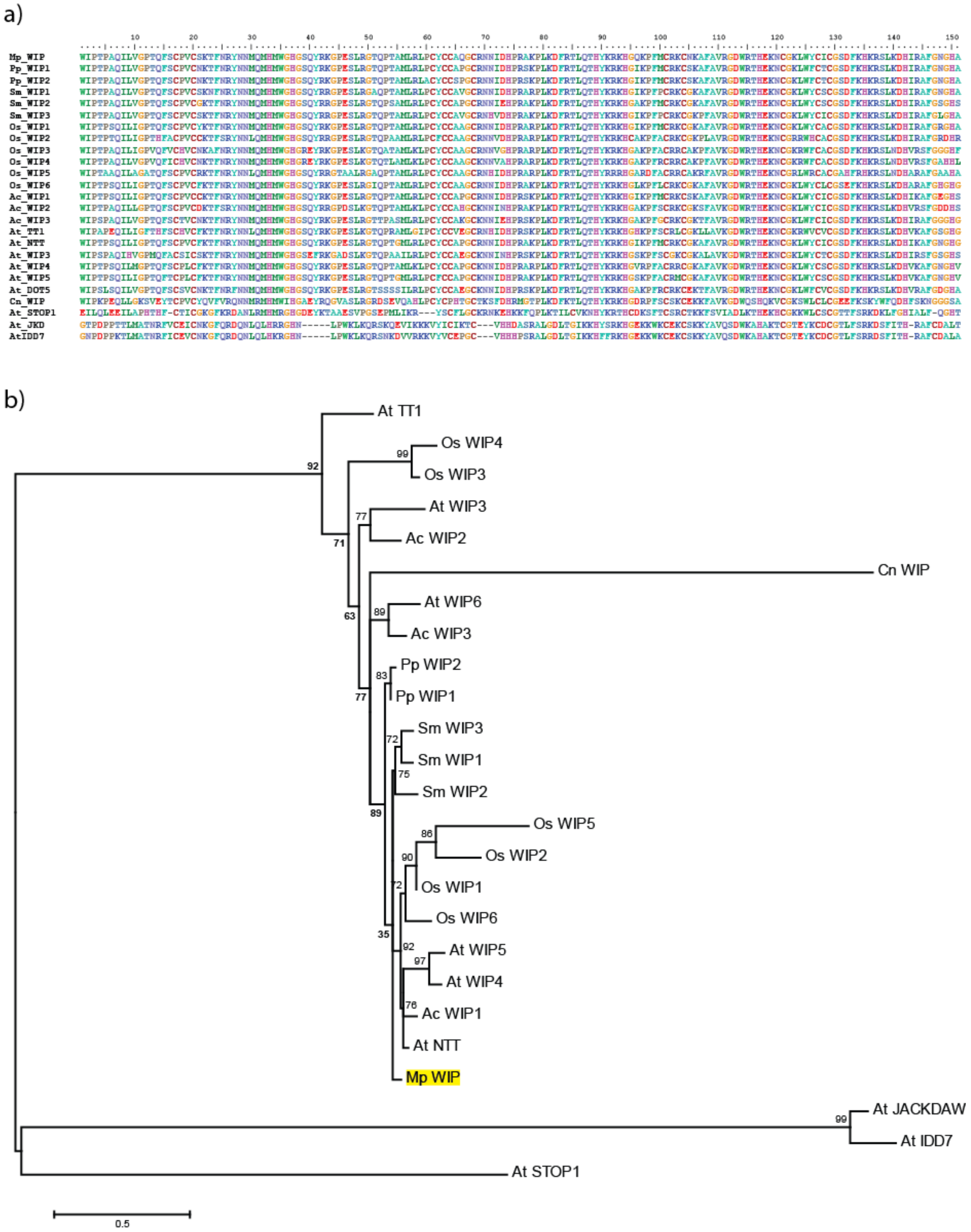


Figure 5.2 MpWIP contains a WIP domain and falls into the WIP clade of zinc finger transcription factors. a) Alignment of the WIP domain of WIP proteins from *Marchantia polymorpha*, *Coleochaete nitellarum*, *Physcomitrella patens*, *Selaginella moellendorffii*, *Oryza sativa*, *Aquilegia caerulea* and *Arabidopsis thaliana*, with AtSTOP1, AtJKD, and AtIDD7. b) Maximum-likelihood phylogeny of WIP proteins. Nodes are marked with aLRT values.

5.3.3 Expression of Mp*WIP* under the 35S promoter causes ectopic development of rhizoids

In the mutant screen I only identified a single Mp*wip* allele, Mp*wip*^{GOF}, in which the level of Mp*WIP* transcript is two thirds higher than in the wild type. To test if increased expression of Mp*WIP* causes the development of ectopic rhizoids I transformed wild type sporelings with a vector carrying $_{pro}$*35S:*Mp*WIP*, to drive the constitutive expression of Mp*WIP*. Approximately 10% of the resulting transformants developed ectopic rhizoids on the dorsal surface of the oldest part of the thallus, as seen in lines $_{pro}$*35S:*Mp*WIP*-1 and $_{pro}$*35S:*Mp*WIP*-2 (Figure 5.3 a). $_{pro}$*35S:*Mp*WIP*-3 is an example of a transformed line that does not produce ectopic rhizoids in this region.

To ascertain whether the development of ectopic rhizoids correlates with increased expression of Mp*WIP*, I carried out RT-qPCR to measure the levels of Mp*WIP* transcript. The mean transcript level of Mp*WIP* is 1.9 to 3.9 times higher in Mp*wip*^{GOF}, $_{pro}$*35S:*Mp*WIP*-1 and $_{pro}$*35S:*Mp*WIP*-2, which produce ectopic rhizoids, than in the wild type Tak-1, while in $_{pro}$*35S:*Mp*WIP*-3, which does not produce ectopic rhizoids, the level of Mp*WIP* is similar to that in Tak-1 (Figure 5.3 b). The fact that an increased mean level of Mp*WIP* transcript correlates with the production of ectopic rhizoids is consistent with the hypothesis that ectopic expression of Mp*WIP* causes the ectopic development of rhizoids. However, the variation in the levels between biological replicates of each line is great, so this conclusion should be treated with caution.

To determine whether Mp*WIP* regulates the expression of Mp*RSL1*, I measured the levels of the transcript of this gene in the wild type and the lines putatively constitutively expressing Mp*WIP* (Figure 5.3 c). Mp*RSL1* transcript levels are increased 2 to 2.6 fold compared to the wild type in the lines with ectopic rhizoids, Mp*wip*^{GOF}, $_{pro}$*35S:*Mp*WIP*-1 and $_{pro}$*35S:*Mp*WIP*-2, while they are similar to the wild type level in $_{pro}$*35S:*Mp*WIP*-3, which does not produce ectopic rhizoids. Thus increased level of Mp*RSL1* transcript correlates with the development of ectopic rhizoids and an increased level of Mp*WIP* transcript in these lines. This suggests that Mp*WIP* might induce rhizoid development by promoting the expression of Mp*RSL1* in the Mp*wip*^{GOF} mutant.

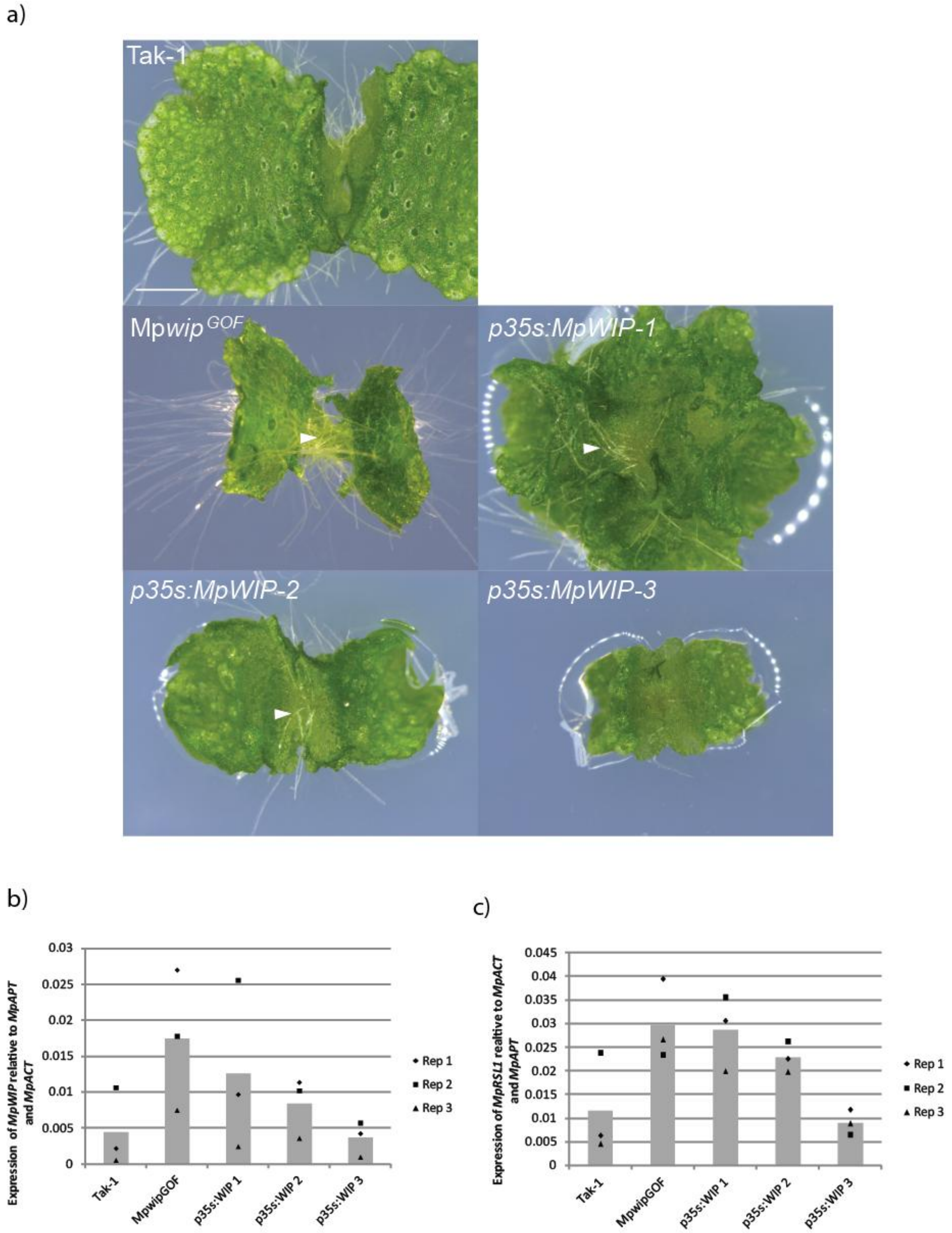


Figure 5.3 Expression of Mp*WIP* under the control of the 35S promoter causes the development of ectopic rhizoids. a) The mutant Mp*wip*^{GOF} and the transformed lines $_{pro}$*35S:*Mp*WIP*-1 and $_{pro}$*35S:*Mp*WIP*-2 develop ectopic rhizoids in the old parts of the thallus (arrowhead), while Tak-1 and $_{pro}$*35S:*Mp*WIP*-3 do not. 10 d old gemmalings, scale 1 mm. b) Transcript levels of Mp*WIP* in wild type and putative Mp*WIP* gain-of-function lines, using Mp*ACT* and Mp*APT* as reference genes. 21 d old gemmalings. c) Transcript levels of Mp*RSL1* in wild type and putative Mp*WIP* gain-of-function lines, using Mp*ACT* and Mp*APT* as reference genes. 21 d old gemmalings.

5.3.4 The Mp*WIP* promoter is active in developing air pores and all cells of the ventral epidermis

Ectopic expression of Mp*WIP* can promote the development of ectopic rhizoids, but if it functions in rhizoid development in the wild type, it should be expressed in the cells that give rise to rhizoids. To investigate where the Mp*WIP* promoter is active in the wild type I expressed 3xYFP-NLS under the control of a 4.7 kb fragment of gDNA upstream of the CDS of Mp*WIP*. In multiple transformed lines YFP expression was strongest near the growing tip of the thallus, especially in the air pores, while in older parts of the thallus it was broadly active at a lower level.

To visualize the pattern of $_{pro}$Mp*WIP* activity with cellular resolution, I imaged $_{pro}$Mp*WIP:3xYFP-NLS* using a confocal scanning laser microscope. Strikingly, on the dorsal side of the thallus the promoter was most active in cells of the developing air pore (Figure 5.4 a), with a low level of activity in the surrounding cells. Near the apex there is a low level of activity in all cells, but the activity increases greatly in a subset of cells that develop into air pores as they differentiate and divide (Figure 5.4 b).

To test whether the higher level of YFP fluorescence seen in the air pores was an artefact, I imaged a line expressing *3xYFP-NLS* under the control of the constitutive Os*ACT* promoter. In this line, fluorescence is seen in all cells, although it is less intense in the cells of the air pores than the cells of the surrounding epidermis (Figure 5.5 a, b). This supports the conclusion that the high level of YFP fluorescence seen in the air pores of $_{pro}$Mp*WIP:3xYFP-NLS 1* reflects the specific activity of the promoter in these cells.

On the ventral surface of the thallus the Mp*WIP* promoter is active in all cell types in the apical region (Figure 5.6 a, b), similar to the pattern seen with the Os*ACT* promoter (Figure 5.7 a, b) . As this is the region where rhizoids develop in the wild type, it is consistent with the hypothesis that Mp*WIP* is a positive regulator of rhizoid differentiation. In addition, Mp*WIP* is preferentially expressed in the cells of developing air pores on the dorsal surface of the thallus, which suggests that it may have a role in the development of this structure.

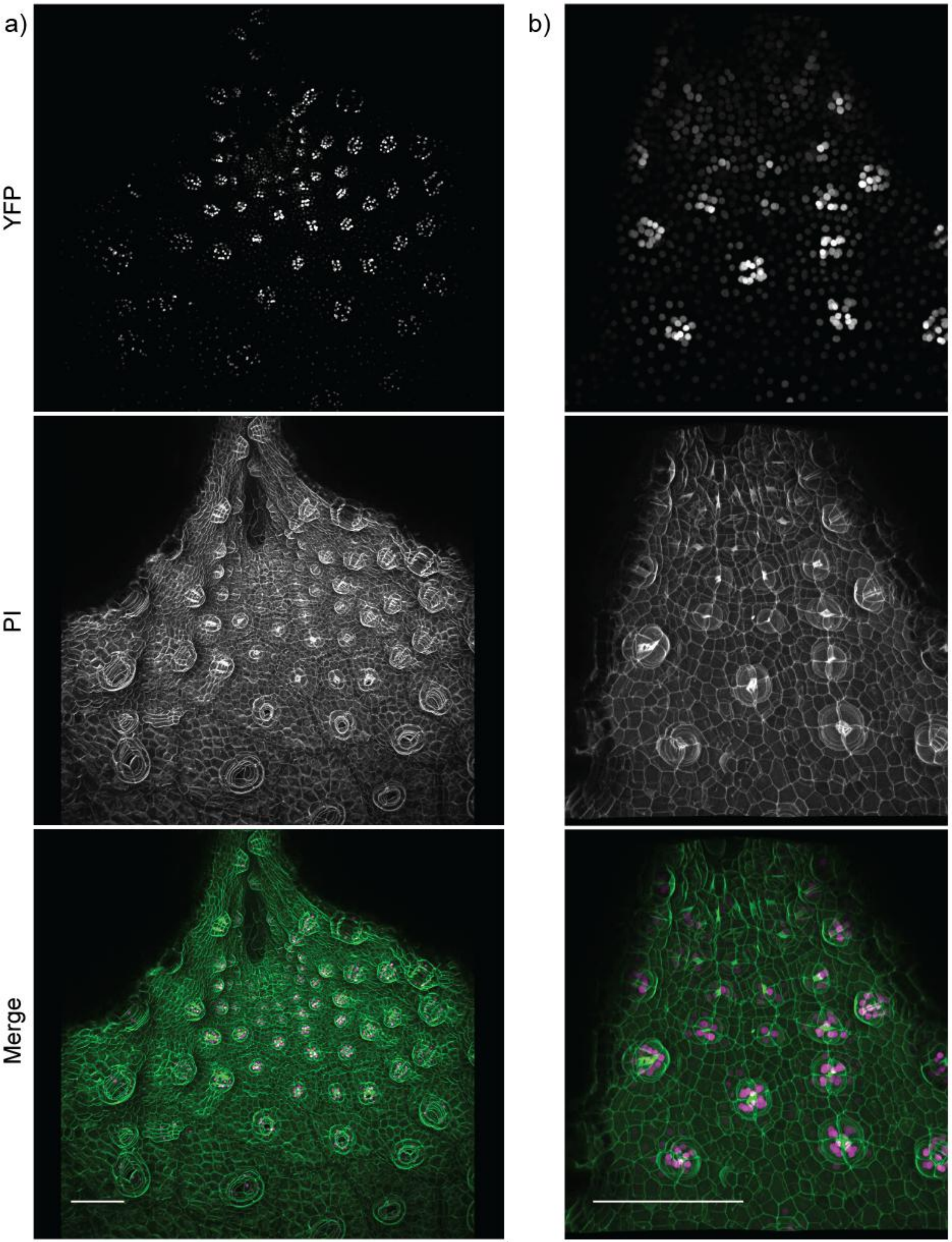


Figure 5.4 The Mp*WIP* promoter is active in developing air pores. a) Apical region of the dorsal surface of the thallus of $_{pro}$Mp*WIP:3xYFP-NLS*. Top panels show YFP fluorescence. In middle panels cell walls are stained with propidium iodide. YFP (magenta) and PI (green) signals are merged in lower panels. 9 d old gemmaling, scale 100 µm, apex at top. b) Detail of a), scale 100 µm.

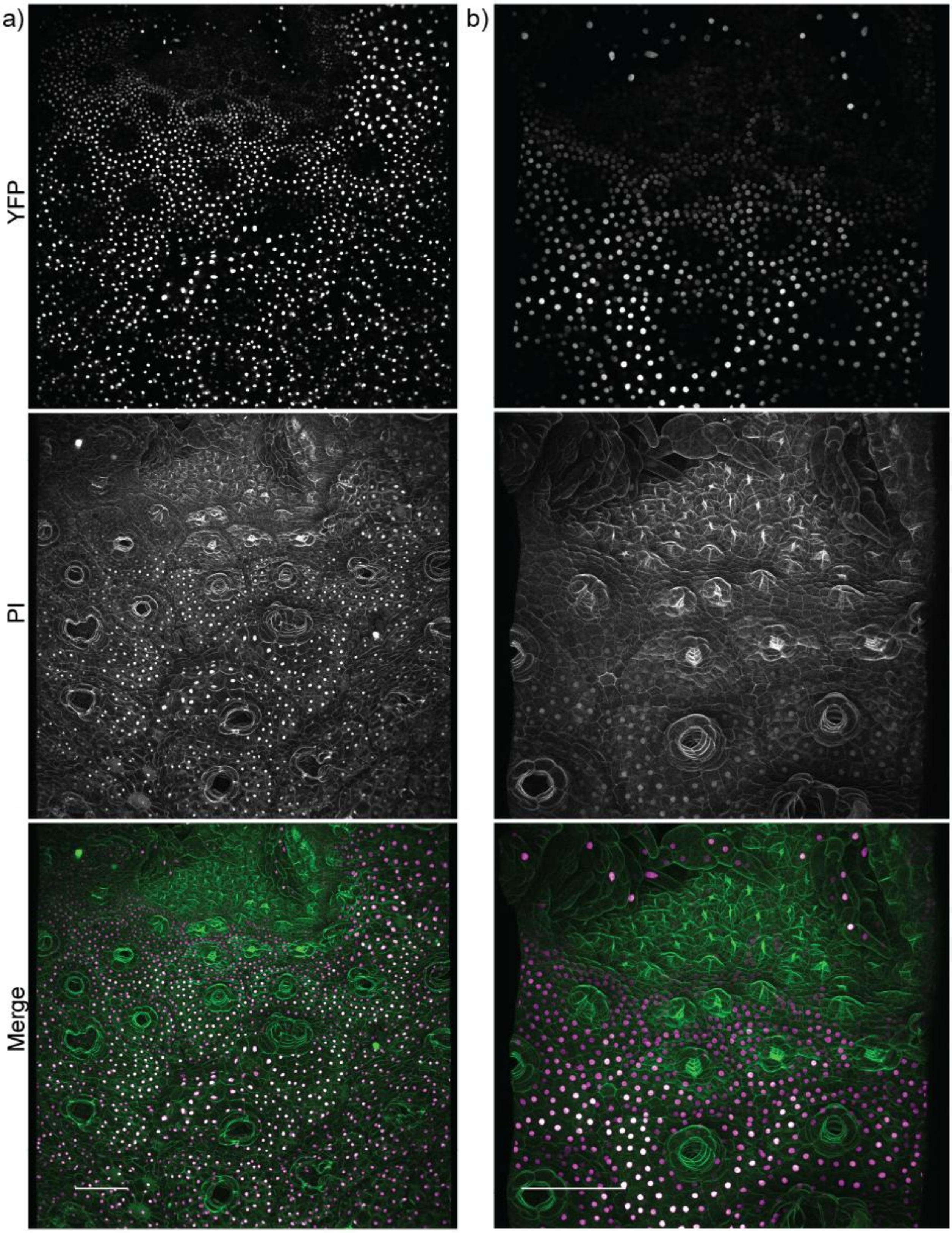


Figure 5.5 The Os*ACT* promoter is active in all dorsal cell types. a) Apical region of the dorsal surface of the thallus of $_{pro}$Os*ACT:3xYFP-NLS*. Top panels show YFP fluorescence. In middle panels cell walls are stained with propidium iodide. YFP (magenta) and PI (green) signals are merged in lower panels. 9 d old gemmaling, scale 100 µm, apex at top. b) Detail of a), scale 100 µm.

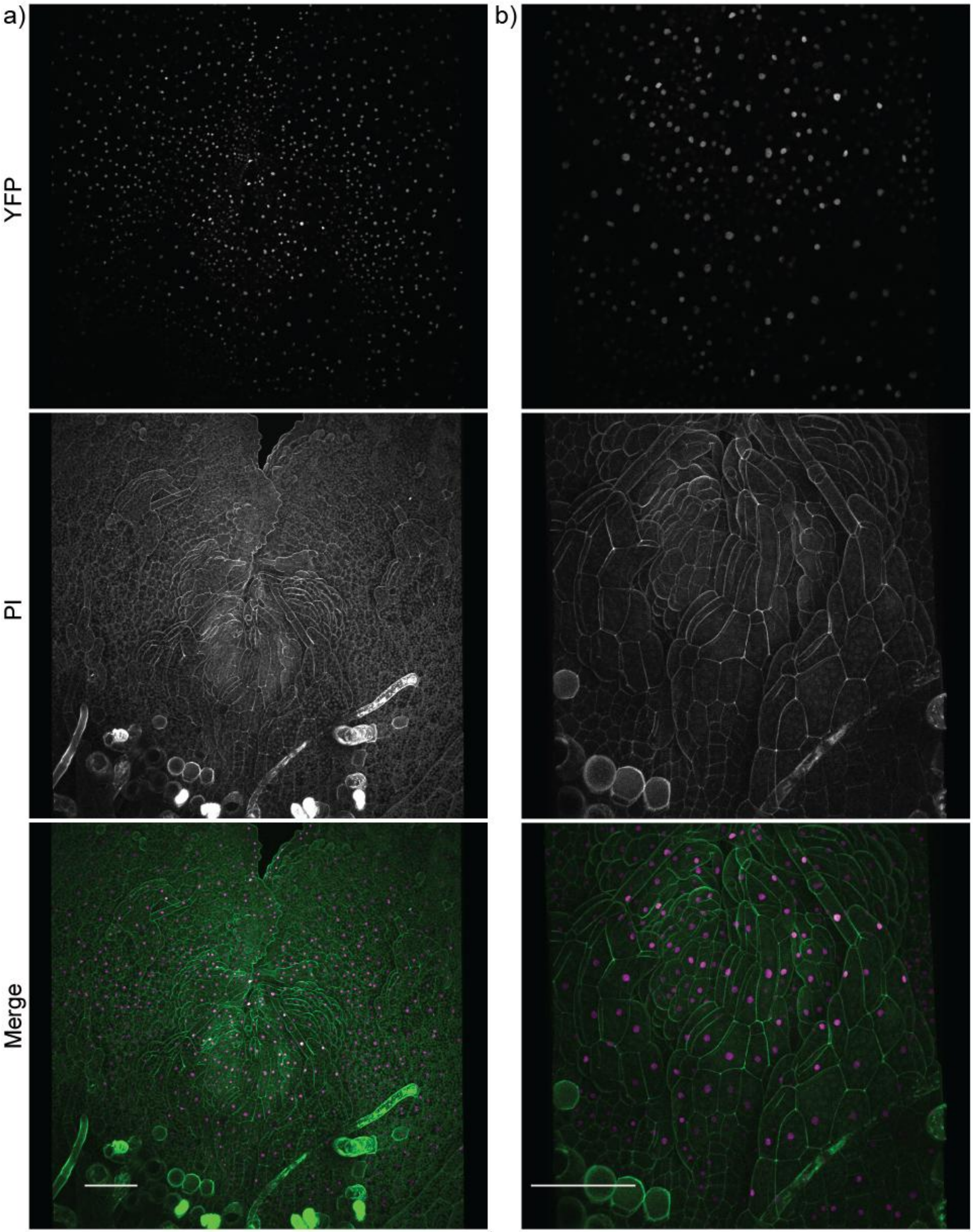


Figure 5.6 The Mp*WIP* promoter is active throughout the ventral surface of the thallus.
a) Apical region of the ventral surface of the thallus of $_{pro}$Mp*WIP:3xYFP-NLS*. Top panels show YFP fluorescence. In middle panels cell walls are stained with propidium iodide. YFP (magenta) and PI (green) signals are merged in lower panels. 9 d old gemmaling, scale 100 µm, apex at top. b) Detail of a), scale 100 µm.

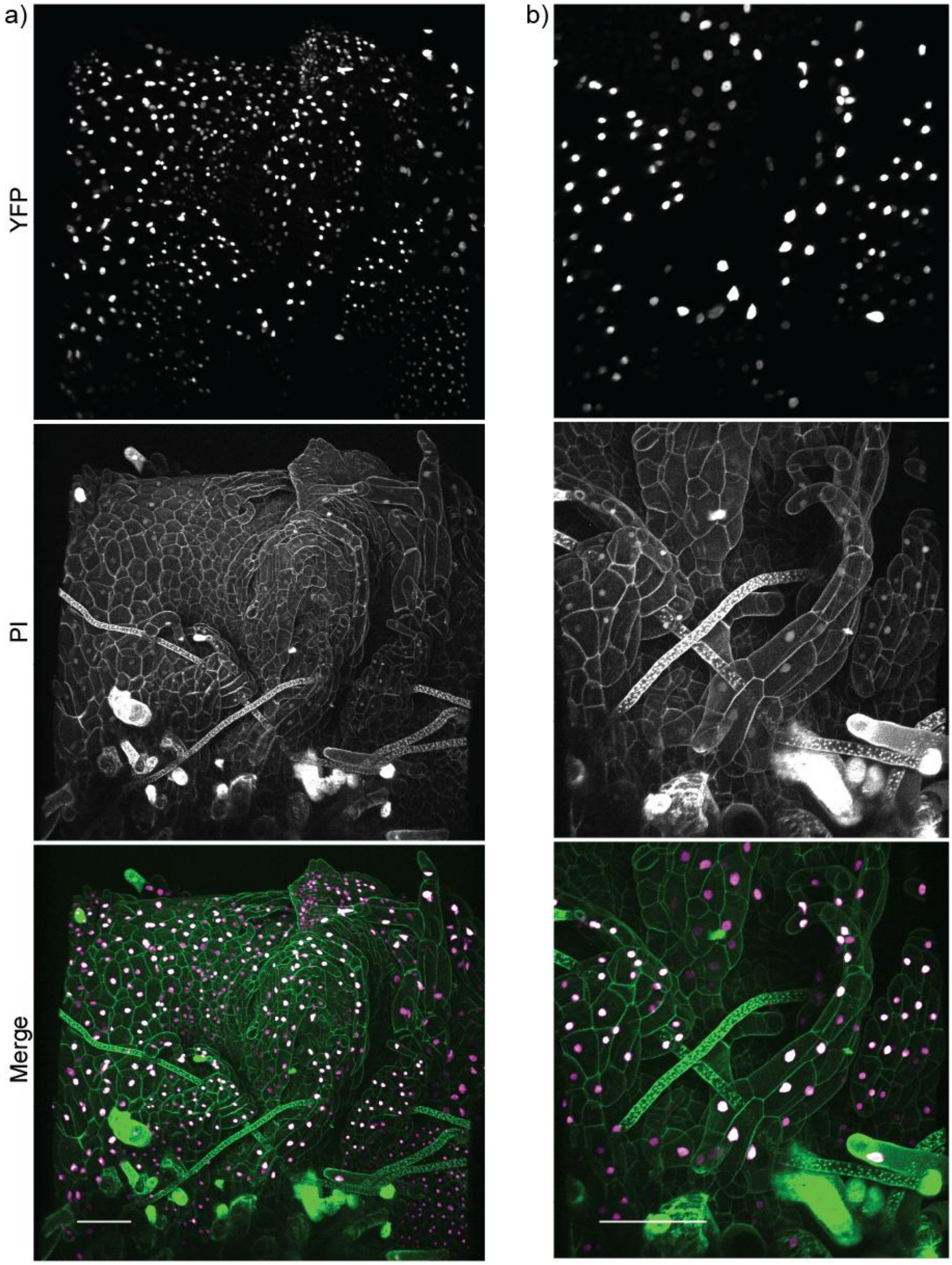


Figure 5.7 The Os*ACT* promoter is active in all ventral cell types. a) Apical region of the ventral surface of the thallus of $_{pro}$Os*ACT:3xYFP-NLS*. Top panels show YFP fluorescence. In middle panels cell walls are stained with propidium iodide. YFP (magenta) and PI (green) signals are merged in lower panels. 9 d old gemmaling, scale 100 µm, apex at top. b) Detail of a), scale 100 µm.

3.5.5 Dominant repression of MpWIP target genes by expression of an MpWIP-SRDX fusion causes thallus defects and the development of ectopic rhizoids

The phenotype of the Mp*wip*^{GOF} mutant and of the $_{pro}$*35S:*Mp*WIP* lines indicated that an excess of Mp*WIP* transcript can induce the development of ectopic rhizoids. The activity of the Mp*WIP* promoter in cells of the ventral surface of the thallus is also consistent with a role in rhizoid development. However, it was also highly active specifically in the developing air pores, the development of which is not affected in mutant *vj7* or the $_{pro}$*35S:*Mp*WIP*. Furthermore, the ability of MpWIP to induce the ectopic development of rhizoids does not necessarily mean that it does perform this function in the wild type. In the absence of an Mp*wip* loss-of-function allele it is difficult to further dissect the function of the gene.

To address this problem, I created a C-terminal protein fusion between MpWIP and the SRDX repressive domain (Hiratsu et al. 2003) to dominantly repress the transcription of the target genes of MpWIP. I expressed this under the constitutive 35S promoter ($_{pro}$*35S:*Mp*WIP-SRDX*). In lines that express this fusion protein ($_{pro}$*35S:*Mp*WIP-SRDX* 1 and 2) a dense growth of ectopic rhizoids develops on the dorsal surface of the thallus, which is otherwise normal (Figure 5.8 a and b). These results indicate that repressing the transcription of the targets of MpWIP can cause the development of ectopic rhizoids, which is consistent with the hypothesis that MpWIP promotes the development of rhizoids by repressing target genes. The fact that ectopic air pores do not form in the $_{pro}$*35S:*Mp*WIP-SRDX* lines indicates that the expression of Mp*WIP-SRDX* is not sufficient to direct cells to develop as air pore cells. However, normal air pore development is not impaired in these lines, as would be expected if the activation of MpWIP target genes was necessary for air pore development. Though not conclusive, this finding is consistent with the hypothesis that MpWIP is required, but not sufficient, for the proper development of air pore complexes, and that this activity relies on the repression of target genes.

In other lines (for example $_{pro}$*35S:*Mp*WIP-SRDX* 3), in addition to the development of a few ectopic rhizoids, the development of the thallus is defective (Figure 5.8 b). The thallus does not expand laterally as much as in the wild type, and appears thickened in the dorsiventral axis. Furthermore, the differentiation of the dorsal surface of the thallus is impaired. The differentiation of the air chambers, filled with dark green photosynthetic filaments in the wild type, and of the air pore at the centre of each chamber, is delayed or abolished. In line $_{pro}$*35S:*Mp*WIP-SRDX* 3, I did not detect Mp*WIP-SRDX* transcript by RT-PCR (Figure 5.8 a). However, in other transformed lines in which Mp*WIP-SRDX* transcript is undetectable ($_{pro}$*35S:*Mp*WIP-SRDX* 4), neither ectopic rhizoids nor thallus developmental defects are seen (Figure 5.8 b). To account for these two different phenotypes in lines in which the transcript of the Mp*WIP-SRDX* fusion is not detectable, I hypothesised that in line $_{pro}$*35S:*Mp*WIP-SRDX 4*, in which development is similar to that of the wild type, the transgene is simply not transcribed. In contrast, in line $_{pro}$*35S:*Mp*WIP-SRDX 3*, which has defects in thallus development, the transgene might be highly transcribed, triggering gene silencing and co-suppression of the native Mp*WIP* gene. If this were the case, targeted gene silencing of Mp*WIP* should give a phenotype similar to that seen in $_{pro}$*35S:*Mp*WIP-SRDX 3*, with defective air pore and thallus development.

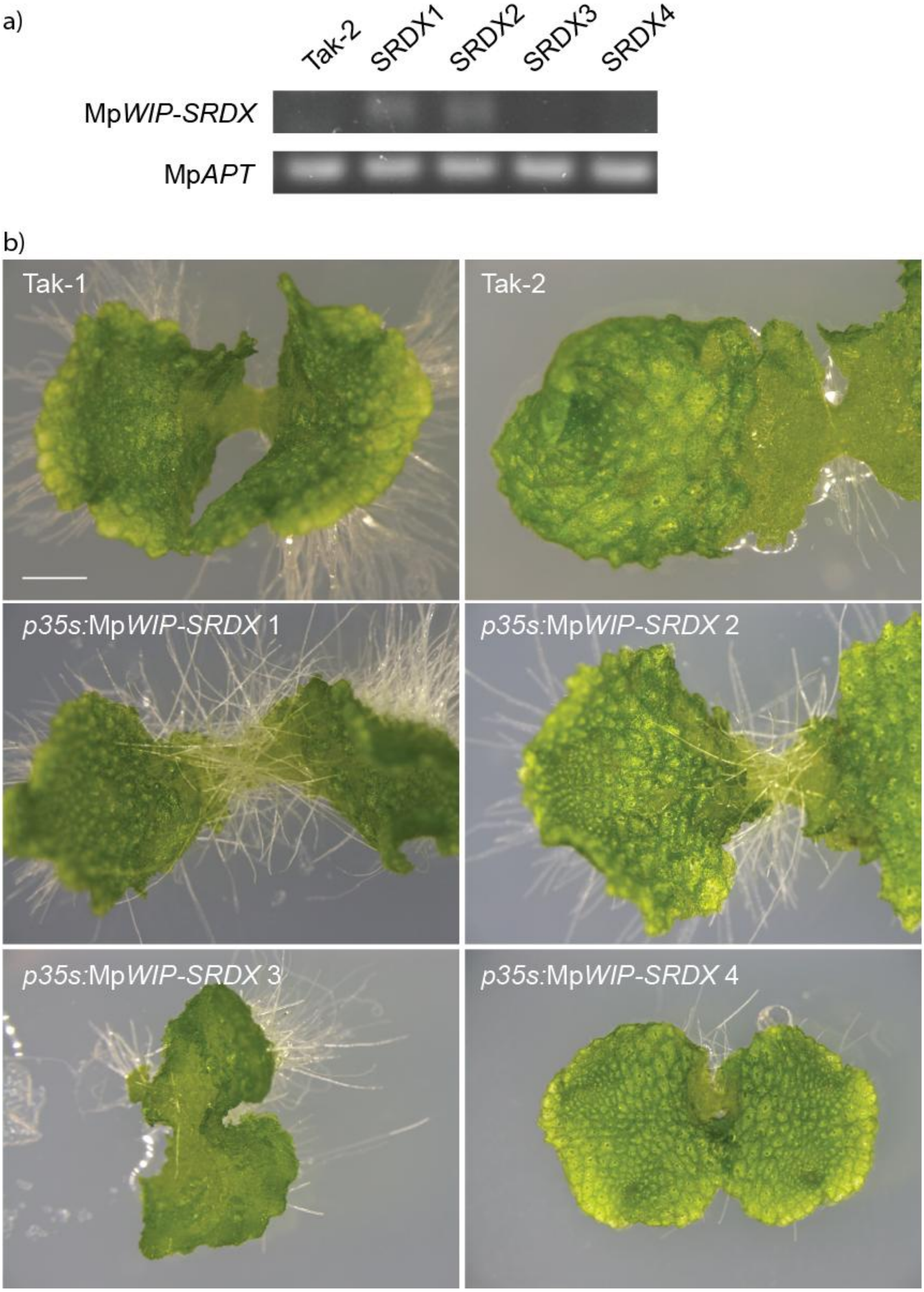


Figure 5.8 $_{pro}$*35S:*Mp*WIP-SRDX* causes development of ectopic rhizoids and defects in thallus and air pore development. a) RT-PCR to detect transcript of Mp*WIP-SRDX*. Reference gene Mp*APT*, 12 d gemmalings. b) Phenotype of wild type and $_{pro}$*35S:*Mp*WIP-SRDX* plants. 12 d gemmalings, scale 1 mm.

5.3.6 Expression of an amiRNA targeting Mp*WIP* causes defects in thallus and air pore development

Dominant repression of its target genes indicates that MpWIP can direct cells to adopt rhizoid fate through repressive activity. In order to investigate the effects of the loss of Mp*WIP* function, and to test the hypothesis that the defective air pore phenotype of $_{pro}$*35S:*Mp*WIP-SRDX 3* (Figure 5.8 b) was due to silencing of Mp*WIP*, I transformed wild type plants with an amiRNA that targets the 3’ UTR of the Mp*WIP* transcript, expressed under the control of the constitutive *OsACT* promoter ($_{pro}$*OsACT:amiR*-Mp*WIP*$^{\mathrm{Mp}\mathit{miR160}}$). In these plants the development of the dorsal thallus is defective (Figure 5.9 a) and is phenotypically similar to line $_{pro}$*35S:*Mp*WIP-SRDX* 3. The development of the air chambers is delayed or abolished, so the thallus is lighter in colour than the wild type, and the lateral expansion of the thallus is impaired. I used SEM and CSLM to image the defects in air pore development of line $_{pro}$*OsACT:amiR*-Mp*WIP*$^{\mathrm{Mp}\mathit{miR160}}$ *1* in more detail. In this line the regular structure of the air pores of the wild type, which consist of 16 cells arranged in a stack of 4 rings of 4 cells each, is absent (Figure 5.9 b, c). Roughly circular openings develop in the dorsal epidermis, but the organized cell divisions surrounding the holes do not take place. Some anticlinal divisions occur but are disorganized, and the periclinal divisions are greatly impaired. I did not observe any clear differences in rhizoid development in the $_{pro}$*OsACT:amiR*-Mp*WIP*$^{\mathrm{Mp}\mathit{miR160}}$ lines in comparison with the wild type. This indicates that the expression of *amiR*-Mp*WIP*$^{\mathrm{Mp}\mathit{miR160}}$ interferes with the differentiation and development of the cells of the air pore.

To test whether Mp*WIP* transcript levels are reduced in these $_{pro}$*OsACT:amiR*-Mp*WIP*$^{\mathrm{Mp}\mathit{miR160}}$ lines, I measured them by RT-qPCR. I also measured the level of Mp*RSL1* transcript, because Mp*RSL1* transcript levels were increased in mutant *vj7* and $_{pro}$*35S:*Mp*WIP* lines (Figure 3.5 d). Contrary to expectations, the levels of both of these transcripts were similar to those of the wild type. This suggests that *amiR*-Mp*WIP*$^{\mathrm{Mp}\mathit{miR160}}$ does not silence Mp*WIP* by reducing the levels of its transcript.

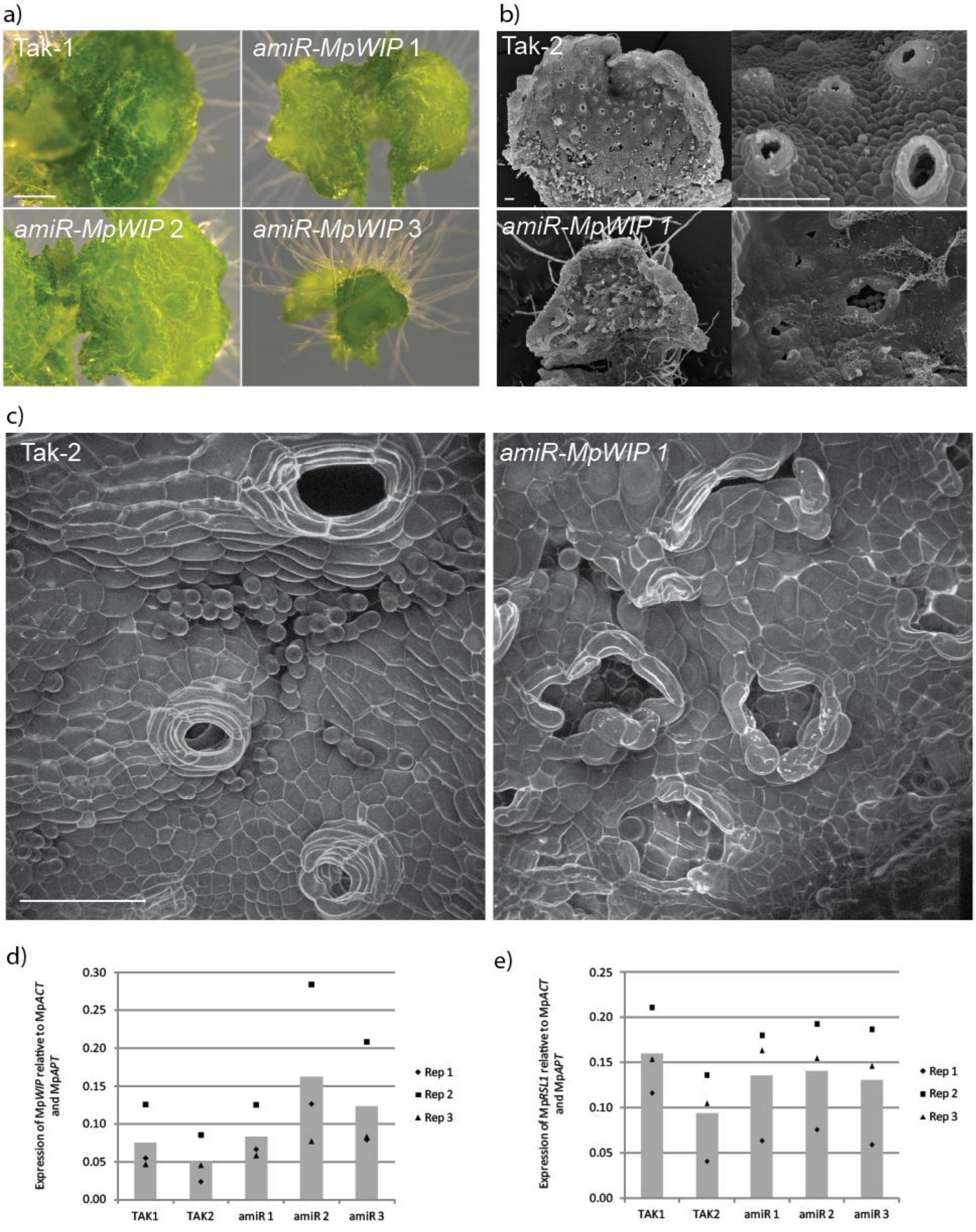


Figure 5.9 Plants transformed with $_{pro}$*OsACT:amiR*-Mp*WIP*$^{\mathrm{Mp}miR160}$ have defects in thallus and air pore development. a) In $_{pro}$*OsACT:amiR*-Mp*WIP*$^{\mathrm{Mp}miR160}$ lines the dark green air chambers seen in the wild type fail to develop, and the lateral expansion of the thallus is impaired. Transmitted light, 10 d old gemmalings, scale 500 µm. b) Air pore morphology is abnormal in $_{pro}$*OsACT:amiR*-Mp*WIP*$^{\mathrm{Mp}miR160}$ lines. SEM, 10 d old gemmalings, scale 100 µm. c) $_{pro}$*OsACT:amiR*-Mp*WIP*$^{\mathrm{Mp}miR160}$ lines do not develop the regular 16 cell air pore structure seen in the wild type. CSLM, PI stained, 9 d old gemmalings, scale 100 µm. d) There is no difference in the level of Mp*WIP* transcript between the wild type and $_{pro}$*OsACT:amiR*-Mp*WIP*$^{\mathrm{Mp}miR160}$ lines. 10 d gemmalings. e) There is no difference in the level of Mp*RSL1* transcript between the wild type and $_{pro}$*OsACT:amiR*-Mp*WIP*$^{\mathrm{Mp}miR160}$ lines. 10 d gemmalings.

5.4 Discussion

5.4.1 Mp*WIP* positively regulates rhizoid identity, perhaps by repressing target genes

I selected mutant *vj7* (Mp*wip*^{GOF}) in the mutant screen because it developed rhizoids on the dorsal surface of the thallus (Figure 5.1 a). The development of rhizoids in ectopic locations suggested that the mutant might harbour a loss-of-function mutation in a negative regulator of rhizoid fate, or a gain-of-function mutation in a positive regulator of rhizoid fate. The fact that the expression of the gene near the insertion linked to the mutant phenotype, Mp*WIP*, was elevated in the mutant (Figure 5.1 b, c) suggested that this was likely to be a gain-of-function mutation, and so that Mp*WIP* might be a positive regulator of rhizoid fate. To test whether ectopic expression of Mp*WIP* is sufficient to promote the development of rhizoids in ectopic locations I expressed Mp*WIP* under the control of a constitutive promoter, which caused the development of ectopic rhizoids (Figure 5.3 a). The mean transcript level of Mp*WIP* in these lines was 1.9 to 3.9 times the wild type level in this experiment, but the variation was great (Figure 5.3 b). Nevertheless, the fact that multiple independent lines carrying this construct produced ectopic rhizoids indicates that expression of Mp*WIP* driven by a constitutive promoter can direct dorsal epidermal cells to adopt the rhizoid developmental fate. This suggests that Mp*WIP* encodes a positive regulator of rhizoid fate. Further experiments using promoters that drive stronger constitutive expression than $_{pro}$*35S* in *Marchantia*, such as $_{pro}$*EF1α* (Holger Breuninger, personal communication), should allow the creation of transformants in which Mp*WIP* transcript level is unambiguously greater than in the wild type.

There is some evidence that WIP proteins in *Arabidopsis* act as transcriptional repressors, and some contain an EAR motif, which confers active transcriptional repression (summarized in section 5.2.3). MpWIP does not contain an EAR motif, as defined by Kagale et al. (2010). However, when I caused constitutive expression of a version of MpWIP fused to an SRDX

(EAR motif), making a dominant repressor of MpWIP target genes (Hiratsu et al. 2003), it lead to the development of a dense growth of ectopic rhizoids on the dorsal surface (Figure 5.8). This is the same phenotype as seen in the Mp*wip*^{GOF} mutant (Figure 5.1, 5.3), and in $_{pro}$*35S:*Mp*WIP* lines (Figure 5.3). This indicates that MpWIP may direct cells to adopt rhizoid fate through repression of target genes. However, although it has been proposed that WIP proteins in other species act as transcription factors, direct binding to DNA has not been shown for any WIP protein. Further work, such as yeast one-hybrid or chromatin immunoprecipitation experiments, will be necessary in order to test whether MpWIP binds to DNA. The hypothesis that MpWIP acts as a transcriptional repressor should also be tested by fusing it to a strong transcriptional activation domain, such as the VP16 domain. If the activity of MpWIP relies on the repression of target genes, a dominant activator version should be unable to promote ectopic rhizoid differentiation when ectopically expressed, and may cause phenotypic defects through interference with the function of native MpWIP.

If MpWIP is a regulator that directs the acquisition of rhizoid fate, it should be expressed in the cells that give rise to rhizoids in the wild type. I found that the promoter of Mp*WIP* is active throughout the apical region of the ventral surface of the thallus, the zone where rhizoids differentiate (Figure 5.7). I did not directly detect Mp*WIP* mRNA or MpWIP protein in these cells, but it is likely that MpWIP is expressed in them, and so the expression pattern of Mp*WIP* is consistent with a role in the regulation of rhizoid fate determination.

The data discussed so far are consistent with Mp*WIP* being sufficient to direct the adoption of rhizoid fate, but they do not provide evidence that it is necessary for rhizoid development. The best evidence of this would be if the loss of MpWIP function impaired the adoption of rhizoid fate. However, I did not identify a loss-of-function Mp*wip* allele, and neither did Suvi Honkanen nor Giulia Morieri, who also conducted screens for mutants with defects in rhizoid development. It may be the case that a complete loss of Mp*WIP* function is lethal.

Arabidopsis possesses six *WIP* genes that are at least in some cases partially redundant, and the loss of the function of a single WIP (At*dot5*) or a clade of three of them (At*ntt/*At*wip4*/At*wip5*) causes severe, wide-ranging developmental defects (Petricka et al. 2008; Crawford et al. 2015). *Marchantia* by contrast has only a single WIP, and so in the absence of functional redundancy between family members the effect of a loss-of-function mutation in Mp*WIP* may be too severe to allow any growth. Even if it were not lethal, such a mutant might have such extensive phenotypic defects that I did not select it as presenting a specific defect in rhizoid development. Alternatively, the loss of MpWIP function might not have any noticeable impact on rhizoid development at all, which would indicate that it is not absolutely required for rhizoid differentiation. In the absence of a loss-of-function Mp*wip* mutant I cannot determine whether MpWIP function is necessary for rhizoid development.

The fact that I did not observe defects in rhizoid development in the *amiR*-Mp*WIP*$^{\mathrm{Mp}\mathit{miR160}}$ lines could be because the defects were too subtle, or because rhizoid development was not affected. If the latter case is true, there may be a number of reasons for this. As already noted, it is possible that the *amiR*-Mp*WIP*$^{\mathrm{Mp}\mathit{miR160}}$ does not reduce MpWIP levels. Alternatively, it may be that rhizoid development is less sensitive to MpWIP levels than is air pore development, and the level is not reduced enough by the amiR to impair the development of rhizoids. Finally, it could be that MpWIP expression is sufficient, but not necessary, to direct the adoption of rhizoid fate, perhaps because there are other factors that act redundantly in this process.

5.4.2 Mp*WIP* may be a positive regulator of air pore development

In some lines that I generated to constitutively express an MpWIP-SRDX fusion protein, air pore and thallus development was abnormal (Figure 5.8 b), which I hypothesized might be due to co-suppression of the endogenous Mp*WIP*. In an attempt to generate plants with a reduced level of functional MpWIP and test this hypothesis, I transformed plants with a construct that drove constitutive expression of an amiRNA designed to target the 3’ UTR of Mp*WIP*. These $_{\mathit{pro}}$*OsACT:amiR*-Mp*WIP*$^{\mathrm{Mp}\mathit{miR160}}$ lines consistently show defects in the development of the thallus and air pores. The stereotypical cell divisions that give rise to a 16 cell air pore structure in the wild type do not occur in the amiR lines (Figure 5.9 b, c), and the morphology of the thallus is abnormal (Figure 5.9 a, b). This suggests that my hypothesis concerning co-suppression may be correct, and that Mp*WIP* is necessary for proper air pore development.

The $_{\mathit{pro}}$*OsACT:amiR*-Mp*WIP*$^{\mathrm{Mp}\mathit{miR160}}$ lines do not display any reduction in the level of Mp*WIP* transcript, but it is still possible that in these lines the level of functional MpWIP protein was reduced. miRNA-directed cleavage at the target site and a reduction of transcript level has been demonstrated for amiRNAs targeting Mp*AUXIN RESPONSE FACTOR1*, Mp*RESPONSE REGULATOR-B*, Mp*ENHANCER OF ZESTE*, and Mp*AUXIN/INDOLE*-3-*ACETIC ACID* (Eklund et al. 2015; Flores-Sandoval, Dierschke, et al. 2015; Flores-Sandoval, Eklund, et al. 2015), but otherwise miRNAs in *Marchantia* have not been studied in great detail. In *Arabidopsis* miRNAs can function through translation repression as well as by cleavage, even if the target site is in the 3’ UTR (Brodersen et al. 2008; Iwakawa and Tomari 2013; Li et al. 2013), so perhaps *Marchantia amiR*-Mp*WIP*$^{\mathrm{Mp}\mathit{miR160}}$ is acting by blocking translation, rather than by cleaving the transcript. It is also possible that *amiR*-Mp*WIP*$^{\mathrm{Mp}\mathit{miR160}}$ does not actually target Mp*WIP* at all, and that the defects seen in the amiR lines are due to silencing of an off-site target. In order to rule this out it will be necessary to show complementation of the defective phenotype by expression of an *amiR*-Mp*WIP*$^{\mathrm{Mp}\mathit{miR160}}$-resistant version of Mp*WIP*, and ideally to show that multiple amiRs targeting different sites on Mp*WIP* cause the same developmental defects.

The Mp*WIP* promoter is active at a low level in the apical region of the dorsal surface of the thallus, but it is strikingly more active in cells that are developing into air pores (Figures 5.4, 5.5). The fact that its promoter is strongly preferentially active in the early stages of air pore development is consistent with the hypothesis that Mp*WIP* is a positive regulator of this process. The low level of expression seen in other cells would be consistent with a role in other processes throughout the thallus that might explain the altered thallus morphology seen in the $_{pro}$*OsACT:amiR*-Mp*WIP*$^{\mathrm{Mp}\mathit{miR160}}$ lines.

Although further work is needed to confirm that the *amiR*-Mp*WIP*$^{\mathrm{Mp}\mathit{miR160}}$ causes developmental defects through a specific effect on Mp*WIP*, the phenotype of the amiR lines and the pattern of Mp*WIP* promoter activity suggest that Mp*WIP* regulates the proper development of air pores. If it is established that the defective phenotype is due to a the targeting of Mp*WIP*, more detailed investigation will be necessary to determine the exact nature of the air pore developmental defect, such as the stage of air pore differentiation at which the wild type and amiR lines diverge. The fact that no ectopic air pores develop in the Mp*wip*^{GOF} mutant or the $_{pro}$*35S:*Mp*WIP* lines indicates that, if Mp*WIP* is involved in air pore development, it is necessary but not sufficient to drive differentiation into air pores. If Mp*WIP* acted in air pore development by activating, rather than repressing, target genes, then aberrant air pore development would be expected in the dominant repressor $_{pro}$*35S:*Mp*WIP-SRDX* lines, but air pore development is normal in the $_{pro}$*35S:*Mp*WIP-SRDX* lines. Bearing in mind the caveats about the amiR experiment and the fact that binding of DNA by MpWIP has not been demonstrated, I hypothesize that Mp*WIP* could promote air pore development through the repression of the transcription of target genes.

5.4.3 Models for the regulation of rhizoid and air pore development by Mp*WIP*

Tentative hypothetical models of the mechanism by which Mp*WIP* might regulate the development of both rhizoids and root hairs can be made that are consistent with the data presented in this chapter (Figure 5.10).

The ectopic-rhizoids phenotype of the original Mp*wip*^{GOF} mutant and the $_{pro}$*35S:*Mp*WIP* lines indicates that MpWIP is able to direct non-rhizoid cells to adopt the rhizoid fate, and its domain of expression in the ventral surface of the thallus is consistent with a role as a positive regulator of rhizoid fate. In addition, levels of Mp*RSL1*, a positive regulator of rhizoid development, are increased in lines with elevated Mp*WIP* expression. The simplest model that explains these findings is that Mp*WIP* promotes rhizoid differentiation by positively regulating the expression of Mp*RSL1*, either directly or indirectly (Figure 5.10 a). The activity of the Mp*WIP* promoter in developing air pore cells, and the defective air pore complexes seen in the *amiR*-Mp*WIP*$^{\mathrm{Mp}\mathit{miR160}}$ lines, suggest that Mp*WIP* may have a role in the promotion of air pore development, but at this point the data are not strong enough to say more than that. Unlike the case with rhizoids, expression of Mp*WIP* is not sufficient to direct cells to differentiate into air pores, as air pore patterning is unaltered in *vj7*, $_{pro}$*35S:*Mp*WIP* lines or lines expressing Mp*WIP-SRDX.*

A slightly more detailed model takes into account the ectopic-rhizoids phenotype of plants expressing Mp*WIP-SRDX* (Figure 5.10 b). Because the phenotype of the $_{pro}$*35S:*Mp*WIP-SRDX* lines is the same as that of *vj7* and the $_{pro}$*35S:*Mp*WIP* lines, I tentatively infer that MpWIP promotes rhizoid fate by the repression of target genes, although further work is necessary to test this hypothesis. As the expression of Mp*RSL1* is elevated in *vj7* and the $_{pro}$*35S:*Mp*WIP* lines, I hypothesize that MpWIP could be a negative regulator of some factor (X in Figure 5.10 b) that negatively regulates the expression of Mp*RSL1*. Expression of

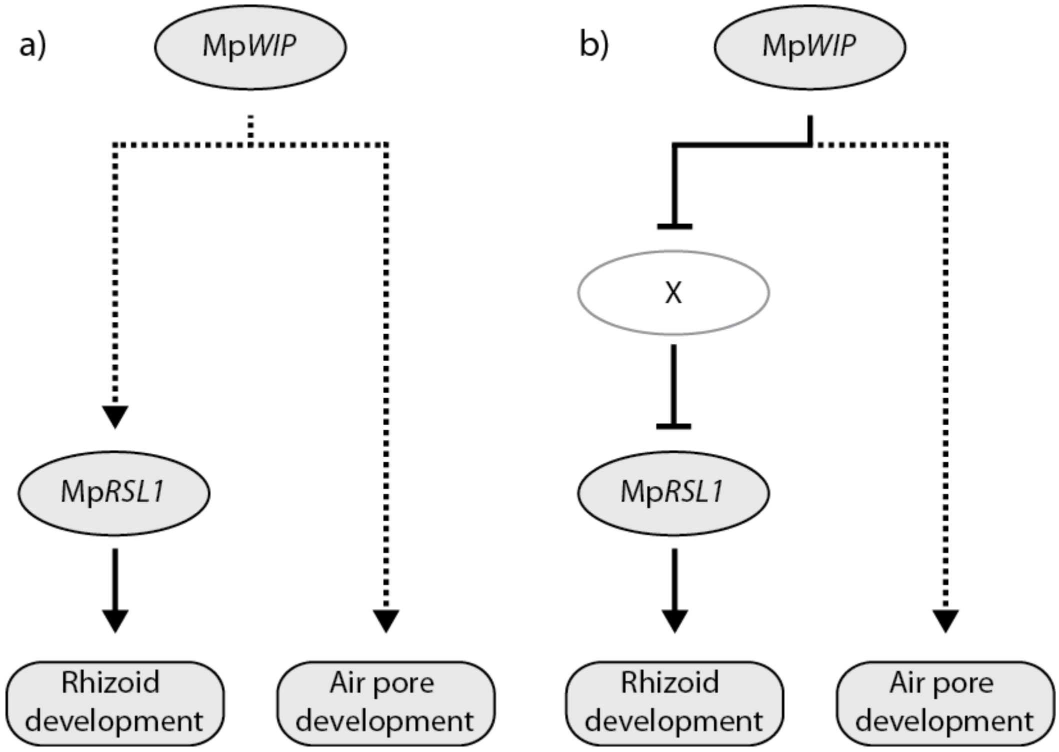


Figure 5.10 Tentative models of Mp*WIP* regulation of rhizoid and air pore development. In a) Mp*WIP* positively regulates rhizoid differentiation by promoting the expression of Mp*RSL1* by unknown means. In b) Mp*WIP* promotes Mp*RSL1* by repressing the expression of a hypothetical repressor of Mp*RSL1* (marked X). In both cases Mp*WIP* promotes air pore development through unknown means.

Mp*WIP* would release the repression of Mp*RSL1*, allowing it to be expressed and promote differentiation of the cell into a rhizoid. Because of the broad expression of Mp*WIP* in the ventral surface of the thallus, other factors would likely be involved in restricting the expression of Mp*RSL1* and rhizoid differentiation to a subset of cells. The existence of factor X in this model is hypothetical, but my colleague Suvi Honkanen has identified an miRNA gene named Mp*FEW RHIZOIDS 1* (Mp*FRH1*) that targets Mp*RSL1* and negatively regulates its transcript levels (unpublished). It is tempting to speculate that Mp*FRH1* may be factor X, but further work is needed to test this hypothesis.

5.4.4 Evolution of the control of rooting cell differentiation by *WIP* genes

The results presented in this chapter suggest a role for Mp*WIP* in the control of the acquisition of rhizoid fate in *Marchantia*, perhaps through the repression of a repressor of Mp*RSL1*. In *Arabidopsis* the factors directly upstream of At*RSL1* in the gene network that regulates root hair development are unknown, though At*RSL1* is negatively regulated by the homeodomain protein AtGL2, whose expression is in turn negatively regulated by a complex of the MYB AtCPC, the bHLH AtGL3 or AtEGL3, and the WD40 protein AtTTG. The topology of this pathway is the same as one I propose for the control of rhizoid differentiation by Mp*WIP* (Figure 5.10 b). This is pure speculation, but given that the *Arabidopsis* WIP AtTT1 protein interacts physically with the MYB members of MYB-bHLH-WD40 complexes (Appelhagen et al. 2011), it is possible that a WIP protein might interact with the root-hair regulating MYB-bHLH-WD40 complexes in *Arabidopsis* to control root hair development. No role in root hair development has been reported for the At*WIP* genes that have been studied, which are all but At*WIP3*. If any of these genes is actually involved in root hair development, the authors may have neglected to observe this aspect of development, or there may be functional redundancy between members of the At*WIP* family. It might therefore be fruitful to measure the expression of the At*WIP* genes in *Arabidopsis* root hairs. At*NTT* and At*WIP4* are expressed in the QC but not elsewhere in the root, while At*DOT5* is not expressed anywhere in the root (Birnbaum et al. 2003; Nawy et al. 2005). Probes for At*TT1*, At*WIP3*, and At*WIP5* are not present on the Affymetrix ATH1 GeneChip used by Birnbaum et al. (2003) and Nawy et al. (2005) to measure gene expression in the different cell types of the root, so information on their expression is not present in these data sets, and targeted measurement of their expression would be necessary. Experiments to test for interactions between the MYB members of the complexes and each AtWIP would be feasible, as the number of bait and prey are each small. Finally, mutants are available for all At*WIP* genes, which could be examined singly or in higher-order combinations for defects in root hair differentiation. Genetic interactions with mutants in other known regulators of root hair development could also be sought.

The evidence presented in this chapter suggests that Mp*WIP* is a novel regulator of rooting cell development. It may be that this role evolved after the divergence of the liverworts from the other land plants, or it may be inherited from the last common ancestor of land plants. *WIP* genes have not been implicated in the regulation of rooting cell development in other land plant lineages to date, but future work may uncover a role for them in this process in other species. If that is the case, it would support the hypothesis that *WIP* genes are part of the ancient, conserved gene network that included *RSL* genes and controlled the development of rooting cells in the earliest land plants.

Chapter 6. General discussion

Filamentous, tip-growing rooting cells develop at the interface between plants and the soil, where they function in anchorage and the uptake of water and nutrients. They were likely one of the keys to the success of the earliest land plants in colonizing the terrestrial surface, and, as rhizoids or root hairs, are found in most extant land plants. Several factors control the development of both the rhizoids of the moss *Physcomitrella* and the angiosperm *Arabidopsis*, namely RSL and LRL transcription factors and the hormone auxin. This suggests the existence of an ancient gene regulatory network that controlled the development of rooting cells in the last common ancestor of mosses and angiosperms, over 420 million years ago.

My aim in this thesis was to investigate the gene regulatory network that controls rooting cell in *Marchantia polymorpha*, a member of the lineage of land plants that likely diverged earliest. This would allow inferences to be made about the evolution of such cells and the genetic control of their development in the first land plants. Were the development of *Marchantia* rhizoids to be regulated by homologues of the genes that control the development of moss rhizoids and angiosperm root hairs, it would suggest that the last common ancestor of all land plants possessed rooting cells. It would also suggest that genes ancestral to the ones that regulate rooting cell development in extant land plants already regulated their development in the earliest land plants. On the other hand, if the genes that regulated *Marchantia* rhizoid development were unrelated to those that carry out this function in other land plant lineages, it would suggest that the rhizoid gene regulatory network in *Marchantia* arose independently of that in other lineages. This would support the hypothesis that the first land plants did not develop rooting cells, and that these arose on land only after the divergence of liverworts and the other embryophtyes.

To discover genes that are part of the network that controls rhizoid development in *Marchantia* I carried out a mutagenesis and screen for mutants with defects in rhizoid development. From the resulting mutant collection I identified 10 genes that mutate to a defective rhizoid phenotype, and so likely belong to the network of genes that directs rhizoid development. This illustrates that large genetic screens are feasible in *Marchantia* and could be used to study various developmental or physiological processes in early-diverging land plants. The fact that I identified four genes about whose *Arabidopsis* homologues there is almost no functional data suggests that *Marchantia* can be useful for studying the functions of genes that have so far resisted characterization in *Arabidopsis*, perhaps because of functional redundancy. Eight of the genes I identified appear not to be part of the core network that regulates the early stages of rhizoid differentiation, but rather to be peripheral genes that are involved in later growth processes. Since two of these are homologues of genes known to be required for root hair elongation in *Arabidopsis*, at least some late-acting parts of the network are likely to have been conserved since liverworts and angiosperms diverged, or perhaps were independently recruited to perform the same function. Some of the genes of unknown function may also yet prove to be involved in *Arabidopsis* root hair elongation.

Among the mutants I isolated in the screen are Mp*rsl1* gain-of-function mutants that develop ectopic rhizoids. These mutants, together with rhizoidless Mp*rsl1* loss-of-function mutants identified by Suvi Honkanen, indicate that Mp*RSL1* is both necessary and sufficient to direct cells to adopt the rhizoid developmental fate. If the liverworts are sister to all other land plants, the most parsimonious explanation for the fact that a key regulator of rooting cell development is conserved between *Marchantia* and the lineage that gave rise to *Physcomitrella* and *Arabidopsis* is that an ancestral *RSL* gene regulated the development of rooting cells in the last common ancestor of all extant land plants. This in turn suggests that rooting cells evolved before or very soon after the conquest of the land by plants, before the divergence of the liverwort lineage. Characterization of the wider *Marchantia* rhizoid development regulatory network will allow us to trace how the network has changed over the course of evolution. The connections between the regulators auxin, *LRL*s and *RSL*s differ between *Physcomitrella* and *Arabidopsis* (Tam et al. 2015). It is already known that auxin is a positive regulator of rhizoid initiation and elongation in *Marchantia* (Tarén 1958; Maravolo and Voth 1966; Maravolo 1980; Ishizaki et al. 2012; Eklund et al. 2015; Flores-Sandoval, Eklund, et al. 2015), though the exact mechanism by which it does so is not yet known. Further studies will discover whether it acts through Mp*RSL1*, or whether it regulates rhizoid development independently. The role of the *LRL* family in *Marchantia* is also unknown; it would be interesting to find out whether this component of the rhizoid developmental network too is conserved in *Marchantia*, or whether it is specific to the lineage that includes moss and angiosperms, and so perhaps was added to the rhizoid network after the divergence of the liverworts.

In addition to Mp*RSL1*, I identified Mp*WIP* as a positive regulator of rhizoid fate determination. My results suggest that it may drive cells to adopt the rhizoid fate by suppressing target genes, though they are not conclusive and further work is needed to test whether this hypothesis is correct. It has been posited that some WIP proteins in *Arabidopsis* function as transcriptional repressors, though again clear evidence of a WIP protein acting by this mechanism has not been reported. I hypothesize that Mp*WIP* may promote rhizoid differentiation by repressing a repressor of Mp*RSL1*, which may be Mp*FRH1*, identified by Suvi Honkanen. The next step towards establishing the network of genes that regulates rhizoid differentiation will be testing the hypothesized regulatory connections between Mp*WIP*, Mp*RSL1* and Mp*FRH1*. Mutants and promoter reporter lines for all three genes are available ,which will facilitate the dissection of the interactions between these genes. It would also be of interest to find out how the patterning of rhizoids is established; the broad expression pattern of Mp*WIP* around the ventral surface of the apex indicates that there must be some other factor that restricts rhizoid fate to only a subset of cells.

Further work would also provide deeper insights into the function of MpWIP. In particular, it would be very useful if a loss-of-function allele of Mp*WIP* could be produced. The gain-of-function mutant and overexpression lines indicate that MpWIP expression is sufficient to induce non-rhizoid cells to develop into rhizoids, but a loss-of-function allele would allow us to test whether it is necessary for the differentiation of rhizoid cells. CRISPR/Cas9 and homologous recombination methods have been reported for *Marchantia* (Ishizaki, Johzuka-Hisatomi, et al. 2013; Sugano et al. 2014), and could be used to create null alleles of Mp*WIP*, unless those prove to be lethal. *Marchantia* may also offer a good system for investigating the molecular mechanism of WIP function. To test the hypothesis that MpWIP functions in a complex with WD40, MYB and bHLH proteins, yeast two hybrid experiments could be performed with potential interacting partners. This would be facilitated by the relatively small gene families in *Marchantia* compared to other species. In particular, MpMYB4 is a promising candidate interactor, as it is involved in epidermal development (Hiroyoshi Kubo, Shinshu University, Japan, personal communication). In combination with experiments to test whether MpWIP or potential interacting partners directly bind DNA, this would give a clearer idea of how MpWIP regulates downstream targets.

The potential role of Mp*WIP* in promoting air pore development is likely to be of restricted phylogenetic distribution. Liverworts lack true stomata, and if liverworts are basal among land plants it is parsimonious to infer that stomata evolved after the divergence of liverworts. Therefore there is no reason to expect that the development of air pores and stomata would be regulated by a conserved genetic mechanism..Within the liverworts, compound air pores are found only in subclass Marchantiidae, while the remaining members of class Marchantiopsida, the Jungermanniopsida and the basally-diverging class Haplomitriopsida lack them (Crandall-Stotler et al. 2009). This suggests that they are an evolutionary novelty of the complex thalloid liverworts of the Marchantiidae, and so the control of their development is likely not inherited from the common ancestor of land plants. Instead, I hypothesize that a pre-existing mechanism involving Mp*WIP* for promoting the differentiation of a specialized epidermal cell type, rhizoids, was co-opted to promote the differentiation of a novel epidermal structure, the air pore, in the complex thalloid liverworts.

The understanding of how developmental networks that determine morphology are established and evolve over time is one of the main goals of evolutionary developmental biology. This thesis contributes to that goal by identifying two regulators of the development of a key adaptation of land plants to terrestrial life, one of which was possibly co-opted to regulate the development of another cell type. Future work that expands on these components and discovers the connections between them will provide insights into a developmental network that operated in the earliest stages of land plant evolution, and how it has changed since then.

Literature cited

Akkerman M, Franssen-Verheijen M a. W, Immerzeel P, Hollander LD, Schel JHN, Emons a. MC. **2012**. Texture of cellulose microfibrils of root hair cell walls of Arabidopsis thaliana, Medicago truncatula, and Vicia sativa. *Journal of Microscopy* **247**: 60–67.

Alonso JM, Stepanova AN, Leisse TJ, Kim CJ, Chen H, Shinn P, Stevenson DK, Zimmerman J, Barajas P, Cheuk R, Gadrinab C, Heller C, Jeske A, Koesema E, Meyers CC, Parker H, Prednis L, Ansari Y, Choy N, Deen H, Geralt M, Hazari N, Hom E, Karnes M, Mulholland C, Ndubaku R, Schmidt I, Guzman P, Aguilar-Henonin L, Schmid M, Weigel D, Carter DE, Marchand T, Risseeuw E, Brogden D, Zeko A, Crosby WL, Berry CC, Ecker JR. **2003**. Genome-wide insertional mutagenesis of Arabidopsis thaliana. *Science* **301**: 653–657.

Apostolakos P, Galatis B. **1985**. Studies on the development of the air pores and air chambers of Marchantia paleacea III . Microtubule organization in preprophase-prophase initial aperture cells - formation of incomplete preprophase microtubule bands. *Protoplasma* **128**: 120–135.

Appelhagen I, Huep G, Lu G-H, Strompen G, Weisshaar B, Sagasser M. **2010**. Weird fingers: functional analysis of WIP domain proteins. *FEBS Letters* **584**: 3116–22.

Appelhagen I, Lu GH, Huep G, Schmelzer E, Weisshaar B, Sagasser M. **2011**. TRANSPARENT TESTA1 interacts with R2R3-MYB factors and affects early and late steps of flavonoid biosynthesis in the endothelium of Arabidopsis thaliana seeds. *Plant Journal* **67**: 406–419.

Augustine RC, Pattavina K a., Tuzel E, Vidali L, Bezanilla M. **2011**. Actin Interacting Protein1 and Actin Depolymerizing Factor drive rapid actin dynamics in Physcomitrella patens. *The Plant Cell* **23**: 3696–3710.

Avisar D, Prokhnevsky AI, Makarova KS, Koonin E V, Dolja V V. **2008**. Myosin XI-K Is required for rapid trafficking of Golgi stacks, peroxisomes, and mitochondria in leaf cells of Nicotiana benthamiana. *Plant Physiology* **146**: 1098–1108.

Baluška F, Salaj J, Mathur J, Braun M, Jasper F, Šamaj J, Chua NH, Barlow PW, Volkmann D. **2000**. Root hair formation: F-actin-dependent tip growth is initiated by local assembly of profilin-supported F-actin meshworks accumulated within expansin-enriched bulges. *Developmental Biology* **227**: 618–32.

Banks JA. **1999**. Gametophyte development in ferns. *Annual Review of Plant Physiology and Plant Molecular Biology* **50**: 163–186.

Banks JA. **2009**. Selaginella and 400 million years of separation. *Annual Review of Plant Biology* **60**: 223–38.

Barko JW, Gunnison D, Carpenter SR. **1991**. Sediment interactions with submersed macrophyte growth and community dynamics. *Aquatic Botany* **41**: 41–65.

Bateman RM, Crane PR, Dimichele WA, Kenrick PR, Rowe P, Speck T, Stein WE, Rowe NP. **1998**. Early evolution of land plants : phylogeny, physiology , and ecology of the primary terrestrial radiation. *Annual Review of Ecology, Evolution, and Systematics* **29**: 263–292.

Bates JW. **1992**. Mineral nutrient acquisition and retention by bryophytes. *Journal of Bryology* **17**: 223–240.

Bates JW, Farmer AM. **1990**. An experimental study of calcium acquisition and its effects on the calcifuge moss Pleurozium schreberi. *Annals of Botany* **65**: 87–96.

Bayer EM, Bottrill AR, Walshaw J, Vigouroux M, Naldrett MJ, Thomas CL, Maule AJ. **2006**. Arabidopsis cell wall proteome defined using multidimensional protein identification technology. *Proteomics* **6**: 301–311.

Baylis GTS. **1975**. The magnoloid mycorrhiza and mycotrophy in root systems derived from it. In: Sanders FE, Mosse B, Tinker PB, eds. *Endomycorrhizae*. New York: Academic Press, 373–389.

Berger F, Linstead P, Dolan L, Haseloff J. **1998**. Stomata patterning on the hypocotyl of Arabidopsis thaliana is controlled by genes involved in the control of root epidermis patterning. *Developmental Biology* **194**: 226–34.

Berner RA. **1997**. The rise of plants and their effect on weathering and atmospheric CO2. *Science* **276**: 544–546.

Bernhardt C, Lee MM, Gonzalez A, Zhang F, Lloyd A, Schiefelbein J. **2003**. The bHLH genes GLABRA3 (GL3) and ENHANCER OF GLABRA3 (EGL3) specify epidermal cell fate in the Arabidopsis root. *Development* **130**: 6431–6439.

Berridge MJ, Irvine RF. **1989**. Inositol phosphates and cell signalling. *Nature* **341**: 197–205.

Bibikova TN, Jacob T, Dahse I, Gilroy S. **1998**. Localized changes in apoplastic and cytoplasmic pH are associated with root hair development in Arabidopsis thaliana. *Development* **125**: 2925–34.

Birnbaum K, Shasha DE, Wang JY, Jung JW, Lambert GM, Galbraith DW, Benfey PN. **2003**. A gene expression map of the Arabidopsis root. *Science* **302**: 1956–1960.

Bischler H. **1986**. Marchantia polymorpha L. S.LAT. karyotype analysis. *J. Hattori Bot. Lab.* **60**: 105–117.

Bowen EJ. **1935**. A note on the conduction of water in Fimbriaria bleumeana. *Annals of Botany* **49**: 844–848.

Box RJ. **1986**. Quantitative short-term uptake of inorganic phosphate by the Chara hispida rhizoid. *Plant, Cell and Environment* **9**: 501–506.

Brady NC, Weil RR. **2008**. *The nature and properties of soils*. Upper Saddle River, NJ: Pearson-Prentice Hall.

Braun M. **1997**. Gravitropism in tip-growing cells. *Planta* **203**: S11–S19.

Brodersen P, Sakvarelidze-Achard L, Bruun-Rasmussen M, Dunoyer P, Yamamoto YY, Sieburth L, Voinnet O. **2008**. Widespread translational inhibition by plant miRNAs and siRNAs. *Science* **320**: 1185–1190.

Bruex A, Kainkaryam RM, Wieckowski Y, Kang YH, Bernhardt C, Xia Y, Zheng X, Wang JY, Lee MM, Benfey P, Woolf PJ, Schiefelbein J. **2012**. A gene regulatory network for root epidermis cell differentiation in Arabidopsis. *PLoS Genetics* **8**: e1002446.

Campbell DH. **1934**. A remarkable development of the sporophyte in Anthoceros fusiformis, Aust. *Annals of Botany* **38**: 473–483.

Cao AJ, Dai X, Zou H, Wang Q, The S, Society B. **2014**. Formation and development of rhizoids of the liverwort Marchantia polymorpha. *Journal of the Torrey Botanical Society* **141**: 126–134.

Carol RJ, Dolan L. **2002**. Building a hair: tip growth in Arabidopsis thaliana root hairs. *Philosophical transactions of the Royal Society of London. Series B, Biological Sciences* **357**: 815–21.

Carol RJ, Takeda S, Linstead P, Durrant MC, Kakesova H, Derbyshire P, Drea S, Zarsky V, Dolan L. **2005**. A RhoGDP dissociation inhibitor spatially regulates growth in root hair cells. *Nature* **438**: 1013–6.

Castillon A, Shen H, Huq E. **2007**. Phytochrome Interacting Factors: central players in phytochrome-mediated light signaling networks. *Trends in Plant Science* **12**: 514–521.

Cavalier DM, Lerouxel O, Neumetzler L, Yamauchi K, Reinecke A, Freshour G, Zabotina O a., Hahn MG, Burgert I, Pauly M, Raikhel N V., Keegstra K. **2008**. Disrupting two Arabidopsis thaliana xylosyltransferase genes results in plants deficient in xyloglucan, a major primary cell wall component. *The Plant Cell* **20**: 1519–1537.

Cavers F. **1903**. On asexual reproduction and regeneration in Hepaticae. *New Phytologist* **2**: 121–133.

Chapin FSI, Oechel WS, Van Cleve K, Lawrence W. **1987**. The role of mosses in the phosphorous cycling of an Alaskan black spruce forest. *Oecologia* **74**: 310–315.

Chen L-Y, Shi D-Q, Zhang W-J, Tang Z-S, Liu J, Yang W-C. **2015**. The Arabidopsis alkaline ceramidase TOD1 is a key turgor pressure regulator in plant cells. *Nature Communications* **6**: 6030.

Chiyoda S, Ishizaki K, Kataoka H, Yamato KT, Kohchi T. **2008**. Direct transformation of the liverwort Marchantia polymorpha L. by particle bombardment using immature thalli developing from spores. *Plant Cell Reports* **27**: 1467–73.

Cho H-T, Cosgrove DJ. **2002**. Regulation of root hair initiation and expansin gene expression in Arabidopsis. *The Plant Cell* **14**: 3237–3253.

Clee DA. **1943**. The morphology and anatomy of Fegatella conica in relation to the mechanism of absorption and conduction of water. *Annals of Botany* **VII**: 185–193.

Clowes FAL. **2000**. Pattern in root meristem development in angiosperms. *New Phytologist* **146**: 83–94.

Cox CJ, Li B, Foster PG, Embley TM, Civan P. **2014**. Conflicting phylogenies for early land plants are caused by composition biases among synonymous substitutions. *Systematic Biology* **63**: 272–279.

Craddock C, Lavagi I, Yang Z. **2012**. New insights into Rho signaling from plant ROP/Rac GTPases. *Trends in Cell Biology* **22**: 492–501.

Crandall-Stotler B, Stotler RE. **2008**. Morphology and classification of the Marchantiophyta. In: Shaw A, Goffinet B, eds. *Bryophyte Biology*. New York: Cambridge University Press, 1–54.

Crandall-Stotler B, Stotler RE, Long DG. **2009**. Phylogeny and classification of the Marchantiophyta. *Edinburgh Journal of Botany* **66**: 155.

Crawford BCW, Ditta G, Yanofsky MF. **2007**. The NTT gene is required for transmitting-tract development in carpels of Arabidopsis thaliana. *Current Biology* **17**: 1101–8.

Crawford BCW, Sewell J, Golembeski G, Roshan C, Long JA, Yanofsky MF. **2015**. Genetic control of distal stem cell fate within root and embryonic meristems. *Science* **347**: 655–59.

Czaja AT. **1936**. Untersuchungen über den Membraneffekt des Absorptionsgewebes und über die Farbstoffaufnahme in die lebende Zelle. *Planta* **26**: 90–119.

Datta S, Kim CM, Pernas M, Pires ND, Proust H, Tam T, Vijayakumar P, Dolan L. **2011**. Root hairs: development, growth and evolution at the plant-soil interface. *Plant and Soil*: 1–14.

Datta S, Prescott H, Dolan L. **2015**. Intensity of a pulse of RSL4 transcription factor synthesis determines Arabidopsis root hair cell size. *Nature Plants* **1**: 15138.

Debeaujon I, Nesi N, Perez P, Devic M, Grandjean O, Caboche M, Lepiniec L. **2003**. Proanthocyanidin-accumulating cells in Arabidopsis testa: regulation of differentiation and role in seed development. *The Plant Cell* **15**: 2514–2531.

Delage E, Ruelland E, Guillas I, Zachowski A, Puyaubert J. **2012**. Arabidopsis type-III phosphatidylinositol 4-kinases β1 and β2 are upstream of the phospholipase C pathway triggered by cold exposure. *Plant and Cell Physiology* **53**: 565–576.

Dittmer HJ. 1949. Root hair variations in plant species. *American Journal of Botany* **36**: 152–155.

Dolan L, Duckett CM, Grierson C, Linstead P, Schneider K, Lawson E, Dean C, Poethig S, Roberts K. **1994**. Clonal relationships and cell patterning in the root epidermis of Arabidopsis. *Plant Science* **2474**: 2465–2474.

Dolan L, Roberts K. **1995**. The development of cell pattern in the root epidermis. *Philosophical Transactions of the Royal Society B* **350**: 95–9.

Driouich A, Follet-Gueye M-L, Bernard S, Kousar S, Chevalier L, Vicré-Gibouin M, Lerouxel O. **2012**. Golgi-mediated synthesis and secretion of matrix polysaccharides of the primary cell wall of higher plants. *Frontiers in Plant Science* **3**: 1–15.

Du F, Edwards K, Shen Z, Sun B, De Lozanne A, Briggs S, Firtel R a. **2008**. Regulation of contractile vacuole formation and activity in Dictyostelium. *The EMBO Journal* **27**: 2064–2076.

Duckett JG. **1994a**. Studies of protonemal morphogenesis in mosses VI. The foliar rhizoids of Calliergon stramineum (Brid.) Kindb. function as organs of attachment. *Journal of Bryology* **18**: 239–252.

Duckett JG. **1994b**. Studies of protonemal morphogenesis in mosses. V. Diphyscium foliosum (Hedw.) Mohr. (Buxbaumiales). *Journal of Bryology* **18**: 223–238.

Duckett JG, Matcham HW. **1995**. Studies of protonemal morphogenesis in mosses. VIII. The perennial rhizoids and gemmiferous protonema of Dicranella heteromalla (Hedw.) Schimp. *Journal of Bryology* **18**: 407–424.

Duckett JG, Renzaglia KS, Pell K. **1991**. A light and electron microscope study of rhizoid-ascomycete associations and flagelliform axes in British hepatics with observations on the effects of the fungi on host morphology. *New Phytologist* **118**: 233–257.

Duckett JG, Russell AJ, Ligrone R. **1996**. Trichomes in the Hymenophyllaceae. In: Camus JM, Gibby M, Johns RJ, eds. *Pteridology in Perspective*. Royal Botanic Gardens, Kew, 511–514.

Duckett JG, Schmid AM, Ligrone R. **1998**. Protonemal morphogenesis. In: Bates JW, Ashton NW, Duckett JG, eds. *Bryology for the twenty-first century*. Leeds, UK: British Bryological Society, 223–246.

Eklund DM, Ishizaki K, Flores-Sandoval E, Kikuchi S, Takebayashi Y, Tsukamoto S, Hirakawa Y, Nonomura M, Kato H, Kouno M, Bhalerao RP, Lagercrantz U, Kasahara H, Kohchi T, Bowman JL. **2015**. Auxin produced by the indole-3-pyruvic acid pathway regulates development and gemmae dormancy in the liverwort Marchantia polymorpha. *The Plant Cell*: tpc.15.00065.

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

Era A, Tominaga M, Ebine K, Awai C, Saito C, Ishizaki K, Yamato KT, Kohchi T, Nakano A, Ueda T. **2009**. Application of Lifeact reveals F-actin dynamics in Arabidopsis thaliana and the liverwort, Marchantia polymorpha. *Plant & Cell Physiology* **50**: 1041–8.

Farrar K, Evans IM, Topping JF, Souter M a., Nielsen JE, Lindsey K. **2003**. EXORDIUM - A gene expressed in proliferating cells and with a role in meristem function, identified by promoter trapping in Arabidopsis. *Plant Journal* **33**: 61–73.

Favery B, Ryan E, Foreman J, Linstead P, Boudonck K, Steer M, Shaw P, Dolan L. **2001**. KOJAK encodes a cellulose synthase-like protein required for root hair cell morphogenesis in Arabidopsis. *Genes & Development*: 79–89.

Finet C, Timme RE, Delwiche CF, Marlétaz F. **2010**. Multigene phylogeny of the green lineage reveals the origin and diversification of land plants. *Current Biology* **20**: 2217–2222.

Finka A, Saidi Y, Goloubinoff P, Neuhaus J-M, Zrÿd J-P, Schaefer DG. **2008**. The knock-out of ARP3a gene affects F-actin cytoskeleton organization altering cellular tip growth, morphology and development in moss Physcomitrella patens. *Cell Motility and the Cytoskeleton* **65**: 769–84.

Finn RD, Clements J, Eddy SR. **2011**. HMMER web server: Interactive sequence similarity searching. *Nucleic Acids Research* **39**: 29–37.

Fitting H. **1936**. Untersuchungen über die Induktion der Dorsiventralität bei den keimenden Brutkörpern von Marchantia und Lunularia. I. Die Induktoren und ihre Wirkungen. *Jahrbücher für wissenschaftliche Botanik* **82**: 333–376.

Fitting H. **1939**. Untersuchungen über den Einfluss von Licht und Dunkelheit auf die Entwicklung von Moosen. I. Die Brutkörper der Marchantieen. *Jahrbücher für wissenschaftliche Botanik* **88**: 633–722.

Flores-Sandoval E, Dierschke T, Fisher TJ, Bowman JL. **2015**. Efficient and inducible use of artificial microRNAs in Marchantia polymorpha. *Plant and Cell Physiology* **0**: 1–10.

Flores-Sandoval E, Eklund DM, Bowman JL. **2015**. A simple auxin transcriptional response system regulates multiple morphogenetic processes in the liverwort Marchantia polymorpha. *PLOS Genetics* **11**: e1005207.

Föhse D, Claassen N, Jungk A. **1991**. Phosphorus efficiency of plants II. Significance of root radius, root hairs and cation-anion balance for phosphorus influx in seven plant species. *Plant and Soil* **132**: 261–272.

Foreman J, Demidchik V, Bothwell JHF, Mylona P, Miedema H, Torres MA, Linstead P, Costa S, Brownlee C, Jones JDG, Davies JM, Dolan L. **2003**. Reactive oxygen species produced by NADPH oxidase regulate plant cell growth. *Nature* **422**: 442–6.

Gahoonia TS, Care D, Nielsen NE. **1997**. Root hairs and phosphorus acquisition of wheat and barley cultivars. *Plant and Soil* **191**: 181–188.

Gahoonia TS, Nielsen NE. **1996**. Variation in acquisition of soil phosphorus among wheat and barley genotypes. *Plant and Soil* **178**: 223–230.

Gahoonia TS, Nielsen NE. **1997**. Variation in root hairs of barley cultivars doubled soil phosphorus uptake. *Euphytica* **98**: 177–182.

Gahoonia TS, Nielsen NE. **1998**. Direct evidence on participation of root hairs in phosphorus (32P) uptake from soil. *Plant and Soil* **198**: 147–152.

Gahoonia TS, Nielsen NE. **2004**. Barley genotypes with long root hairs sustain high grain yields in low-P field. *Plant and Soil* **262**: 55–62.

Gahoonia TS, Nielsen NE, Joshi PA, Jahoor A. **2001**. A root hairless barley mutant for elucidating genetic of root hairs and phosphorus uptake. *Plant and Soil* **235**: 211–219.

Galway ME, Heckman JW, Schiefelbein J. **1997**. Growth and ultrastructure of Arabidopsis root hairs: the rhd3 mutation alters vacuole enlargement and tip growth. *Planta* **201**: 209–18.

Galway ME, Masucci JD, Lloyd AM, Walbot V, Davies RW, Schiefelbein J. **1994**. The TTG gene is required to specify epidermal cell fate and cell patterning in the Arabidopsis root. *Developmental Biology* **166**: 740–754.

Gamsjaeger R, Liew C, Loughlin F, Crossley M, Mackay J. **2007**. Sticky fingers: zinc-fingers as protein-recognition motifs. *Trends in Biochemical Sciences* **32**: 63–70.

Gensel PG. **2008**. The earliest land plants. *Annual Review of Ecology, Evolution, and Systematics* **39**: 459–477.

Gerald NJ, Siano M, De Lozanne A. **2002**. The Dictyostelium LvsA protein is localized on the contractile vacuole and is required for osmoregulation. *Traffic* **3**: 50–60.

Gheysen G, Montagu M V, Zambryski P. **1987**. Integration of Agrobacterium tumefaciens transfer DNA (T-DNA) involves rearrangements of target plant DNA sequences. *Proceedings of the National Academy of Sciences of the United States of America* **84**: 6169–6173.

Goffinet B, Buck WR, Shaw AJ. **2008**. Morphology, anatomy and classification of the Bryophyta. In: Shaw A, Goffinet B, eds. *Bryophyte Biology*. New York: Cambridge University Press, 55–138.

Graham LE, Graham JM, Wilcox LW. **2009**. *Algae*. San Francisco, CA, USA: Pearson Benjamin Cummings.

Grebe M, Friml J, Swarup R, Ljung K, Sandberg G, Terlou M, Palme K, Bennett MJ, Scheres B. **2002**. Cell polarity signaling in Arabidopsis involves a BFA-sensitive auxin influx pathway. *Current Biology* **12**: 329–34.

Grubb PJ. **1970**. Observations on the structure and biology of Haplomitrium and Takakia, hepatics with roots. *New Phytologist* **69**: 303–326.

Guindon S, Dufayard J-F, Lefort V, Anisimova M. **2010**. New alogrithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biology* **59**: 307–321.

Haberlandt GFJ. **1914**. *Physiological plant anatomy [Translated from the 4th German ed. by M. Drummond]*. London: MacMillon & Co.

Hammond GR V, Fischer MJ, Anderson KE, Holdich J, Koteci A, Balla T, Irvine RF. **2012**. PI4P and PI(4,5)P2 are essential but independent lipid determinants of membrane identity. *Science* **337**: 727–730.

Hannun Y a, Obeid LM. **2008**. Principles of bioactive lipid signalling: lessons from sphingolipids. *Nature Reviews Molecular Cell Biology* **9**: 139–150.

Harries P a, Pan A, Quatrano RS. **2005**. Actin-related protein2/3 complex component ARPC1 is required for proper cell morphogenesis and polarized cell growth in Physcomitrella patens. *The Plant Cell* **17**: 2327–2339.

Van Der Heijden MG a., Klironomos JN, Ursic M, Moutoglis P, Streitwolf-Engel R, Boller T, Sanders IR, Wiemken A, Sanders IR. **1998**. Mycorrhizal fungal diversity determines plant biodiversity, ecosystem variability and productivity. *Nature* **396**: 69–72.

Hepler P, Vidali L. **2001**. Polarized cell growth in higher plants. *Annual Review of Cell and Developmental Biology* **17**: 159–187.

Hiratsu K, Matsui K, Koyama T, Ohme-Takagi M. **2003**. Dominant repression of target genes by chimeric repressors that include the EAR motif, a repression domain, in Arabidopsis. *Plant Journal* **34**: 733–739.

Hofmeister W. **1851**. *Vergleichende Untersuchungen der Keimung, Entfaltung und Fruchtbildung höherer Kryptogamen (Moose, Farrn, Equisetaceen, Rhizocarpeen und Lycopodiaceen) und der Samenbildung der Coniferen.* Leipzig: Verlag F. Hofmeister.

Humphreys CP, Franks PJ, Rees M, Bidartondo MI, Leake JR, Beerling DJ. **2010**. Mutualistic mycorrhiza-like symbiosis in the most ancient group of land plants. *Nature Communications* **1**: 103.

Hussey PJ, Ketelaar T, Deeks MJ. **2006**. Control of the actin cytoskeleton in plant cell growth. *Annual Review of Plant Biology* **57**: 109–125.

Ikeda Y, Men S, Fischer U, Stepanova AN, Alonso JM, Ljung K, Grebe M. **2009**. Local auxin biosynthesis modulates gradient-directed planar polarity in Arabidopsis. *Nature Cell Biology* **11**: 731–8.

Ischebeck T, Seiler S, Heilmann I. **2010**. At the poles across kingdoms: Phosphoinositides and polar tip growth. *Protoplasma* **240**: 13–31.

Ishida T, Kurata T, Okada K, Wada T. **2008**. A genetic regulatory network in the development of trichomes and root hairs. *Annual Review of Plant Biology* **59**: 365–86.

Ishizaki K, Chiyoda S, Yamato KT, Kohchi T. **2008**. Agrobacterium-mediated transformation of the haploid liverwort Marchantia polymorpha L., an emerging model for plant biology. *Plant & Cell Physiology* **49**: 1084–91.

Ishizaki K, Johzuka-Hisatomi Y, Ishida S, Iida S, Kohchi T. **2013**. Homologous recombination-mediated gene targeting in the liverwort Marchantia polymorpha L. *Scientific Reports* **3**: 1532.

Ishizaki K, Mizutani M, Shimamura M, Masuda A, Nishihama R, Kohchi T. **2013**. Essential role of the E3 ubiquitin ligase nopperabo1 in schizogenous intercellular space formation in the liverwort Marchantia polymorpha. *The Plant Cell* **25**: 4075–84.

Ishizaki K, Nonomura M, Kato H, Yamato KT, Kohchi T. **2012**. Visualization of auxin-mediated transcriptional activation using a common auxin-responsive reporter system in the liverwort Marchantia polymorpha. *Journal of Plant Research*.

Iwakawa H, Tomari Y. **2013**. Molecular insights into microRNA-mediated translational repression in plants. *Molecular Cell* **52**: 591–601.

Jang G, Yi K, Pires ND, Menand B, Dolan L. **2011**. RSL genes are sufficient for rhizoid system development in early diverging land plants. *Development* **138**: 2273–2281.

Johnson CM, Stout PR, Broyer TC, Carlton a. B. **1957**. Comparative chlorine requirements of different plant species. *Plant and Soil* **8**: 337–353.

Jones MA, Raymond MJ, Yang Z, Smirnoff N. **2007**. NADPH oxidase-dependent reactive oxygen species formation required for root hair growth depends on ROP GTPase. *Journal of Experimental Botany* **58**: 1261–70.

Jones MA, Shen J, Fu Y, Li H, Yang Z, Grierson CS. **2002**. The Arabidopsis Rop2 GTPase is a positive regulator of both root hair initiation and tip growth. *The Plant Cell* **14**: 763–776.

Jungk A. **2001**. Root hairs and the acquisition of plant nutrients from soil. *Journal of Plant Nutrition and Soil Science* **164**: 121–129.

Jungkunz I, Link K, Vogel F, Voll LM, Sonnewald S, Sonnewald U. **2011**. AtHsp70-15-deficient Arabidopsis plants are characterized by reduced growth, a constitutive cytosolic protein response and enhanced resistance to TuMV. *Plant Journal* **66**: 983–995.

Kagale S, Links MG, Rozwadowski K. **2010**. Genome-wide analysis of ethylene-responsive element binding factor-associated amphiphilic repression motif-containing transcriptional regulators in Arabidopsis. *Plant Physiology* **152**: 1109–1134.

Kamerling Z. **1897**. Zur Biologie und Physiologie der Marchantiaceen. *Flora* **84**: 1–68.

Karol KG, McCourt RM, Cimino MT, Delwiche CF. **2001**. The closest living relatives of land plants. *Science* **294**: 2351–3.

Katoh K, Standley DM. **2013**. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution* **30**: 772–780.

Kenrick PR, Crane PR. **1997a**. The origin and early evolution of plants on land. *Nature* **389**: 33–39.

Kenrick PR, Crane PR. **1997b**. *The origin and early diversification of land plants: a cladistic study*. Washington, D. C.: Smithsonian Institution Press.

Kessler S a, Shimosato-Asano H, Keinath NF, Wuest SE, Ingram G, Panstruga R, Grossniklaus U. **2010**. Conserved molecular components for pollen tube reception and fungal invasion. *Science* **330**: 968–971.

Ketelaar T, Galway ME, Mulder BM, Emons a. MC. **2008**. Rates of exocytosis and endocytosis in Arabidopsis root hairs and pollen tubes. *Journal of Microscopy* **231**: 265–273.

Kiefer CS, Claes AR, Nzayisenga J, Pietra S, Stanislas T, Hu A, Kiefer CS, Claes AR, Nzayisenga J, Pietra S, Stanislas T. **2015**. Arabidopsis AIP1-2 restricted by WER-mediated patterning modulates planar polarity. *Development* **142**: 151–161.

Kim S-J, Brandizzi F. **2014**. The plant secretory pathway: an essential factory for building the plant cell wall. *Plant and Cell Physiology* **55**: 687–693.

Kirik V, Simon M, Huelskamp M, Schiefelbein J. **2004**. The ENHANCER of TRY and CPC1 gene acts redundantly with TRIPTYCHON and CAPRICE in trichome and root hair cell patterning in Arabidopsis. *Developmental Biology* **268**: 506–513.

Knox K, Grierson CS, Leyser O. **2003**. AXR3 and SHY2 interact to regulate root hair development. *Development* **130**: 5769–77.

Kny L. **1890**. Bau und Entwickelung von Marchantia polymorpha. *Botanische Wandtafeln mit Erläuterndem Text*. Berlin: Wiegandt, Hempel & Parey, .

Kobiyama Y, Crandall-Stotler B. **2011**. Studies of specialized pitted parenchyma cells of the liverwort Conocephalum Hill and their phylogenetic implications. *International Journal of Plant Sciences* **160**: 351–370.

Krinke O, Ruelland E, Valentová O, Vergnolle C, Renou J-P, Taconnat L, Flemr M, Burketová L, Zachowski A. **2007**. Phosphatidylinositol 4-kinase activation is an early response to salicylic acid in Arabidopsis suspension cells. *Plant Physiology* **144**: 1347–1359.

Kusano H, Testerink C, Vermeer JEM, Tsuge T, Shimada H, Oka A, Munnik T, Aoyama T. **2008**. The Arabidopsis phosphatidylinositol phosphate 5-kinase PIP5K3 is a key regulator of root hair tip growth. *The Plant Cell* **20**: 367–380.

Kwak S-H, Schiefelbein J. **2008**. A feedback mechanism controlling SCRAMBLED receptor accumulation and cell-type pattern in Arabidopsis. *Current Biology* **18**: 1949–1954.

Lambers H, Brundrett MC, Raven JA, Hopper S. **2010**. Plant mineral nutrition in ancient landscapes: high plant species diversity on infertile soils is linked to functional diversity for nutritional strategies. *Plant and Soil* **334**: 11–31.

Lee MM, Schiefelbein J. **1999**. WEREWOLF, a MYB-related protein in Arabidopsis, is a position-dependent regulator of epidermal cell patterning. *Cell* **99**: 473–483.

Lee JH, Schöffl F. **1996**. An Hsp70 antisense gene affects the expression of HSP70/HSC70, the regulation of HSF, and the acquisition of thermotolerance in transgenic Arabidopsis thaliana. *Molecular and General Genetics* **252**: 11–19.

Lee BM, Xu J, Clarkson BK, Martinez-Yamout M a., Dyson HJ, Case D a., Gottesfeld JM, Wright PE. **2006**. Induced Fit and “Lock and Key” Recognition of 5S RNA by Zinc Fingers of Transcription Factor IIIA. *Journal of Molecular Biology* **357**: 275–291.

Leliaert F, Smith DR, Moreau H, Herron MD, Verbruggen H, Delwiche CF, De Clerck O. **2012**. Phylogeny and Molecular Evolution of the Green Algae. *Critical Reviews in Plant Sciences* **31**: 1–46.

Lenton TM, Crouch M, Johnson M, Pires N, Dolan L. **2012**. First plants cooled the Ordovician. *Nature Geoscience* **5**: 86–89.

Lewis LA, McCourt RM. **2004**. Green algae and the origin of land plants. *American Journal of Botany* **91**: 1535–1556.

Li S, Liu L, Zhuang X, Yu Y, Liu X, Cui X, Ji L, Pan Z, Cao X, Mo B, Zhang F, Raikhel N, Jiang L, Chen X. **2013**. MicroRNAs inhibit the translation of target mRNAs on the endoplasmic reticulum in Arabidopsis. *Cell* **153**: 562–574.

Lin BL, Wang JS, Liu HC, Chen RW, Meyer Y, Barakat A, Delseny M. **2001**. Genomic analysis of the Hsp70 superfamily in Arabidopsis thaliana. *Cell Stress & Chaperones* **6**: 201–208.

Lipka V, Kwon C, Panstruga R. **2007**. SNARE-ware: The role of SNARE-domain proteins in plant biology. *Annual Review of Cell and Developmental Biology* **23**: 147–174.

Lloyd J, Meinke D. **2012**. A comprehensive dataset of genes with a loss-of-function mutant phenotype in Arabidopsis. *Plant Physiology* **158**: 1115–1129.

Lorbeer G. **1938**. Über das Vorkommen von drei verschiedenen Geschlechtsrealisatoren bei den Lebermoosen. *Planta* **27**: 708–717.

Mägdefrau K. **1938**. Reviews of recent research 2. Der Wasserhaushalt der Moose. *Annales Bryologici* **10**: 141–150.

Maravolo NC. **1980**. Control of development in hepatics. *Bulletin of the Torrey Botanical Club* **107**: 308–324.

Maravolo NC, Voth PD. **1966**. Morphogenic effects of three growth substances on Marchantia gemmalings. *Botanical Gazette* **127**: 79–86.

Marschner P. **2012**. *Marschner's mineral nutrition of higher plants*. London: Academic Press.

Martin A, Troadec C, Boualem A, Rajab M, Fernandez R, Morin H, Pitrat M, Dogimont C, Bendahmane A. **2009**. A transposon-induced epigenetic change leads to sex determination in melon. *Nature* **461**: 1135–1138.

Masucci JD, Rerie WG, Foreman DR, Zhang M, Galway ME, Marks MD, Schiefelbein JW. **1996**. The homeobox gene GLABRA2 is required for position-dependent cell differentiation in the root epidermis of Arabidopsis thaliana. *Development* **122**: 1253–1260.

Masucci JD, Schiefelbein J. **1994**. The rhd6 mutation of Arabidopsis thaliana alters root-hair initiation through an auxin- and ethylene-associated process. *Plant Physiology* **106**: 1335–1346.

Masucci JD, Schiefelbein JW. **1996**. Hormones act downstream of TTG and GL2 to promote root hair outgrowth during epidermis development in the Arabidopsis root. *The Plant Cell* **8**: 1505–1517.

McConaha M. **1939**. Ventral surface specializations of Conocephalum conicum. *American Journal of Botany* **26**: 353–355.

McConaha M. **1941**. Ventral structures effecting capillarity in the Marchantiales. *American Journal of Botany* **28**: 301–306.

McManus HA, Qiu Y-L. **2008**. Life cycles in major lineages of photosynthetic eukaryotes, with a special reference to the origin of land plants. *Fieldiana Botany*: 17–33.

Menand B, Yi K, Jouannic S, Hoffmann L, Ryan E, Linstead P, Schaefer DG, Dolan L. **2007**. An ancient mechanism controls the development of cells with a rooting function in land plants. *Science* **316**: 1477–1480.

Miller MW, Garber ED, Voth PD. **1962**. Biosynthetic pathways in nutritionally deficient mutants of Marchantia polymorpha L. *Nature* **195**: 1220–1221.

Mirbel M. **1835**. Recherches anatomiques et physiologiques sur le Marchantia polymorpha, pour servir a l'histoire du tissue cellulaire, de l'épiderme et des stomates. *Mém. de l'Acad. Royale des Sci. de l'Inst. de France* **Tome XIII**: 337–373.

Molendijk AJ, Bischoff F, Rajendrakumar CS, Friml J, Braun M, Gilroy S, Palme K. **2001**. Arabidopsis thaliana Rop GTPases are localized to tips of root hairs and control polar growth. *The EMBO journal* **20**: 2779–88.

Monshausen GB, Bibikova TN, Messerli M a, Shi C, Gilroy S. **2007**. Oscillations in extracellular pH and reactive oxygen species modulate tip growth of Arabidopsis root hairs. *Proceedings of the National Academy of Sciences of the United States of America* **104**: 20996–1001.

Morimoto RI. **1998**. Regulation of the heat-shock transcriptional response: cross talk between a family of heat-shock factors, molecular chaperones, and negative regulators. *Genes & Development* **12**: 3788–3796.

Nacry P, Camilleri C, Courtial B, Caboche M, Bouchez D. **1998**. Major chromosomal rearrangements induced by T-DNA transformation in Arabidopsis. *Genetics* **149**: 641–650.

Nadeau JA. **2009**. Stomatal development: new signals and fate determinants. *Current Opinion in Plant Biology* **12**: 29–35.

Nawy T, Lee J-Y, Colinas J, Wang JY, Thongrod SC, Malamy JE, Birnbaum K, Benfey PN. **2005**. Transcriptional profile of the Arabidopsis root quiescent center. *The Plant Cell* **17**: 1908–25.

Nielsen E, Cheung a. Y, Ueda T. **2008**. The regulatory RAB and ARF GTPases for vesicular trafficking. *Plant Physiology* **147**: 1516–1526.

Nye PH. **1966**. The effect of the nutrient intensity and buffering power of a soil, and the absorbing power, size and root hairs of a root, on nutrient absorption by diffusion. *Plant and Soil* **25**: 81–105.

Odu EA. **1978**. The adaptive importance of moss rhizoids for attachment to the substratum. *Journal of Bryology* **10**: 163–181.

Odu EA. **1989**. Extracellular adhesive substances on bryophyte rhizoids. *Acta Botanica Hungarica* **35**: 273–277.

Odu EA, Richards P. **1976**. The stimulus to branching of the rhizoid tip in Lophocolea cuspidata (Nees) Limpr. *Journal of Bryology* **9**: 93–05.

Okada S, Fujisawa M, Sone T, Nakayama S, Nishiyama R, Takenaka M. **2000**. Construction of male and female PAC genomic libraries suitable for identification of Y-chromosome-specific clones from the liverwort , Marchantia polymorpha. *The Plant Journal* **24**: 421–428.

Ovecka M, Berson T, Beck M, Derksen J, Samaj J, Baluska F, Lichtscheidl IK. **2010**. Structural sterols are involved in both the initiation and tip growth of root hairs in Arabidopsis thaliana. *The Plant Cell* **22**: 2999–3019.

Park S, Szumlanski AL, Gu F, Guo F, Nielsen E. **2011**. A role for CSLD3 during cell-wall synthesis in apical plasma membranes of tip-growing root-hair cells. *Nature Cell Biology* **13**: 973–980.

Parker JS, Cavell a C, Dolan L, Roberts K, Grierson CS. **2000**. Genetic interactions during root hair morphogenesis in Arabidopsis. *The Plant Cell* **12**: 1961–1974.

Parniske M. **2008**. Arbuscular mycorrhiza: the mother of plant root endosymbioses. *Nature Reviews Microbiology* **6**: 763–75.

Parton RM, Dyer AF, Read ND, Trewavas AJ. **2000**. Apical Structure of Actively Growing Fern Rhizoids Examined by DIC and Confocal Microscopy. *Annals of Botany* **85**: 233–245.

Pearson HM. **1969**. Rhizoids and root hairs of ferns. *American Fern Journal* **59**: 107–122.

Pemberton L. **2001**. Epidermal patterning in seedling roots of eudicotyledons. *Annals of Botany* **87**: 649–654.

Pena MJ, Kong Y, York WS, O’Neill M a. 2012. A galacturonic acid-containing xyloglucan is involved in Arabidopsis root hair tip growth. *The Plant Cell* **24**: 4511–4524.

Peremyslov V V., Klocko AL, Fowler JE, Dolja V V. 2012. Arabidopsis myosin XI-K localizes to the motile endomembrane vesicles associated with F-actin. *Frontiers in Plant Science* **3**: 1–10.

Peremyslov V V, Morgun E a, Kurth EG, Makarova KS, Koonin E V, Dolja V V. **2013**. Identification of myosin XI receptors in Arabidopsis defines a distinct class of transport vesicles. *The Plant Cell* **25**: 3022–38.

Peremyslov V V, Prokhnevsky AI, Avisar D, Dolja V V. **2008**. Two class XI myosins function in organelle trafficking and root hair development in Arabidopsis. *Plant Physiology* **146**: 1109–1116.

Peremyslov V V, Prokhnevsky AI, Dolja V V. **2010**. Class XI myosins are required for development, cell expansion, and F-Actin organization in Arabidopsis. *The Plant Cell* **22**: 1883–1897.

Perroud P-F, Quatrano RS. **2006**. The role of ARPC4 in tip growth and alignment of the polar axis in filaments of *Physcomitrella patens*. *Cell Motility and the Cytoskeleton* **63**: 162–171.

Petersen TN, Brunak S, von Heijne G, Nielsen H. **2011**. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nature Methods* **8**: 785–786.

Peterson RL, Farquhar ML. **1996**. Root hairs: specialized tubular cells extending root surfaces. *Botanical Review* **62**: 1–40.

Petricka JJ, Clay NK, Nelson TM. **2008**. Vein patterning screens and the defectively organized tributaries mutants in Arabidopsis thaliana. *The Plant Journal* **56**: 251–63.

Pillitteri LJ, Torii KU. **2012**. Mechanisms of stomatal development. *Annual Review of Plant Biology* **63**: 591–614.

Pires N, Dolan L. **2010**. Origin and diversification of basic-helix-loop-helix proteins in plants. *Molecular Biology and Evolution* **27**: 862–74.

Pires ND, Yi K, Breuninger H, Catarino B, Menand B, Dolan L. **2013**. Recruitment and remodeling of an ancient gene regulatory network during land plant evolution. *Proceedings of the National Academy of Sciences of the United States of America* **110**: 9571–6.

Pitts RJ, Cernac A, Estelle M. **1998**. Auxin and ethylene promote root hair elongation in Arabidopsis. *The Plant Journal* **16**: 553–60.

Pocock K, Duckett JG. **1985**. On the occurence of branched and swollen rhizoids in British hepatics: their relationships with the substratum and associations with fungi. *New Phytologist* **99**: 281–304.

Pressel S, Duckett JG. **2009**. Studies of protonemal morphogenesis in mosses. XII. Ephemeropsis, the zenith of morphological differentiation. *Journal of Bryology* **31**: 67–75.

Pressel S, Ligrone R, Duckett JG. **2008**. Cellular differentiation in moss protonemata: a morphological and experimental study. *Annals of Botany* **102**: 227–45.

Preuss ML, Schmitz AJ, Thole JM, Bonner HKS, Otegui MS, Nielsen E. **2006**. A role for the RabA4b effector protein PI-4Kβ1 in polarized expansion of root hair cells in Arabidopsis thaliana. *Journal of Cell Biology* **172**: 991–998.

Proctor MCF. **1984**. Structure and ecological adaptation. In: Dyer AF, Duckett JG, eds. *The experimental biology of bryophytes*. London: Academic Press, 9–37.

Proctor MCF. **2000**. The bryophyte paradox : tolerance of desiccation, evasion of drought. *Plant Ecology* **151**: 41–49.

Prokhnevsky AI, Peremyslov V V, Dolja V V. **2008**. Overlapping functions of the four class XI myosins in Arabidopsis growth, root hair elongation, and organelle motility. *Proceedings of the National Academy of Sciences of the United States of America* **105**: 19744–19749.

Pusapati G V., Luchetti G, Pfeffer SR. **2012**. Ric1-Rgp1 complex is a guanine nucleotide exchange factor for the late Golgi Rab6A GTPase and an effector of the medial Golgi Rab33B GTPase. *Journal of Biological Chemistry* **287**: 42129–42137.

Qiu Y-L, Li L, Wang B, Chen Z, Dombrovska O, Lee J, Kent L, Li R, Jobson RW, Hendry TA, Taylor DW, Testa CM, Ambros M. **2007**. A nonflowering land plant phylogeny inferred from nucleotide sequences of seven chloroplast, mitochondrial, and nuclear genes. *International Journal of Plant Sciences* **168**: 691–708.

Qiu Y-L, Li L, Wang B, Chen Z, Knoop V, Groth-Malonek M, Dombrovska O, Lee J, Kent L, Rest J, Estabrook GF, Hendry T a, Taylor DW, Testa CM, Ambros M, Crandall-Stotler B, Duff RJ, Stech M, Frey W, Quandt D, Davis CC. **2006**. The deepest divergences in land plants inferred from phylogenomic evidence. *Proceedings of the National Academy of Sciences of the United States of America* **103**: 15511–6.

Rahman A, Hosokawa S, Oono Y, Amakawa T, Goto N, Tsurumi S. **2002**. Auxin and ethylene response interactions during Arabidopsis root hair development dissected by auxin influx modulators. *Plant Physiology* **130**: 1908–1917.

Raven JA, Edwards D. **2001**. Roots: evolutionary origins and biogeochemical significance. *Journal of Experimental Botany* **52**: 381–401.

Raviol H, Sadlish H, Rodriguez F, Mayer MP, Bukau B. **2006**. Chaperone network in the yeast cytosol: Hsp110 is revealed as an Hsp70 nucleotide exchange factor. *The EMBO Journal* **25**: 2510–2518.

Read DJ, Duckett JG, Francis R, Ligrone R, Russell A. **2000**. Symbiotic fungal associations in “lower” land plants. *Philosophical Transactions of the Royal Society of London. Series B, Biological sciences* **355**: 815–30; discussion 830–1.

Reiss H-D, Herth W. **1979**. Calcium gradients in tip growing plant cells visualized by chlorotetracycline fluorescence. *Planta* **146**: 615–621.

Remy W, Taylor TN, Hass H, Kerp H. **1994**. Four hundred-million-year-old vesicular arbuscular mycorrhizae. *Proceedings of the National Academy of Sciences of the United States of America* **91**: 11841–11843.

Renzaglia KS, Villarreal JC, Duff RJ. **2008**. New insights into morphology, anatomy, and systematics of hornworts. In: Shaw A, Goffinet B, eds. *Bryophyte Biology*. New York: Cambridge University Press, 139–171.

Ridge RW, Uozumi Y, Plazinski J, Hurley UA, Williamson RE. **1999**. Developmental transitions and dynamics of the cortical ER of Arabidopsis cells seen with green fluorescent protein. *Plant & Cell Physiology* **40**: 1253–61.

Ringli C, Baumberger N, Diet A, Frey B, Keller B. **2002**. ACTIN2 is essential for bulge site selection and tip growth during root hair development of Arabidopsis. *Plant Physiology* **129**: 1464–1472.

Rounds CM, Bezanilla M. **2013**. Growth mechanisms in tip-growing plant cells. *Annual Review of Plant Biology* **64**: 243–65.

Rubinstein C V, Gerrienne P, de la Puente GS, Astini RA, Steemans P. **2010**. Early Middle Ordovician evidence for land plants in Argentina (eastern Gondwana). *New Phytologist* **188**: 365–9.

Ruhfel BR, Gitzendanner M a, Soltis PS, Soltis DE, Burleigh JG. **2014**. From algae to angiosperms - inferring the phylogeny of green plants (Viridiplantae) from 360 plastid genomes. *BMC Evolutionary Biology* **14**: 23.

Ruijter JM, Ramakers C, Hoogaars WMH, Karlen Y, Bakker O, Van Den Hoff MJB, Moorman a. FM. **2009**. Amplification efficiency: Linking baseline and bias in the analysis of quantitative PCR data. *Nucleic Acids Research* **37**.

Russell J, Bulman S. **2005**. The liverwort Marchantia foliacea forms a specialized symbiosis with arbuscular mycorrhizal fungi in the genus Glomus. *New Phytologist* **165**: 567–79.

Saedler R, Jakoby M, Marin B, Galiana-Jaime E, Hülskamp M. **2009**. The cell morphogenesis gene SPIRRIG in Arabidopsis encodes a WD/BEACH domain protein. *Plant Journal* **59**: 612–621.

Sagasser M, Lu G-H, Hahlbrock K, Weisshaar B. **2002**. A. thaliana TRANSPARENT TESTA 1 is involved in seed coat development and defines the WIP subfamily of plant zinc finger proteins. *Genes & Development* **16**: 138–49.

Saint-Marcoux D, Proust H, Dolan L, Langdale J a. 2015. Identification of reference genes for real-time quantitative PCR experiments in the liverwort Marchantia polymorpha. *Plos One* **10**: e0118678.

Sakakibara K, Nishiyama T, Sumikawa N, Kofuji R, Murata T, Hasebe M. **2003**. Involvement of auxin and a homeodomain-leucine zipper I gene in rhizoid development of the moss Physcomitrella patens. *Development* **130**: 4835–46.

Sasaki G, Katoh K, Hirose N, Suga H, Kuma K, Miyata T, Su Z-H. **2007**. Multiple receptor-like kinase cDNAs from liverwort Marchantia polymorpha and two charophycean green algae, Closterium ehrenbergii and Nitella axillaris: Extensive gene duplications and gene shufflings in the early evolution of streptophytes. *Gene* **401**: 135–44.

Savage NS, Walker T, Wieckowski Y, Schiefelbein J, Dolan L, Monk NAM. **2008**. A mutual support mechanism through intercellular movement of CAPRICE and GLABRA3 can pattern the Arabidopsis root epidermis (D Weigel, Ed.). *PLoS Biology* **6**: 11.

Scheller HV, Ulvskov P. **2010**. Hemicelluloses. *Annual Review of Plant Biology* **61**: 263–289.

Schellmann S, Schnittger A, Kirik V, Wada T, Okada K, Beermann a., Thumfahrt J, Jürgens G, Hülskamp M. **2002**. *TRIPTYCHON* and *CAPRICE* mediate lateral inhibition during trichome and root hair patterning in *Arabidopsis*. *The EMBO Journal* **21**: 5036–5046.

Schiefelbein JW. **2000**. Constructing a plant cell. The genetic control of root hair development. *Plant Physiology* **124**: 1525–1531.

Schiefelbein J, Kwak S-H, Wieckowski Y, Barron C, Bruex A. **2009**. The gene regulatory network for root epidermal cell-type pattern formation in Arabidopsis. *Journal of Experimental Botany* **60**: 1515–1521.

Schiefelbein J, Shipley A, Rowse P. **1992**. Calcium influx at the tip of growing root-hair cells of Arabidopsis thaliana. *Planta* **187**: 455–459.

Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, Tinevez J-Y, White DJ, Hartenstein V, Eliceiri K, Tomancak P, Cardona A. **2012**. Fiji: an open-source platform for biological-image analysis. *Nature Methods* **9**: 676–682.

Schröder F, Lisso J, Lange P, Müssig C. **2009**. The extracellular EXO protein mediates cell expansion in Arabidopsis leaves. *BMC Plant Biology* **9**: 20.

Shih H-W, Miller ND, Dai C, Spalding EP, Monshausen GB. **2014**. The Receptor-like kinase FERONIA is required for mechanical signal transduction in Arabidopsis seedlings. *Current Biology* **24**: 1887–1892.

Simon M, Lee MM, Lin Y, Gish L, Schiefelbein J. **2007**. Distinct and overlapping roles of single-repeat MYB genes in root epidermal patterning. *Developmental Biology* **311**: 566–578.

Simpson P. **1990**. Lateral inhibition and the development of the sensory bristles of the adult peripheral nervous system of Drosophila. *Development* **109**: 509–519.

Siniossoglou S, Peak-Chew SY, Pelham HR. **2000**. Ric1p and Rgp1p form a complex that catalyses nucleotide exchange on Ypt6p. *The EMBO Journal* **19**: 4885–4894.

St John T V. **1980**. Root size, root hairs and mycorrhizal infection: a reexamination of Baylis's hypothesis with tropical trees. *New Phytologist* **84**: 483–487.

Steffens A, Bräutigam A, Jakoby M, Hülskamp M. **2015**. The BEACH Domain Protein SPIRRIG Is Essential for Arabidopsis Salt Stress Tolerance and Functions as a Regulator of Transcript Stabilization and Localization. *PLOS Biology* **13**: e1002188.

Stenmark H. **2009**. Rab GTPases as coordinators of vesicle traffic. *Nature Reviews Molecular Cell Biology* **10**: 513–525.

Stenzel I, Ischebeck T, König S, Hołubowska A, Sporysz M, Hause B, Heilmann I. **2008**. The type B phosphatidylinositol-4-phosphate 5-kinase 3 is essential for root hair formation in Arabidopsis thaliana. *The Plant Cell* **20**: 124–141.

Strasburger E. **1894**. Uber periodische Reduction der Chromosomenzahl im. Entwicklungsgang der Organismen. *Biologisches Centralblatt* **14**: 8I7–838; 849–866.

Sugano SS, Shirakawa M, Takagi J, Matsuda Y, Shimada T, Hara-Nishimura I, Kohchi T. **2014**. CRISPR/Cas9-mediated targeted mutagenesis in the liverwort Marchantia polymorpha L. *Plant and Cell Physiology* **55**: 475–481.

Synek L, Schlager N, Eliáš M, Quentin M, Hauser M-T, Žárský V. **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.

Takano M, Egawa H, Ikeda JE, Wakasa K. **1997**. The structures of integration sites in transgenic rice. *The Plant Journal* **11**: 353–361.

Takeda S, Gapper C, Kaya H, Bell E, Kuchitsu K, Dolan L. **2008**. Local positive feedback regulation determines cell shape in root hair cells. *Science* **319**: 1241–4.

Tam THY, Catarino B, Dolan L. **2015**. Conserved regulatory mechanism controls the development of cells with rooting functions in land plants. *Proceedings of the National Academy of Sciences of the United States of America*: 201416324.

Tanimoto M, Roberts K, Dolan L. **1995**. Ethylene is a positive regulator of root hair development in Arabidopsis thaliana. *The Plant Journal* **8**: 943–948.

Tarén N. **1958**. Factors regulating the initial development of gemmae in Marchantia polymorpha. *The Bryologist* **61**: 191–204.

Van Tooren BF, Van Dam D, During HJ. **1990**. The relative importance of precipitation and soil as sources of nutrients for Calliergonella cuspidata (Hedw .) Loeske in chalk grassland. *Functional Ecology* **4**: 101–107.

Trachtenberg S, Zamski E. **1979**. The apoplastic conduction of water in Polytrichum juniperinum Willd. gametophytes. *New Phytologist* **83**: 49–52.

Tzafrir I, Pena-Muralla R, Dickerman A, Berg M, Rogers R, Hutchens S, Sweeney TC, Mcelver J, Aux G, Patton D, Meinke D, Oklahoma IT. **2004**. Identification of genes required for embryo development in Arabidopsis. *Plant Physiology* **135**: 1206–1220.

Ueda M, Koshino-Kimura Y, Okada K. **2005**. Stepwise understanding of root development. *Current Opinion in Plant Biology* **8**: 71–76.

Vermeer C, Escher M, Portielje R, de Klein JJM. **2003**. Nitrogen uptake and translocation by Chara. *Aquatic Botany* **76**: 245–258.

Vidali L, Burkart GM, Augustine RC, Kerdavid E, Tuzel E, Bezanilla M. **2010**. Myosin XI is essential for tip growth in Physcomitrella patens. *The Plant Cell* **22**: 1868–1882.

Vissenberg K, Fry SC, Verbelen JP. **2001**. Root hair initiation is coupled to a highly localized increase of xyloglucan endotransglycosylase action in Arabidopsis roots. *Plant Physiology* **127**: 1125–1135.

Völker H, Halbsguth W. **1966**. Zur Feinstruktur der Rhizoidinitialen bei den Brutkörpern von Marchantia polymorpha L. *Zeischrift für Pflanzenphysiologie* **55**: 158–164.

Wada T, Tachibana T, Shimura Y, Okada K. **1997**. Epidermal cell differentiation in Arabidopsis determined by a Myb homolog, CPC. *Science* **277**: 1113–1116.

Wang X, Cnops G, Vanderhaeghen R, De Block S, Van Montagu M, Van Lijsebettens M. **2001**. AtCSLD3, a cellulose synthase-like gene important for root hair growth in Arabidopsis. *Plant Physiology* **126**: 575–586.

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

Wang W, Vinocur B, Shoseyov O, Altman A. **2004**. Role of plant heat-shock proteins and molecular chaperones in the abiotic stress response. *Trends in Plant Science* **9**: 244–252.

Wang B, Yeun LH, Xue J-Y, Liu Y, Ané J-M, 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. *The New Phytologist* **186**: 514–25.

Wasteneys GO, Ambrose JC. **2009**. Spatial organization of plant cortical microtubules: close encounters of the 2D kind. *Trends in Cell Biology* **19**: 62–71.

Wellman CH, Osterloff PL, Mohiuddin U. **2003**. Fragments of the earliest land plants. *Nature* **425**: 282–285.

Wodniok S, Brinkmann H, Glöckner G, Heidel AJ, Philippe H, Melkonian M, Becker B. **2011**. Origin of land plants: do conjugating green algae hold the key? *BMC Evolutionary Biology* **11**: 104.

Wolfe SA, Nekludova L, Pabo CO. **2000**. DNA recognition by Cys2His2 zinc finger proteins. *Annual Review of Biophysics and Biomolecular Structure* **29**: 183–212.

Wong K, Cantley LC. **1994**. Cloning and characterization of a human phosphatidylinositol 4-kinase. *Journal of Biological Chemistry* **269**: 28878–28884.

Wu S-Z, Ritchie J a., Pan A-H, Quatrano RS, Bezanilla M. **2011**. Myosin VIII regulates protonemal patterning and developmental timing in the moss Physcomitrella patens. *Molecular Plant* **4**: 909–921.

Wymer CL, Bibikova TN, Gilroy S. **1997**. Cytoplasmic free calcium distributions during the development of root hairs of Arabidopsis thaliana. *The Plant Journal* **12**: 427–439.

Yamaoka S, Takenaka M, Hanajiri T, Shimizu-Ueda Y, Nishida H, Yamato KT, Fukuzawa H, Ohyama K. **2004**. A mutant with constitutive sexual organ development in Marchantia polymorpha L. *Sexual Plant Reproduction* **16**: 253–257.

Yi K, Menand B, Bell E, Dolan L. **2010**. A basic helix-loop-helix transcription factor controls cell growth and size in root hairs. *Nature Genetics* **42**: 264–7.

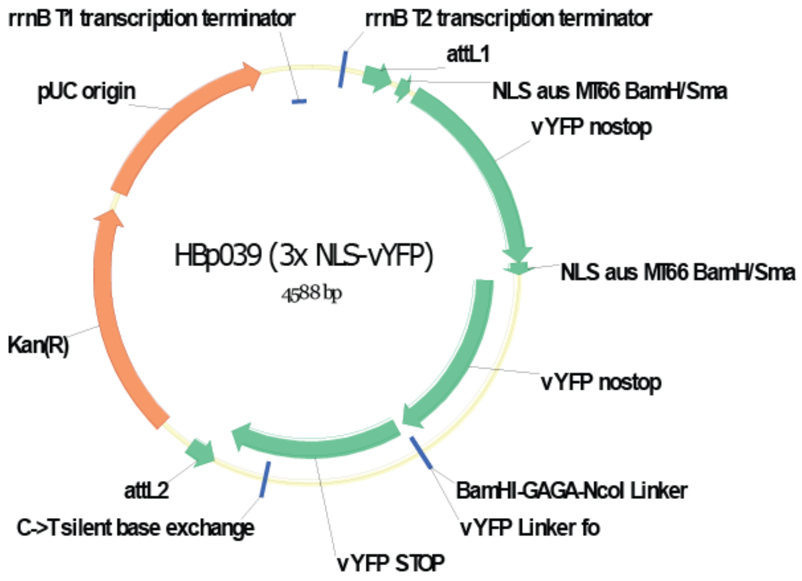
Yoo CM, Quan L, Cannon AE, Wen J, Blancaflor EB. **2012**. AGD1, a class 1 ARF-GAP, acts in common signaling pathways with phosphoinositide metabolism and the actin cytoskeleton in controlling Arabidopsis root hair polarity. *Plant Journal* **69**: 1064–1076.

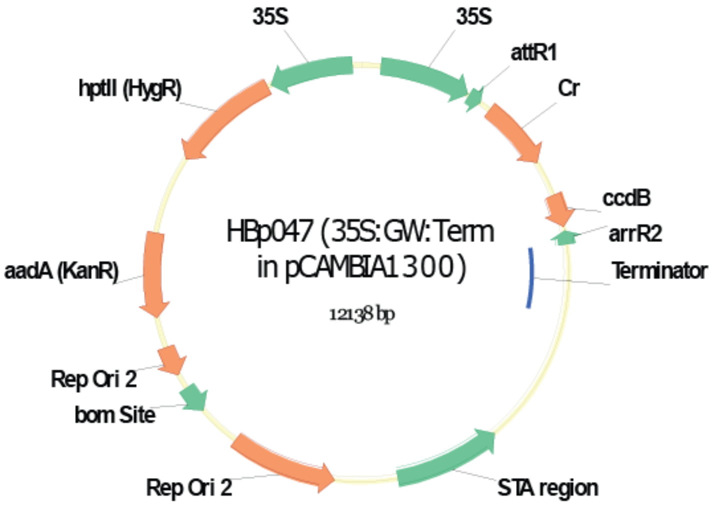
Appendices

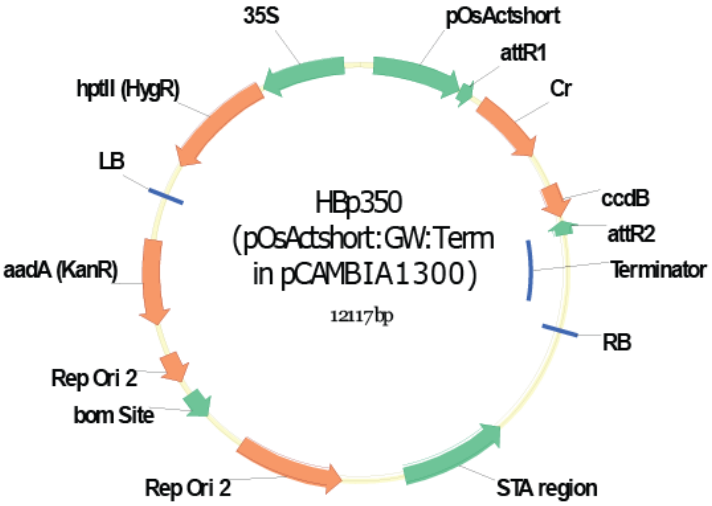
Appendix 1 Primer sequences

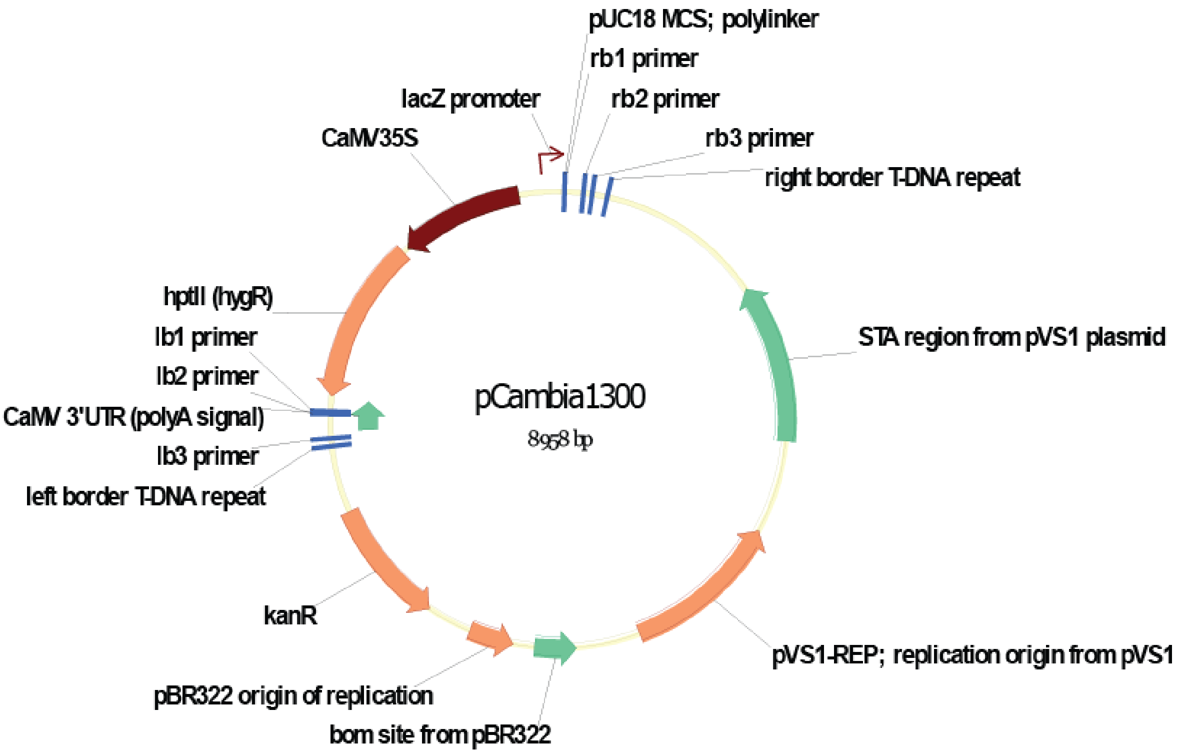
AD1	NGTCGASWGANAWGAA
AD2	TGWGNAGSANCASAGA
AD3	AGWGNAGWANCAWAGG
AD4	GTNCGASWCANAWGTT
AD5	NTCGASTWTSGWGTT
AD6	WGTGNAGWANCANAGA
LB1	CAGATAAGGGAATTAGGGTTCCTATAGG
LB2	TATAGGGTTTCGCTCATGTGTTGAGC
LB3	AGTACATTAAAAACGTCCGCAATGTG
MpCUL F	AGGATGTGGACAAGGATAGACG
MpCUL R	GTTGATGTGGCAACACCTTG
MpEF1α F	CCGAGATCCTGACCAAGG
MpEF1α R	GAGGTGGGTACTCAGCGAAG
MpELF5 F	AGGTTTCCACCTCCAAGACC
MpELF5 R	AACGACCTCAGGGACATCAC
MpRSL1 F	AGATGAGTCTGGGGCAACC
MpRSL1 R	GGATGAGCGCTTTAGAGTG
RB1	CCTGCAGGCATGCAAGCTTGG
RB2	GCTGGCGTAATAGCGAAGAGG
RB3	CCTGAATGGCGAATGCTAGAG
VJ_45	GCAGACATTGATTGAAGTCGC
VJ_46	AAATCACTGACTGCAATTGAAGG
VJ_66	CGCTACAACAACATGCAGATGC
VJ_67	GGGTGGTCGATGTTGTTCC
VJ_81	TGTGTCAAACGAATACTCGAGAG
VJ_82	TCTCTCTCTCTCTCTCTCTATC
VJ_83	ATGAGCTCGAGGTACCATGCC
VJ_84	TCATTGCGATCCGGCCGAC
VJ_141	ACCCGAAGCTTCTGAACGTA
VJ_143	GCCCATCATAACAGCACCAC
VJ_144	TTATTTCAAGCTCCGCGACG
VJ_186	CGTGCTCAGACCCCCTTC
VJ_187	CACCATCTGTCATCTGAGTCTGC
VJ_188	GGGATGATGGGCAAGGCT
VJ_190	AACGAAAGCTCGATCCTCTAGACCCGAAGCTTCTGAACGTAA
VJ_191	TTCGAGCTCGGTACCCGGGTTAACGTGCTCAGACCCCCTTC
VJ_213	AGGCATCTGGTATCCACGAG
VJ_214	ACATGGTCGTTCCTCCAGAC
VJ_215	CGAAAGCCCAAGAAGCTACC
VJ_216	GTACCCCCGGTTGCAATAAG

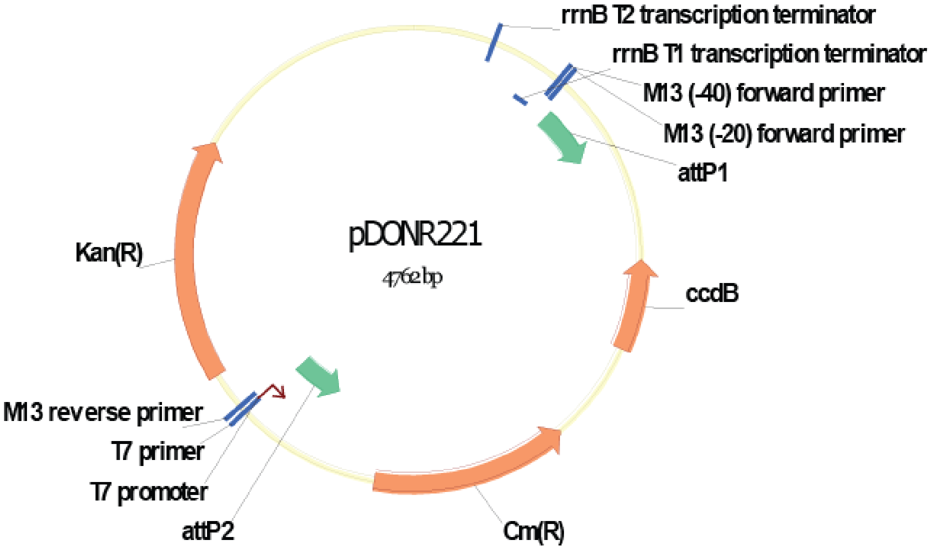
Appendix 2 Plasmid maps

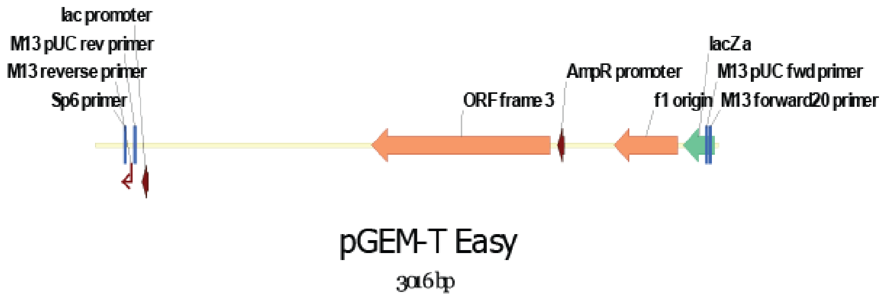


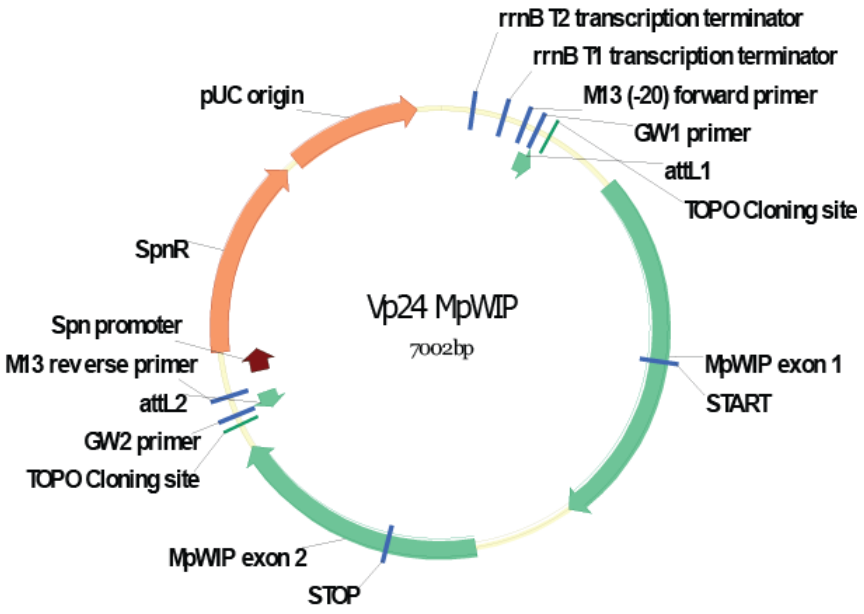


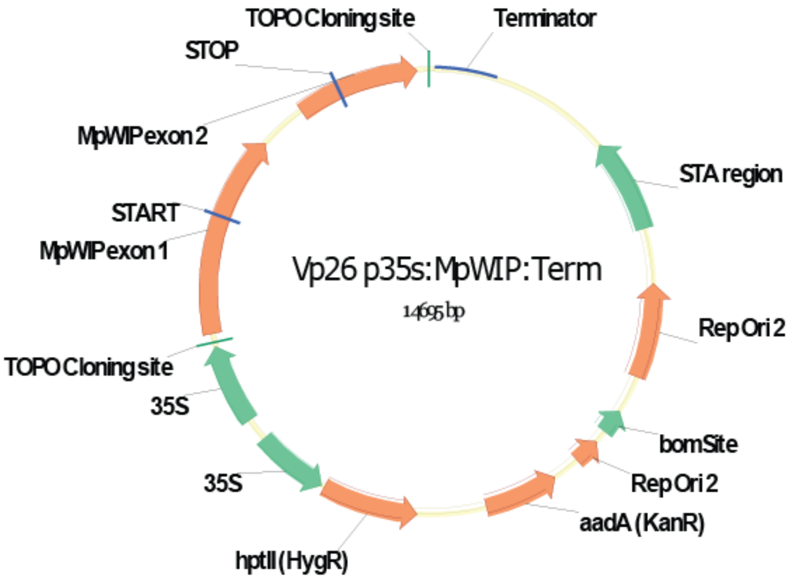


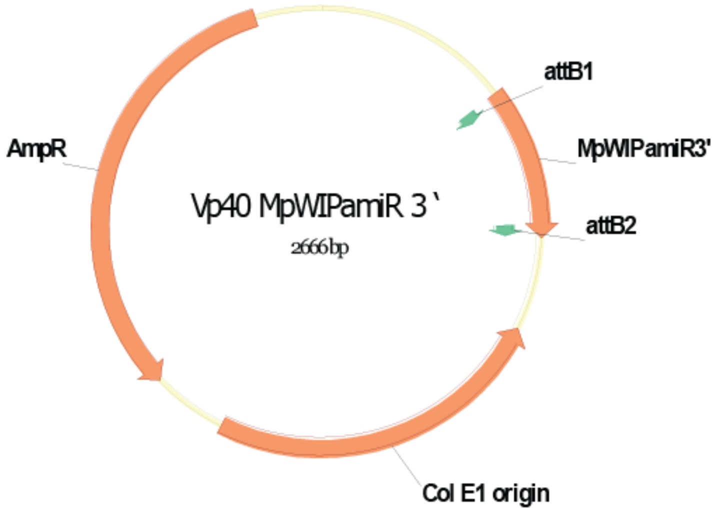

pUC18 MCS; polylinker
rb1 primer
lacZ promoter
rb2 primer
CaMV35S
rb3 primer
right border T-DNA repeat
hptII (hygR)
lb1 primer
STA region from pVS1 plasmid
lb2 primer
pCambia1300
8958 bp
CaMV 3'UTR (polyA signal)
lb3 primer
left border T-DNA repeat
kanR
pVS1-REP; replication origin from pVS1
pBR322 origin of replication
bom site from pBR322

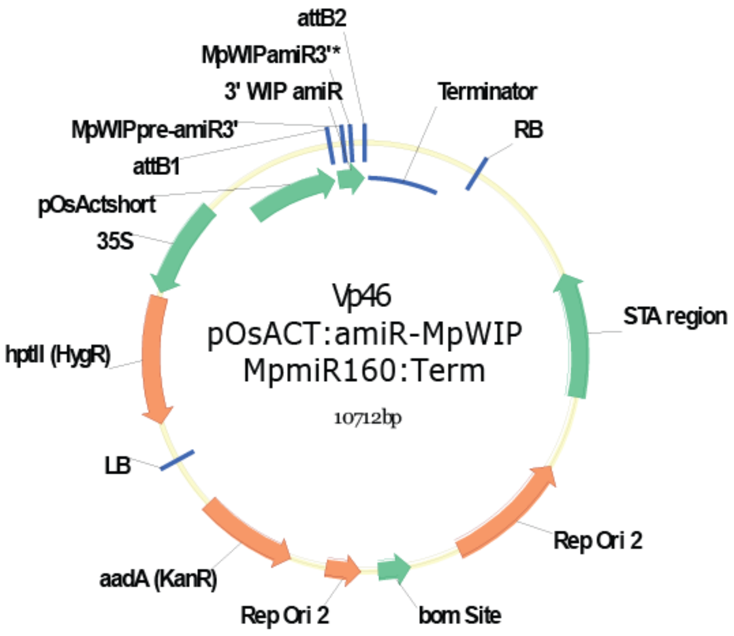

rrnB T2 transcription terminator
rrnB T1 transcription terminator
M13 (-40) forward primer
M13 (-20) forward primer
attP1
pDONR221
4762 bp
Kan(R)
ccdB
M13 reverse primer
T7 primer
T7 promoter
attP2
Cm(R)

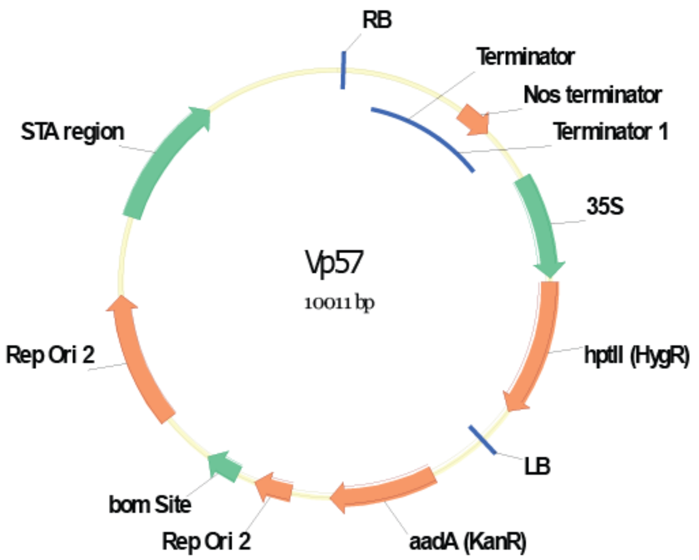

lac promoter
M13 pUC rev primer
M13 reverse primer
Sp6 primer
lacZ a
AmpR promoter
M13 pUC fwd primer
ORF frame 3
f1 origin
M13 forward20 primer
pGEM-T Easy
3016 bp

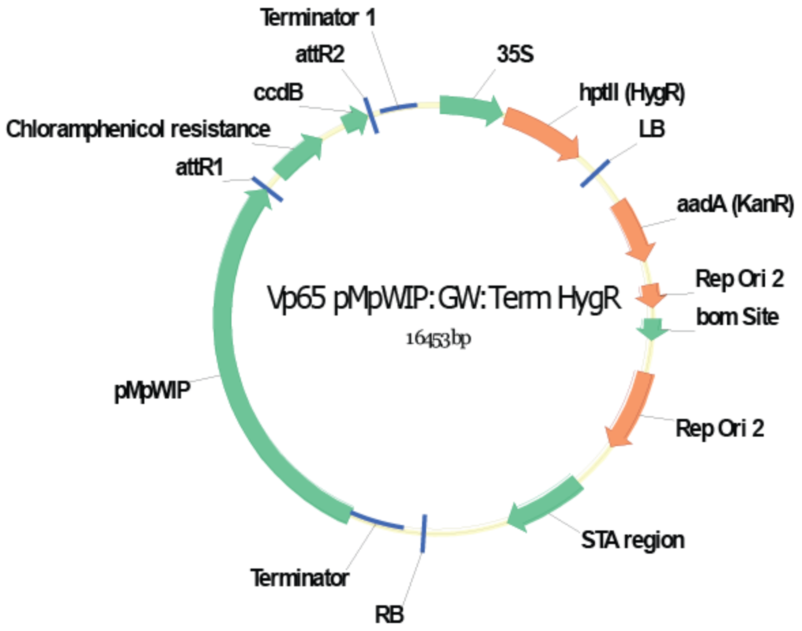

rrnB T2 transcription terminator
rrnB T1 transcription terminator
M13 (-20) forward primer
GW1 primer
attL1
TOPO Cloning site
pUC origin
SpnR
Spn promoter
M13 reverse primer
attL2
GW2 primer
TOPO Cloning site
Vp24 MpWIP
7002bp
MpWIP exon 1
START
MpWIP exon 2
STOP

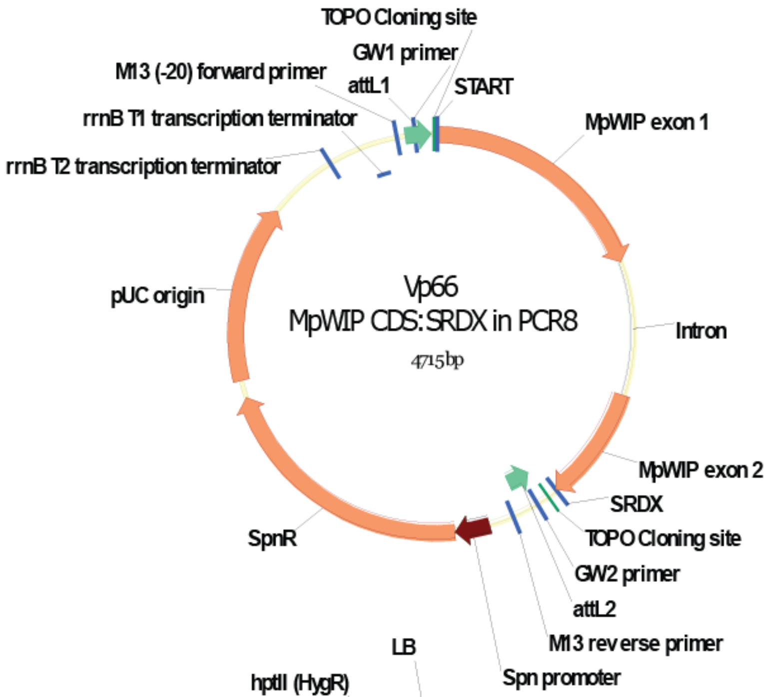

TOPO Cloning site
Terminator
STOP
MpWIPexon 2
STA region
START
MpWIPexon 1
Vp26 p35s:MpWIP:Term
14695bp
RepOri 2
TOPO Cloning site
35S
35S
bomSite
RepOri 2
aadA (KanR)
hptII (HygR)

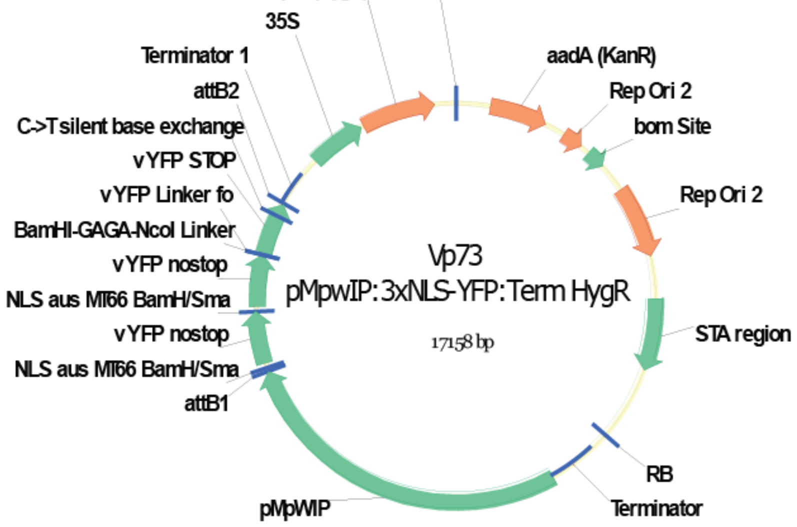

attB1
AmpR
MpWIPamiR3'
Vp40 MpWIPamiR 3`
2666bp
attB2
Col E1 origin

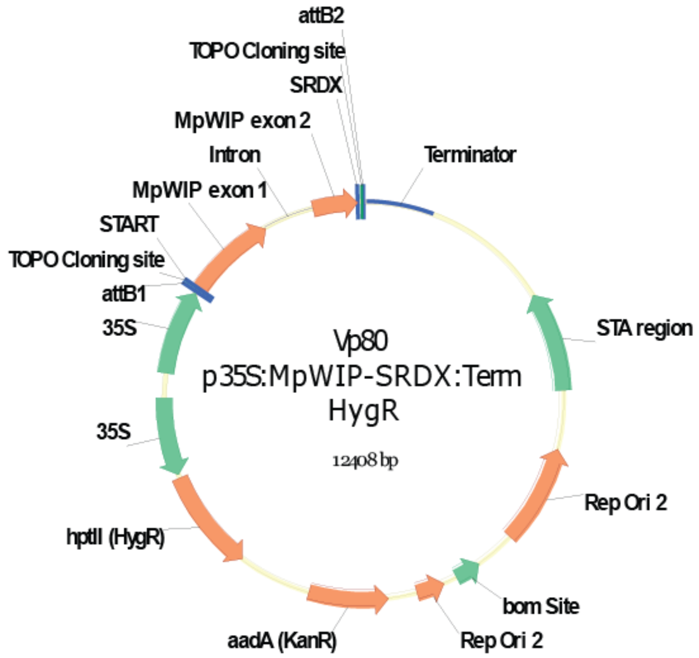

attB2
MpWIPamiR3'*
3' WIP amiR
Terminator
MpWIPpre-amiR3'
RB
attB1
pOsActshort
35S
Vp46
pOsACT:amiR-MpWIP
MpmiR160:Term
10712bp
STA region
hptII (HygR)
LB
Rep Ori 2
aadA (KanR)
Rep Ori 2
bom Site


RB
Terminator
Nos terminator
STA region
Terminator 1
35S
Vp57
10011 bp
Rep Ori 2
hptII (HygR)
LB
bom Site
Rep Ori 2
aadA (KanR)


Terminator 1
attR2
35S
ccdB
hptII (HygR)
Chloramphenicol resistance
LB
attR1
aadA (KanR)
Rep Ori 2
Vp65 pMpWIP:GW:Term HygR
bom Site
16453bp
pMpWIP
Rep Ori 2
STA region
Terminator
RB


TOPO Cloning site
GW1 primer
M13 (-20) forward primer
attL1
START
rrnB T1 transcription terminator
MpWIP exon 1
rrnB T2 transcription terminator
Vp66
MpWIP CDS:SRDX in PCR8
4715 bp
pUC origin
Intron
MpWIP exon 2
SRDX
SpnR
TOPO Cloning site
GW2 primer
attL2
M13 reverse primer
Spn promoter


LB
hptII (HygR)
35S
Terminator 1
aadA (KanR)
attB2
Rep Ori 2
C->T silent base exchange
bom Site
vYFP STOP
vYFP Linker fo
Rep Ori 2
BamHI-GAGA-NcoI Linker
Vp73
vYFP nostop
pMpwIP:3xNLS-YFP:Term HygR
NLS aus MT66 BamH/Sma
STA region
vYFP nostop
17158 bp
NLS aus MT66 BamH/Sma
attB1
RB
pMpWIP
Terminator


attB2
TOPO Cloning site
SRDX
MpWIP exon 2
Intron
Terminator
MpWIP exon 1
START
TOPO Cloning site
attB1
35S
STA region
Vp80
p35S:MpWIP-SRDX:Term HygR
35S
12408 bp
Rep Ori 2
hptII (HygR)
bom Site
aadA (KanR)
Rep Ori 2

Appendix 3 TAIL PCR products

```
>EMB2756_vj24_LB
TTTATTAGTGTCTAGCGTCATTTGTTTACACCACAATATATCCTGAATCTATTTCGCTTCATGAAACCATGAGATCAACATTCTGAATAACTTC
ATCAATTCATGATTGTGTGTTTAACTACTTAAAAAGTACAAATCATTCGAATCGAATGATTCATATTCGGGTTGTGTCAGGACTGTGAGCGGAA
GGCGATGGCAAAGTTGTACCGACACAAGACAGAAGATGCATTGGTGAAAGGTCGCAGGCACTGAGAGGAAGAATGTGGTCTCTTCAGATGATTT
GCCAGCAACATTCGAACATGTCTTGCATGTTCCAGGTTGAGGTTTATTTTTTTCCAAATACTGCAGGAGCAGGTTTTGTAAAGTTAGGCTTCTT
CCTTTGTATCTTCCAGTTTTGTATCTTGACCATTCTTTTTCCGAGTACAGTAGAGAGGCAAATACCGGCGGCTTTGATGTTCATAAAGCTCATA
GGAGATTTGAAGAGAAGATTTGTTCCATTAGATTGTACACAACACGTTTCTGATTTACTTGTAGAAGTATTACATTTTAGTCTCCGACTAGCCG
TTATCAGTCTATTCATTGGGCTCCAGAGTCCACAACTGAATCTGCAACAATTTTATATTCAATTTAGTCCTTCTAATGGATGCCACCTGCAAGG
CAAAGTCAGTTGGATCATTCAATCTTATCATCTCTTGTTTTGTAATAACTGACTTCAATAGTTTTCTTTTTTGGCAAGCGTCATGATAATTTGT
GCACAGGCGAGGTGTCTCCTGGAATAGCTTTAGCTCATGGATATGCACCTTTTATGCACTGTCACTCTTTCTTTGGAGTTGCGTGAGTGCTGGA
AGAGGACGCCGTTGAATGTTTAGTTCTATTGTATAGTATATACTTACAGTGGATGGAAATTTTGCTACAAAACCTTGGGTTCGCAATAGCTCAA
ATCCTGAGATTTGATTCGAACTGTTCTTGAAACGGTGATTGAAGAGTGAGTGCTTTACAAGTTCAATGTGTGCACCTACCTTTCTCAAAAAAAG
CTGCCCAAAGGTTTTGCAACAAACCTCGT
> EMB2756_vj24_RB
GCCCGGGCTGAAAATTGTCGTTTCCGCCTTCAGTTTAACTATCAGTGTTTGATTAGTAAAAATGAATAAGTTCACTAACCTTAAACATGTGAAG
CCGAAACTGAGAGTTTGGGTTCGTGCGCACAAGCTTCATGTACGTGTAAGCGAAACTCAACTGATCTCTTGGCGTAAACCGGTCAACTTCGTTA
AACCAAAGACAAGAAAATAAATTGGACATAGGAGTATGTGCCCTAACAATGAAGGATCCTTCCGGCACATCTGCATTTCAAGAAGCTGGATGGT
GAGATGGATAGAAAAACAAGGCAGGCACTTCAAAAATAGCTCCCATATTTTTTTAACAACCGTCCTGACAATTTTCTATGTTTGCAGTCCAAA
AAAACTCAAACCAGACCAATTGCGCATGACTTTAAATTCGTCACATTTAGGAAATTGTTGTAGCACTCTTAAACTCTGGTGAGGTTTGAGTGGG
ATATGGAAATGGTGGTGGAATAAAACAGACGTAGAAATCTATGCACCAAGTTTGCTACGAGATTTAGTCCCGAGCCATCAGACACGAAAGACGT
CAGACTTACGGCTTGGTAGGAGCTTCTCAGGGTCGGATGCATTGAAGCGAATCAGTCCATCAGCCTGATAGAACGCAAATTGCTGATCGATGAC
AGAGTGATTAAACTTGTTTAATTTTTTGTTTTGTTCTACTTCTTCCCACACACAGTGCCGATCATAGTGATTTGAAATAGCGTATTCATGATGT
CCACGCCACAAAAAGTACTCCAAAATCAGGAGAGGATCACTTTGCAGGCGTAGCTTGCTGTCCAGCCACACAGAATACCTGCACCCAAGGGAGA
CGGTAGACCACACTGGAAAGATGTTCATCGAAACTAAAAAGACGACCCGTATTGCACATTTGACGGCGGGAACAATCGGACACACTACCGCTAA
AAGAAAAGTACCAAGGGCCACACCACTTTTGAGGGTCATGATATTAAGGTTTAGAACTGCAAACAGGGAATACCTTTCCCATTGGAAAAACCAA
ACTGTAAAAATACGGTACCCTGGCATTTAGGAA
> EMB2756_vj68_LB
GTGCACGCGTCATTTGTGTTCTCGTCAGGAGATTCACCCTCCTCGATCATGCCATCAATACTCGGTTGGTCCACAAACATTGCAAAGCATACTT
TCATC
> EXL_vj19_RB
NNNNNNNNNTGNNTCNCTTCTGTCNCCNCAAAATGTACATCGTCGCGTCATTTATTATTATGTTAATGTCAAACCACTTCATTTGAACGACTGC
AGTAACCCCATGAAGCAGTTGCAAACATCCATAATGATAAGCTGGGGTGAACATTGGAGAAGTCTGTTTCTGCAGACATTCAGATGTGACGAGG
AGACTGAAAGATTTTTCCCGTAAGCAATCACCTATCTAAATAATCTCCCTAACGGCAACCACCATCAACTGCTTCTATGACACCAACCTCCGTA
CATTGGCAAGTATCTCCGTCTTTCCTTGTGACTAAAGTATTCGGCTTCTAGCATAACATTTGCTCACAATTCAATGCCCCTATCTTTACTCCAC
TCNAAAGCAATGTATAGATGTNTGAAAAATCCTGTAGGAAATACCANGCCACTCGGGGTATGTATAACTCTTGNNTAGGAGANGAAAGAANATT
ATAANCNACTTNTGGAAGNGGTTGGNGAAANTGCNTCCATTCTTTTTTTTTTCTANCACTCCCTTTTTNATTTTTTTTTTTTTTTTTTTTTTTAA
AA
> PI4Kα_vj39_LB
NNNNNANNTNNCTAGCGTCATTTGTTTACNCCACAATATATTTTAAAGACTTCATGCTCGCCGTGGTGAAATTCACTAAGTCGAGGGCGGTGCT
CGCTGCAGAACTTAGCACTATTATAGCTGACATTATCCAGAAGATCACCGCGCTGTCGAATCGGGAGGTCAGCGCACCAATTGCTCCAGGAGTA
AAGTTTTTCCTCGTTGCTCTCTCTGAGAATTGTCCTCATCTGTCCGCCGATGATGTCGAGAGAGTGATCAAGGCCCTGCTGGAGCAATGGGCTG
TGAACTTTCAACCGGGAAACACTCAGCCTTCATCCTCTGACACTTCATCTCACTCGACN
> PI4Kα_vj39_RB
NNNNNNNNNNNNNNNANATTGTCGTTNCNNCCTTCAGTTTAAACTATCAGTGTTTGTTCCTGAGCATAAGCAGTCGCCCAGAGCGAGCTAAATGC
CGATGACCAGATATCATCTTGACCATAATAAGGAATTGTGCGAAGAAACTCCAGAATGACATGCCAAAGATCCATTTGAATTTCCTCAGCAGAC
GAACTCGATTGCGAGAGAAACTGAGCCACAGCAAGAATGGCACTGAGTTGCCGGGGGTGAAGGCGAGTGGATCCATCATGCAGCGAACCCGACG
GGGGACACTGACACAGGATTACGCCGAGCTTTTCAGCCGCGAGCTCTGCATTGCCGGCCAACAGCGAAGACAGCTCTCGCAGCGACGCGCCACC
GCCTCCAGCCATCGAGGTAAGACGCAAGACGGCAGCAGCAGCAGCAGCACTCCTCCCGCCGGCGCCTGGCCGTGCTGCCCCGTCCGGCCGACTG
CAACGTCTATGCGGATTGGCGCGGCAGCACCTCGTGCATCCTCCTCCTCTTGCTCCTTCATCTCACTCGACGA
> PI4Kα_vj32 _LB
NNNNNNNNNNNNANCGTCATTTGTTTACNCCNCAATATATCCTGCATATTCCGAAATTGTTCCTGGAATCACAGAAAGAGATCCTTATTTATCTG
TAGTGGAAAATGAAGTAAAACAAACAGTATTAATGAATAAACAATCCAGCTGCTGTACTCAGCTGAATATTTCAGAACTCTATGTACAAAGGGT
AAGGGAAAGTAGTATATGAATCAAGAGTGAACGCAATAATTCAGATGTACCTGTTTCAGCAGAACTAATTACCATTACCTACAATCCAAGCTTA
TGGTATGAAATACATCTTTGAAAGGCAGGAGTTTTTTCGTCTGAGTAGTACCTGAAGAAGTCCTTGTACAGTACAAGGAGCCAGTACAATAGCA
TAAGCTATCCACTCTCTAACTCTTTGATTCCAAAGTCCTCGCGAAGAGGCGGGTGCCCCAGCATCAACACCNAAACTCNNN
>  PI4Kα_vj32 _RB
NNNNNNNNNNNNNANNGTCGTTTCCGCCTTCNGTTTNACTATCAGTGTTTGAAAGCAGCACAGACGAGTGATCTTCTGTCTCTGCTCCTNCNCAA
A
> PI4Kα_vj13_RB
NNNNNNNNNNNNNATTGTCGTTTCCGCCTTCNGTTTAACTATCAGTGTAAGGTGTACTACAAGCTGAGGCAAAGGCTATGGGAATATTTGTGCAG
TACCTGATGTCGAATAGGCCACAAGGCAGTGCCAACATCGAAGGGTCTTTTAGACACACTCAGGAAAATCCCTTTTCAGACAGTGTACAGGTGT
GAAGATCGGCGTAACACATTCTATGGAAATATTCTTAGTAGGGAAGAAAGTTGTTTTCGACTGTTCTAGACCTTGATATTGNAATTGACCACTT
ATGCAAAGAGAACCCGACGGACGGATGNNGAGTTGGCCTACGGCCTGNTGGGGTACTGGGGANATCNGCATGCCNANTNATNCTTCCCTNNAAC
AGGANGAANTGNGTTGGTAGTTNTNCTCNNCACGACCCTCCATCCNCCATGCATCTCTCCGCTGCTTTCTNCTCCGCTCCTTCCGTTCTTTANC
TTCNAGTGAAGGTGNNCTTCCTNNNTNNTTNTTTCNCNCATCCCCANNNCTGCTTTTTNTGNNTCCTGTTTTTTTGTCNNANANNTAANAACCC
CATATCTTTTATTNNNTCATCGCGGCAANNCNCTCTCTTATCNNAGAGCCCCNCTTTTCTCTTNTNNATGTATATGTGTTTACTCTCCGTTCAT
ATATACTCTCNATATTNCCTNTNNCNCACNGCTCTCCTGTNGAAACAACATCTCCTTCTCTTTCNTCTCTACTTTATTNATCATGTGTCTCCCG
AAAGGTGTCTTTNNNNN
> SPI_vj69_RB
NNNNNNNNNNNNNNNNNATTGTCGTTTNCNNCTTCAGTTTAAACTATCAGTGTTTGAGACAGTGGCAAGGGCGTCTTATGATGTTCAATATATGA
GTGATTTCTGAAGTGGAGAAATGCGAACAGTATTGGAAGTAAAATCGAATAAACATGAAACCTGCTTGGACAGACAACTCTTTACGAAAACCTG
CAATTTTTTTCTAGGAATCATGGTTAGTCTAACTTTAGGCTCCACAATTACTTGTAAGAATGAGAAATGCTGACTGATAACTAAAGATAGAGCT
TCAGGTGGAAATCTTACAAGCTGTTGCGTGGTAGATTCATAGCAAGAAGATTCTTGAACACCTCGTCTCGCTCAGCTTTGTGGAAAACCAGGAG
CTCGTTGCATCCATCCATGCTGAAAAGCTCAATGGCAACGGGCCGAAGCTGGTATCGACGCTTTAAGAGCTCGTGTACACTCTCGAGCTTCCAC
ATCCTCCAAGTATGGGACATCTGCTTCCCAGCTTTAGCTGTCTCCTTACCCCAAGCCCCACCACTGTAAGCCCAAGCTCTGCCTCCAGGCCATG
CCTCACCAGACCGCTCATTCATGCTGTTTGCCCTCTCAGGCTCATGGAGACTGGCGGCTCCGGCAGAATTAATCTTCACTCCCAATGCCCTGTC
```

```
AATTACCGAGAGTTCACCCTCCCAGGTTTTCTCGTTGATACGGCCATCAGCATCTATGAAGTAGTGTTCTATGACGTACAGGCACAGTTCGCCA
ATCAGAAATATCCCGTCATGCTTGTCCAGGCCCACAACGCGTTCACAATTATAGCGGAACCTAATCTTCTCGCCGGGTACGAGGTGAGGTCGAA
TCAAGTACTCTCCATCGTCAGTGATGACTTTGTCAAGCTTTTTCTCCTCATTCCTTTGGAGATCCTCAGCAGCATCCGGTGAGGAAGGNCTTGA
CGAATCAACTCCTTGGGTCCCAGTTCGGCAGCTTCCATCAATAGAGACTGCACTCTTGTTNCCATTGACGGAGTTTNCAGANGTANTCNNNNTNT
CNNNNCGNNN
> SRI1_vj25_RB
NNNNNNNNNNNNNNNGTCGTTTCCGCCTTNNGTNNNNNNATCNGTGTTTGAAAGCCGCAGTAGAGTAGCGAGAACCACGACCCACTCTTTAAAT
CCCACAGCTTTCATATGACCCAAAAGAAATTCAGCATCTTGTCGACTTTCCGCCACATCCATATCCAGCTATGATCACCCATATACAGATTAAT
CAGCTACAACTACCACAATACAACGGAATGCGCAAAAAAGGACTTCTTGACTACATAAAACATCGATTCCAGTTTTCACCCATGCAACCAGCCT
GAGGTGTTCCTACAAAACTTACCTTTTGGCCAATGATCCGAAGCGCCGAAGAAAAGTCCTCCAGTCGGGCTGACTTACCCTTCTCTGATTTCAG
GAAGTCCTACAAAGATATGACGCAAGAATGAACAAAAGAGTACATACGTGAAGAAACTCTTGAATCTGAGAACGAGGACTAAGTCTTCAAAACA
CGTTCATGAGTAGATGAGTGCTAACTATCTAAGCTTACTGTGATATCAAATTGGGTCCCCCTGGCGAAGGTCACAAGTTCTTGAAGTTCCTTCA
ATTCCATAAGATGAGCCGCGTGCTTCTCAAGAATATGCTTTACGGTCGCATGGTGCATATCGTCACCCATAGCAGGACTGGTACTGCTGACAAA
GAAGAGACAGCCCATAAGCAATAATATCTACGGATGGGATCGCAGGTAGTTTTAAGTCCGAAATATTGTTCAATCAGGTATGCAAATTCAACAC
TAGTTTGCCGAACAACAACACAGCATGACAAGCATCCAGCTTNGTGAATTCAACACAACATAAAATTCCTCCCCTGGCAACATTAAAGAAAATA
CATAAATAGGCTCACAGTGAGTGAACCTACGGCGAACTNNNCCCGANGAANAACGCTGAAAAGNNNNTCCGTTCANCTTCCTTGNCTGCTTNNT
NNNNNNNNCNNNNNN
> SRI1_vj14_LB
CNNNNNNNNNNNNNATCTTGTATCTGTACGCGGGACAAGATCTTCCAGTGTTATCAGCAAATCTTAATAAAGTCATCCTTTCTCGAACTCTGAGT
ACGTCCCAGCACAAGTTGTCAACGTCTGTTGTTTCTGTGCTCCACCCTGCTATCAACAGACGTGCCAGCGTCATGGACTTGCTTATCATCAGGT
GGGCTTCAGCCGTTCACCTCCTTTGAATCACAATTTAAGACACGATCAGTATTCAGTGACTCATGGCTGGGAGTTAAAGTTTCGTTCTTCTGCC
CAGTTGACTCTCTACTCCTACTAGGCCCGGTGTTTGCTTCTGTGACTTCTTCCGCATTCGTTGTCAGTGCTTAGATTAAGTTAAAATTATTACA
TCTTTTAAAGAACCTGGTGCGAAGAAGTGAAGTATAAGTTCATAATTTGTAAGTATTCAAATCGAACAAAGTCTGTACAAATCAAAGTGCATGT
GTCATGTATGTTTATCCGCGAGCCAAATCTGGACGAATCTTGAGATATTACCTGCAGATTTCTCTACGAGTCGATCAAAGTTCAGATCTTTCTT
CATCAAATCGTAACTAGGACTTTCCAAAGCAGTCCCGGTATAAACCCTGACGTTGCGAAATCGTTCCCAAAACATACAATGAATTATATTGTCC
GTAGCTTAGATATACCCAAAGTTTATGTCAATATAATGAGTGGCTGGAATGGAGAGACGCTTGAATAATATTTTATTTTACTTCATCATATGCA
AGCAAAAAAATTACTGCGAATATCTGGAAGTAGCCAATAAAAAGTAAAGAGAATTCTTCCTTGAATGCGAGCAGTTTCATCTNCACTCGACNAN
> SRI1_vj14_RB
NNNNNNNNNNNNNNNATTGTCGTTTCCNCCTTCNGTTTAAACTATCAGTGTTTGTGGATGGCCGGTCACTTTTTAAGAGTAAAAGAATCTAGACA
TAGAACTGGCCATGGAGTACTATGCGTGGGCAGATACAAAATTGAAGACACTAAGACTCTTACCTGCTGTACACGATGCAGGGCAGGGGATGCA
AACCTCAGCTCACTCATACCCTTAACGTGACACCTCCTCATACTGCAAGACTGTTTCATCTCACTCGACGA
> SRI3_vj67_LB
NNNNNNNNNNCNACNNTGTTACGTGCGCATCGTGATAAAACTAAAACATGAACAACAAAAAAAGGGGGAATTCGAGAAAAGCCCAGATCGTCTC
GTTAGAGTTCATCGCAACGAAATCTACACCAGGCTACTACTCCTATTGCTAAATATGATTCTAAGTCTTGGAAGAATGATCTGGATTGACTAGA
CAAGCAAGCAACAGATAAATTTCTGGAAAATTACCTTCCTCTTTGCCACCAACTTTCCTCATAGGCATGTCCATGTTATCCAGATTCATCTCAC
TCGACCN
>SRI3_vj67_RB
GCCCGTCCTAAAATTGTCGTTTCCGCCTTCAGTTTAAACTATCAGTGTTTGATATTCTTCTCAGCATGTCTAAAACTCCTCCCACCAGAAGATC
TTGATAAGATCGAATTACCTGAGGCAAGCTTCAGTGGAGTCAGGATCCTAAGTTATGTCCATAATAATATTGATTCTAAAGAAAGTAAGGTACA
AGACAGCTCCATGGCCAGTGCATTTGGAGGATATCAAACGATGGATGAGAGAGCTAAGACCTTTGAGGTATTTACTAGTGCTTGTCACCCTAGT
CCCTTGTGTGTGCAGCTCATTCATTTTGCTTCCAGGTTTGACTGTATGAGCCGCCAGACAAGAGAAGTATTAAGGCTTAGAAAGTGACTTCCAA
AGCAATCGTGGCTGAGGCCGAGATTCGCATGCTTGTTTGATCTTTGTGCTGAATATATATATGACGATTTATGATCCACATTTGCCTGTTTTCG
TTGTGAAGTTGAGAGATACCATGGTGGTGCCTGCGGATTTTGTGAACATGGCATCTGTTCACATTTACTCTCCGACAAGTCTTCAAGGAGAGCT
GTTCATTTAATTTAATAGCACATTCTAAT
> XI_vj21_LB
TCTAGCGTCATTTGTTTACACCACAATATATATAAGGTCACAGCTGATGTGCATGACGTGGTTGGGAACTGAGCGCGGTCACGTAGAAGATGTC
CATATTGTGTTTGGGGATGACGATGATATGTCAAATCAATGAAAATTGGATTGATTAGTGCATTGGGTGATGAGATTTCTCCGGGGTCGAAGTT
TGTTTCATTTACATATGTAGCTAGATTGCTGCGGAGAGGTGAGGAATGCCATTATTTAACGATGTAAGATTATCGCCACGGACGGTGAAGACGA
GTTCAACGTTGAGCCGAGAGGCGGAGAGGAGGAGAGGAGAGGATAGTGGAGGCCTAGTGCATAGGACAGTAGCTGATTGACATTCACAGGGAGG
AACATCTTCCCCATTCCAGGGACCGGCCGCTCGTCCGCTGATCAAGTAGTAGGGAGGAAGACACGGGAGTCAAGGTAACGTGTGCTTGAAGATG
CTTTGATATTTTTGGGCCCCAGACCGGACATGAACTTGGGTAGCGGCGACACATGTTGCTCGGATCCGACAGGCAGGTACTCCAGGAGTGGAGT
GAGGGAGTGCTCGTGCTCGTGCTCTTGTAAGTGGCTGGATGGCTGGATGGCTATTAGCTGACCAACAGAACGTGCGTGTGAAGTTAAGTTGGGG
GTCACGGTGCCTGGGGTAATGTAAGTACGGAAATGAAGGTAAAGCCGTGAATAAACGTAGGCCATGACGAGAATTCTATCCGGGCTTTCTCATA
GAATGTGGTATTAACGAACCATGTAAAAGGAACAAGCACATTAAGTTTCCCCCGATCGCTAAGCACATGTGCTGCTCTGGCACCCTTTGATCCC
TAGGAGAATGAGATTCTTTATATAATTCCATTCATCTTCAGCAACTCATATTCTTTTCACCCTTTGACAGTTGGACCATCTTCACTCAA
> XI_vj11_RB
NNNTNNNNNNNANTGTCGTTTCCGCCTTCNGTTTAACTATCATGTTTGGTCTGGGGCCCAAAAATATCAAAGCATCTTCAAGCACACGTTACCT
TGACTCCCGTGTCTTCCTCCCTACTACTTGATCAGCGGACGAGCGGCCGGTCCCTGGAATGGGGAAGATGTTCCTCCCTGTGAATGTCAATCAG
CTACTGTCCTATGCACTAGGCCTCCACTATCCTCTCTGCTNNNNCNCNN
> RSL1_vj3_RB
NNNNNNNNNNGGCGATTGGGCCGACGTCGCATGCTCCCGGCCGCCATGGCGGCCGCGGGAATTCGATTCCTGAATGGCGAATGCTAGAGCAGCT
TGAGCTTGGATCAGATTGTCGTTTCCCGCCTTCAGTTTAAACTATCAGTGTTTGACGACCAATTAGAAGGGGCCCACGTCGCCCACCTCCTCCA
ATGGCTAGATAGAGGCGTATTCTCCAGGAAGTCGAACGAATTCCCCTTGACTGAGATGTTTTACTTTGTCTTCCCAAAGTAGAGGAGCATACTG
TGTCCTGTTCGCAGAGGAGCACAGAAAGACCTTCCCGAAGGTAGTTTTCCGTATCGTTACTACTATCTGTAATCAAAGCTTATCCTTCCACTTT
CTACTTACTACCAATCCCGAGCTTTGACACGTTGCTATCACTGAACGCATGATGGATGTCCATTATGGCTGTTTAACTGTGGGTAGGCTTCGTT
TAAGTAGACACCTAGAATGTGTGTGGACGGACGGATGGAAGAAAGCTTTCGAACCGAGAGCCTGCTCTAATCGAATTTCTAATCTTCACCAACT
AGAATATACTATATCACTGATTGAGGTTCCTTCTTCACGGAAGAAAATTAGTACACTGTATGTTAGCTCACCAGTAAGTATCTTACGCGCTACT
TTTCAATTTACTAATTAAACACTAAGCATGCCGCTAATAAGTTCTCATGGTGCAAACCGCGTGCTCCTCGCGTGCTACGGTTGCGGCGACTATA
GTGACACGTCCAACGAGGCACTTACTCCGGTTAGAGTTGCAGAGGCGGTGGATCTTGTGGAATGGCGAAAGTAAGGCAAGTTTTAGCTGGACGT
GGGATGCTCTCTCCATGTTGCCAGACTGTCGGTCTCGACGCGTAGCCAACAGTAGCGACAGTCATGGTCAATCAAAAGCTTGCAAGTTCCTGTA
CCTATTTCGTTCAATTAACATCTGACTCGCACAATCACTAGTGAATTCGCGGCCGCCTGCAGGTCGACCATATGGGAGAGCTCCCAACGCGTTG
GATGCATAGCTTGAGTATTCTATAGTGTCACCTAAATAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCANN
ATTCCNNANANATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGNGCCTAATGAGTGAGCTAACTCNCATTAATTGNGTTGCGNNNNCTNNC
NNCTTTNCAGTCGGGNAANCTGTCGNGCNNCTGCNTNNGANCGNNNANNNNNNGGGGNNNNNGGCNNNGNGNNNTGGGGNGNNNTTCCNNTTCC
TNNNTCANNGANNNNNCNNNNNNNTNNGGNNNNTNNGGNTNNNNGNNNNNNNNN
> RSL1_vj4_RB
CCGGGAATTGGAAAATGGCGTTTCCGCCTTCCGTTTAACTATCAGTTGTTCATTTTGCCTTCTCTCAGGCATAGCCTTTAAGGGCCCCTCGCAG
CCTCTTCTCCCTTTCTATGTGGATTTAGGGGGTAGCGGTAGAATGAGGAGGGGGTCAACATAGAATTGAGCTTGACTTCATACTCTACTTCAC
ATTCATCCCCTCGCTCAAGTAGATTGCCATTGGCAGTTCATGTTCTAAGGCATGGAAGTGCCTAGATAACGGTGACCGAGGAAGACCAGAGAGC
```

```
GAACACTGATATGGCTTGAGATGAGAAGAGATCAACTTGTTTATCGAACGAACTGACCGTCACGTTTGCGCTGCGATGGTACCGCTGGTGTGTT
TCCACCTAATCACATCGATTCTCGGAGTCAAGCTTCCAATAATCCATGTTTTCTGCCGACTGTATCCTATCAAGGTCTTGCATGCAGTATGAAT
CCAATCAAAAATGGTTTCATGTACTATATAGACGATGGACAGGAATGAAAACTCACACACCTGCCTTCCTTCATCCTTCGGCTTGATACGATAC
TGGCCTTGGATTGAGAAGTTTAGCATGTGCATATCTGGGCACCTCCGAAGTCCGTACTTGTCCGTCCAAATTGTTCAGGAAGGTCCTGGCAAGC
AGTTTGCGTCTAGATGACAGTAAGAACTAGCCCATTCCCTCCTGCATCTAGAACTCTCAGTGTTTTTCCCGGACATGTGTTCTTCCCGCAATAT
GTATTTGGGAGCAATTGGTATCAACTAATAAGTTGAGCATTGTCACAACAGTACGACTGCGATTGAATTGTATTTATGTTTACTTTACTATATA
CTATCTTTATTACTGAAGTCATTATTCTGCCTTATTCTTATATCATCTTTTTTTATCACTTGATCATTTTGAATATTGTTCATATTCCGATTAT
TAATATTATACATT
> RSL1_vj5_LB
NNNNNNNNNNNCNNGCTGGTAGACGATGTCTGATTAATCTCATGTGCTATTGGAAAGGAATCGGTACCATACGTGTAGTCACTGTAGAGAGTAA
TTCTCTGTCTATCATCACTCGGCTTGTCCCTCACGAGCTGGGAGTCCACATCAAGCTGAACCAAACCAGAAAGGACTAGATGTACTAGAATCGT
AAAAACCTTTTAGCAGTACCTAGTACCTACATCAATCGTCAGAAGTGCTGATCTGGGGAATCTTTTTTGCTCACATACTCCCGGATGNCAAGCC
TTAAGATTCGATATCTGTTACTTTCTTTNATGNCCT
> RSL1_vj6_LB
NNNNNNNNGTCNAGCGTCATTTNTTTACNCCACAATAATGGACATCCATCATGCGTTCAGTGATAGCAACGTGTCAAAGCTCGGGATTGGTAGT
AAGTAGAAAGTGGAAGGATAAGCTTTGATTACAGATAGTGGTNGCGATACGGAAAACAACGGTCNGNAAGGTCCTNNNGTNTNTNNNTGGCTAA
ACGGANNCNTATTNNTCA
> RSL1_vj6_RB
NNNNNNNNNNNNNNNNNNNNNTTGTCGTNNNNNNCTTCAGTTTAAACTATCAGTGTTTGACTGATAGTTTAACTGTGGGTAGGCTTCGTTTAAG
TAGACACCTAGAATGTGTGTGGACGGACGGATGGAAGAAAGCTTTCGAACCGAGAGCCTGCTCTAATCGAATTTCTAATCTTCACCAACTAGAA
TATACTATATCACTGATTGAGGTTCCTTCTTCACGGAAGAAAATTAGTACACTGTATGTTAGCTCACCAGTAAGTATCTTACNCGCTACTTTTN
AATTTACTAATTAAACACTAAGCTTGCCGCTGATAAGT
> RSL1_vj8_LB
NNNNNNNTGTCTAGCGTCATTTNTTTACNCCACAATATATCCGGGTATACAGGAAGAGTTCCCCATGCAACTTTACTGCCTGGGTCCCGCTCAC
GCGAAATTGTGTATCAGGGTCCCCGCATCTATCGTCACCAGATTTGGAGGANTNNACTTTCTCCCGACGTTCTGTTCNNCCTACATTCACATAT
ATTTNGCCNACNTTNTC
> RSL1_vj8_RB
GCCCTACTGGAAGATTGTCGTTTCCCGCCTTCAGTTTAAACTATCAGTGTTTGAATGCATGCATCAATCCCACATCATGCCTGGCCACCCCATG
CAAGGCGCATCATGGTCTAGTCGTGCAGAATCTGCCCTGGCTGCCCGACCAATTAGAAGGGGCCCACGTCGCCCACCTCCTCCAATGGCTAGAT
AGAGGCGTATTCTCCAGGAAGTCGAACGAATTCCCCTTGACTGAGATGTTTTACTTTGTCTTCCCAAAGTAGAGGAGCATACTGTGTCCTGTTC
GCAGAGGAGCACAGAAAGACCTTCCCGAAGGTAGTTTTCCGTATCGTTACTACTATCTGTAATCAAAGCTTATCCTTCCACTTTCTACTTACTA
CCAATCCCGAGCTTTGACACGTTGCTATCACTGAACGCATGATGGATGTCCATTATGGCTGTTTAACTGTGGGTAGGCTTCGTTTAAGTAGACA
CCTAGAATGTGTGTGGACGGACGGATGGAAGAAAGCTTTCGAACCGAGAGCCTGCTCTAATCGAATTTCTAATCTTCACCAACTAGAATATACT
ATATCACTGATTGAGGTTCCTTCTTCACGGAAGAAAATTAGTACACTGTATGTTAGCTCACCAGTAAGTATCTTACGCGCTACTTTTCAATTTA
CTAATTAAACACTAAGCATGCCGCTAATAAGTTCTCATGGTGCAAACCGCGTGCTCCTCGCGTGCTACGGTTGCGGCGACTATAGTGACACGTC
CAACGAGGCACTTACTCCGGTTAGAGTTGCAGAGGCGGTGGATCTTGTGGAATGGCGAAAGTAAGGCAAGTTTTAGCTGGACGTGGGATGCTCT
CTCCATGTTGCCAGACTGTCGGTCTCGACGCGTAGCCAACAGTAGCGACAGTCATGGTCAATCAAAAGCTTGCAAGTTCCTGTACCTATTTCGT
TCAATTAACATGTGAGAGCTCAATGCTTGACATGGATCATATTTAATTGAGGGGAATCGGGGGGCATCGAGAAAGGACCCAAAAAAACTTCCCG
GGGCCAAACCAGGGACCCCCGGTGCTTGCCCCTCCGTTCCCCTTCCCC
> WIP_vj7_LB
NNNNNNNNTGNCNAGCGTCATTTGTTTACACCACAAACATGCATGTCATTTGTCACATCTGTCCTAGCCTACGGTCCTTGAACAAGCTTTTCCC
ATACAAATATGAGCCGCGGCGGTGGGCTGTCCCTTTGCCTTTCACGTGCGCAGCCCGGCACTTGGGCCGGTAATAAGTATGGGGCT
```

Appendix 4 MpWIP and MpRSL1 sequences

```
>MpWIP_genomic
acttcactgcttcgctttcaacagcacacctcctatctgctcctctcccccctccctccctccctctctctctctccctctctctctatctct
ctctctctctctctatctccctctctcccccctctcccatcctttcttctccctccctctctccctcctgtcctgttcgctgacctgcctgtccg
agggtcgtcttcctccatatgagtgccaagctctccctcactcatttctccccagaggacccatcattctctctctctctccctccctctctct
ctctctcttctcttctcttctcttctcttctcctctcttctctccttcaccatctgtcatctgagtctgcaatttgccgatcgtcgcggagctt
gaaataaaagagccgtgcgcctcggtagcaagctcgctcagagtcgcatctgcggccgccggtgcatgccgagagatctagcaggctcatcgca
ccagcactagcactgtgtattattactactactacaaggtctgctcggatcactgaggctgctgtcattcgcaactgttgcactattagctgga
actccgacacgagcagcggaggaaagcacgggtccagagaactgatccgtcaatcgtcttcccatcggcggggacctcgctcgctcgctcgctc
gctcgcattcctcagacaagtgtacaggaaggttatcattttagcgcaccaacaccagcagcattggaattggtgctgctcgagtaattctcat
ctccgacgttctccgagaatgtttatacgcagtcgactcagctgtcatggattggggtcggcgggtagtcggagcttttgaatagcgatcctct
atctgcatcagcccccgcatcttccatcgcatcgcatcacatcacatcacgtccgcctctctctatatctcgagcagacagctgcttctcgcct
ctcgcggacaggataactcttcggctccgtgggaaaagtgagagaagacgagtcagtgattgagtgagagagagagagtggaagggaggaagga
aggaaggaagcaatcaatcaaggaaggaaagaactgctttgctcagtgcagtgcagtgcagtggagtggagtccagtggagtccagtgcagcgc
acttgagagctcgggggctcgaaagggatagatccacaggagaatgcgagtgcatcgagggggagtatccatctaagcgcaaggcaggaggatc
aggcatcagagagtcgagcagcggagtgagagggcaagtcagcgaatctggggctcctcacgaaagcgacagtcgaggtgggctcgcgaggga
cgaagaagggggtctgagcacgatgagctcgaggtaccatgcccagagtagcaccgtggcggtgccgcgggctctgagctccatcacggcgcac
cctctcccgttctgcgaatggctaaaaccgcacctggccccgcccgtgacgcaattgcccgtggccatcccattcaccaccgaggacaatgccg
gcacatcattcagcgtgtcggggccttcgtctggcgccgattcgcccgcctcgggcgcttcccattacaacatgttcggtggcgctcaccacca
attcttccaaactgccaccacgcactacaatcaccaaaggcgatgcccatcgccaggcccctctcgcaccgcccagcccctcccgtttctgagt
ctcagccccccgaccaacaacaacaacaacaacaccagcggcggcagcagcagcaacaaccaccatcaccagcaccacaacagccacaacagcc
acggctcgtcgggggcggcggcggcgtcgacgacgacggccagcaaataccagcagagccagaaccagaatcacaatcacaatcacaatcacaa
tcaggaagaggtgccgctgcgcaccgccatcctgctgcccagcgcttcgtcgtccgggcccgtggacgacgagccggacgagaatgtgacggtg
gcgctgcacatcggcccgcccacctcgtccgactgctcgctcatgagcagcctgcgcagctccgtggtcgagggcatgatgggcaacctgaact
acgagagccccatgggcggcctcgacaaggaccgcttcaaccgagctcgcggccgcctgcccgacggccagtactggatccccacccccgctca
gatcctcgtcggccccacgcagttctcgtgccccgtttgcagcaagaccttcaaccgctacaacaacatgcaggtactgcctccatatctctct
ctctctctctcttcgcttctctctctctctttgtcaggcttttctcgttcattattctttctttnntncatgcttcntttcttctcgtctca
gtgtccgacatgggctcggatcattntatgcnaaaggctcacgngctcgaaaattagcacgtactcatcgtggacatgaatctgtcgacccgnc
cgcagcttagctagctgcgtggcggcgagagtccacggtcagaaaatctgggctgaattccgantctgaggaaacatcagaggctcggcccatc
gcgctcatcatcagatgagatgagtcgagatgatgctcctcgtggggctcatggggcctgagcagcagcagcagcagcagctggctcttggtct
gtccgtgtgtctgagcgtaattatcagaatcgtgtcgattgtttgtgtgagagatggaggcaattgtgtgggatgagtgaggggtgggactgac
gatggcgattcgtggtgggcacggcagatgcacatgtgggggcacgggtcgcaatacaggaaagggccggagtcgctgcggggcacgcagccga
cggcaatgctgcggctgccgtgctactgctgcgcgcaagggtgccggaacaacatcgaccacccgcgcgccaagccgctcaaggatttccgcac
gctgcagacgcactacaagcggaagcacgggcagaagccgttcatgtgccgcaaatgcaacaaggccttcgccgtgcgcggcgactggcggacg
cacgagaagaattgcggcaagctgtggtactgcatatgcgggtcggatttcaagcacaagcgctcgctcaaggatcacatccgcgccttcggca
atgggcacgcgccgcatggcgtcgactcgtgcgaggacgaggaggagaacgcgtcggaggacgatgcggacgaggacatgtcggccggatcgca
atgacgagagtgccctgggagccagtgaatttcttcgccatccacatccatttcatccgccctcgcaataagacttgctttgtaacaagccctc
tcttatctttgtatcccctacacacatttaagagtggtttgctactcagcagcaatgtaccttgcaaagattagatcggatcgtagacgtaaaa
ctcagtgcgtggggctaatgagaaatgctctctccatgatcagcccgcccagcccactcgagacattcttgtagtcttagccttttaaaagaag
tactgctgaggtcgagtgacctttgatcgaacacacaccttcgacttcagatgggcatcaagtcatcgccctgaccaactatccagccagccac
caaccaaccgacgagccatccagtcagtcagtcagtcaatcagtctgccagcccggttagcctctccagcaacaacagcagcaacaacaacaac
aacgacaacaacagcattcacactttacacttttacccgatcaagcgtttagacagaccatgacagcagcagcagctcagactatccttcgac
ctcgcattctcgaccttgctcaacccttcataccttcaatacggcatttcttgtaagtagaacactcttttgaagggctttacagatcaggga
gaagcaattgagaagaacagattagacccacctgtgaaaaccccccttatagggctcttgtgtttacagtcacattcttttcattagacctaatg
cagatctgcagttttcaaggcaagaatcgagttgaaggctgtacctataagcaacatgtaaatgagtcatttgggtatcttcacttattactct
cgagtattcgtttgacacaatctctatcgttgtctctcgtcacattcttcgtcgtgatgaggtctgtgtatctgtgatgattttcttctgatct
gcatatcttcatctatgattctcgagtatttt
```

```
>MpWIP_transcript
tctctctctctctcattccccttcaaagactccagtcgcatttctcttcttttcacttcactgcttcgctttcaacagcacacctcctatctgc
tcctctccccctccctccctccctctctctctcctcaagaagcttgtttttttttttttttttttttttttttttttttttttttttttttt
tttttttttttttctctctctctctctctatctccctctctcccccctctcccatcctttcttctccctccctctctccctcctgtcctgttc
gctgacctgcctgtccgagggtcgtcttcctccatatgagtgccaagctctccctcactcatttctccccagaggacccatcattctctctctc
tctccctccctctctctctctctcttctcttctcttctcttctcttctcctctcttctctccttcaccatctgtcatctgagtctgcaatttgc
cgatcgtcgcggagcttgaaataaaagagccgtgcgcctcggtagcaagctcgctcagagtcgcatctgcggccgccggtgcatgccgagagat
ctagcaggctcatcgcaccagcactagcactgtgtattattactactactacaaggtctgctcggatcactgaggctgctgtcattcgcaactg
ttgcactattagctggaactccgacacgagcagcggaggaaagcacgggtccagagaactgatccgtcaatcgtcttcccatcggcggggacct
cgctcgctcgctcgctcgctcgcattcctcagacaagtgtacaggaaggttatcattttagcgcaccaacaccagcagcattggaattggtgct
gctcgagtaattctcatctccgacgttctccgagaatgtttatacgcagtcgactcagctgtcatggattggggtcggcgggtagtcggagctt
ttgaatagcgatcctctatctgcatcagcccccgcatcttccatcgcatcgcatcacatcacatcacgtccgcctctctctatatctcgagcag
acagctgcttctcgcctctcgcggacaggataactcttcggctccgtgggaaaagtgagagaagacgagtcagtgattgagtgagagagagaga
gtggaagggaggaaggaaggaaggaagcaatcaatcaaggaaggaaagaactgctttgctcagtgcagtgcagtgcagtggagtggagtccagt
ggagtccagtgcagcgcacttgagagctcgggggctcgaaagggatagatccacaggagaatgcgagtgcatcgagggggagtatccatctaag
cgcaaggcaggaggatcaggcatcagagagtcgagcagcggagtgagagggcaagtcagcgaatctggggctcctcacgaaagcgacagtcgag
gtgggctcgcgaggggacgaagaagggggtctgagcacgatgagctcgaggtaccatgcccagagtagcaccgtggcggtgccgcgggctctga
gctccatcacggcgcaccctctcccgttctgcgaatggctaaaaccgcacctggccccgcccgtgacgcaattgcccgtggccatcccattcac
caccgaggacaatgccggcacgtcattcagcgtgtcggggccttcgtctggcgccgattcgcccgcctcgggcgcttcccattacaacatgttc
ggtggcgctcaccaccaattcttccaaactgccaccacgcactacaatcaccaaaggcgatgcccatcgccaggcccctctcgcaccgcccagc
ccctcccgtttctgagtctcagccccccgaccaacaacaacaacaacaacaccagcggcggcagcagcagcaacaaccaccatcaccagcacca
caacagccacaacagccacggctcgtcgggggcggcggcggcgtcgacgacgacggccagcaaataccagcagagccagaaccagaatcacaat
cacaatcacaatcacaatcaggaagaggtgccgctgcgcaccgccatcctgctgcccagcgcttcgtcgtccgggcccgtggacgacgagccgg
acgagaatgtgacggtggcgctgcacatcggcccgcccacctcgtccgactgctcgctcatgagcagcctgcgcagctccgtggtcgagggcat
gatgggcaacctgaactacgagagccccatgggcggcctcgacaaggaccgcttcaaccgagctcgcggccgcctgcccgacggccagtactgg
atccccacccccgctcagatcctcgtcggccccacgcagttctcgtgccccgtttgcagcaagaccttcaaccgctacaacaacatgcagatgc
acatgtgggggcacgggtcgcaatacaggaaagggccggagtcgctgcggggcacgcagccgacggcaatgctgcggctgccgtgctactgctg
cgcgcaagggtgccggaacaacatcgaccacccgcgcgccaagccgctcaaggatttccgcacgctgcagacgcactacaagcggaagcacggg
cagaagccgttcatgtgccgcaaatgcaacaaggccttcgccgtgcgcggcgactggcggacgcacgagaagaattgcggcaagctgtggtact
```

gcatatgcgggtcggatttcaagcacaagcgctcgctcaaggatcacatccgcgccttcggcaatgggcacgcgccgcatggcgtcgactcgtg
cgaggacgaggaggagaacgcgtcggaggacgatgcggacgaggacatgtcggccggatcgcaatgacgagagtgccctgggagccagtgaatt
tcttcgccatccacatccatttcatccgccctcgcaataagacttgctttgtaacaagccctctcttatctttgtatcccctacacacatttaa
gagtggtttgctactcagcagcaatgtaccttgcaaagattagatcggatcgtagacgtaaaactcagtgcgtggggctaatgagaaatgctct
ctccatgatcagcccgcccagcccactcgagacattcttgtagtcttagccttttaaaagaagtactgctgaggtcgagtgacctttgatcgaa
cacacaccttcgacttcagatgggcatcaagtcatcgccctgaccaactatccagccagccaccaaccaaccgacgagccatccagtcagtcag
tcagtcaatcagtctgccagcccggttagcctctccagcaacaacagcagcaacaacaacaacaacgacaacaacagcattcacacttttacac
ttttacccgatcaagcgtttagacagaccatgacagcagcagcagctcagactatccttcgacctcgcattctcgaccttgctcaacccttcat
accttcaatacggcattttcttgtaagtagaacactcttttgaagggctttacagatcagggagaagcaattgagaagaacagattagacccac
ctgtgaaaaccccccttatagggctcttgtgtttacagtcacattcttttcattagacctaatgcagatctgcagttttcaaggcaagaatcgag
ttgaaggctgtacctataagcaacatgtaaatgagtcatttgggtatcttcacttattactctcgagtattcgtttgacaca

>MpWIP_CDS
atgagctcgaggtaccatgcccagagtagcaccgtggcggtgccgcgggctctgagctccatcacggcgcaccctctcccgttctgcgaatggc
taaaaccgcacctggccccgcccgtgacgcaattgcccgtggccatcccattcaccaccgaggacaatgccggcacgtcattcagcgtgtcggg
gccttcgtctggcgccgattcgcccgcctcgggcgcttcccattacaacatgttcggtggcgctcaccaccaattcttccaaactgccaccacg
cactacaatcaccaaaggcgatgcccatcgccaggcccctctcgcaccgcccagcccctcccgtttctgagtctcagcccccgaccaacaaca
acaacaacaacaccagcggcggcagcagcagcaacaaccaccatcaccagcaccacaacagccacaacagccacggctcgtcggggggcggcggc
ggcgtcgacgacgacggccagcaaataccagcagagccagaaccagaatcacaatcacaatcacaatcacaatcaggaagaggtgccgctgcgc
accgccatcctgctgcccagcgcttcgtcgtccgggcccgtggacgacgagccggacgagaatgtgacggtggcgctgcacatcggcccgccca
cctcgtccgactgctcgctcatgagcagcctgcgcagctccgtggtcgagggcatgatgggcaacctgaactacgagagccccatgggcggcct
cgacaaggaccgcttcaaccgagctcgcggccgcctgcccgacggccagtactggatccccacccccgctcagatcctcgtcggccccacgcag
ttctcgtgccccgtttgcagcaagaccttcaaccgctacaacaacatgcagatgcacatgtgggggcacgggtcgcaatacaggaaagggccgg
agtcgctgcggggcacgcagccgacggcaatgctgcggctgccgtgctactgctgcgcgcaagggtgccggaacaacatcgaccacccgcgcgc
caagccgctcaaggatttccgcacgctgcagacgcactacaagcggaagcacgggcagaagccgttcatgtgccgcaaatgcaacaaggccttc
gccgtgcgcggcgactggcggacgcacgagaagaattgcggcaagctgtggtactgcatatgcgggtcggatttcaagcacaagcgctcgctca
aggatcacatccgcgccttcggcaatgggcacgcgccgcatggcgtcgactcgtgcgaggacgaggaggagaacgcgtcggaggacgatgcgga
cgaggacatgtcggccggatcgcaatga

>MpWIP_protein
MSSRYHAQSSTVAVPRALSSITAHPLPFCEWLKPHLAPPVTQLPVAIPFTTEDNAGTSFSVSGPSSGADSPASGASHYNMFGGAHHQFFQTATT
HYNHQRRCPSPGPSRTAQPLPFLSLSPPTNNNNNNTSGGSSSNNHHHQHHNSHNSHGSSGAAAASTTTASKYQQSQNQNHNHNHNHNQEEVPLR
TAILLPSASSSGPVDDEPDENVTVALHIGPPTSSDCSLMSSLRSSVVEGMMGNLNYESPMGGLDKDRFNRARGRLPDGQYWIPTPAQILVGPTQ
FSCPVCSKTFNRYNNMQMHMWGHGSQYRKGPESLRGTQPTAMLRLPCYCCAQGCRNNIDHPRAKPLKDFRTLQTHYKRKHGQKPFMCRKCNKAF
AVRGDWRTHEKNCGKLWYCICGSDFKHKRSLKDHIRAFGNGHAPHGVDSCEDEEENASEDDADEDMSAGSQ

>MpRSL1_genomic
ctcgtcttcaatattttggttataaccctcaatgcatcgcccgtcctctctttcttccgtcagagtctctctccgccccggcctttcctcgttc
ggggatcctttcccttctgtgttctctctctctctttctctctctcctgccggagaacccagtagtctctttctcccgagaattttttactcggt
ctcttgcgcattttgccttctctcaggcataggaccccggtgggtcctcgcagcctcttctccctttctatgtgcaagacttgaagtagcggta
gaatgaggaggggggtcaacatagaattgagcttgacttcatactctacttcacattcatcccctcgctcaagtagattgccattggcagttcat
gttctaaggcatggaagtgcctagataacggtgaccgaggaagaccagagagcgaacactgatatggcttgagatgagaagagatcaacttgtt
tatcgaacgaactgaccgtcacgtttgcgctgcgatggtaccgctggtgtgtttccacctaatcacatcgattctcggagtcaagcttccaata
atccatgtgagctgccgactgtatcctatcaaggtcttgcatgcagtatgaatccaatcaaaaatggtttcatgtactatatagacgatggaca
ggaatgaaaactcacacacctgccttccttcttccttcggcttgatacgatactggccttggattgagaagtttagcatgtgcatatctgggca
cctccgaagtccgtacttgtccgtccaaattgttcaggaaggtcctggcaagcgttttgcgtctagatgacagtaagaactagcccattcccac
ctgcatccagaactctcagtgtttttcccggacatgtgttcttcccgcaatatgtatttgggagcaattggtatcaactaataagttttgcatt
gttacaacagtacgactgcgagtgaactgcagtaatgtgtacaccactaaacaccatctttatacgtgaagtcagaattctgccctattcttat
atcagctgcctcgaccacctgaccatttggaagatgcatcatattccgatgatagacattagacaatttgagcaatcagcctagtacagatacg
gtcgtccgacagacgtttcactgtaaatctcgtgtctcagtgtcaattgtgtaagttgtttgacgcgcatagatcattcggcagctcgaagttg
tctttcgattgaatcaaccagctgacccatacgcgtccatccagttagcccagagtcccgcgttggtgcgattgaagggcctcgggttgtgata
catcagtttttgagcaaccttctcgatgaacgtgatctgctgtctcaaacgagctcttcttacctgccttaaggctcgttagactgttatcgtg
gatctaatcgtgatcttggaaagattatgtccatctatttttttttctgcgggtgaatatcttacagcgctatacacttttgtgtctgtcgag
gtagatgtatccaataacttggtttcgtgatttctcacgactctggtggaccaccagaagactatattggcaccatttgatccacattgtgcgt
tctttgacttccagcttacctcttctttcaatttctgtgtgtagattcaactctctgcgtaatacgagctcactgcgatcttcgacggagctgc
gttcggctattgtggcggccaatgcacgggttttactggctgcattccacaaagttacagctttatatggcttgcattgccagattcagacgac
aaactttgctgttttacccacattctgtcgcattcgtggatccggcttgtactcaagctcaagtgtaggtgcagtatgatgagctcgagaacgc
tgacactgagcaggaggaacaacatggagggagttcttcaagtggcgaaaagtccgaatcacgaaaaccaatcctcgttacaaacatggcaatg
tgatcaagagtctaccgagttcgatacagggtcagctggaaatcttctgagcgcacaagttgctgcagccagtcttgccagttcgtcgcctgag
tctgggcctgagagtgatctgctcccttccgagttgcagaacctagaatcgggtaaccttctttctaacgatgttccctcctccaatccaacgg
caaactttctattacccaagtcatcaatatggggaacaggtgactccgagtatagctgggcacaagaaagtctgtcgcggtcttccggctccat
ggatgatagttctcttttgaccagtgacactctcgggaatgacacttccaggttgstggagacgccgaaaactgatcctctyttgccccagtgg
gcgagtcagctacccaatacaactcagtcatcggcatctatacaaccgcctacgcagatgttccagcgacgatcatggcagacgtcggatgatc
acacttctttgggtctcatacccactatcaatatgcaggcgtacaacacgtcgtttccagctctctcgtcttcttctggaggggggtatgagtct
cgatgcgacgaaattcaacatggactacttgggtggagaaaggagctgcaatatgggaaacttcgaggggacatcggtctacgactctgtggca
aggaagctacactgtcagatgtccggtttagaaacaagctctggaagtgatttttctagcggatctgggatttcaacgcttggtttcgctagtc
ttcagatggggcaaagtcagggtataaccgattcgagcgagaataatgtttctagcggattgttcttacagcagggactcatgcagcagaagaa
ttataatggtgtatctacgatgaatcacgacaaaagcagctacttgagcagtaattcccgtgtgtacttcaaccaatcttctagttccacttta
gatactactaacactattatgaatacatctacttcgaatcagaggtttggtgagaacatgcacagcccaaaattgttgctcgagtcattgtccg
gtaactcagtgttgcgtaacttggagacaaatatctccggtagtcataatccatcggagtccgtcttaggttcacatggctatttgtattctac
gcagaaaaaagtggacacccttgaaagcttgaggaagtggccgaacaatgccagtgggaatggggtgaaagatgtcactgctaacccaaccata
ggtaatttgggaaggttttcgtcctacaatagtaacggagataaggaagatgctttgtacgacacttggccaccatcgagcgtcggacaaaacc
atgccttggatccggtacccgagaccgattctaggaagcgatttctccagatgagtctggggcaacccaacgatacgaatcacaagcgcctgca
tttggtgagttcgcagagactaatttgattttgcacatttgtggtggacatcgattcgcgcctcatactctttcaattgctctttgagctctgg
ctggagtatgaatcagttgacaatcttccattccagttttccagtttggccatcatgttgttttccatgcgtgtccggtttctgctgaacaatc
gtccaccagcaagttacgacaccacttcaagaatggcctgtgtgagagtaggaatgtgattaacctgaccaaaatccgaaatttgtttaggaat
ttttcttcgtctactgacattctcgaaacgatcaaacccatgcaggacacgtctccaaacgtgcagacaacggttgcgcctaccactctaaagc
gctcatcctcatgggcctcttctctcacatccacctgttctaagactgtagacgccattaaccctgggcctctaccttacaccaagtcgaattc
acgacatgccgtcgggccggctttgaacaccaatctgaagccgcgagcccggcagggcagtgcgaacgatcctcagagcatagcagctcgcgta

```
agtttccattatctgtgctaaagccgacatcgagttctgagctcacagtagacgtttcaatgagaagtggatgtccaatatgttaagtccgttg
tcgtgcttctccgtatgtgtatgggccgaaagcgtatcatgcatctaacgacctattttcgcgacatttgtttgcagcatcggagggaacgcat
cagcgagcgactgaagactctacaagatcttgttcctaatggttccaaggtaatcgatcaggaaaggacgaagattgtttctgctgctgggtta
agagcgaatatcagtggaactgagcaagttgtcgcgcggagtagcctataatctcaatttgaaatacttttccccgattcttgcaattgagat
gtgcaatggatgttgcactttaatttcttgtagctgttgacattatcttcttctactcgcaggtagacttggtcaccatgcttgaaaaggccat
caactacgtcaagtttctgcaacttcaagtgaaagtaggtactcttatacagacttaaatcagcaagcgccgaaccgaagtagctgcatttctt
gtgacacacgcatccttagttcttgagattttacgaggatggctttttcttttcttttcttttgtccttgccaacgttwagtacgttttcttc
gcttcagttcagagcactttctttgttttatggccacaattgttcgttttgcaagagaagtctggaaaattgaagctagtaggtactgaaaaaa
ggatgtctggaaatcatcatctcaatggagaaatctcgtgttggaggaaaacatcatctgattttcacaagcagcttccacgaaatatgaacgt
ggcttctgcagtcctagatctgaattacctttactgtactacatgggcagtgccgagttaccactcggcataatcctgttcagtgtcaaggatg
tgctcattcagcatcttccatcgtgaaccttctttttcatctgcgaggtcgttgccgatacggcattgatagattgrgctgtttatcaatacca
agggaaggcggacgttggttaagacctacttatcccaatccttaacattgttcgcacttgaagtggacgagttttaagagatttttttgttcatg
taggtgctgactactgacgactattggccttccaacgagaagggcgctgccccatgtactgctcttcccaactttgaagatatggagggcgctc
tgaaagttctggtcgcgcaggctgagaaagaggcaaattccaatttgagtaacaccactagcaacaacggtggcaattccagcagtagcagtca
agagaacaaatcaccgaatcaagcagacaacaactcgtcctgacagttgatgtgagaaacaccgaataatagaggcttcgaagtcaagaaktcg
cgattgacgtcgaggctttggagaaaacctcctcccactgtttgaaatctgaaggatctcgagaggcctgccgcgactgtacgcctacatgaac
ttgttaatgcgagccatcctcatggcttggtctgcatgatgagatgatggccttaactagatgtagtgtagtgtacgtacatctcggaatcgaa
ttgcatggaggatttaccagtcgctttacagcataagattgtcctgaagtatactgctgtttgagattcatacaagttagacttccacgaggat
tgtctaatgagtgaagtggaaatcaatattagcacagctgtactagtgctactagactaaggacaggactcgacgctttaggacatttgattta
ttcagaaacttaaagctcttgctttactagcataagggcctttgttcaatcccctaattagcttctgattagctcgtgattagcggcttgaaca
gagatagggaacccttagttaattaattaattaattaggtgaaccggtaattaattaattaattattctcacatggttggtgtgatgacatgta
cacgcatcatgtctgcaacgccctaaggagtagtggggaa

>MpRSL1_transcript
ctcgtcttcaatatttggttataaccctcaatgcatcgcccgtcctctctttcttccgtcagagtctctctccgccccggcctttcctcgttc
ggggatcctttcccttctgtgttctctctctctttctctctctcctgccggagaacccagtagtctctttctcccgagaatttttactcggt
ctcttgcgcattttgccttctctcaggcataggaccccggtgggtcctcgcagcctcttctccctttctatgtgcaagacttgaagtagcggta
gaatgaggaggggggtcaacatagaattgagcttgacttcatactctacttcacattcatcccctcgctcaagtagtatagaattgagcttgact
tcatactctacttcacattcatcccctcgctcaagtagattgccattggcagttcatgttctaaggcatggaagtgcctagataacggtgaccg
aggaagaccagagagcgaacactgatatggcttgagatgagaagagatcaacttgtttatcgaacgaactgaccgtcacgtttgcgctgcgatg
gtaccgctggtgtgtttccacctaatcacatcgattctcggagtcaagcttccaataatccatattcaactctctgcgtaatacgagctcactg
cgatcttcgacggagctgcgttcggctattgtggcggccaatgcacgggttttactggctgcattccacaaagttacagctttatatggcttgc
attgccagattcagacgacaaactttgctgttttacccacattctgtcgcattcgtggatccggcttgtactcaagctcaagtgtaggtgcagt
atgatgagctcgagaacgctgacactgagcaggaggaacaacatggaggtttttttagtatgatgagctcgagaacgctgacactgagcaggagg
aacaacatggagggagttcttcaagtggcgaaaagtccgaatcacgaaaaccaatcctcgttacaaacatggcaatgtgatcaagagtctaccg
agttcgatacagggtcagctggaaatcttctgagcgcacaagttgctgcagccagtcttgccagttcgtcgcctgagtctgggcctgagagtga
tctgctcccttccgagttgcagaacctagaatcgggtaaccttctttctaacgatgttccctcctccaatccaacggcaaactttctattaccc
aagtcatcaatatggggaacaggtgactccgagtatagctgggcacaagaaagtctgtcgcggtcttccggctccatggatgatagttctcttt
tgaccagtgacactctcgggaatgacacttccaggttgctggagacgccgaaaactgatcctcttttgccccagtgggcgagtcagctacccaa
tacaactcagtcatcggcatctatacaaccgcctacgcagatgttccagcgacgatcatggcagacgtcggatgatcacacttctttgggtctc
atacccactatcaatatgcaggcgtacaacacgtcgtttccagctctctcgtcttcttctggaggggggtatgagtctcgatgcgacgaaattca
acatggactacttgggtggagaaaggagctgcaatatgggaaacttcgaggggacatcggtctacgactctgtggcaaggaagctacactgtca
gatgtccggtttagaaacaagctctggaagtgatttttctagcggatctgggatttcaacgcttggtttcgctagtcttcagatggggcaaagt
cagggtataaccgattcgagcgagaataatgtttctagcggattgttcttacagcagggactcatgcagcagaagaattataatggtgtatcta
cgatgaatcacgacaaaagcagctacttgagcagtaattcccgtgtgtacttcaaccaatcttctagttccactttagatactactaacactat
tatgaatacatctacttcgaatcagaggtttggtgagaacatgcacagcccaaaattgttgctcgagtcattgtccggtaactcagtgttgcgt
aacttggagacaaatatctccggtagtcataatccatcggagtccgtcttaggttcacatggctatttgtattctacgcagaaaaaagtggaca
cccttgaaagcttgaggaagtggccgaacaatgccagtgggaatggggtgaaagatgtcactgctaacccaaccataggtaatttgggaaggtt
ttcgtcctacaatagtaacggagataaggaagatgctttgtacgacacttggccaccatcgagcgtcggacaaaaccatgccttggatccggta
cccgagaccgattctaggaagcgatttctccagatgagtctggggcaacccaacgatacgaatcacaagcgcctgcatttggacacgtctccaa
acgtgcagacaacggttgcgcctaccactctaaagcgctcatcctcatgggcctcttctctcacatccacctgttctaagactgtagacgccat
taaccctgggcctctaccttacaccaagtcgaattcacgacatgccgtcgggccggctttgaacaccaatctgaagccgcgagcccggcagggc
agtgcgaacgatcctcagagcatagcagctcgccatcggagggaacgcatcagcgagcgactgaagactctacaagatcttgttcctaatggtt
ccaaggtagacttggtcaccatgcttgaaaaggccatcaactacgtcaagtttctgcaacttcaagtgaaagtgctgactactgacgactattg
gccttccaacgagaagggcgctgccccatgtactgctcttcccaactttgaagatatggagggcgctctgaaagttctggtcgcgcaggctgag
aaagaggcaaattccaatttgagtaacaccactagcaacaacggtggcaattccagcagtagcagtcaagagaacaaatcaccgaatcaagcag
acaacaactcgtcctgacagttgatgtgagaaacaccgaataatagaggcttcgaagtcaagaagtcgcgattgacgtcgaggctttggagaaa
acctcctcccactgtttgaaatctgaaggatctcgagaggcctgccgcgactgtacgcctacatgaacttgttaatgcgagccatcctcatggc
ttggtctgcatgatgagatgatggccttttttttttttttttttttttttttggtctgcatgatgagatgatggccttaactagatgtagtg
tagtgtacgtacatctcggaatcgaattgcatggaggatttaccagtcgctttacagcataagattgtcctgaagtatactgctgtttgagatt
catacaagttagacttccacgaggattgtctaatgagtgaagtggaaatcaatattagcacagctgtactagtgctactagactaaggacagga
ctcgacgctttaggacatttgatttattcagaaacttaaagctcttgctttactagcataagggcctttgttcaatcccctaattagcttctga
ttagctcgtgattagcggcttgaacagagatagggaacccttagttaattaattaattaattaggtgaaccggtaattaattaattaattattc
tcacatggttggtgtgatgacatgtacacgcatcatgtctgcaacgccctaaggagtagtgggga

>MpRSL1_CDS
atggagggagttcttcaagtggcgaaaagtccgaatcacgaaaaccaatcctcgttacaaacatggcaatgtgatcaagagtctaccgagttcg
atacagggtcagctggaaatcttctgagcgcacaagttgctgcagccagtcttgccagttcgtcgcctgagtctgggcctgagagtgatctgct
cccttccgagttgcagaacctagaatcgggtaaccttctttctaacgatgttccctcctccaatccaacggcaaactttctattacccaagtca
tcaatatggggaacaggtgactccgagtatagctgggcacaagaaagtctgtcgcggtcttccggctccatggatgatagttctcttttgacca
gtgacactctcgggaatgacacttccaggttggtggagacgccgaaaactgatcctctcttgccccagtgggcgagtcagctacccaatacaac
tcagtcatcggcatctatacaaccgcctacgcagatgttccagcgacgatcatggcagacgtcggatgatcacacttctttgggtctcataccc
actatcaatatgcaggcgtacaacacgtcgtttccagctctctcgtcttcttctggaggggggtatgagtctcgatgcgacgaaattcaacatgg
actacttgggtggagaaaggagctgcaatatgggaaacttcgaggggacatcggtctacgactctgtggcaaggaagctacactgtcagatgtc
cggtttagaaacaagctctggaagtgatttttctagcggatctgggatttcaacgcttggtttcgctagtcttcagatggggcaaagtcagggt
ataaccgattcgagcgagaataatgtttctagcggattgttcttacagcagggactcatgcagcagaagaattataatggtgtatctacgatga
atcacgacaaaagcagctacttgagcagtaattcccgtgtgtacttcaaccaatcttctagttccactttagatactactaacactattatgaa
tacatctacttcgaatcagaggtttggtgagaacatgcacagcccaaaattgttgctcgagtcattgtccggtaactcagtgttgcgtaacttg
gagacaaatatctccggtagtcataatccatcggagtccgtcttaggttcacatggctatttgtattctacgcagaaaaaagtggacacccttg
```

aaagcttgaggaagtggccgaacaatgccagtgggaatggggtgaaagatgtcactgctaacccaaccataggtaatttgggaaggttttcgtc
ctacaatagtaacggagataaggaagatgctttgtacgacacttggccaccatcgagcgtcggacaaaaccatgccttggatccggtacccgag
accgattctaggaagcgatttctccagatgagtctggggcaacccaacgatacgaatcacaagcgcctgcatttggacacgtctccaaacgtgc
agacaacggttgcgcctaccactctaaagcgctcatcctcatgggcctcttctctcacatccacctgttctaagactgtagacgccattaaccc
tgggcctctaccttacaccaagtcgaattcacgacatgccgtcgggccggctttgaacaccaatctgaagccgcgagcccggcagggcagtgcg
aacgatcctcagagcatagcagctcgccatcggagggaacgcatcagcgagcgactgaagactctacaagatcttgttcctaatggttccaagg
tagacttggtcaccatgcttgaaaaggccatcaactacgtcaagtttctgcaacttcaagtgaaagtgctgactactgacgactattggccttc
caacgagaagggcgctgccccatgtactgctcttcccaactttgaagatatggagggcgctctgaaagttctggtcgcgcaggctgagaaagag
gcaaattccaatttgagtaacaccactagcaacaacggtggcaattccagcagtagcagtcaagagaacaaatcaccgaatcaagcagacaaca
actcgtcctga

>MpRSL1_protein
MGQSQGITDSSENNVSSGLFLQQGLMQQKNYNGVSTMNHDKSSYLSSNSRVYFNQSSSSTLDTTNTIMNTSTSNQRFGENMHSPKLLLESLSGN
SVLRNLETNISGSHNPSESVLGSHGYLYSTQKKVDTLESLRKWPNNASGNGVKDVTANPTIGNLGRFSSYNSNGDKEDALYDTWPPSSVGQNHA
LDPVPETDSRKRFLQMSLGQPNDTNHKRLHLDTSPNVQTTVAPTTLKRSSSWASSLTSTCSKTVDAINPGPLPYTKSNSRHAVGPALNTNLKPR
ARQGSANDPQSIAARHRRERISERLKTLQDLVPNGSKVDLVTMLEKAINYVKFLQLQVKVLTTDDYWPSNEKGAAPCTALPNFEDMEGALKVLV
AQAEKEANSNLSNTTSNNGGNSSSSSQENKSPNQADNNSS

Appendix 5 Index of mutants

Gene	Mutants
RSL1	vj3, vj4, vj5, vj6, vj8
WIP	vj7
PI4Ka	vj13, vj32, vj39
MYO-XI	vj11, vj21
SPI	vj69
SRI1	vj25, vj14
EMB2756	vj24, vj68
SRI3	vj67
EXL1	vj19
HSP70-1	vj12