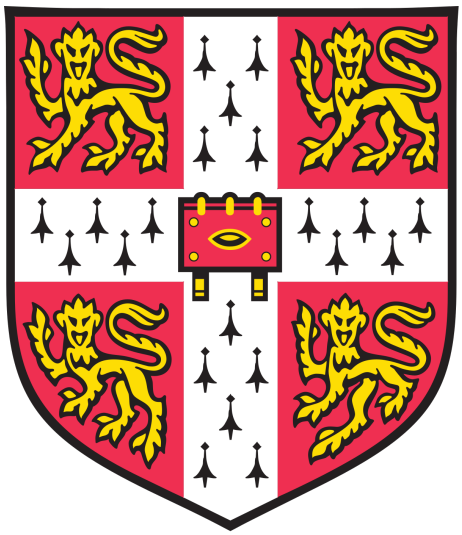
Cell markers for air chamber development in *Marchantia polymorpha*



Marta Tomaselli

Department of Plant Sciences

University of Cambridge

This dissertation is submitted for the degree of Doctor of Philosophy

Lucy Cavendish College

September 2020

Declaration

I hereby declare that this thesis is the result of my own work, unless otherwise specified in the text. The content presented here is original and it has not been submitted in any form for any degree or other qualification. This dissertation contains fewer than 60,000 words exclusive of tables, footnotes, bibliography and appendices.

Marta Tomaselli, September 2020

Abstract

Marchantia polymorpha, thanks to its small genome and its simple cellular architecture, is an excellent model for the study of developmental processes. *Marchantia* can propagate asexually producing small flat disc-like propagules, called gemmae, that develop into adult plants. Gemmae have an open mode of development, which facilitates the live observation of developmental processes.

The fundamental processes underlying morphogenesis in plants are poorly understood and epidermal patterning represents one of the key events during plant development. *Marchantia* produces characteristic epidermal structures called air chambers that represent an example of a self-organising, three-dimensional morphological unit that has a modular and repetitive arrangement, similar to the one observed in stomata meristemoid cells in vascular plants. To date, the cellular scale and the genetic description of air chamber development are poorly resolved.

Therefore, in this dissertation I applied modern genetic and microscopy tools to better map the process of air chamber morphogenesis at a cellular scale and to generate genetic markers to understand the steps involved in this developmental process. Similar tools have been used in *Arabidopsis thaliana*, and they have shown that patterning and cell differentiation can be dissected into discrete steps that provide insights into the mechanisms controlling the process and can also represent valuable targets for engineering the entire process.

In my thesis, I developed high resolution imaging techniques to produce a systematic description of air chamber development. I generated genetic resources to identify potential candidate genes involved in air chamber development by comparing the transcriptome of wild type plants and the *nopperabo1* (*nop1*) mutants that are unable to produce air chambers. Candidate genes were used to build a collection of cellular markers to be screened for their potential involvement in air chamber patterning. The screening led to the identification of two potential transcription factors, Mp*bHLH24* and Mp*WRKY10*.

To my family:
To grandad for supporting me since my very first essay.
To grandma for teaching me strength and persistency.
To dad for teaching me that girls can do anything.
To mum for being my wonder woman.

Acknowledgments

My PhD experience here in Cambridge would have not been the same without the people I met in the last four years. They all helped and supported me in many ways.

I would like to thank Professor Jim Haseloff for guiding me towards the beauty of Marchantia and confocal microscopy.

I would like to thank Dr Eftychios Frangedakis for his patience and his support through every day lab struggle and every page of this manuscript. The lab fairy has once more helped Cinderella to get to the ball on time.

I would like to thank Dr Susana Sauret Gueto for her help with the design of the initial microscopy experiments and for managing the synthesis of all DNA parts, which were developed in the Lab.

I would like to thank Linda Silvestri who taught me the basic of Marchantia tissue culture and helped me through many challenging times. I did appreciate every single word of comfort and every slice of pizza.

I would like to thank Marius Rebmann for his help screening and storing all the generated transformed lines and for sharing his knowledge on data analysis.

I would like to thank Dr Natasha Elina for helping me to troubleshoot the last experiments of my PhD and being an inspiring scientist and ballet enthusiast.

I would also like to thank past and present members of the Haseloff Lab for all the interesting and productive conversations: Adolfo Aguilar, Dr Alan Marron, Alexandra Ting, Alicja Szalapak, Dr Anton Kan, Dr Bernardo Pollak, Camillo Moschner, Dr Christian Boehm, Dave Preston, Emanuele Orsini, Dr Fernando Guzman-Chavez, Jenna Rever, Dr Jenny Molloy, Kasey Markel, Dr Lukas Muller, Dr Mihails Delmans, Massimo Guazzini, Owen Male, Dr Stephanie Norwood, William Boxall.

I would like to thank Dr Sebastian Schornack for his guidance throughout my entire PhD, Dr Philipp Carella and Dr Giulia Arsuffi for their help troubleshooting Marchantia transformations.

On a less academic note...

A huge thanks to Miss Katjana Lange, who has been an incredible housemate, a perfect co-captain and a fantastic submission buddy. Thank you for all the little gestures and the great moments we shared.

Thank you to the Cambridge University Dancesport Team (CUDT) for keeping me sane throughout my PhD.

A special thanks to my Yellowpool flock, Hendrik, Maddi and Rachel, who have been the most incredible and supportive humans I could have asked for. Thank you for all the food, the dancing and the board game nights, which filled these past two years with great memories.

A big thanks goes also to Carrie, who always brought a smile on my face at every training and at every brunch.

Thank you to all the friends who have been supporting me from far away: Elisa and Giulia from Switzerland, Lalla from Trento, Moreno from Moscow/Longiano, Marghe and her little Allegra from home.

Last, but not least, thank you to my amazing family. I would have never been able to do all of this without your love and support.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Plants develop from a single cell, which via cell division and differentiation gives rise to a range of tissues with distinct functions and shapes, in a process called morphogenesis. The fundamental processes underlying morphogenesis in plants are poorly understood. Epidermal patterning, one of the key events during plant morphogenesis, has been extensively studied in angiosperms in an attempt to decipher the underlying mechanisms. However, angiosperms, such as *Arabidopsis thaliana*, have a series of limitations as models for the study of epidermal patterning. For example, the genetic networks controlling these processes are often very complex and redundant. In addition, the accessibility of epidermal structures for observation at the early stages of development is limited. Simpler model systems, like *Marchantia polymorpha*, with less redundant and simple regulatory networks, as well as an open form in its early development, could help to understand the basic mechanisms regulating patterning in plants.

1.1 *Marchantia polymorpha* is a simple model for plant development

Marchantia polymorpha is a thalloid liverwort. It belongs to the group of bryophytes, which includes the closest living relatives to early divergent land plants that colonised terrestrial environments more than 470 million years ago [1, 2]. The dominant phase in the life cycle of *Marchantia* is the haploid gametophyte (Figure 1.1).

The gametophyte has a simple structure, it lacks specialised conducting tissue and water and nutrients are absorbed from the environment by osmosis and active trans-

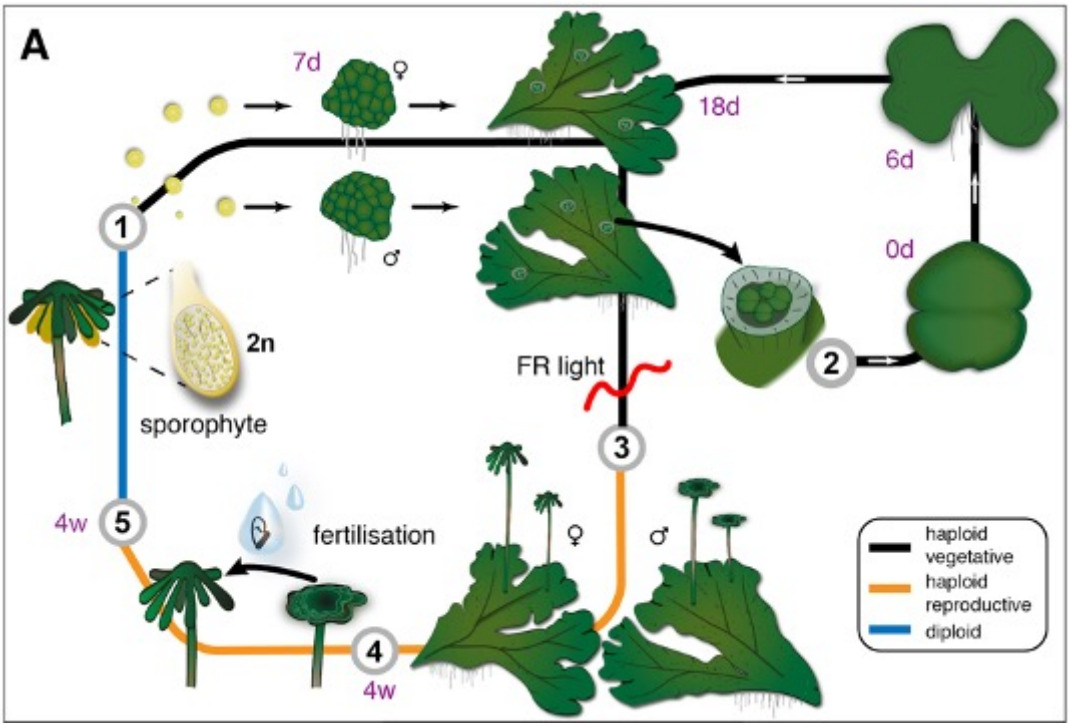


Figure 1.1 **Life cycle of *Marchantia polymorpha*.**
Adapted from [3]. *Marchantia* develops as a flat thallus, which can grow from spores (1) or gemmae (2). *Marchantia* is a dioecious plant, meaning that there are both male and female plants, and it can undergo both sexual and asexual reproduction. Sexual reproduction (yellow line) is induced by far red light (FR light), which prompts the development of the reproductive structures, antheridia and archegonia, in male and female plants, respectively. Female plants produce eggs inside the archegonia, while male plants produce sperm in their antheridia. Sperm exploits rain droplets to move onto archegonia, where it fertilises the eggs (4). A diploid sporophyte is produced (5) and spores are generated inside it by meiosis. Once the sporophyte is mature, the haploid spores are released into the environment and they will develop into new flattened structures called thalli. Vegetative reproduction (2) is carried out through the formation of structures on the dorsal surface of *Marchantia* thalli, called gemma cups. Gemma cups contain clonal multicellular disc-like structures called gemmae that can be released into the environment and produce new thalli.

port, through dedicated structures called rhizoids, which develop on the ventral side of the thallus. *Marchantia* propagates both sexually and asexually [4]. Sexual reproduction occurs via reproductive organs called antheridia and archegonia produced by male and female individuals, respectively. Archegonia produce the egg, which is fertilised by the sperm produced in antheridia. The fertilised egg then develops into the sporophyte that, via meiosis, produces spores that will germinate and develop into new gametophytes.

Asexual propagation occurs via gemmae. Gemmae are multicellular clonal propagules, half millimetre in size, which are produced inside dedicated structures called gemma cups, which develop on the dorsal surface of the thallus around two weeks after spore germination [5, 6, 7]. Gemmae develop from single cells on the floor of gemma cups; they undergo a transverse division to separate the developing gemma from the stalk, attaching the gemma to the thallus surface. They are maintained in an undifferentiated state until they are released into the environment. Gemmae are symmetrical, and they present two indentations (Figure 1.2), called notches, containing the apical cell, from which the gemma will grow to develop an adult branched thallus [8]. In the central part of the gemma, rhizoid initials (large rounded cells) can be observed and once the dorso-ventral axis is defined under the effect of gravity, rhizoids will develop on the ventral side of the gemma, to anchor the thallus to the ground. A few days after gemma deposition on soil, dark green pyramid-like structures with round cell-rimmed openings can be observed on the dorsal side of the gemmae: they are called air chambers (Figure 1.2). As the thallus grows, new apical notches are created and thalli undergo a series of ordered bifurcations, while new air chambers continue to be generated.
The ability of gemmae to develop in an open and modular fashion, provides an advantage for the microscopic observation of *Marchantia* development and morphogenesis.
In developmental biology, the control of cell patterning and differentiation are two of the main questions to be answered. The plant epidermis (thanks to its restricted number of cell types, its accessibility for observation and manipulation) represents a simple system to understand the key regulators involved in patterning. The root epidermis of the Brassicaceae plant model *Arabidopsis* shows only two cell-types: root hairs and non root-hair cells. Similarly, the leaf epidermis is composed of three cell-types: pavement cells, stomata and trichomes. Genetic screenings have led to the identification of key mechanisms involved in the cell fate determination of the

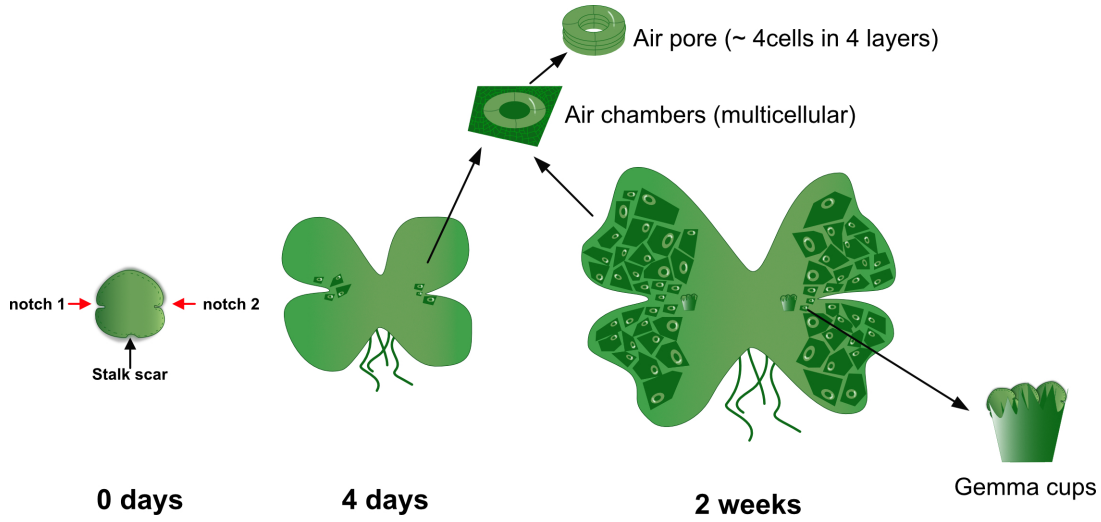


Figure 1.2 **Development of dorsal features in gemmae**.
Once gemmae are released from the gemma cup (0 days) only the notches (red arrows) and the stalk scar (small black arrow) are recognisable. 4 days later the first air chambers are visible by naked eye and they continue to cover the thallus as the plant grows. Gemma cups appear on gemmalings dorsal epidermis approximately two weeks after plating.

above mentioned cell-types in *Arabidopsis*. Analogous mechanisms have been shown to be conserved in other plant clades, including monocots as well as early divergent land plants [9, 10, 11, 12].
Currently, the debate on bryophytes phylogeny is still open, but recent studies support the monophyly of liverworts and mosses, while hornworts could be a sister group to them or to all land plants (see Figure 1.3 b) and c)) [13, 14, 15] .

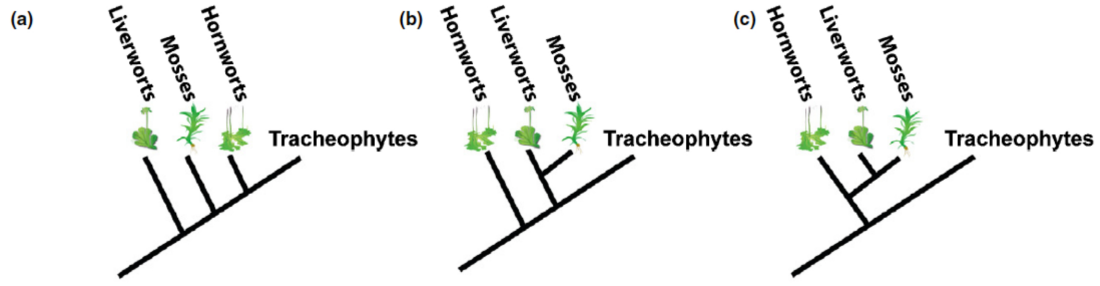


Figure 1.3 **Phylogeny of land plants**.
Figure adapted from [16]. a) the hypotesis from [17] suggested that liverworts were a sister group to all land plants. b) hornworts are sister to all land plants or c) to other bryophytes. There is mounting evidence that all bryophytes are monophyletic as shown in c).

The hypothesis of a monophyletic origin is supported by evidences obtained through the analysis of over 1000 plant transcriptomes [18] and genome studies [12, 19, 20],

rejecting the previous hypothesis considering liverworts a sister group to all land plant lineages (Figure 1.3 a))[17].
Considering the phylogenetic relationship between the three groups, it would be highly informative to study epidermal patterning in *Marchantia*. This could shed light on the origin of mechanisms regulating epidermal patterning in land plants as well as clarifying the relationship (or lack of) between air chambers and stomata.

Additionally, the study of a structure that appears to be lost in other plant organisms could help to better understand the evolution of genetic circuits, in general. Finally, *Marchantia*, thanks to its morphological simplicity and low genetic redundancy, represents the ideal platform to engineer these circuits to understand their potential and employ them to encode for new and improved functions.

1.2 Patterning in plants

Patterning includes all those mechanisms that allow a group of identical cells to adopt different cell fates and arrangements during morphogenesis. Biochemical gradients, physical processes and gene regulatory networks contribute to morphogenesis [21]. Biochemical gradients rely on molecules, called morphogens, that distribute in an organism establishing a gradient, which provides cells with positional information. Cells can then respond to the morphogen activating different genetic networks to undertake distinct cell fates and initiate morphogenesis. Cell patterning can be seen as a self-propagating system, where secreted morphogens and cell-cell interactions lead to the formation of cellular domains, where cells undertake an identity (cell differentiation) as a function of distance from the source of the signal [22]. Differentiated cells can, in turn, act as new sources of signals with different chemical nature. In fact, patterning can also be regulated by short-range signals (peptides, proteins, miRNAs etc.), which move between neighbouring cells contributing to the differntiation of new cell types [23]. Therefore, cell patterning and differentiation are two distinct processes, but they are tightly linked and dependent on each other. Although physical processes have been shown to play an important role in morphogenesis, the investigation of genetic networks and biochemical gradients can be considered a starting point to describe morphogenesis.
In angiosperms, most of these developmental decisions are taken in regions localised at the apical meristems of shoots and roots. The aerial structures of the plant will develop from the shoot apical meristem, while underground organs will originate

from the root apical meristem.

1.3 Epidermal patterning in plants

The plant epidermis plays a key role in plant adaptation on land: it ensures protection against dehydration and UV damage, while allowing gas exchange and also providing a first line of defence against herbivores and pathogens. To perform all these tasks, once the epidermal layer has been formed, it has to produce specialised cell types to carry out all these functions. For example, stomata regulate gas exchanges, trichomes protect leaves from UV damage and predators, and root hairs optimise nutrient absorption and establish interactions with microorganisms. The development and correct distribution of these structures requires further patterning processes to take place on the epidermal layer. Being at the interface between the plant and the outer environment, the epidermis is easily accessible for observation, and therefore, it represents the ideal platform to study the differentiation and patterning processes of these specialised cell types, which have been shown to be regulated by conserved mechanisms. One of the major features of the characterisation of the epidermis is the ease of genetic analysis, with which specified epidermal features can be mutated without lethal effects. The study of the development of three different epidermal cell types, trichomes, stomata and root hairs 1.4, has shed light onto the mechanisms and gene families regulating epidermal patterning in shoots and roots of land plants.

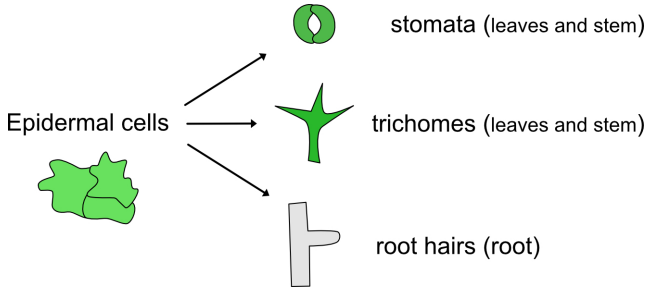


Figure 1.4 **Epidermis derived cell types in *Arabidopsis thaliana*.**
Arabidopsis thaliana epidermis shows morphologically diverse cells. The study of stomata, trichomes and root hairs has allowed to shed light on the mechanisms regulating cell differentiation and patterning.

Arabidopsis has been widely used as a model plant since the 1980s [24]. Compared to other angiosperms, including crops, it has a relatively fast life cycle, it is a diploid organism, easy to transform and propagate. Thanks to its ease of genetic manip-

ulation, it soon became the favoured organism to identify genes involved in plant development. Consequently, *Arabidopsis* has been employed for quantitative and qualitative studies on the morphogenesis of the embryo [25], the apical meristem [26], leaves [27] and root [28]. All epidermal structures in *Arabidopsis* develop from the outer layer of the embryo, which is defined from the 16-cell stage of embryogenesis. By the late globular stage (Figure 1.5 C), when the apical and basal domain of the embryo have been established, the shoot apical meristem outer layer (Figure 1.5 C in orange) can be distinguished from the other embryonic epidermal tissues, as it retains the ability to differentiate into the plant vegetative and reproductive organs (leaves and flowers).

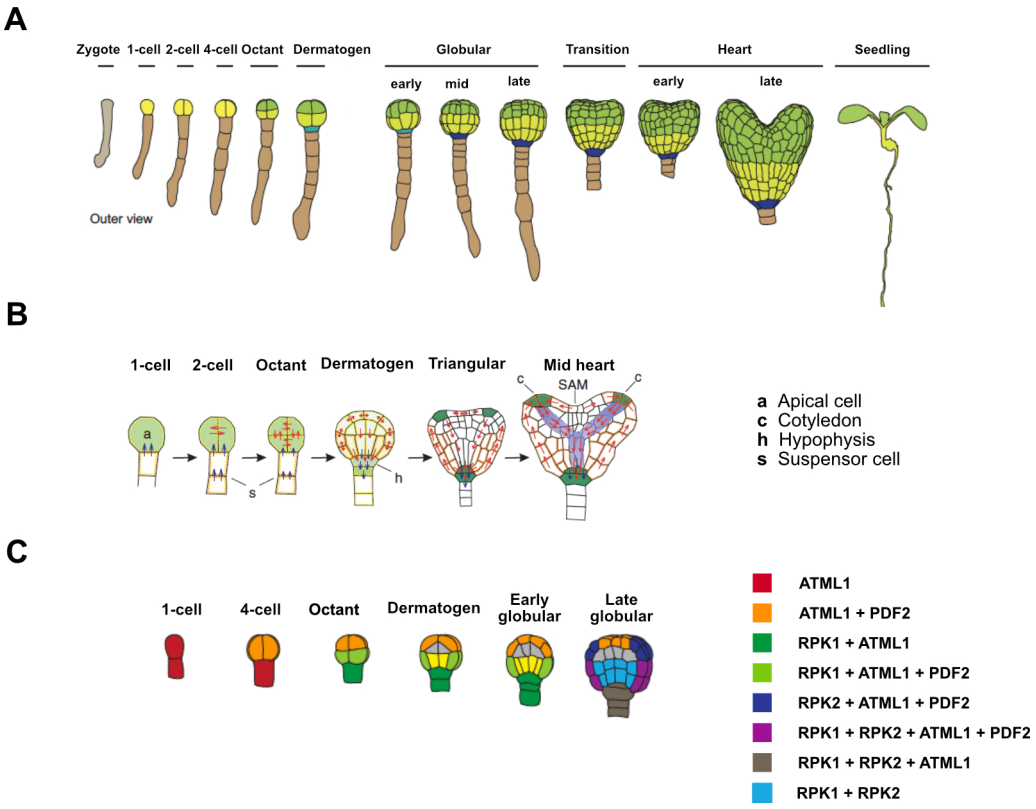


Figure 1.5 **Embryogenesis in *Arabidopsis thaliana*.**
Diagram adapted from [25] and [29]. A) Key stages of embryogenesis in *Arabidopsis*. *Arabidopsis* embryo develops from a single cell, called zygote. The first cell division (1-cell stage) defines a suspensor (elongated brown cell), which will support the embryo (yellow round cells). The embryo will undergo a series of cell divisions to reach the final late heart stage before germination, where different cell type and cell layers are defined, determining the radial and longitudinal axes of the seedling. B) Auxin flux during embryogenesis. The accumulation of auxin is shown in green and the arrows indicate the direction of the flow, corresponding to the polar localisation of PIN proteins. C) From the early stages, the expression of *ATML1* and *PDF2* localises in the outer layer of the embryo. Different colours correspond to different gene combinations. The layer 1 of SAM is visible in orange at the late globular stage. All the cell types expressing both the *ATML1* and *PDF2* genes will form the epidermal layer of the plant.

1.3.1 Establishment of epidermis during embryogenesis

Auxins play a major role as morphogens in *Arabidopsis* and their action starts from the very first divisions of the embryo. As the zygote divides asymmetrically forming a small cell at the top (apical cell) and an elongated one at the bottom, auxin is transported to the apical cell through a polarised subcellular distribution of *PIN-formed* (*PIN*) transporters, creating a transient auxin maximum in the top part of the embryo [29] (green areas in Figure 1.5 B). At the 8-cell stage of embryo development, the expression of two *WUSCHEL RELATED HOMEOBOX* genes, *WOX2* and *WOX8/9* defines the apical and the basal domains [30]. The transient accumulation of auxin at the apex of the embryo induces the transcription of two auxin responsive factors (ARFs), *MONOPTEROS* (*MP*) and *NONPHOTOTROPIC HYPOCOTYL4* (*NPH4*), which in turn promote the expression of the AP2-domain *PLETHORA* (*PLT*) family in the basal domain [31, 32]. *MP* genes and *WOX2* and *8* promote as well, in a redundant manner, the transcription of PIN1 proteins, which redirect the auxin flux towards the basal cells, forming a new maximum [30]. The basal cell, where the *PLT* gene is expressed and the auxin maximum forms, will give rise to the quiescent centre (QC) of the root, and will coordinate the maintenance of the root apical meristem (RAM) [32].

While *PLT* genes define the region of the embryo from which the root will develop, the upper region, expressing class III homeodomain ZIP proteins (HD-ZIP III), will give rise to the shoot [33]. As the embryo grows, the polar transport of auxin creates two maxima on the sides of the apical region of the embryo and then auxin moves downwards, generating a low auxin area at the centre of the embryo apex. This domain will become the shoot apical meristem (SAM). In this area, HD-ZIP III proteins induce the expression of *WUSCHEL* (*WUS*), a gene expressed from the 16-cell stage of embryo development which is required for shoot meristem maintenance [34, 35]. When the embryo enters the transition phase (Figure 1.5 A), the NAC transcription factor *CUP-SHAPED COTYLEDON* (*CUC2*) starts to be expressed between the two developing cotyledons, and it defines a boundary zone between developing leaves and the SAM [36]. At the same time, *SHOOT MERISTEMLESS* (*STM*), a Class I KNOTTED-like homeobox protein, starts to be expressed in the central portion of the SAM, where, together with *WUS*, it contributes to SAM maintenance [37].

The SAM can be anatomically divided into three different zones (Figure 1.6 A) and three separate layers (Figure 1.6 B). The central part of the SAM is called the central zone (CZ) and it contains undifferentiated cells which divide rarely (red cells in Figure 1.6 A). The surrounding area is called peripheral zone (PZ) (orange cells in Figure 1.6 A). Here, cells divide and can be recruited to undertake different cell-fates. Finally, the rib zone (RZ, green cells in Figure 1.6 A) is located underneath the CZ and it will give rise to the internal tissues (vasculature and pith) of the shoot. A cross section of the shoot apical meristem reveals three different layers, layer 1 and 2 (L1 and L2 also referred to as tunica) and layer 3 (L3, also called corpus), shown in Figure 1.6 B. Cells belonging to L3 divide in all directions, while L2 cells tend to divide mostly anticlinally; both layers contribute to the formation of vasculature and internal tissue of the plant lateral organs. Conversely, L1 will give rise to the shoot epidermis [38].

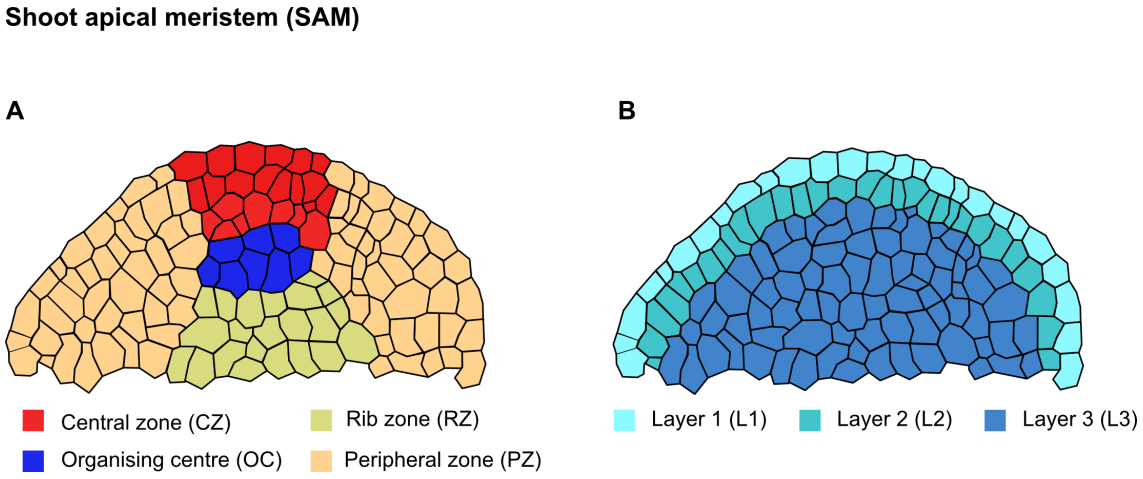


Figure 1.6 **Structure of the shoot apical meristem in *Arabidopsis thaliana*.** A) Schematic representation of the three different functional zones in the shoot apical meristem (SAM). The central zone (CZ), in red, contains undifferentiated cells, whose division is controlled by the organising centre (OC), in blue. The peripheral zone (PZ), in orange, surrounds the CZ, while the rib zone (RZ), in green, gives rise to the vasculature and pith of the shoot. B) Schematic representation of the three layers of the SAM. Layer 1 (light blue) will generate the shoot epidermis, while L2 (jade) and L3 (blue) will contribute to the formation of the vasculature and internal tissues of the plant lateral organs.

ARABIDOPSIS THALIANA MERISTEM LAYER1 (*ATML1*) is a homeodomain gene expressed during the first divisions of the embryo and its expression is restricted to the outer layer of the protodermal tissue (L1) starting at the octant stage of the embryo (Figure 1.5 C). The *ATML1* gene belongs to the homeodomain leucine zipper

class IV (HD-ZIP IV) of transcription factors. Its closest homologue, *PROTODERMAL FACTOR 2* (*PDF2*) is also expressed exclusively in the outer layer at the early globular stages [39]. While the *atml1* single mutant does not show a clear phenotype, the *atml1 pdf2* double mutant shows epidermal defects and it cannot produce cotyledons [39, 40]. Both ATML1 and PDF2 are able to bind to a cis-regulatory region called L1-box region, which is present in various epidermis-specific genes including *ATML1* and *PDF2* themselves, suggesting that they reinforce each others' activity in establishing protodermal tissue identity [41, 42]. Recently, the ectopic expression of *ATML1* has been shown to induce epidermis-specific genes in the inner tissue of the embryo, showing that this gene is sufficient to trigger protodermal identity and it can be considered a master transcriptional regulator of epidermal cell fate in the shoot [43].The *ATML1* gene expression in the epidermal layer ensures the correct differentiation of the plant epidermis, where specialised structures, such as stomata, trichomes and root hairs will develop during the plant vegetative growth.

1.3.2 Trichome and root hair patterning

Trichomes are structures that grow on plant shoot derived epidermal tissue (leaves, petals and stems). Their name comes from the Greek word $\tau\rho\acute{\iota}\chi\omega\mu\alpha$ (trichoma), which means hair. Trichomes present themselves with different morphologies: they can be single celled or multicellular, branched or not branched and they can be connected to glands producing defence related substances. Trichomes have been used as a model system to study the patterning of epidermal cells in the leaf of the model plant *Arabidopsis*. They develop from protodermal cells of leaves and stem epidermis, which stop mitotic cell division and undergo endoreduplication cycles, increasing their nuclear DNA content. This event leads to an increase in cell size and a change in the direction of the cell growth from parallel to perpendicular to the leaf surface. In the case of branched trichomes, the enlarged cells undergo two consecutive branching events to form the mature trichome. The entire process is illustrated in Figure 1.7.

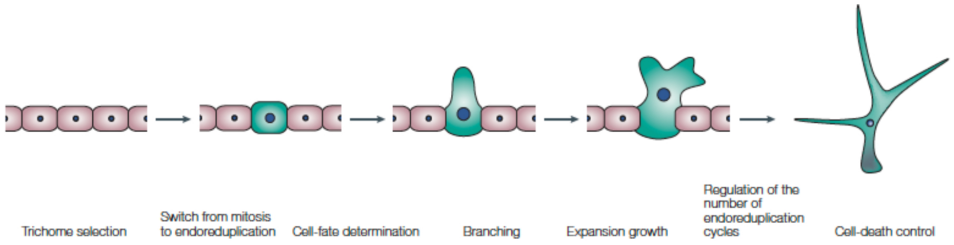


Figure 1.7 **Stages of unicellular branched trichomes differentiation in *Arabidopsis thaliana*.**
Image from [44]. The protodermal cells of the epidermis (pink) can undergo endoreduplication (green cells) and develop into trichomes. Once trichome morphogenesis is initiated, the cell starts branching, expanding and undergoes additional cycles of endoreduplication.

Trichomes are initially separated by three cells from other developing initials, indicating that a mechanism is regulating their spacing. Mutagenesis and clonal analysis experiments led to the identification of a set of genes regulating trichome initiation and spacing. A MYB- basic helix-loop-helix (bHLH)-WD40 (MBW) complex determines which cells will become trichomes and which ones will remain protodermal cells. The R2-R3 MYB *GLABRA1* (*GL1*) is part of the MBW complex, along with the two redundant bHLHs, *GLABRA3* (*GL3*) and *ENHANCER OF GL3* (*EGL3*) and a WD40 protein *TRANSPARENT TESTA GLABRA1* (*TTG1*) [45]. This conformation of the MBW complex acts as a positive regulator of trichome morphogenesis, inducing the transcription of *GLABRA2* (*GL2*). The complex also induces

the transcription of three negative regulators of trichome development: a set of MYB-like proteins, which include *TRIPTYCHON* (*TRY*), *CAPRICE* (*CPC*), and *CAPRICE TRIPTYCHON1* (*ETC1*). TRY, CPC and ETC1 travel to neighbouring cells, where they substitute GL1 in the MBW complex. This change in the MBW composition causes an inhibition of *GL2* transcription and prevents the cell from developing into a trichome (Figure 1.8 A) [45, 46].

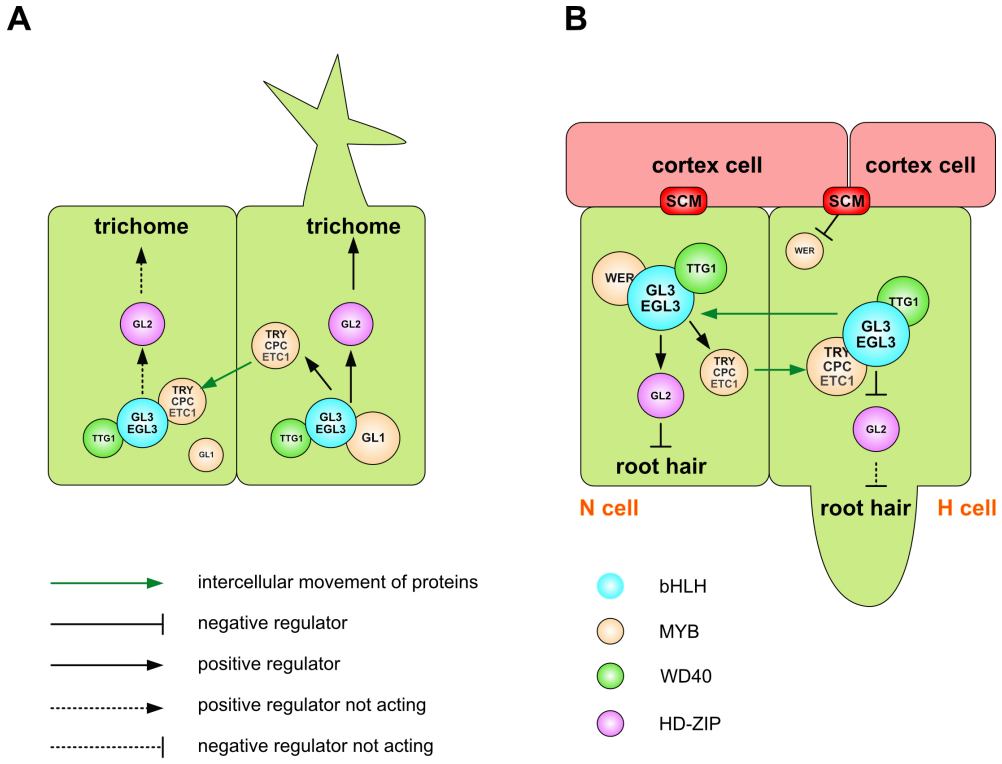


Figure 1.8 **Genetic control of trichomes and root hairs development *Arabidopsis thaliana*.**
Adapted from [47]. A) Trichomes develop when TRY/CPC/ETC1 are replaced in the MBW complex by GL1 and the transcription of *GL2* initiates trichome morphogenesis. Green arrows indicate the proteins' movement between neighbouring cells. Black arrows illustrate the positive activators and black bars negative activators. MYBs transcription factors are highlighted in yellow, WD 40 proteins in green, bHLH in blue and HD-ZIP in magenta. B) Root hair development is determined by positional cues coming from the cortical layer (SCRAMBLE, SCM, receptor-like kinase in red) which prevents WER from binding to the MBW protein complex formed by GL3/EGL3, TTG1 and TRY/CPC/ETC1. The MBW complex prevents the transcription of *GL2*, a negative regulator of root hair development (H cell). When WER substitutes TRY/CPC/ETC1 in the MBW complex, *GL2* is transcribed preventing the cell to enter the root hair fate (N cells).

The overexpression of any of the negative regulators, *TRY*, *CPC* and *ETC1*, is sufficient to abolish trichome development. Loss of function mutants for individual

genes only form small clusters of trichomes, which increase in size in the double *try cpc* and triple *try cpc etc1* mutants. This reveals the redundant function of the three inhibitors and their crucial role in controlling the correct trichome patterning [48]. Similarly, the loss of function mutants of the positive regulators *TTG, GL3/EGL1* and *GL1* lead to glabrous plants, with little or no hair at all [48].

Root hairs are filamentous structures developing on the root epidermis of vascular plants. They absorb water and nutrients from the soil, anchor the plant to the ground and can form symbiotic interactions with microorganisms present in the environment. In *Arabidopsis*, a positional cue mediated by the SCRAMBLED (SCM) leucine-rich receptor-like kinase triggers the initiation of the root hair cell fate depending on the position of the epidermal cell in respect to the underlying cortical layer cells [49]. There are some conserved elements between trichome and root hair initiation. The MBW complex is formed by the WD40 protein TTG1 and GL3/EGL3 bHLH, and in the root, they interact with a R2-R3 MYB protein called WEREWOLF (WER) to induce the transcription of *GL2*. In this instance, *GL2* acts as a negative regulator, preventing the cell from entering the root hair fate (N cells in Figure 1.8 B). The MBW complex induces the expression of *TRY/CPC/ETC1*, which move to neighbouring cells, where they compete with WER in assembling the trimeric MBW complex and inhibit *GL2* transcription to allow the cell to differentiate into a root hair (H cells), Figure 1.8 B. In roots, it has also been shown that the bHLH genes *GL3* and *EGL1* are transcribed and translated in H cells and they are then transported as proteins into non root hair cells, showing another example of positional regulation of patterning [49]. The high degree of conservation of this regulatory network, which functions in both shoot and root derived epidermis, suggests that it might be conserved in plants' L1 derived tissues.

1.3.3 Stomatal patterning

Stomata, from the Greek word $\sigma\tau\acute{o}\mu\alpha$ for "mouth", are small apertures that can be found on the aerial epidermis of plants. They allow exchange of CO_2 with O_2 and water vapour, and they can open and close in response to environmental stimuli. These apertures are surrounded by two cells called guard cells, and each aperture is separated from the other by at least one multi-lobed epidermal cell, called a pavement cell. Stomata development involves several stages. They originate from protodermal cells, which develop into meristemoid mother cells (MMC). MMCs divide asymmetrically to generate a large stomatal-lineage ground cell (SLGC) and

a small meristemoid cell. This initial step is called "entry division", to indicate the cell commitment to the stomatal lineage. Meristemoid cells undergo a second asymmetric cell division in which guard mother cells (GMCs) are produced. Finally, the GMCs divide symmetrically to form the two guard cells (GCs of stomata). All the stages are schematically represented in Figure 1.9. The meristemoid cell can also continue to divide asymmetrically ("amplifying divisions") producing more SLGCs, which could later divide to generate more pavement and meristemoid cells [50].

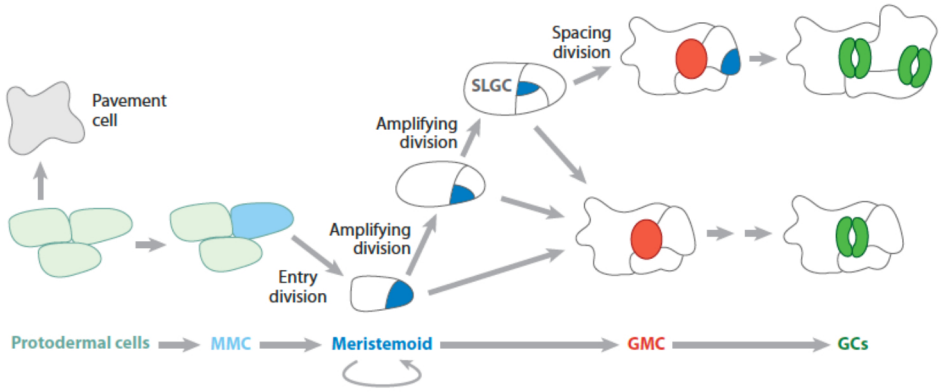


Figure 1.9 **Stages of stomata development in *Arabidopsis thaliana*.** Image from [51]. Diagram of cell state transitions from protodermal to differentiated guard cells. MMC (in light blue) is the meristemoid mother cell, meristemoids are shown in blue, SLGC (in white) is a stomatal-lineage ground cell, GMC (in red), stands for guard mother cell and GC (in green) for guard cell. Meristemoids can either differentiate into guard cells (lower part of the graph) or undergo a series of amplification divisions before differentiating and generating a new meristemoid cell distal to the original one (upper part of the graph). This system contributes to the correct patterning of stomata.

A series of intracellular and extracellular signals that regulate stomata development have been identified [52, 53, 54, 55]. Three bHLH proteins are involved in cell fate determination from the establishment of meristemoid mother cells to the differentiation of guard cells: *SPEECHLESS* (*SPCH*), *MUTE* and *FAMA*. *SPCH* is responsible for initiating the stomatal lineage, coordinating the transition from meristemoid mother cells (MMCs) to meristemoid. Loss of function mutants for this gene are not able to divide asymmetrically to generate a meristemoid and a stomatal lineage ground cell [56]. *MUTE* terminates the asymmetric divisions that occurs in meristemoid cells, allowing the transition to guard mother cells [57]. Finally, *FAMA* controls the symmetric division of GMCs leading to the formation of two fully differentiated guard cells. Plants with non functional copies of *FAMA*

Table 1.1 List of genes involved in trichome and/or root hair patterning in *Arabidopsis thaliana.*

Gene name	**Class**	**Effect in trichomes**	**Effect in root hairs**
TTG1	TF, WD40	number and spacing	number and spacing
GL1	TF, MYB	number and spacing	NA
WER	TF, MYB	NA	number and spacing
GL3	TF, bHLH	number and spacing	number and spacing
EGL1	TF, bHLH	number and spacing	number and spacing
GL2	TF, HD-ZIP	morphogenesis	morphogenesis
TRY	TF, MYB	number and spacing	number and spacing
CPC	TF, MYB	number and spacing	number and spacing
ETC1	TF, MYB	number and spacing	number and spacing
SCM	LRR-RLK	NA	root hair initiation

show excessive division of GMCs and they never differentiate into guard cells [58]. The expression of these three bHLH is specific to each cell fate transition, but two more bHLH proteins are involved throughout the entire differentiation process of guard cells. SCREAM and SCREAM2 form heterodimers with SPCH, MUTE and FAMA to promote the transition from MMCs to meristemoid to GMCs and finally to GCs [52] (Figure 1.10, blue circles). When mutated, *scream/2* plants show phenotypes similar to the ones observed in *spch*, *fama* or *mute* loss of function mutants. The final differentiation step of GCMs to guard cells is also coordinated by two more transcription factors belonging to the MYB family: *FOUR LIPS* (*FLP*) and *MYB88* (yellow circles in Figure 1.10) [59]. Despite members of the MYB family have been known to interact with bHLH protein, there is no experimental evidence of physical interaction between FLP and MYB88 and FAMA [58].

As previously mentioned, stomata development is also regulated by extracellular signals, providing cells with positional cues. Small cysteine-rich peptides [53, 52] called EPIDERMAL FACTOR1 (EPF1), EPIDERMAL FACTOR 2 (EPF2) and CHALLAH (CHAL/EPFL6) [60] bind to the ERECTA family of leucine-rich-repeat receptor-like kinases (LRR-RLKs) [55] and TOO MANY MOUTHS (TMM), another LRR receptor [50, 59]. The binding of these peptides to the receptors activates the components of a mitogen-activated protein kinase (MAPK) cascade that inactivates SPCH through phosphorylation, preventing the differentiation of MMCs into meristemoids and SLGCs. STOMATOGEN/EPFL9 [61, 62] is another cysteine rich peptide, belonging to the EPF family. It has been hypothesized that STOM-

ATOGEN competes with the other peptides to bind LRR-receptors [63], but it acts instead as a positive regulator of stomata development. While loss of function of *EPF1*, *EPF2*, *CHAL*, *TMM* and *ERECTA* leads to an increase in the number of stomata on the epidermis, *stomatogen* mutants show decreased numbers of stomata.

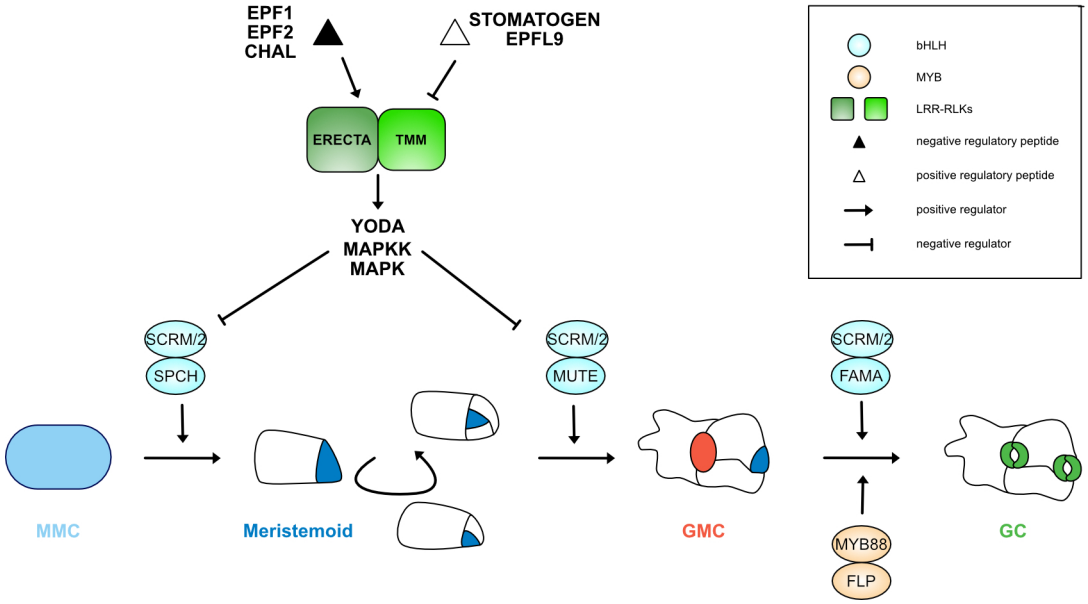


Figure 1.10 **Internal and external signals regulating stomata development in *Arabidopsis thaliana*.**
Adapted from [51]. Diagram of external and internal cues regulating cell fate transitions in stomata development. The EPF1, EPF2, CHAL and STOMATOGEN peptides (black and white triangles respectively) bind to the LRR receptors (ERECTA and TMM) in the green squares. The bHLH heterodimeric complex (light blue circles) formed by SPCH and SCRM/2 forms in MMC cells and it is responsible for their transition to meristemoid. The expression of the MUTE protein in meristemoid cells and its interaction with SCRM/2 lead to the differentiation into GMC. Finally, the formation of the heterodimer by FAMA and SCRM/2 in GMC leads to fully differentiated guard cells. MYB88 and FLP are two MYB proteins (yellow circles) that are required for the correct specification of GC.

Regulation of asymmetric cell divisions in stomata development

Asymmetric cell divisions (ACD) play a fundamental role in cell fate specification in the early stages of stomata development to form distinc meristemoid and stomata-lineage ground cells (SLGCs). For this type of division to take place, cells need to generate asymmetry in the progenitor cell. This is achieved by changing the physical shape of the cell and unequally distributing intracellular components to daughter

cells [63].

In plants, internal and external cues contribute to the process. The asymmetric division leading to the first meristemoid cells is triggered by internal signals, while secondary meristemoids are oriented in response to external signals coming from neighbouring guard mother cells and guard cells [64]. The asymmetry of these divisions is not physically determined by the action of these internal and external cues [65]. In plants a series of proteins (called polar proteins) have been identified to play a role in the generation of asymmetry [66, 67, 68, 69]. *BREAKING OF ASYMMETRY ON THE STOMATAL LINEAGE* (*BASL*) controls the asymmetric cell division leading to the formation of meristemoid [67] (Figure 1.11 A). During stomata development BASL localises both in the nucleus and in the membrane distal to the plane of cell division, organising the outgrowth of the stomatal cell lineage before and after the asymmetric division takes place [67]. The dynamic localisation of BASL proteins is possible thanks to the presence of a nuclear localisation sequence for nuclear import and FxFP and CCR/KF motives at the C-terminus of the protein [70]. The FxFP domain mediates the interaction between Mitogen-Activated Protein Kinases (MAPKs), such as YDA (see Figure 1.11 B), and their substrate, while the CCF/KF motif is important for the protein retention at the membrane [71, 72, 73]. POLAR LOCALIZATION DURING ASYMMETRIC DIVISION AND REDISTRIBUTION (POLAR) [68] is another polar protein which transiently localises in the cell cortex distally to the cell division plane in asymmetrically dividing cells. POLAR loss of function mutants do not show clear phenotypes, implying that the gene function might be redundant. *POLAR* expression seems to be enhanced by the presence of functional BASL proteins, pointing to possible interaction between the two. Potential additional players are the proteins belonging to the BREVIS RADIX (BRX) family that have been shown to co-localise with BASL during asymmetric cell division in meristemoid differentiation [69]. BRX proteins have been widely studied for their role in cell proliferation and elongation in the root [74], as well as a component integrating auxin and brassinosteroid hormone signalling [75]. Quadruple mutants of *BRX* genes show an increase in cell divisions in the stomatal lineage and no physical asymmetry in the produced cells, similarly to *basl* mutant plants [69]. Moreover, it has been shown that BASL and BRX interact with each other and share the same cellular localisation, suggesting that they are both part of a core polarity generating system (1.11 B).

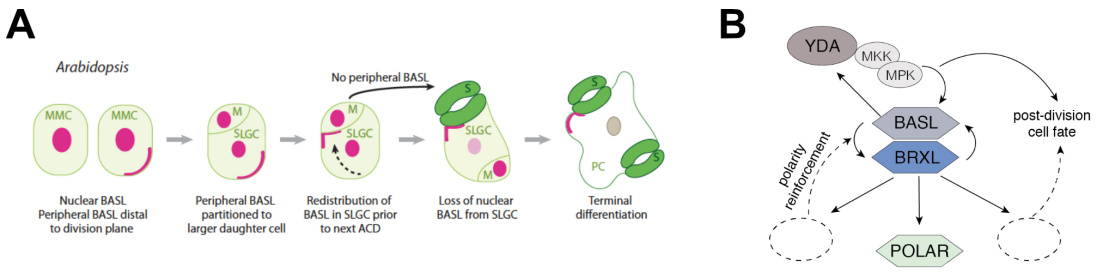


Figure 1.11 **Establishment of cell polarity in stomata asymmetric cell division in *Arabidopsis thaliana*.**
Adapted from [51] and [69]. A) The expression of *BASL* (in magenta) highlights its cellular re-localisation to promote the asymmetric cell division leading to the formation of a meristemoid. B) Diagram describing how the interaction between BASL and BRXL promotes asymmetric cell division.

Table 1.2 List of genes required for stomatal patterning in *Arabidopsis thaliana*.

Gene name	Class	Affected process	Reference
ERECTA	LRR-RLK	stomata lineage initiation	[55]
TMM	LRR-RLK	stomata lineage initiation	[50]
EPF1	cysteine rich peptide	stomata lineage initiation	[53]
EPF2	cysteine rich peptide	stomata lineage initiation	[52]
CHAL	cysteine rich peptide	stomata lineage initiation	[76]
STOMATOGEN	cysteine rich peptide	stomata lineage initiations	[62]
SPCH	TF, bHLH	MMC transition	[56]
MUTE	TF, bHLH	GMC transition	[57]
FAMA	TF, bHLH	GC transition	[58]
SCRM/2	TF, bHLH	MMC to GMC to GC transition	[76]
FOUR LIPS	TF, R2R3 MYB	GMC division	[59]
MYB88	TF, R2R3 MYB	GMC division	[59]
BASL	polar protein	meristemoid asymmetric division	[67]
POLAR	polar protein	meristemoid asymmetric division	[77]
BRX	polar protein	meristemoid asymmetric division	[69]

1.4 Epidermal patterning in *Marchantia*

After the colonisation of land, a series of different morphologies of unicellular and multicellular structures evolved on the plant outer layer (the epidermis) and they mediated the establishment of interactions with the environment.
Bryophytes, despite their structural simplicity when compared to vascular plants, show a variety of epidermal structures, such as rhizoids, gemma cups and air chambers, whose development have been shown to be controlled by the same gene families regulating epidermal patterning in vascular plants. These findings suggest that some of the regulatory modules present in the gametophytes of early divergent land plants have been recruited by the sporophyte of land plants.

The ability to study these processes in fast growing organisms, with low genetic redundancy, makes early divergent land plants the optimal model to study the key regulators of patterning. *Marchantia*, as described in Section 1.1, is one of these and its vegetative propagules, gemmae, show distinct epidermal structures, which start to appear once they are released from gemma cups: rhizoids develop on the ventral side and air chambers and gemma cups on the dorsal side.

Rhizoids are filamentous cells, that resemble in their function the root hairs of vascular plants [78, 79] [53]. Rhizoid development is controlled by ROOT-HAIR DEFECTIVE SIX-LIKE (RSL) class I of bHLH transcription factor proteins, which positively regulate the protrusion of single cells from the epidermal surface of *Marchantia*. Mutant thalli, which have lost *RSL* gene function are not able to produce rhizoids, as well as other specialised cell-types: slime papillae, mucilage papillae or gemmae (producing empty gemma cups) [11]. The same class of genes regulates the development of axillary hairs in the moss model plant *Physcomitrella patens*, and homologous genes in *Arabidopsis*, *AtRHD6* and *AtRSL1*, positively regulate root hair development [79]. The Arabidopsis *rhd6 rsl-1* double mutant, in fact, could be rescued by the complementation with the *Marchantia RSL1* gene, demonstrating the conservation of the *RSL1* gene family function throughout the evolution of land plants [11].

A series of regulators of gemma cup patterning have also been identified in *Marchantia*. The R2-R3 MYB GEMMA CUP-associated MYB 1 (*MpGCAM1*) transcription factor regulates the initiation of gemma cups [80]. Loss of function mutants of

MpGCAM1 are unable to produce either gemma cups or gemmae, while its overexpression leads to the formation of undifferentiated masses of tissue on the thallus surface. This result suggests a role for *MpGCAM1* in maintaining cell pluripotency to produce cups and develop gemmae. The Arabidopsis homologue, indeed, *REGULATOR OF AXILLARY MERISTEM 3*, is involved in the formation of axillary meristems. *Marchantia GCAM1* could complement the *Arabidopsis* triple mutant genotype only after the deletion of 28 N-terminal amino acids upstream the R2-R3 domain, which are not present in angiosperms. This supports the hypothesis that *GCAM1* might be involved in the maintenance of cells in an undifferentiated state. A ROP guanine nucleotide exchange factor (RopGEF) has also been shown to be required for gemma development in *Marchantia* [81]. The loss of function mutant of this gene, called *KARAPPO* (*KAR*) from the Japanese word for "empty", produces empty gemma cups. *Marchantia* has a single ROP GTPase gene, *MpROP*, and the interaction with *KAR* was confirmed *in vitro* and *in vivo*. The expression of *KAR* and *MpROP* was not restricted to gemma cups but it could be observed on the entire thallus surface, suggesting a role for ROP GTPase in thallus growth and development, supported as well by the impaired development of *Mprop* mutant plants [81]. RopGTPase proteins have been shown to have a role in the control of polar cell growth in root hairs in *Arabidopsis*, as well as regulating the asymmetrical division of stomata mother cells in monocots [82, 83, 84].

Table 1.3 List of genes required for epidermal patterning in *Marchantia polymorpha*.

Gene name	**Class**	**Affected process**	**Reference**
Mp*RSL*	TF, bHLH	rhizoids, gemmae, mucillage papillae	[11]
Mp*GCAM1*	TF, MYB	gemma cups	[80]
Mp*KAR*	guanine nucleotide exchanging factor	gemmae	[81]
Mp*ROP*	GTP binding protein	entire thallus	[81]
Mp*NOP*	E3 ubiquitin ligase	air chambers	[85]
Mp*WIP*	TF, zinc finger	air pore, rhizoids	[86]

1.5 Air chamber biology

Air chambers are epidermal features covering the entire dorsal epidermis of *Marchantia* [87]. These structures are present in all members of the *Marchantiales* order, including organisms living in water, such as *Riccia natans*[88]. In most cases, air chambers are localised on the epidermal surface, but they can also develop in different sub epidermal layers (e.g. *Riccia natans*, *Plagiochasma sp*) [88]. Air chambers are a typical trait of liverworts. Hornworts' gametophyte has cavities located at the end of small pores called mucilage clefts. They are filled with mucilage and they are colonised by cyanobacteria, mainly belonging to the *Nostoc* genus, hence their name, *Nostoc* cavities [89]. On their sporophyte instead, similarly to mosses, hornworts present structures that are similar to the stomata of tracheophytes. The stomata cavities are also filled with liquid and they are used to favour spores dispersal [16]. The role of air chambers, instead, has not been clarified, yet. Some studies have shown that they might favour the interaction with microbial organisms, such as pathogens [90], and there is fossil evidence from the Late Devonian of liverworts tissue associated with fungi and cyanobacteria [91]. Furthermore, mutants lacking air chamber are still able to develop viable thalli [85].

Marchantia air chambers open to the outer environment through a pore, usually made of 16 cells organised in 4 rings stacked one above the other. The pore leads to a chamber where filamentous chloroplast-rich cells, called photosynthetic filaments, protrude from the floor (Figure 1.12).

At the beginning of the 20^{th} century, Barnes and Land formulated a new theory on air chambers' formation based on the observations collected by several researchers during the 19^{th} century and comparing them with their results to document the formation of intercellular spaces in different liverwort species, including *Marchantia polymorpha* [88]. Barnes and Land refused the previous explanation of air chamber development in *Marchantia* proposed by Leitgeb [92], according to which air chamber generated by overgrowth of adjacent cells over a depression present on the thallus surface close to the apical notch. Barnes and Land documented the early splitting of cells in close proximity of the apical cell, which reaches the surface of the thallus. They also postulated the existence of a "mother cell of the air chamber", stating that the splitting takes place at the centre of cells derived from the same segment of the apical cell (segments are the cells obtained during the first divisions of the apical cell). All the observations were made using anatomical sections and light microscopy.

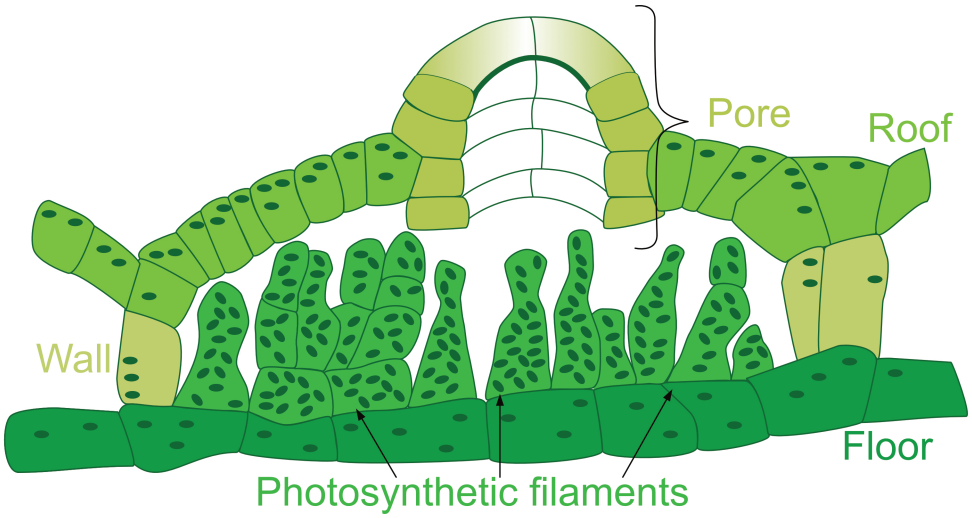


Figure 1.12 **Anatomy of an air chamber.**
Schematic cross section showing the different features of air chambers, which are identified with labels and colours. The model is based on a microscopy cross section of an air chamber. Cell sizes are proportional to their real size, but not the number of photosynthetic filaments.

In the 80s, Apostolakos described the origin of air chambers in *Marchantia paleacea* using electron microscopy in combination with anatomical sections and light microscopy, and he proposed the current theory on air chamber development [93, 94]. Apostolakos excluded the existence of a "mother cell of the air chamber" as air chambers are formed from cells derived from different segments of the apical cell. Moreover, the initial splitting appears to be exogenous (from the outside to the inside) and it appears at the cell junction between 3-4 cells derived from segments S_1, S_3, S_4 of the apical cell. This result was confirmed by clonal analysis experiments performed by Suzuki *et al.* [95]. He also denied the presence of a mother cell of the pore, as it is formed by periclinal divisions of cells at the junction where the initial space is formed.

Apostolakos' model suggests that at first an intercellular opening is formed by the splitting of adjacent anticlinal walls at a four-way cell junction. The initial aperture is firstly closed by the overgrowth of cells on the dorsal layer, which will give origin to the pore and the roof. Asymmetric divisions of the cells around the initial aperture define four mother cells of the pore (PMCs), which are arranged in a cruciform pattern (Figure 1.13). Meanwhile, in the sub-epidermal layer, walls are separated by the dividing and expanding floor cells, from which photosynthetic filaments will protrude.

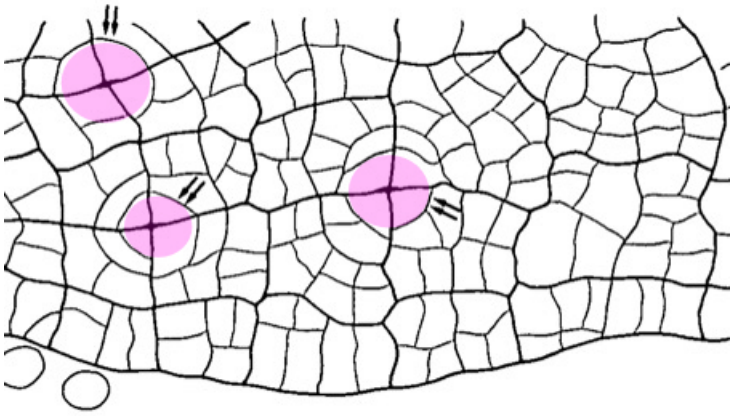


Figure 1.13 **Cruciform structures formed at early stages of air chambers development.**
Adapted from Apostolakos *et al.* [93]. The double arrows point to the cruciform structures (highlighted with pink circles) formed after asymmetric cell divisions.

Despite their analogous physiological function in exchanging gases, air chambers appear not to be related to stomata. Liverworts, indeed, are the only bryophytes to have lost or not possess the Ia subfamily of bHLH transcription factors [9, 10], which includes the three key regulators of cell fate transition in stomata development in *Arabidopsis*: *SPCH, MUTE* and *FAMA*. The bHLH subfamily Ia has been shown to regulate stomata development on the sporophyte of the moss *Physcomitrella patens*. *Physcomitrella* has two Ia bHLH homologues, *PpSFM1* and *PpSFM2* and one IIIb bHLH, *PpSCRM1*, which is a homologue of *Arabidopsis SCRM/2* genes. *PpSFM1* and *PpSCRM1* coordinate the correct division of the two guard cells, allowing the desiccation of the sporophyte to favour spore dispersal [10, 96]. The hornwort *Anthoceros punctatus*, likewise, has homologues of SFM and SCRM bHLH proteins and it develops stomata on its sporophyte, similar to the *Physcomitrella*. There is no experimental evidence for these genes being involved in stomata development in *Anthoceros* and no phylogenetic analysis showing evidence of the presence of the Ia family in other bryophyte families other than mosses, but the high degree of conservation suggests that they might have a similar function in hornworts [97].

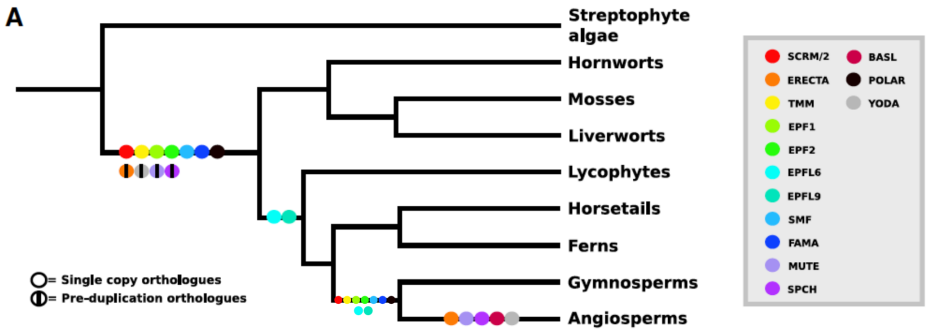


Figure 1.14 **Phylogeny tree showing the presence of orthologs genes involved in stomata function and development.**
Adapted from [12]. On the basis of detailed phylogenetic analysis, it has been hypothesised that the common ancestor of embryophytes possessed complex stomata and the core toolbox of genes necessary for stomata development in angiosperms. Bryophytes have lost some of the genes by secondary reduction, while several genes of tracheophytes have further diverged via duplication events that led to the set of genes known from *Arabidopsis*.

A current theory, based on phylogenetic analysis of genes involved in stomata development and function, suggests that a common toolkit of genes for stomata development was present in the ancestral embryophyte (Figure 1.14), making its stomata more similar to those of tracheophytes than the air pores of bryophytes [12]. Bryophytes appear to have a reduced number of genes compared to the ones included in the ancestral toolkit, hence their stomata evolved by secondary reduction, through the loss of some genes regulating their development and function. In this scenario, liverworts' air pores are considered different for their origin and function from those of stomata. The presented evidence that the ancestor of liverworts did have aerial epidermal structures more similar to tracheophyte stomata than air pores, suggests that selective pressure might have played a role in the appearance of air chambers and air pores in liverworts with the aim to favour gas exchanges and/or interactions with symbionts [12]. Nevertheless, the current knowledge is based on genes regulating stomata in *Arabidopsis* and a deeper understanding of the mechanisms controlling stomata or similar structures in other plant species could help explain the phylogeny of different aerial epidermal structures. It could also be possible that, the same mechanism is present and carries out different functions in early land plants and it has been recruited for stomata differentiation at a later time in plant evolution.

In the past 5 years, two air chamber mutants have been identified, showing that correct patterning of air chambers is not essential for *Marchantia* survival in laboratory

conditions. Ishizaki *et al.* identified an E3 ubiquitin ligase mutant, *NOPPERABO1* (Mp*NOP1*) [85], which completely lacks air chambers (Figure 1.15 A). Mp*nop1* mutant plants are still able to grow and progress through their entire life cycle and they do not seem to show any other defect. Following the discovery of the Mp*nop1* mutant, a new model for air chamber development has been proposed [98], suggesting a role for Mp*NOP1* in the degradation of factors involved in the inhibition of cell wall modifications to allow the formation of the initial splitting at four-ways cell junctions. This model suggests the possibility of spatial signalling components, such as the ones responsible for the correct spacing of stomata (Section 1.2.2), being in place to ensure that separation occurs only in the presence of the cruciform structure [85]. Nevertheless, it is worth underling, that correct spacing of air chamber could still take place, even if the splitting of the walls does not occurs at four-ways cell junctions.

Jones *et al.* found a mutant defective in pore formation. These plants have been generated using a miRNA targeting the 3'untranslated region (3'UTR) or the coding sequence (CDS) of the *WIP* gene, a zinc-finger transcription factor [86]. These plants are still able to produce the chamber, but no pore is formed, leaving larger apertures on the roof and affecting the shape of the chamber itself (Figure 1.15 B), suggesting that Mp*WIP* could have a role in the morphogenesis of pores of air chambers. The overexpression of Mp*WIP*, instead, causes ectopic development of rhizoids on the dorsal surface of the thallus. Mp*WIP* expression appears to be regulated, as well, by auxin, suggesting a broader role of this gene in controlling epidermal patterning.

The bHLH and MYB families of transcription factors regulate many epidermal patterning processes (see Section 1.2) in vascular plants, and their function is often shared in early divergent land plants, supporting the theory that they may have common origins. Rhizoid development is regulated by a homolog bHLH involved in root hair development, while a cell pluripotent state can be maintained thanks to a R2-R3 MYB, whose homologue regulates axillary meristems in vascular plants (see Section 1.3) [80]. No transcription factors regulating air chamber development have been identified yet, other than Mp*WIP* involvement in pore development. The identification of genes involved in air chamber morphogenesis could also shed light on the origin of these structures.

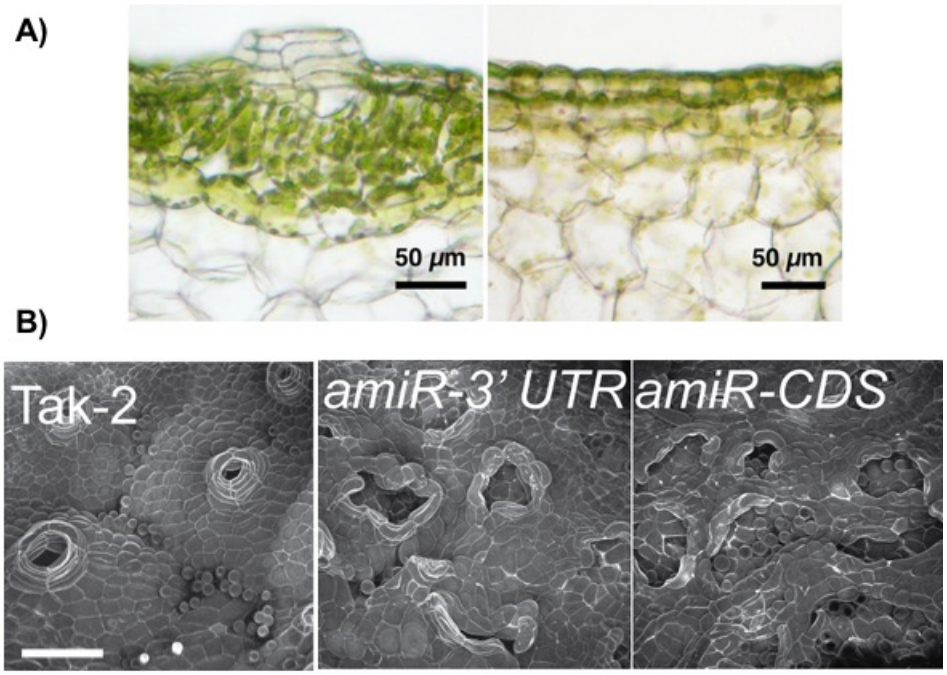


Figure 1.15 **Published *Marchantia polymorpha* mutants for air chamber development.**
Adapted from [85] and [86]. A) On the left a section of a wild-type plant air chamber. On the right, the Mp*nop1* mutant shows an even and continuous layer of photosynthetic cells.
B) Image from Jones *et al.*. Wild type Tak-2 plants' dorsal layer is compared with two knock-down Mp*wip* mutants obtained targeting the 3'UTR (amiR-3'UTR) and CDS (amiR-CDS) of the gene with miRNAs.

1.6 *Marchantia polymorpha* as a simple model for epidermal patterning

Marchantia polymorpha, as already described in Section 1.1, thanks to its ability to clonally reproduce through gemmae, disc-like clonal propagules, which develop in an open fashion, represents an ideal model for epidermal patterning. Moreover, the nuclear genome of *Marchantia* is small when compared to other land plant models: the total size of the nuclear genome is approximately 226 Mb encoding for 19,421 protein-coding loci [99]. The first complete assembly of the nuclear genome of *Marchantia* was released in 2017 [100] and a newly curated version is now available [99]. Mitochondria and chloroplast genomes have also been published, initially from *Marchantia paleacea* [101, 102], but now an assembly of *Marchantia* chloroplast genome has been released [3]. All the 47 transcription factor families found in the *Marchantia* genome can also be found in other land plants, but the number of representatives for each family is indicative of the low gene redundancy [100].
This makes *Marchantia* an ideal system to try to unravel the complex and often redundant genetic circuits regulating cell patterning. In addition, a whole set of molecular biology tools have been developed in the past few years for easy manipulation of *Marchantia*. There are protocols for efficient *Agrobacterium tumefaciens* mediated transformation of thalli and sporelings [103, 104], biolistic procedures for chloroplast transformation [105, 3], vectors to produce CRISPR/Cas9 mutants [106] or induce conditional gene expression [107]. An entire collection of *Marchantia* DNA parts is now available as part of the OpenPlant toolkit [3] and these can be used to explore this new model system with a synthetic biology approach.

The simplicity of the biological system and the broad range of available tools are encouraging the use of *Marchantia* to answer complex developmental questions. The degree of conservation of genetic networks regulating epidermal patterning amongst vascular and early divergent plants, as well as the possibility of observing a fast developing flat organism with a restricted number of genes, makes *Marchantia* gemmae the ideal platform to rigorously document and dissect epidermal patterning mechanisms.

1.7 PhD objectives

Marchantia is a plant system of unparalleled simplicity in regard to its genetic and cellular system. Its open mode of development and the access to clonal propagules allow direct observation of live developmental processes. Epidermal development is characterised by cellular patterning and differentiation and *Marchantia* shows specific cell types distinct from the ones found on the epidermis of late diverging sporophytic plants, but they still share common mechanisms, which may be ancient. *Marchantia* air chambers are an example of self-organising morphological units that have a modular form and a repetitive arrangement on the dorsal epidermis (in a similar way to meristemoids). The cellular scale and genetic details of air chambers are poorly resolved. Hence, I aim to apply modern genetic and microscopy tools to better map the process of air pore morphogenesis at the cellular scale and generate genetic markers in order to understand the steps involved in the process. The application of similar tools in *Arabidopsis* has shown that such patterning and differentiation processes can be broken into discrete steps that provide an insight into mechanisms and tools for genetic engineering.

1. **Systematic structural description of air chamber development.** The study of developmental processes requires an accurate description of key stages, which can be identified in a unique and objective way. High resolution microscopy techniques were used to identify stages in the development of air chambers, to determine characteristic landmarks and provide quantitative characterisation of the different developmental stages.

2. **Development of tools for imaging plant tissue without affecting the three dimensional structure.** Air chambers are three-dimensional epidermal structures that extend below the dorsal epidermal cell layer. Clearing techniques can help to access cell layers below the plant surface and they can be employed in combination with nuclear fluorescent markers as well as chemical stains. This allows the simultaneous observation of fluorescent markers and tissue architecture without altering the three-dimensional environment.

3. **Creation of genetic resources for the identification of potential candidate genes.** The generation of a transcriptomic dataset comparing the gene expression of wild type and Mp*nop1* mutant plants, unable to produce air chambers, at key developmental stages can help to identify genes, which

might be involved in air chamber patterning and differentiation. This can be used to guide the choice of candidate genes for the generation of cell-type specific markers.

4. **Generating a collection of cell specific markers for air chamber development.** The identification of markers, which can be visualised in live tissue, provides a valuable tool to study the dynamics of developmental processes. A library of transcription factor promoters that can be fused to genes encoding fluorescent proteins will be generated in a systematic way and used to identify markers specific for air chamber specific cell-types. These markers will be used to refine the model for cellular development and provide candidates for air chambers development.

Chapter 2

Material and Methods

2.1 Molecular biology

2.1.1 *Escherichia coli* transformation

Chemically competent *Escherichia coli* cells, TOP10 strain (Thermo Fisher Scientific), were taken out of the -80℃ freezer and thawed on ice. Plasmid DNA was diluted to a concentration of 50-100 ng/μl and 1 μl was mixed in 10 μl to 50 μl of cells. The cell-plasmid mix was left on ice for 10 minutes, heat shocked in a water bath at 42℃ for 45 seconds and then moved back on ice for 5 minutes to recover. A volume of 250 μl of liquid LB media was added to the cells and they were left in a shaking incubator at 37℃, 200rpm for 2 hours. An aliquot of 100 μl was plated on Petri dishes containing LB media, Miller's Broth Base (Invitrogen), with 1 % w/v agar (Melford capsules, cat no A20021) and the appropriate antibiotic selection (see table 2.2 for concentrations). Plates were left overnight in a 37℃ incubator.

Table 2.1 List of plasmid backbones and antibiotic resistance.

Backbone name	**Antibiotic resistance**	**Working concentration**
pUAP4	Chloramphenicol	50 μg/μl
pCk 1 to 4	Kanamycin	50 μg/μl
pCs A/B/C/E	Spectinomycin	100 μg/μl

2.1.2 *E.coli* growth

In order to recover plasmids from *E.coli*, bacteria were grown overnight in 14 ml round-bottom Falcon tubes (Thermo Fisher Scientific) filled with 3 ml of liquid LB,

Miller's Broth Base (Invitrogen), with the addition of the appropriate antibiotic selection (see Table 2.1 for reference). Tubes were incubated overnight at 37℃ on a shaker set at 200 rpm.

2.1.3 Loop assembly

Loop assembly is a DNA assembly strategy based on the recursive use of two type IIS restriction enzymes: *BsaI* and *SapI* [108]. This protocol can be used to join single DNA parts into a transcriptional unit. Multiple transcriptional units can then be combined together and transformed into an organism of interest. All Loop assembly backbones contain an expression cassette encoding the α subunit of the β-galactosidase gene (here referred to as LacZ cassette), which is excised when the desired fragments have been inserted. In the presence of the β-galactosidase substrate X-gal (5-bromo-4-chloro-3-indolyl-β-D-halactopyranoside), *E.coli* cells expressing the α subunit of β-galactosidase produce a blue precipitate. If the cassette has instead been replaced with the desired insert, *E.coli* colonies appear white.

Level 0 parts design and assembly

Level 0 (L0) parts consist of DNA parts, which were either obtained by chemical synthesis from Genewiz® or cloned from genomic DNA of *Marchantia polymorpha*, Cambridge accession Cam-1. L0 sequences were designed in accordance with the scheme shown in Figure 2.1 A.

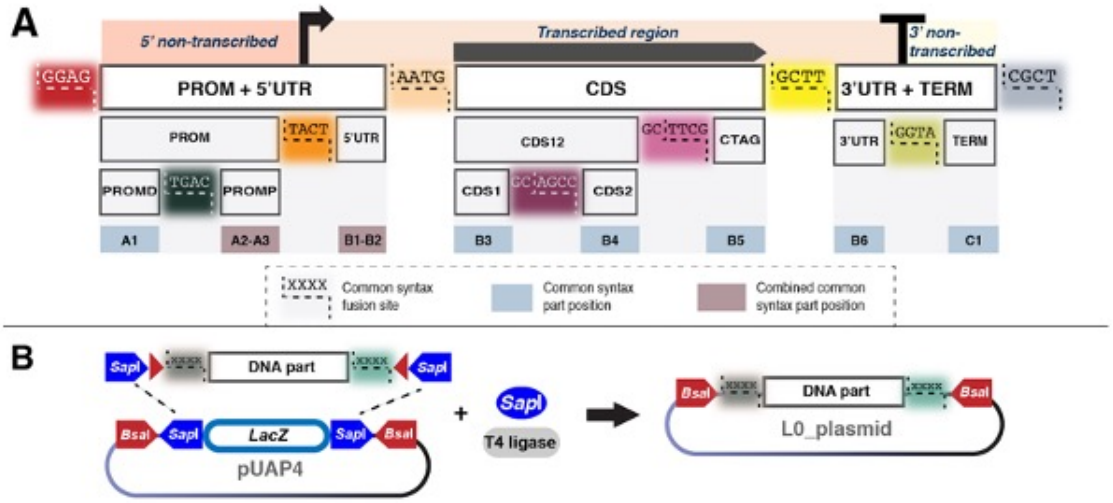


Figure 2.1 **Diagram of L0 parts design and assembly.**
Adapted from [3]. Figure A) shows the overhangs for DNA parts included in the OpenPlant toolkit and compares them with the Phytobrick [109] annotation for reference. Figure B) describes the mechanism through which L0 parts are inserted into the pUAP4 backbone [3].

Each part contains specific 5' and 3' overhang sequences next to a *SapI* recognition site, which allows their cloning into the pUAP4 backbone [3]. L0 parts obtained by chemical synthesis were already cloned into the pUAP4 backbone, while parts cloned from genomic DNA were amplified using polymerase chain reaction (PCR) with primers already containing the overhang sequence. PCR products were loaded on 1% agarose (Sigma Aldrich) w/v gels and run in an electrophoresis machine. Fragments with the expected size were extracted from the gel and purified using the QIAquick Gel Extraction Kit (QIAGEN). Purified fragments were prepared in aliquots with a concentration of 15 nM and the pUAP4 acceptor backbone was diluted to 7.5nM. In a 0.2 ml tube, the following reagents were added to 1 μl of PCR product 15 nM and 1 μl of pUAP4 7.5 nM: 5 μl of nuclease-free H_2O, 1 μl of 10x Tango Buffer (Thermo Fisher Scientific), 1 μl of 10mM ATP (Sigma Aldrich), 0.5 μl of 1 mg/ml bovine serum albumin (BSA, NEB), 0.25 μl *SapI* 5U/μl (Thermo Fisher Scientific), 0.25 μl T4 DNA ligase 5U/μl (Thermo Fisher Scientific). The tube content was mixed and incubated in a thermal cycler with the program shown in Table 2.2.

One microlitre of the reaction was then transformed into chemically competent *E.coli* cells, as described in 2.1.1 and plated on LB agar plates containing 50 μg/μl chloramphenicol and 40 μg/ml X-gal.

Table 2.2 Loop assembly thermal cycler programme.

Num of cycles	Step	Temperature	Time
x26	Digestion	37℃	3 min
	Ligation	16℃	4 min
x1	Termination	50℃	5 min
x1	Inactivation	80℃	10 min

Level 1 plasmids assembly

Level 1 (L1) reactions were used to assemble transcriptional units. The protocol described below has been published in [3]. The backbone plasmids used in L1 reactions are pCambia vectors resistant to kanamycin [110]. There are four L1 backbones called pCk1, pCk2, pCk3 and pCk4, which allow for the assembly of transcriptional units into level 2 (L2) plasmids in an ordered sequence. To assemble L1 plasmids, the required L0 parts were prepared at a concentration of 15 nM and the desired backbone diluted to 7.5 nM. One microlitre of each 15 nM DNA part was mixed with 1 μl of the desired backbone 7.5 nM, 1 μl of 10x T4 DNA ligase buffer (NEB), 0.5 μl of 1 mg/mL bovine serum albumin (BSA, NEB), 0.25 μl of T4 DNA ligase 400U/μl (NEB), 0.25 μl *BsaI* 10U/μl (NEB) and nuclease free water to reach a final volume of 10 μl. L1 reactions were incubated in the thermal cycler using the same program described in Table 2.2. One microlitre of the reaction product was transformed into *E.coli* cells according to section 2.1.1 and the reaction was then plated onto LB 1% w/v agar plates containing 50 μg/μl kanamycin and 40 μg/ml X-gal.

Level 2 assembly

Level 2 (L2) plasmids combine together multiple transcriptional units and they were used to transform plant tissue. The backbones used for L2 vectors are pCambia [110] with a spectinomycin resistance cassette. The four different construct pCsA, pCsB, pCsC and pCsE can be used to locate fragments in specific positions in higher level constructs (Level 3). Each L1 part to be assembled was diluted to 15 nM, while the desired pCs backbone was diluted to 7.5 nM. One microlitre of all the diluted transcriptional units and backbone were mixed with 1 μl of 10x Tango Buffer (Thermo Fisher Scientific), 0.5 μl of 1 mg/mL bovine serum albumin (BSA, NEB), 0.25 μl T4 DNA Ligase 5U/μl (Thermo Fisher Scientific), 1 μl of 10 mM ATP (Sigma Aldrich), 0.25 μl *SapI* 5U/μl (Thermo Fisher Scientific) and 2 μl of nuclease-

free water. The reaction was then incubated in the thermocycler using the program described in Table 2.2. One microlitre of the reaction product was transformed into *E.coli* cells according to section 2.1.1 and reactions were then plated onto LB 1% w/v agar plates containing 100 μg/μl spectinomycin and 40 μg/ml X-gal.

2.1.4 gRNA design

Single guide RNA (sgRNA) target sequences were predicted using the online software CRISPOR [111]. Once the genomic region of interest was uploaded (<2000 bp), sequences of 20 nt flanking an NGG protospacer adjacent motif (PAM), which is recognised by *Streptococcus pyogenes* Cas9, were selected. The chosen sequences were targeting gene's exons, had a specificity score=100 and a predicted efficiency>60. The selected gRNA targeting sequences did not have off-target sites with less than 4 mismatches. For each gRNA target sequence, a forward and a reverse primers were synthesised. They both contained the 20 nucleotide sequence and specific 5' and 3' overhangs for cloning in the Loop acceptors L1_lacZgRNA-Ck2, L1_lacZgRNA-Ck3 or L2_lacZgRNA-Cas9-CsA [3], see Table 2.3 for overhang sequences.

Table 2.3 Overhang sequences for gRNA targeting sequences cloned into Loop acceptor plasmids.

Overhangs are shown in capital letters.

Acceptor plasmid	**Primer**	**Sequence 5'- 3'**
L1_lacZgRNA-Ck2	FW	CTCG-$(n)_{20}$-GT
L1_lacZgRNA-Ck3	RV	TAAAAC-$(n)_{20}$
L2_lacZgRNA-Cas9-CsA	FW	CTCG-$(n)_{20}$-GT
	RV	AAAAC-$(n)_{20}$

2.1.5 gRNAs annealing and cloning

The forward and reverse oligos were diluted to 100 μM and 1 μl of the forward and 1 μl of the reverse oligos were mixed with 8 μl of nuclease-free water. The reaction was then incubated in a thermal cycler with the following programme: 30 min at 37℃, 5 min at 95℃ followed by a ramp down to 25℃ at a rate of 5℃ per minute.
The annealed oligos were cloned into L1_lacZgRNA-Ck2 or L1_lacZgRNA-Ck3 mixing 1 μl of annealed oligos with 1 μl of 50 ng/μl of the acceptor plasmid, 6 μl of nuclease free water, 1 μl of T4 ligase buffer (NEB), 0.5 μl of 1 mg/ml bovine serum albumin (BSA, NEB), 0.25 μl of *BbsI* 5U/μl (Thermo Fisher Scientific) and

0.25 μl of T4 DNA ligase 400U/μl (NEB). The reaction mix was then incubated in a thermal cycler with the programme described in Table 2.4.

Table 2.4 Thermal cycler programme for gRNAs oligos cloning in Loop acceptor vectors.

Num of cycles	Step	Temperature	Time
x15	Digestion	37℃	3 min
	Ligation	16℃	4 min
x1	Termination	50℃	5 min
x1	Inactivation	80℃	10 min

One microlitre of the reaction product was transformed into *E.coli* cells according to section 2.1.1 and, later, the samples were plated onto LB 1% w/v agar plates containing 50 μg/μl kanamycin and 40 μg/ml X-gal.
When oligos were cloned into the L2_lacZgRNA-Cas9-CsA acceptor, 1 μl of annealed oligos were mixed with 1 μl of 50 ng/μl of the acceptor plasmid, 5 μl of nuclease free water, 1 μl of 10x Tango buffer (Thermo Fisher Scientific), 1 μl of 10 mM ATP (Sigma Aldrich) ,0.5 μl of 1 mg/ml bovine serum albumin (BSA, NEB), 0.25 μl of *SapI* 5U/μl (Thermo Fisher Scientific) and 0.25 μl of T4 DNA ligase 5U/μl (Thermo Fisher Scientific). The reaction mix was then incubated in the thermal cycler with the programme described in Table 2.4. *E.coli* cells were transformed with one microlitre of the reaction following the procedure described in 2.1.1 and they were plated on LB 1% w/v agar plates containing 100 μg/μl spectinomycin and 40 μg/ml X-gal.

2.1.6 Polymerase Chain Reaction (PCR)

Polymerase chain reactions were performed using two different polymerases and following the manufacturer instructions. For DNA amplification with the Thermo Fisher Scientific Phusion™ High-Fidelity DNA polymerase, 10 ng/μl of plasmid DNA were used as template or 50-200 ng/μl of genomic DNA. The used thermal cycler programme is indicated in Table 2.5.

Alternatively, KOD Hot Start DNA polymerase (Merck cat. no. 71086) was used, following the manufacturer's directions and using 100 ng/μl of genomic DNA as template and 4.5 μl of $MgSO_4$ 25 mM. The used thermal cycler programme is indicated in Table 2.6.

Table 2.5 Thermal cycler programme for PCR with Phusion™ High-Fidelity DNA polymerase.

Num of cycles	Step	Temperature	Time
x1	Denaturation	98℃	30 s
x35	Denaturation	98℃	10 s
	Annealing	55-60℃	10 s
	Extension	72℃	15-30 s/kb
x1	Final extension	72℃	10 min
x1	Hold	4℃	∞

Table 2.6 Thermal cycler programme for PCR with KOD DNA polymerase.

Num of cycles	Step	Temperature	Time
x1	Activation	95℃	2 min
x35	Denaturation	95℃	20 s
	Annealing	55-60℃	10 s
	Extension	70℃	10-25 s/kb
x1	Final extension	70℃	10 min
x1	Hold	4℃	∞

2.1.7 Gel electrophoresis

PCR reaction products were mixed with 6x Gel Loading Dye Purple (NEB, cat no B7024S) and loaded on 1% w/v agarose (Sigma Aldrich) gel prepared in TRIS-acetate-EDTA (TAE) and stained with SYBR™ Safe DNA gel stain (Invitrogen cat no S33102). The Bioline HyperLadder™ 1kb or 100 bp markers were used as molecular weight references depending on the size of the PCR product of interest. Loaded gels were run on an electrophoresis machine set at 90V for 30-60 min and PCR products were visualised in a gel imaging machine (SYNGENE).

2.1.8 Minipreps

Plasmid DNA extraction from *E.coli* was performed using the QIAprep Spin Miniprep Kit (QIAGEN) following the manufacturer's instructions. Single colonies were grown overnight in 3ml of liquid LB media containing the appropriate antibiotic selection as described in Section 2.1.2. The bacterial cultures were pelleted at 13 000 rpm on a table top centrifuge and DNA was extracted following the QIAprep Spin Miniprep protocol. Automated plasmid DNA extraction was instead performed loading the pelleted bacterial culture into the QIAcube platform (QIAGEN) and running the

built-in plasmid DNA purification protocol without buffer PE. All plasmids were eluted in 30 μl of nuclease free water and concentration was verified using a spectrophotometer.

2.1.9 Sanger sequencing

Plasmid DNA and purified PCR products were sent for sequencing to the Genewiz® sequencing facility following their sample submission guidelines. A list of the sequencing primers used in this dissertation is shown in Appendix .1.

2.2 Plant methods

2.2.1 *Marchantia* plant propagation and maintenance

Marchantia plants were grown onto Petri dishes filled with half strength Gamborg B5 media with vitamins (Duchefa, cat no G0210) and containing 1.2% w/v of agar (Melford, cat no A20021) with pH=5.8. Plates were then grown at 21℃ under continuous light (150 mol/m/s). Plants were propagated plating gemmae or fragments of the plant thallus onto new plates. For long term storage (2-3 months), plants were grown under continuous light (150 mol/m/s) at 14℃.

2.2.2 Genomic DNA extraction

The protocol described here allows the extraction of DNA from plant tissue with high content of polysaccharides and polyphenols. It is based on the cetyltrimethylammonium bromide (CTAB) DNA extraction protocol published in [112, 113]. The extraction buffer was prepared as indicated in Table 2.7 and it was incubated at 60℃ before usage.

Table 2.7 CTAB DNA extraction buffer.

Reagent	**Required concentration**	**Supplier**
Tris-HCl	100 mM, pH=8	Melford
NaCl	1.4 M	Thermo Fisher Scientific
EDTA	20 mM	Thermo Fisher Scientific
CTAB	2% w/v	Sigma Aldrich
β-mercaptoethanol	0.3% v/v	Merck

Marchantia gemmalings were grown for two weeks and 1 g of tissue was collected and flash frozen in liquid nitrogen. The frozen samples were ground into a fine powder with mortar and pestle, and then transferred to 30 ml Falcon tubes containing 10 ml of pre-warmed extraction buffer, with the addition of 100 mg of PVP-40 (Sigma Aldrich) per gram of tissue and 5 μl of RNase A 100 mg/ml (QIAGEN). The plant tissue and the extraction buffer were mixed by inversion and tubes were incubated at 60 ℃ for 30 min. In each tube, 12 ml of a 24:1 v/v solution of chloroform (Thermo Fisher Scientific) and 3-indol acetic acid (Sigma Aldrich), respectively, were added and then samples were spun at 7 000 rpm for 10 min at room temperature. After centrifugation, the aqueous phase was transferred to new 50 ml Falcon tubes and half volume of 5M NaCl (Sigma Aldrich) was added to the samples and mixed. Two volumes of cold (-20℃) 95% ethanol (Thermo Fisher Scientific) were added to the tubes and they were incubated at -20℃ for 10 min. Samples were spun at 13 000 rpm for 15 min and the supernatant was discarded. The obtained pellet, containing DNA, was washed with cold (0-4℃) ethanol (Thermo Fisher Scientific) 70% v/v. The resuspended pellet in ethanol was transferred to 1.5 ml sterile Eppendorf tubes cutting the end of a 1000 μl pipette's tip. Tubes were then spun down at 13 000 rpm for 5 min and the supernatant discarded again, they were then left open to allow the pellet to dry for 5 min. The pellet was finally dissolved in a 100 μl of TE buffer (EDTA pH=8, 1 mM; Tris-HCl, pH=8, 10mM). To facilitate pellet resuspension, tubes can be left at 4℃ over night.

2.2.3 Fast DNA extraction protocol for genotyping

For routine genotyping experiments, small squares of thalli (3 mm x3 mm) were crushed with a sterile micro-pestle into an 1.5 Eppendorf tube containing 100 μl of the extraction buffer described in Table 2.8.

Table 2.8 DNA extraction buffer for genotyping.

Reagent	**Required concentration**	**Supplier**
Tris-HCl	100 mM	Melford
KCl	1 M	Thermo Fisher Scientific
EDTA	10 mM, pH=9.5	Thermo Fisher Scientific

The ground tissue was incubated at 80℃ for 10 min and then 400 μl of sterile water were added to each tube. One microlitre of the mixture containing the extracted DNA was used for PCR amplification.

2.2.4 RNA extraction

A 100 mg of *Marchantia* tissue was collected in RNase free 2 ml Eppendorf tubes containing two sterile metal beads and it was flash frozen in liquid nitrogen. The tissue was disrupted using the TissueLyser II (QIAGEN) at 30 Hz for 30-60 s. After grinding, 1 ml of Trizol (Ambion) was added to the samples and they were incubated at room temperature for at least 5 minutes. For each millilitre of Trizol, 200 μl of chloroform (Thermo Fisher Scientific) were added to the sample and the tubes were mixed for 20 second on a vortex. Samples were left at room temperature for 2-3 min and then spun down in a centrifuge set at 4℃ for 18 min at 10 000 xg. The aqueous phase was recovered and moved to clean RNase free Eppendorf tubes and an equal volume of 100% RNase free ethanol (Thermo Fisher Scientific) was added and mixed to the samples. The mix was loaded into RNeasy columns from the QIAGEN RNeasy Plant Mini Kit (QIAGEN, cat no 74903) and RNA was washed and eluted following steps 5 to 9 of the kit's Quick-Start protocol.

2.2.5 Spores sterilisation and germination

Sterile *Marchantia* spores were prepared accordingly to the published protocol [3] and stored at -80℃. For each plasmid to transform, two sporangia were placed into a 1.5 ml Eppendorf tube and they were crushed with sterile tweezers into 500 μl of 0.05% w/v Milton solution (Milton®, Mini Sterilising Tablets, 1 tablet for 25 ml of sterile water) to obtain a fine pulp. Crushed sporangia were filtered through a 40 μm cell strainer on a sterile Falcon tube. The emptied tubes were washed a second time with 1ml of 0.05%w/v Milton solution and filtered through the cell strainer to recover residual spores. The filter was washed with 2.5 ml of 0.05% w/v Milton solution and filtered spores were aliquoted into four new 1.5 Eppendrof tubes containing the same volume of samples. The tubes were left at room temperature for 10 min and then spun at 16 000 xg for 2 min on a table top centrifuge. The supernatant was discarded and pelleted spores were resuspended in 150 μl of 0.05% w/v Milton solution. Spores were spread onto half strength Gamborg B5 with 1.2% w/v agar plates, sealed with micropore tape and placed upside down at 21℃ under continuous light (150 mol/m/s) for five days. After 5 days, sporelings were ready to be transformed.

2.2.6 *Agrobacterium tumefaciens* growth

Agrobacterium tumefaciens cells from the GV2260 strain [114] were transformed with the pSoup helper plasmid [115] and grown on LB with 1.2% w/v agar plates containing 50 μg/μl of carbenicillin, 10 μg/μl of rifampicin and 5 μg/μl of tetracyclin.

2.2.7 *A.tumefaciens* transformation

Aliquots of 50 μl of electrocompetent *A.tumefaciens* cells (GV2260 strain transformed with the pSoup helper plasmid) were thawed on ice for 10 min and 50 ng/μl of the plasmid of interest were added and mixed gently into the tube. The mix was loaded into pre-chilled electroporation cuvettes with a 2 mm gap size (VWR, cat no 732-1136) and shocked at 2.5 kV with an electroporator. One millilitre of liquid LB media was added to the cuvettes and they were incubated for 2 h in a shaker set at 28℃ and 120rpm, allowing bacteria to recover. A 100 μl aliquot of transformed cells was spread on LB plates containing 50 μg/μl of carbenicillin, 10 μg/μl of rifampicin, 5 μg/μl of tetracyclin and an antibiotic specific to the transformed plasmid (in this manuscript 100 μg/ml of spectinomycin).

2.2.8 Spores transformation

A.tumefaciens cells, transformed with the plasmid of interest, were inoculated in a 50 ml Falcon tube containing 5 ml of liquid LB media with the appropriate antibiotic selection (50 μg/μl of carbenicillin, 10 μg/μl of rifampicin, 5 μg/μl of tetracyclin and the plasmid specific antibiotic selection) and incubated for two days on a shaker set at 28℃ and 150 rpm. This step is usually performed three days after spores sterilisation, so that *A.tumefaciens* is ready to be co-cultivated with spores when these have been growing for 5 days as described in 2.2.5. After two days, *A.tumefaciens* cultures were spun down at 3 000 xg for 15 min. The supernatant was discared and the pellet was resuspended in 3 ml of liquid half strength Gamborg B5 media plus supplements (see Table 2.9 for composition), with the addition of acetosyringone to reach a final concentration of 100 μM.

Five days old sporelings were collected using a sterile L-shaped spreader onto a 50 ml falcon tube and resuspended in half strength Gamborg B5 with supplements solution (1 ml per used wells).

Three millilitres of half strength Gamborg B5 with supplements were added to the wells of a 6-wells plate (one well per transformed plasmid, Costar®, Corning cat

Table 2.9 Half strength Gamborg B5 with supplements, pH=5.8.

Reagent	Required concentration	Supplier
Gamborg B5 plus vitamin	3.16 g/L	Duchefa cat no 60210
N-Z amino acids	0.1 % w/v	Sigma Aldrich cat no C7290
L-Glutamine	0.03% w/v	Thermo Fisher Scientific cat no BP379
Sucrose	2% w/v	Thermo Fisher Scientific

no 3516) and then 1 ml of the spores' suspension was added to them. Eighty μl of the induced *A.tumefaciens* cultures were added to the corresponding wells with 4 μl of acetosyringone 100mM. The 6-wells plate (Costar®, Corning cat no 3516) was sealed with microtape and placed for two days on a shaker set at 21℃ and 120 rpm under continuous light (150 fmol/m/s).

After two days, sporelings were removed from the wells and filtered through 70 μm strainers (one per each used well) placed on a 50 ml Falcon tube and rinsed with 50 ml of sterile water supplemented with 100 μg/ml of cefotaxime to kill residual *A.tumefaciens*. The spores left inside the strainer were washed multiple times and then plated onto half strength Gamborg B5 with 1.2% w/v agar plates containing 100 μg/ml cefotaxime and the appropriate plant antibiotic selection (in this manuscript 20 μg/ml hygromycin). Successfully transformed sporelings were moved to new selection plates after 2 weeks and they were maintained until they produced gemmae from gemma cups emerged in successfully transformed region of the thallus. The above mentioned gemmae (giving rise to so called G1 thalli) were then plated onto new selection plates containing 100 μg/ml cefotaxime and the appropriate plant antibiotic selection, to grow isogenic mutant lines.

2.3 Microscopy

2.3.1 *Marchantia* tissue clearing in chloral hydrate, staining with propidium iodide and mounting in Hoyer's solution

Marchantia tissue was collected in a cell culture plate and submerged with a fixative solution containing 50% v/v methanol and 10 % v/v acetic acid. Samples were

incubated overnight in the fixative solution at 4℃ and subsequently washed twice with water. For *Marchantia* gemmae older than 0 day, waxes were removed washing them twice in each of the following aqueous solutions containing increasing amounts of ethanol (EtOH): 40% v/v EtOH, 60% v/v EtOH and 80% v/v EtOH. While incubating in the 80% v/v EtOH solution, samples were heated at 80℃ in a water bath for 2-10 min, depending on the size of the sample. The tissue was subsequently rehydrated with a series of washes in a decreasing gradient of ethanol solutions. Samples were washed twice in each of the following solutions: 60% v/v EtOH, 40% v/v EtOH, 20% v/v EtOH and then left in water for 10 minutes. *Marchantia* tissue was treated with α-amylase to remove starch. Ten millilitres of α-amylase buffer (see Table 2.10 for composition) were mixed with 15U/ml of α-amylase (Sigma Aldrich, cat no A4551). Samples were covered with the α-amylase solution and incubated overnight at 37℃.

Table 2.10 α-amylase buffer solution.

Reagent	**Required concentration**	**Supplier**
NaCl	2mM	Sigma Aldrich
CaCl	0.25 mM	Thermo Fisher Scientific
PB buffer pH=7 see table 2.11	20 mM	

Table 2.11 Phosphate buffer (PB) 0.1M composition for 1L, pH=7.

Reagent	**Required concentration**	**Supplier**
Na_2HPO_4-$7H_2O$	0.05 M	Finson
$NaH_2PO_4H_2O$	0.04 M	Sigma Aldrich
HCl or NaOH	adjust pH to 7	Thermo Fisher Scientific

All the following steps were performed under the fume hood. After α-amylase treatment, samples were washed three times with water and then moved into 1% w/v periodic acid solution. Samples were incubated in the solution for 40 min to 1h at room temperature and then they were washed twice with water. After washing, Schiff's reagent solution (see table 2.12 for composition) with the addition of 10 μg/ml of propidium iodide (Sigma Aldrich, cat no P7875), was used to cover the samples, which were left at room temperature for 1-2 hours or until they turned pink in colour.

The Schiff's reagent solution was removed, samples were washed twice with water and moved onto microscopy slides. A drop of chloral hydrate clearing solution (see

Table 2.12 Schiff's reagent solution.

Reagent	Required concentration	Supplier
Sodium metabisulfite	0.33 M	Thermo Fisher Scientific
HCl 5M	10% v/v	Thermo Fisher Scientific
Water	as required	

Table 2.13 for composition) was added to cover the samples on glass slides, which were incubated overnight at room temperature inside closed and sealed dishes to prevent evaporation. In the morning, excess clearing solution was removed from the slide.

Table 2.13 Chloral hydrate clearing solution.

Reagent	Required concentration	Supplier
Chloral hydrate	4 M	Sigma Aldrich
Glycerol	6% v/v	Sigma Aldrich

A Gene Frame® (Thermo Fisher Scientific, cat no AB0557) was placed on the glass slides containing the samples to avoid squashing the tissue. The mounting Hoyer's solution (see table 2.14 for composition) was used to cover the plant tissue and size 0 cover slips were stuck to the Gene Frame. The slides were left in the dark on a flat surface for 2 days to a week to let the mounting solution dry.

Table 2.14 Hoyer's solution.

Reagent	Required amount	Supplier
Chlroal Hydrate	400 g	Sigma Aldrich
Gum Arabic	60 g	Fluka
Glycerol	40 ml	Sigma Aldrich
Water	100 ml	

The Hoyer's solution needs to be stored in the dark. Avoid the formation of bubbles and let the solution settle, allowing bubbles to disappear before usage.

2.3.2 ClearSee clearing protocol

The fixative solution was prepared with 4% w/v paraformaldehyde (Sigma Aldrich) dissolved in phosphate-buffered saline solution, PBS (see table 2.15 for composition), and stored in 20 ml aliquots for single use. *Marchantia* tissue was placed in a cell

culture plate, submerged with the fixative solution and incubated over night at 4℃. In the morning, samples were washed twice with 1% v/v PBS.

Table 2.15 Phosphate-buffered saline solution (PBS), pH=7.4.

Reagent	**Required concentration**	**Supplier**
NaCl	0.137 M	Sigma Aldrich
KCl	2.7 mM	Termo Scientific
Na_2HPO_4	10 mM	Finson
KH_2PO_4	1.8 mM	Thermo Fisher Scientific

Table 2.16 ClearSee clearing solution.

Reagent	**Required concentration**	**Supplier**
Urea	25% w/v	Thermo Fisher Scientific
Sodium deoxycholate	15% w/v	Sigma Aldrich
Xylitol	10% w/v	Sigma Aldrich

The ClearSee clearing solution was prepared in accordance with the published protocol [116] and the solution components are described in table 2.16. *Marchantia* gemmalings were submerged with the clearing solution and incubated on a shaker at 120 rpm at room temperature for 3 days. The clearing solutions was changed every day. After clearing, plant tissue was stained with Calcofluor White M2R (Sigma Aldrich, cat no 18909) diluted in ClearSee clearing solution to reach a concentration of 0.01% w/v. Samples were left in the staining solution overnight, incubating at room temperature on a shaker set at 120rpm. In the morning the staining solution was removed, samples were washed with ClearSee solution and mounted on a glass slide with the addition of Gene Frames® (Thermo Fisher Scientific, cat no cat no AB0557). The ClearSee clearing solution was used as mounting media, using enough volume to fill the Gene Frame® and cover the sample. Slides were sealed with size 0 cover slips.

2.3.3 Epifluorescence microscopy

Marchantia plants were screened inside microtape sealed Petri dishes under a Leica M205 FA fluorescence stereo microscope. The microscope was equipped with the filters indicated in table 2.17 and in this dissertation filters for the observation of mTurquoise, eGFP, mVenus and chlorophyll fluorescent signal were used.

Table 2.17 List of Leica M205 FA filters.

Filter name	Excitation	Emission
ET CFP	426-446 nm	460-500 nm
ET GFP	450-490 nm	500-550 nm
ET YFP	490-510 nm	520-550 nm
ET Chlorophyll LP	460-500 nm	610 nm LP

2.3.4 Confocal microscopy

Sample preparation for live tissue imaging

A Gene Frame® (Thermo Fisher Scientific, cat no AB0577) was placed on a microscope slide and a thin layer of half strength Gamborg B5 media with 1.2 % w/v agar was spread inside it. One to four gemmae or gemmalings were placed onto the media with an inoculation loop and a drop of sterile water was added before placing a size 0 cover slip to seal the slide. Alternatively, the plant tissue was directly placed between the microscope slide and the cover slip with the addition of a drop of water.

Microscope imaging settings

Images were acquired using a Leica SP8 Spectral Confocal Microscope upright system equipped with a super continuum white light laser (WLL) and two diode lasers (405 nm and 442 nm). Most of the images acquired in this dissertation were taken with the following objectives:

- HC PL APO 10x 0.40 CS2 dry
- HC PL APO 20x/0.75 CS2 dry
- HC PL APO 40x/1.10 W CORR CS2 water
- HC PL APO 63x/1.20 W CORR CS2 water

The settings used to image fluorescent molecules are indicated on table 2.18.
If multi-spectral imaging was required and fluorophores could not be imaged at the same time the following sequential settings were used:

- Scan 1: Calcofluor
- Scan 2: mTurquoise2 + Chlorophyll

- Scan 1: mTurquoise2 + Chlorophyll
- Scan 2: eGFP+ Chlorophyll
- Scan 3: mVenus

Table 2.18 Leica SP8 imaging settings for fluorescent proteins and stains.

FP/Stain	**Laser Line**	**Excitation**	**Emission**
Calcofluor white	Diode	405 nm	420-440 nm
mTurquoise	Diode	442 nm	460-485 nm
mVenus	WLL	515 nm	520-540 nm
eGFP	WLL	488 nm	409-516 nm
Chlorophyll	Diode/WLL	442 nm o r448 nm	650-700 nm
Propidium Iodide	WLL	555 nm	600-630 nm

2.3.5 Multi-well plate imaging set up

A 384 wells plate with clear bottom was filled with half strength Gamborg B5 media with 1.2% w/v agar in each well used for imaging purpose. Air bubbles were removed with a pipette and extra rows and columns surrounding the imaged wells were filled with media to prevent evaporation. Once solidified, the surface of the plate was levelled with a sterile scalpel and rectangular Gene Frames ® (Thermo Fisher Scientific, cat no AB0578) were placed around the imaging area. Each well contained a single gemma placed at the centre. Cover slips were cleaned with 70% v/v ethanol and sprayed with anti-fog spray (B-Clean, Bollé safety cat no B250) to prevent condensation. The cover slip was placed with the anti-fog treated side facing down towards the sample. The position of each imaged well (up to a maximum of 106, due to the limitations of the microscope's stage travelling range) was recorded using the 'Mark and Find' option on the Leica Application Suite X software (LAS X). Z-stacks of 400 μm with 10μ step size were acquired for each position.

2.3.6 Pipeline for image processing on ImageJ

The LIF files produced during image acquisition were imported in the ImageJ software, Fiji distribution, for basic image processing [117, 118]. Maximum intensity projections were obtained for Z-stacks for each imaged channels and merged images were produced to show all channels in a single file.

For datasets produced during multi-well plate imaging a Macro was used to streamline the generation of maximum intensity projections for each imaged positions. The commands performed by the Macro are shown in Appendix B.

2.3.7 Measuring pore size and distance from the notch

The diameter of *Marchantia* pores was measured using the ImageJ measuring tool on maximum intensity projection of acquired images and tracing a straight line with the Line Selection Tool through the diameter of the pore complex and using the ImageJ Measure function[117]. The notch was identified as a single point at centre of the gemmalings notch area, in the lower XY plane acquired in the Z-stacks, in correspondence of the apical cell. A straight line was traced from the notch to the centre of the pore complex using the Line Selection Tool and the distance was calculated with the ImageJ Measure function [117].

2.4 RNAseq data analysis

RNA was extracted from *Marchantia* day 1, day 4 and day 8 gemmalings and it was sent to Genewiz® for polyA library preparation and single-index sequencing on an Illumina HiSeq platform.

2.4.1 Filtering and mapping

The obtained reads were filtered based on their quality using the FastQC software [119] and adpaters were removed using Trim Galore [120]. Obtained reads were mapped against the *Marchantia* Tak genome v5.1 [99] and transcripts' abundance was retrieved using Kallisto [121].

2.4.2 Counts normalisation

Transcripts counts were imported in R using the *tximport* package [122] and they were normalised using the *vst* function, which perform a variance stabilising transformation for data visualisation purposes and normalises counts to the library size, while removing dependence of the variance from the mean.

2.4.3 Identification of differentially expressed genes

The DEseq function [123] was applied to the raw counts and a table containing the log2 of genes expression fold change, p-values and adjusted p-values (p-adj) was returned, for each possible comparison. The p-adjusted values were calculated using the Benjamini & Hochberg method [124], which is also known as false discovery rate. All the plots were generated using the ggplot2 or the pheatmap packages [125, 126].

All the commands used for the analysis of the RNA-seq data in R are shown in Appendix A.

Chapter 3

Development of high-resolution microscopy techniques for *Marchantia polymorpha*

3.1 Introduction

The study of the morphogenesis of organs or entire organisms requires knowledge of cellular processes that take place in the three dimensional space and time. In order to obtain information about all three dimensions, there is a need to access the inner layers of the sample [127]. This can be achieved by cutting sequential thin sections from a sample to image them individually and reconstruct the three dimensional structure by overlapping the acquired images [128]. This approach requires much manual work and accuracy in aligning the images to reconstruct the sample volume. Alternatively, with the "blockface" strategy [129], the surface of the sample is imaged and then removed to reveal the layer underneath to image it. The process is repeated for the entire thickness of the sample. Both these approaches can lead to satisfactory results, but they are laborious and in the case of the blockface approach, the sample will be lost by the end of the imaging process.

Biological samples can also be imaged without the need for physical sections. Optical sectioning aims to collect the light from single focal planes, while removing any out of focus signal, which would affect the image contrast and resolution [130]. This principle is applied in confocal laser scanning microscopy, where the use of a pinhole allows the exclusion of out of focus light. Nowadays, this approach is also

used in two-photon microscopy, light sheet approaches and applied for other imaging modalities [131, 132, 133].

Even so, optical sectioning has its drawbacks when looking at thick biological samples due to the presence of pigments, the presence of fluorescent molecules and the translucency caused by light scattering [127]. Pigments can absorb part of the light used to image a structure of interest, reducing the amount of excitation light entering the tissue and, therefore, decreasing the amount of signal emitted by the sample. Autofluorescence is the background signal that could result from the excitation of intrinsically fluorescent molecules present in the sample or molecules introduced during sample preparation. Pigments and autofluorescent compounds can be removed using different strategies (for reference see [134, 135, 136]).

The different optical properties of the molecules present in a biological sample cause incident light to be scattered. In order to prevent light scattering, a series of clearing techniques have been developed with the aim of minimising the amount of light lost through lateral scattering, allowing most of the light to pass through the sample. Older techniques required one to dehydrate the samples and to embed them in hydrophobic rigid materials, such as resins. Many of these materials are intrinsically fluorescent, and the dehydration step quenches the signal from fluorescent proteins,the native structures of which require an aqueous environment [127].
The last decade has seen the development of techniques that are instead based on solutions that allow solvation of lipids, which cause the majority of light scattering in biological samples, and match the tissue refractive index (RI). The RI is a dimensionless number that describes how much slower light travels through a certain material compared to vacuum. When light travels through media with different RI, it changes its travel path and this leads to scattering. Hence, having a constant RI throughout the sample reduces the amount of light scattering. Aqueous-based clearing techniques use solutions with high RI to match the sample's one. They can preserve the emission of fluorescent protein and avoid changes in the tissue architecture due to the dehydration steps that take place in solvent-based clearing techniques.

More recent aqueous-based techniques added mild detergents to remove lipid molecules from the sample. Sca*l*e [137] and CUBIC [138] hyper hydrate the sample using urea-based aqueous solutions, which preserve protein structures and decrease the sample

RI. SeeDB [139] exploits the presence of sugars in the clearing media to adjust the sample RI and it avoids the use of detergent and denaturing reagents. Other aqueous-based techniques embed samples in hydrogel matrices to use stronger detergents to extract lipids and they are combined with the use of high RI mounting media, such as Focus Clear (CLARITY [140] and PEA-CLARITY [141], adapted for plant tissues).

Table 3.1 Comparison of different aqueous based clearing techniques.

	Sca*l*e[136]	**CUBIC** [138]	**SeeDB** [139]	**CLARITY**[140]	**ClearSee**[116]
Refractive Index	1.38	1.48	1.49	1.45	1.41
Clearing Time*	weeks	weeks	weeks	days/week 4-6 weeks in plants**	days 2 weeks in plants**
Tested in Plants	N	N	N	Y	Y

*Clearing Time is calculated for clearing a whole adult mouse brain, unless specified otherwise.
**tested in *Arabidopsis thaliana*.

The combination of the use of cleared samples with confocal microscopy optimised the quantity and quality of the collected signals, enabling researchers to reconstruct their sample in three-dimensions. However, all the above mentioned techniques require weeks to clear the sample (Table 3.1). Kurihara et al. developed a new clearing solution called ClearSee, which contains a mixture of urea, sugar (xylitol) and detergent (sodium deoxycholate) [116]. This mixture allows the plant tissue to become transparent in only a few days after immersion in the clearing solution and samples can be immediately imaged, without the need of additional mounting solutions, as the ClearSee solution has a RI of $\sim$1.4. This technique preserves nuclear fluorescent proteins and it is also compatible with chemical stains. The ClearSee protocol has other benefits: speed and avoiding the use of toxic chemical compounds, such as chloral hydrate.

Air chamber morphogenesis in *Marchantia polymorpha* gemmae starts in close proximity to the notch [93] at the interface between the dorsal epidermis and the subepidermal cell layers. A chamber is formed between the two layers and new cell types appear inside the chamber as well as on the dorsal epidermis. In order to optimally visualise all of them, *Marchantia* gemmae need to be cleared of pigments and fluorescent compounds, such as chlorophyll. Moreover, the chambers are filled with air, which contributes to light scattering and imaging of the deeper layers requires the immersion in a solution to match the sample refractive index. Moreover, the ability

to preserve nuclear localised protein fluorescence is necessary to visualise GFP cell-markers in sub-epidermal layers, which would be otherwise difficult to document in live microscopy.

ClearSee and thiodiethanol () based clearing protocols have been tested in *Arabidopsis thaliana* with successful results on different types of tissue [116, 142], preserving the signal of fluorescent proteins expressed in the nucleus. These two techniques were tested in *Marchantia* gemmalings. The TDE clearing protocol was discarded as it did not lead to successful tissue clearing, since the clearing solution contains a percentage of TDE, which does not allow to match the sample RI (97% v/v solution of TDE is required to reach an RI of 1.5 and it causes loss of fluorescent proteins signal).

In this chapter, I will describe how I tested and optimised the original ClearSee protocol to clear *Marchantia* gemmae at different developmental stages and combined it with chemical cell wall stains. My results were compared with those obtained for samples cleared with chloral hydrate, stained with propidium iodide (PI) and mounted on Hoyer's solution [17]. Gemmae removed from the gemma cup (0 days, d0) were used to test the ClearSee protocol, as well as 5 days (d5) gemmalings to ensure the entire air chamber structure could be accessed with optical sections after ClearSee treatment. *Marchantia* lines expressing constitutive nuclear markers were also tested to confirm the ability of the ClearSee solution to preserve fluorescent proteins signal.

3.2 Results

3.2.1 Testing the ClearSee protocol in *Marchantia*

Marchantia gemmae have been previously imaged in three dimensions by combining chloral hydrate clearing with periodic acid treatment to modify the cell wall carbohydrates and produce aldehydes, which will react with the PI stain [143]. Samples were then mounted in a chloral hydrate and gum arabic based solution called Hoyer's solution. The details of the protocol are described in Chapter 2, section 2.3.1, and it will be referred to as mPAS/PI technique. A new clearing protocol, called ClearSee, was tested on *Marchantia* gemmae of the Cam-1 accession, and the results were compared to those obtained with the mPAS/PI protocol. Gemmae were taken out of the gemma cup (d0) or grown on half strength Gamborg B5 media plates with 1.2% agar v/v for 5 days (d5) and then moved into the fixative solution containing 4% w/v paraformaldehyde (Chapter 2, section 2.3.2).

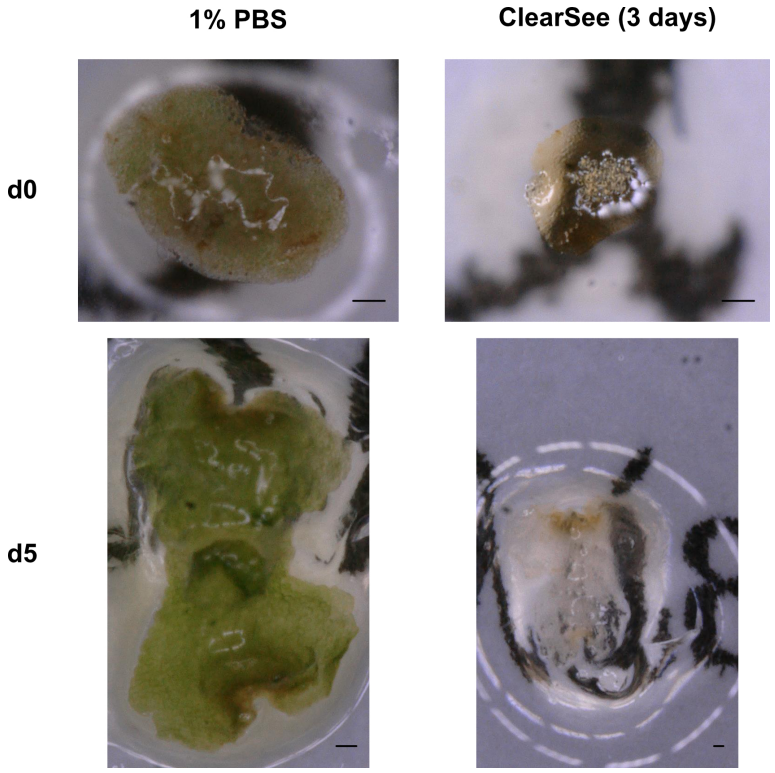


Figure 3.1 **ClearSee clears gemmae tissue in 72 hours.**
The d0 and d5 samples were cleared after being immersed in ClearSee solution for 3 days. Samples were laid onto paper printed text (Odyssey, Homer) to assess the tissue transparency. The first column displays control samples incubated for 3 days in 1% PBS: the letters underneath the gemmae cannot be read. The second column shows the cleared samples and the text can be seen through the plant tissue. Top row has d0 gemmae and the bottom one d5 gemmae. Scale bars are all 100 μm.

Day 0 gemmae were used for comparison with the mPAS/PI approach, while d5 gemmae were used to test if the protocol allows the imaging of the entire air chamber structure. After fixation, samples were incubated in the ClearSee solution for 3 days and the clearing media changed every 24h. The transparency of cleared samples was tested against gemmae incubated in 1% v/v phosphate saline buffer (PBS) for the same amount of time.

3.2.2 ClearSee is as efficient as mPAS/PI in clearing d0 gemmae

In order to assess the penetration depth reached by the clearing solution, ClearSee samples were stained with Calcofluor White M2R 0.01% w/v to allow the visualisation of the cell outlines. Cleared and stained ClearSee samples were compared with samples cleared and stained according to the mPAS/PI protocol. Live d0 gemmae

stained with 10 μM PI were used as a negative control for the mPAS/PI protocol and fixed, Calcofluor White M2R stained d0 gemmae as control for the ClearSee protocol. A schematic description of the two protocols is shown in Figure 3.2.

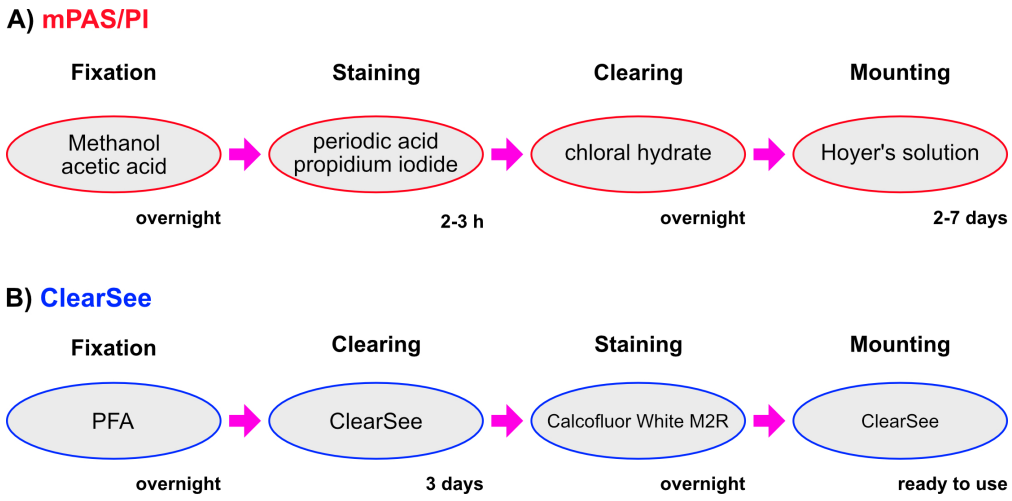


Figure 3.2 **Diagram showing the steps and timing of the mPAS/PI and ClearSee protocol.**
The ellipses outlined in red refer to the steps of the mPAS/PI protocol, while the one in blue are the ones used in the ClearSee clearing protocol. Inside the ellipses, the chemicals, or solutions used in each step is indicated. Underneath the ellipses, the duration of each step is indicated in bold. A) The mPAS/PI protocol requires periodic acid treatment to allow propidium iodide to bind the cell wall. The Hoyer's solution contains gum arabic and additional chloral hydrate to match the sample RI and it needs to dry before imaging. B) The ClearSee protocol requires less time and fewer steps. The mounting media is the same as the clearing media and samples can be imaged immediately.

Live d0 gemmae stained with PI (Figure 3.3 A) showed homogeneous staining on the surface, but only the first cell layer from the surface could be visualised in the XZ plane. A similar result was observed when d0 gemma fixed in 4% w/v paraformaldehyde (PFA) were stained with 0.01% w/v Calcofluor White M2R (Figure 3.3 B). Gemmae cleared with mPAS/PI and ClearSee allowed visualisation cell layers underneath the epidermis (Figure 3.3 C and D) and 5 cell layers, between the dorsal and ventral epidermis in the thickest part of the 0 day gemmae.

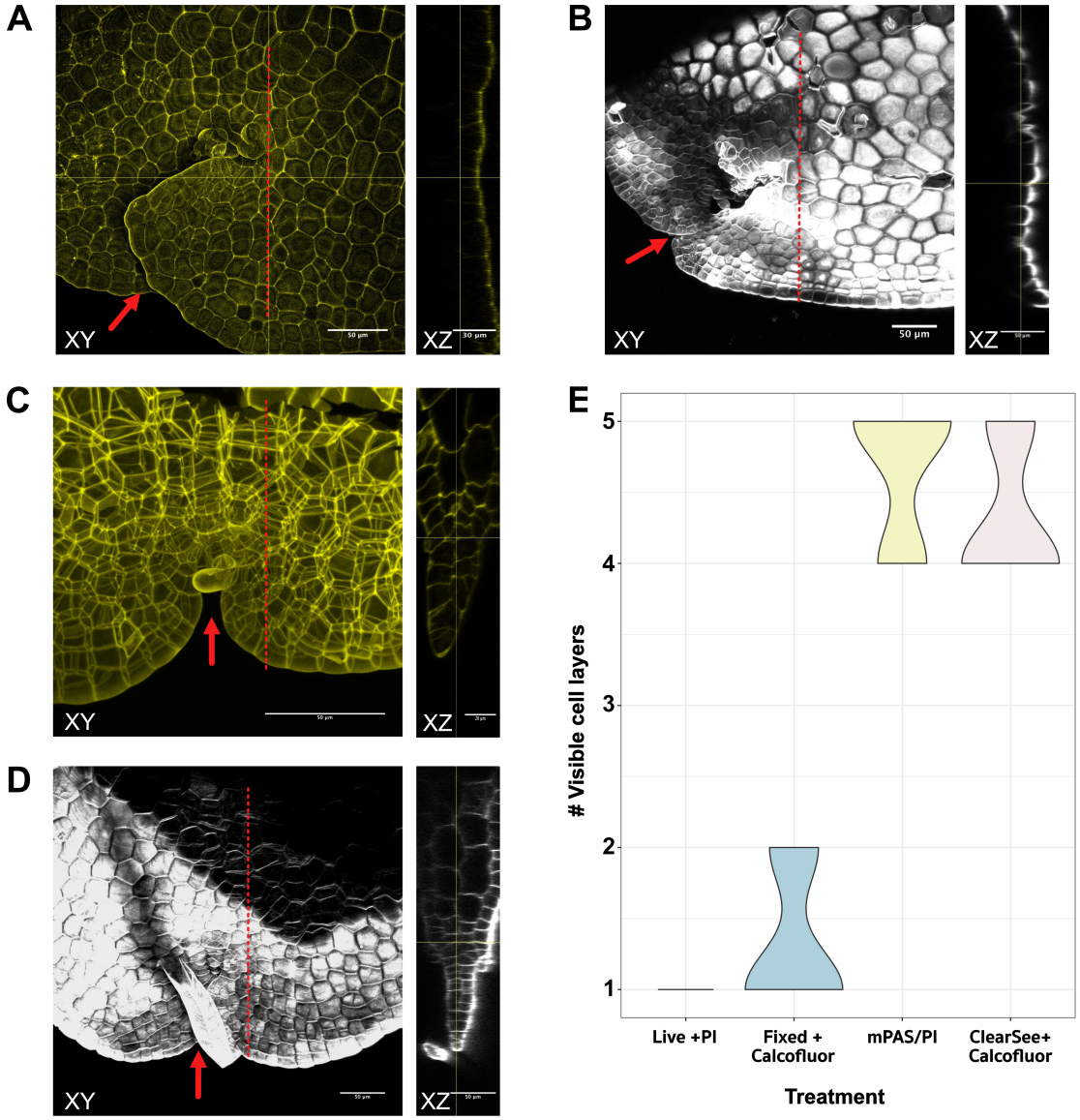


Figure 3.3 **ClearSee and mPAS/PI allow imaging of cells 5 layers deep in d0 gemmae.**
Cam-1 d0 gemmae were cleared and stained using the mPAS/PI and ClearSee protocol to compare the performance of the two techniques. A) Live tissue of a d0 gemma stained with PI 10 μM. B) d0 gemma fixed in 4% w/v PFA and stained with 0.01% w/v Calcofluor White M2R. C) d0 gemma cleared with mPAS/PI and stained with PI. D) d0 days gemma cleared with ClearSee and stained with 0.01% w/v Calcofluor White M2R. E) The violin plot shows the number of cell layers accessible for each treatment (n=5 per treatment). All images were taken with 40x objective immersed in water, while C was imaged with a 40x oil immersion objective. Red arrows point at notches and dotted lines indicate the XZ plane shown in the side sections.

3.2.3 ClearSee preserves nuclear localised CFP fluorescence

The ClearSee formulation has shown to preserve the signal of fluorescent proteins localised in the cell nuclei of Arabidopsis root and apical shoot [116]. Clearing solutions that contain detergents, indeed, cannot preserve membrane localised fluorescent proteins, as they get removed during lipid solvation. To test whether ClearSee preserve the signal of fluorescent proteins in *Marchantia*, a line transformed with a construct for the expression of the mTurquoise2 fluorescent protein fused to the nuclear localisation N7 signal peptide (from *Arabidopsis* ankyrin-like protein) [144], under the control of the *Marchantia* promoter *ELONGATION FACTOR-1α* (Mp*EF1α*) [145] was used. This *Marchantia* line was generated by a former member of the lab, Dr. Bernardo Pollak.

Gemmae were removed from the gemma cup and they were cleared following the ClearSee protocol (see section 2.3.2, Chapter 2) and the mPAS/PI one (see section 2.3.1, Chapter 2) for comparison. Samples cleared in the ClearSee solution were subsequently stained with Calcofluor White M2R 0.01% w/v. Fluorescent labelled nuclei could be observed throughout the entire depth of the optical sections when the tissue was cleared using the ClearSee solution (Fig 3.4 B and C), while no signal could be detected when gemmae were cleared with chloral hydrate following the mPAS/PI protocol (Fig 3.4 A). The signal through the YZ sections in Figure 3.3 became dimmer as a low magnification and low numerical aperture objective was used (20x) to capture the entire thickness of gemmae.

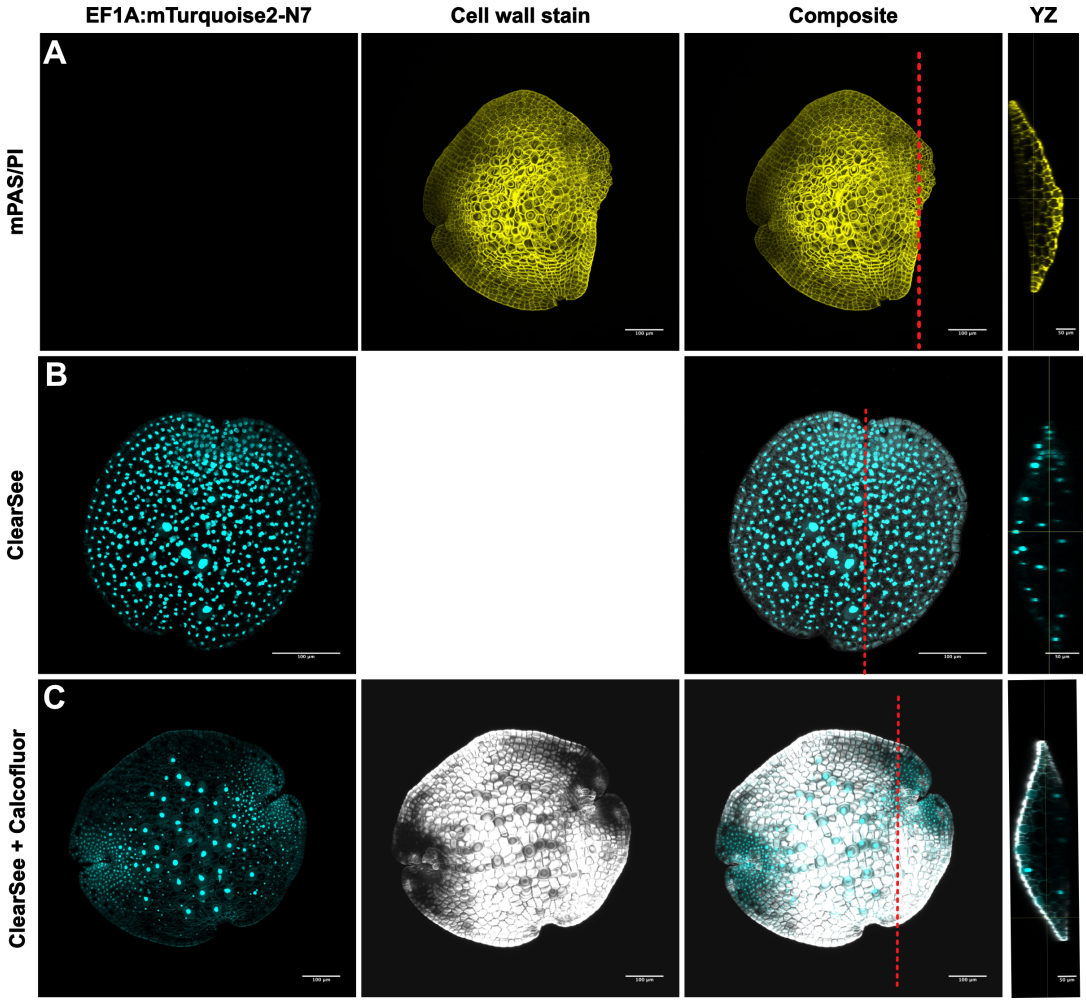


Figure 3.4 **ClearSee preserves nuclear localised mTurquoise2 fluorescence.** Cleared and stained d0 gemmae expressing mTurquoise2 in the nucleus under the control of the Mp*EF1α* promoter. Z plane projections of the collected z-stack images are shown (Stack = 150 μm, step size = 2.5 μm), as well as the orthogonal views of the YZ planes indicated by the red dotted line in the projections. A) Gemma cleared and stained following the mPAS/PI protocol. B) Gemma cleared with ClearSee. C) Gemma cleared with ClearSee and stained with 0.01% v/v Calcofluor White M2R.

3.2.4 ClearSee immersion for 3 days leads to best clearing results in d0 gemmae

In order to assess the timing of the clearing procedure, different incubation times in ClearSee solution were tested on d0 gemmae to identify the optimal timing to ensure access to all cell layers without losing the fluorescent protein signal. After fixation, gemmae from the line expressing mTurquoise2 in the nucleus under the control of the Mp*EF1α* gene's promoter, were incubated for 1 or 3 days in ClearSee solution. The prepared samples were then imaged under the Leica SP8 confocal microscope with a 20x air objective using the XZY function that captures 1 μm thick optical sections out of a 100 μm section in the XZ plane (Figure 3.5 A). Gemmae, which were fixed in 4% w/v PFA and incubated in 1% PBS v/v for 1 day, showed a single cell layer of fluorescent nuclei on the dorsal side of the gemmae (Figure 3.5 B). Samples that were fixed in 4% w/v PFA and incubated in ClearSee for 1 day, allowed to visualize nuclei belonging to two to three distinct cell layers. (Figure 3.5 C). PFA fixed gemmae incubated in ClearSee for 3 days showed CFP expression in nuclei belonging to the 5 distinct cell layers, which could be visualised in d0 gemmae.

Older gemmae (5 day) could also be successfully cleared in 3 days, allowing access to up to 5 cell layers. Clearing of older tissue (1 week thalli) requires up to a week and this could be easily evaluated by eye looking at the transparency of the sample (data not shown).

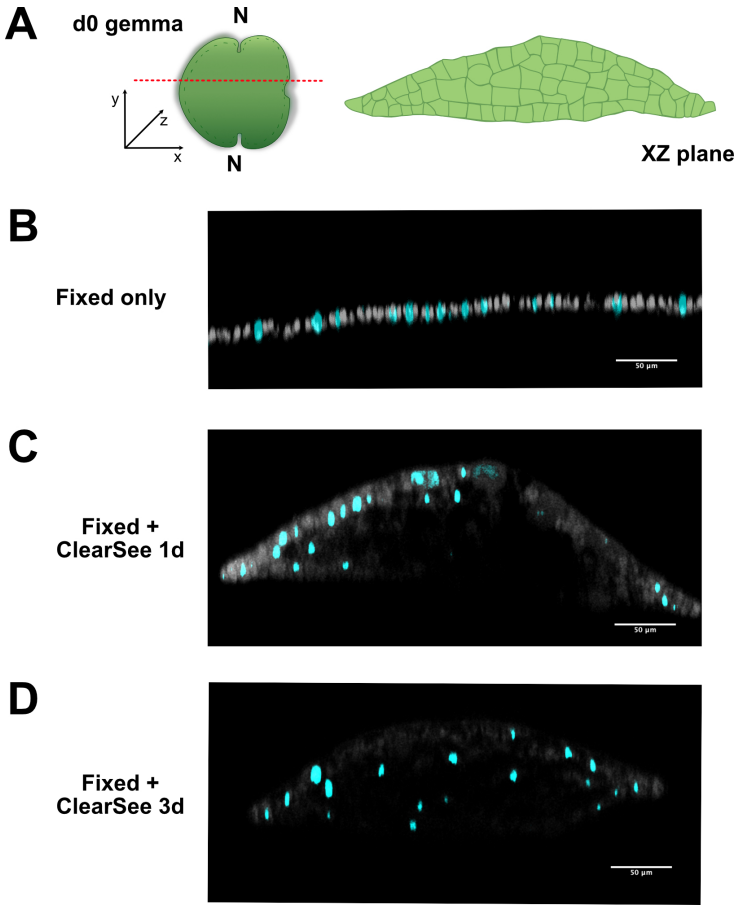


Figure 3.5 **Incubation in ClearSee for 3 days leads to optimal clearing of d0 gemmae.**
Marchantia d0 gemmae constitutively expressing mTurquoise2 in the cell nucleus under the control of the Mp*EF1α* promoter, were incubated in ClearSee solution for different timings to gauge the penetration of the clearing agent. One micrometre thick XZ planes of d0 gemmae were reconstructed using the XZY option on a Leica SP8 confocal microscope, using a 20x air objective. The fluorescent nuclei are shown in cyan and chlorophyll autofluorescence in grey. A) The diagram shows a schematic representation of the plane imaged on a d0 gemmae indicated by a red dotted line. The cutting plane is perpendicular to the gemma surface and it shows a single XZ plane. N indicates the gemma meristematic notches. The XZ plane shows a schematic view of a typical XZ section. B) Gemmae that were not cleared show a single layer of fluorescent nuclei. C) Gemmae that have been infiltrated for 1 day in ClearSee showed nuclei through to 2-3 cell layers deep. However, nuclei in the central region of the cross section could not be visualised. D) After three days in the ClearSee solution, d0 gemmae showed nuclei from all the cell layers present in the XZ plane.

3.2.5 Paraformaldehyde fixation in combination with ClearSee preserves cell shape

Before samples could be cleared to make them more transparent, it was important to fix them to preserve a 'life-like' state of the cell. Fixative solutions, indeed, stop enzymatic reactions and harden the sample, preventing its decay. Fixation and clearing protocols, though, can lead to shrinkage or expansion of the plant tissue [127]. Air chambers are large structures, which include large air spaces and tissue distortion could cause its collapse, affecting the visualisation of distinct cell types and their physical relationship. Hence, the mPAS/PI and the ClearSee clearing protocols were tested on 5 day old gemmalings (d5), which have fully developed air chambers on their dorsal surface.

Alcoholic fixation followed by dehydration and rehydration steps to remove waxes, pigments and chlorophylls, as well as treatments with low pH solutions, such as in the mPAS/PI protocol (Chapter 2 Section 2.3.1), can lead to tissue distortion. *Marchantia* d5 gemmae were fixed using either an alcohol based fixative followed by the mPAS/PI clearing protocol or with 4% w/v PFA in combination ClearSee clearing. Samples were mounted on slides with up to four gene frames stacked on top of each other to prevent damaging the sample and preserve its 3D structure and air chamber images were taken to assess their structural integrity.

The mPAS/PI technique cleared samples were fixed with 50% methanol and 10% acetic acid, and subsequently treated with an ascending and a descending gradient of ethanol solutions, and they show a lack of turgor in their cells (Figure 3.6 A). Moreover, the periodic acid treatment at low pH modifies the cell wall composition, affecting its stiffness, causing the cell wall to collapse. The air chamber pore shrunk (Figure 3.6 A) and the chamber roof collapsed (Figure 3.6 C) on top of the photosynthetic filaments emerging from the floor of the chamber. ClearSee, instead, retained the three dimensional structures of the air chambers, maintaining the spatial relationships between the different cell types. In Figure 3.6 D, the cells of the pore have preserved their turgor and the space between the roof and the photosynthetic filaments is maintained (Figure 3.6 F).

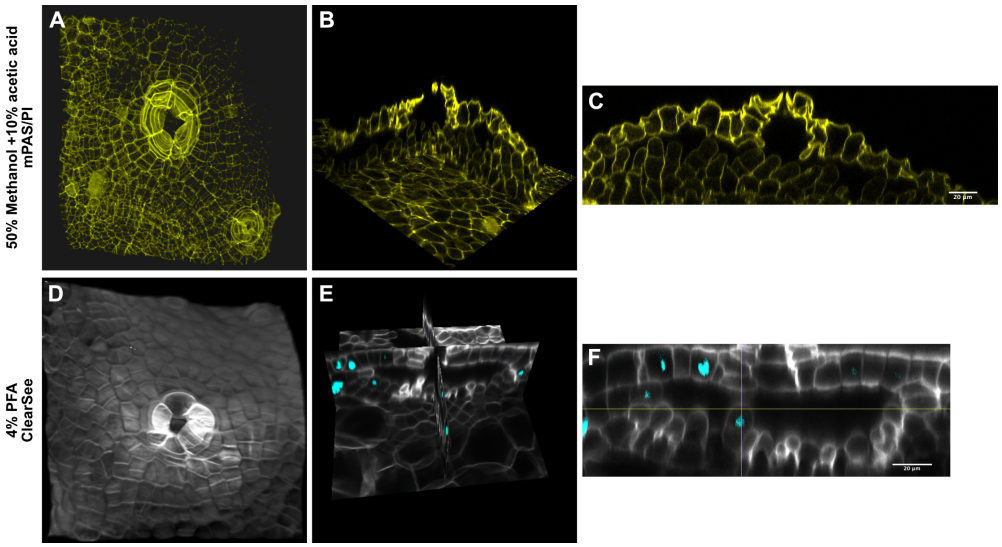


Figure 3.6 **Paraformaldehyde fixation and ClearSee clearing preserve the three dimensional structure of air chambers.**
Marchantia d5 gemmae were fixed and cleared following the mPAS/PI or the ClearSee protocol. Air chamber three dimensional structure is better preserved in ClearSee treated samples. A) Surface view of an air chamber fixed with 50% methanol and 10% acetic acid and cleared with mPAS/PI. The pore cells have shrunk and they look distorted. B) Orthogonal view of the air chamber, showing the sub-epidermal tissues. C) XZ plane of the chamber showing the collapsed roof of the chamber. D) Surface view of an air chamber fixed with 4% w/v PFA, cleared with ClearSee and stained with Calcofluor White M2R 0.01% w/v. E) Orthogonal view of the air chamber, showing the sub-epidermal tissues. F) XZ plane of the chamber showing that the three dimensional relationships between the epidermal and sub-epidermal layer are maintained.

3.3 Discussion

3.3.1 ClearSee is a safer and faster method to clear *Marchantia* samples

The visualisation of biological specimens in three dimensions is one of the main problems that developmental biologists face when studying morphogenesis [127]. Light scattering poses the main obstacle to access to deeper layers of a sample. Hence, a variety of techniques have been developed to reduce this phenomenon [146].

ClearSee is a solution containing urea, detergents and sugars, which allows the clearing of pigments and autofluorescent compounds present in the samples, as well as lipids, which are the main source of light scattering [127]. The presence of sugars in the clearing solutions allows the maintenance of a RI closer to the one of the cleared sample, hence further reducing light scattering. Moreover, ClearSee is an aqueous solution, which decreases the risk of osmotic shock and tissue shrinkage.

In this chapter, I compared the performance of ClearSee clearing protocol on *Marchantia* gemmae with that of the mPAS/PI protocol, which is based on chloral hydrate clearing and sample embedding in materials with high RI (chloral hydrate in the Hoyer's solution). The performance of the two techniques was comparable as they both allowed access to the same number of cell layers in d0 and d5 gemmae. The mPAS/PI protocol involved alcohol based fixative and a series of dehydration and hydration steps that led to tissue shrinkage, while ClearSee samples were fixed in a 4% PFA aqueous solution and did not include any dehydration steps, leading to better preservation of three dimensional structures. Furthermore, the mPAS/PI required the use of many hazardous chemicals (methanol, periodic acid, sodium metabisulfite, chloral hydrate, propidium iodide) and sample preparation can take up to a week in order to allow the sample to solidify (Hoyer's mounting solution). The ClearSee solution instead employs non-hazardous chemicals and samples are ready for imaging in less than a week and it can be combined with cell wall stains, such as Calcofluor White M2R, for the visualisation of the cell outline.

3.3.2 ClearSee protocol can be used on samples expressing fluorescent markers

Another advantage of the ClearSee solution compared to mPAS/PI, is the ability to preserve nuclear fluorescent proteins. This allows cellular markers to be visualised at cellular resolution, even if they are not expressed on outer cell layers. Membrane fluorescent markers are not preserved as the ClearSee solution contains detergents that wash out the membrane lipids and the embedded proteins. Nevertheless, chemical stains can be used in combination with the ClearSee clearing protocol to visualise the cell outline.

Amongst the acqueous based clearing techniques mentioned in this chapter, only PEA-CLARITY [141], TDE [142] and ClearSee [116] were tested in plants and all the available results where retrieved from testing these techniques on *Arabidopsis thaliana* tissues (lives, roots and shoots). No literature is currently available on the testing of the above mentioned techniques to clear bryophytes tissue, with the only exception of ClearSee, which was tested on the moss *Physchomitrella patens* gametophore [116]. All the microscopy data available for *Marchantia* is obtained imaging live tissue and three dimensional reconstructions are obtained using physical sections or electron microscopy. The protocols developed in this Chapter will allow for the visualisation of *Marchantia* tissue in three dimension, with no need to destroy the biological sample.

The ClearSee technique can be used to study the morphogenesis of *Marchantia* air chambers at cellular resolution, accessing its sub-epidermal features as well as surface landmarks. This allows a researcher to document their development at the cellular scale and could be combined at later stages with cellular markers to improve the description of the steps involved in air chamber morphogenesis.

Chapter 4

A model for air chamber morphogenesis

4.1 Introduction

The generation of maps with cellular resolution to study the development of an organism provides a valuable tool to dissect the cellular processes underlying tissue morphogenesis [147, 148]. Advances in genetics, molecular biology, biochemistry and especially microscopy have allowed the documentation of many biological processes at cellular and subcellular level. The generation of a plant cell atlas has been proposed with the aim of describing plants' features at a micro and nanoscale, integrating information related to cellular architecture, nucleic acids, protein and metabolites [149]. This type of tool could also be used to simulate biological processes, including tissue growth and differentiation [149]. Hence, the systematic description of cell morphology represents the first step towards the characterisation of differentiated tissue and the identification of distinct cell types.

Marchantia polymorpha air chambers display 5 distinct cell types:

1. Air pore cells (a pore is composed of 4 rings of 4 cells), on the dorsal surface, that form the opening of the chamber to allow for gas exchanges and prevent water loss from transpiration.
2. Cells on the dorsal epidermis form the roof of the chamber.
3. Wall cells that delimit the perimeter of the chamber.
4. Floor cells form the bottom of the chamber.

5. Elongated and chloroplast rich cells called photosynthetic filaments, protrude from the cells on the floor and they fill the cavity of the chambers.

All these cell types have been identified since the early 19^{th} century by many botanists who observed the growth of *Marchantia* tissue [150, 151, 152, 153, 154]. Air chambers are continuously produced from the distal boundary of the meristem, and are initiated by schizogenous processes that produce sub-epidermal spaces at the junction of four or five cells. These sub-epidermal (and intercellular) spaces are not only present in early divergent land plants, but they can also be found in late divergent ones as well. Intercellular spaces (ICSs) can form through lisogeny or schizogeny processes. ICSs formed via lisogeny require spatially specified cell death to leave a space, while schizogenous spaces are formed by cell separation [98].

The first description of their development dates back to the beginning of the 20^{th} century [88, 92], but it was only in the 1980s, that Apostolakos provided a cellular description of the events leading to air chamber development by imaging anatomical sections of *Marchantia paleacea* with light and electron microscopy [93].

The model for air chamber development proposed by Apostolakos recognises the schizogenous formation of the intercellular space at four-way cell junctions as the event triggering air pore and air chambers formation [94], confirming the observations from Barnes and Land [88]. None of the cells belonging to four-way junctions come from one mother cell, but they are derivatives of initial segments produced by anticlinal and periclinal divisions of the apical cells. According to his study, the broadening of the chambers is due to a coordinate rate of anticlinal divisions between the protodermal cells of the roof and the sub-protodermal cells forming the floor of the chamber. The air pore is not produced from one mother cell. It derives from the cells surrounding the entrance of the chamber, where the initial intercellular space was formed. These cells shift their division plane from anticlinal to periclinal, becoming mother cells of the pore, and generating a four cell-layer pore in fully developed air chambers. Finally, the photosynthetic filaments, which protrude inside the chamber, derive from polarised growth and periclinal divisions taking place in the floor cells of the chamber.

After the identification of the Mp*nop1* mutant [85], which lacks intercellular spaces, Ishizaki proposed that the formation of this initial aperture is induced by extracel-

lular signals, which initiate a signalling cascade, which activates genes involved in cell-remodelling processes [98]. Mp*NOP1* encodes a plant U-box armadillo repeat (PUB-ARM) E3 ubiquitin ligase and they have been shown to be signalling proteins involved in signalling cascades initiated by leucine rich receptor-like kinases (RLK) [155, 156, 157, 158, 159, 160]. Despite the discovery of an additional *Marchantia* mutant affecting only air pore morphogenesis [86], the description of air chambers' development remains largely unexplored.

Air chambers are three dimensional structures, which develop on the dorsal surface of gemmae. Thanks to the open development of gemmalings, surface landmarks can easily be accessed by microscopy, although sub-epidermal structures require access to underlying cell layers. In order to map the key stages in the development of the different cell types composing air chambers in *Marchantia*, I developed tissue-clearing protocols, described in Chapter 3, in combination with stains marking cell outlines to capture the appearance of the different cell types both in the outer epidermal and sub-epidermal layers. The creation of a model describing the stages of air chamber development would help to characterise future genetic experiments attempting to provide further insight into this process.

4.2 Results

4.2.1 The formation of intercellular apertures at apical segment derivatives junctions is required for the development of air chambers

It has been shown that *Marchantia* gemmae bear fully developed air chambers 5 days after being removed from the gemma cups (d5 gemmae), when they are grown on gel based media [8]. The development of air chamber, though, starts within 48h after germination [161] and this process is then carried out continuously as gemmalings expand (white arrows in Figure 4.3 A). Apostolakos [93] identified the initial step of air chamber development consists in the formation of intercellular apertures through a schizogenous process taking place at a distance of 4 to 5 cells from the apical cell. These openings develop at the intersection of 3, 4 or 5 cells originated from the so called segments S_1, S_3 and S_4 of the apical cell (first paradermal, S_1, and first two anticlinal divisions, S_3 and S_4, Figure 4.1).

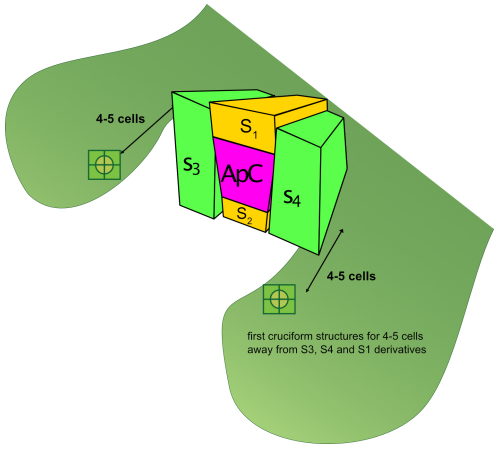


Figure 4.1 **Initial divisions of the apical cell in *Marchantia polymorpha*.** The apical cell (ApC) is represented at the centre of the notch in magenta. Its first paradermal division generates the two segments in yellow (S_1 and S_2), while the first anticlinal division of the apical cell generates S_3 and S_4, shown in green. Cells deriving from subsequent divisions of S_1, S_3 and S_4 give rise to the four cell junctions, from which air chambers will develop.

The intercellular apertures initially reach the surface (Figure 4.2) and they subsequently close due to thickening of the cell walls of the cells facing the newly formed intercellular aperture [93, 162].

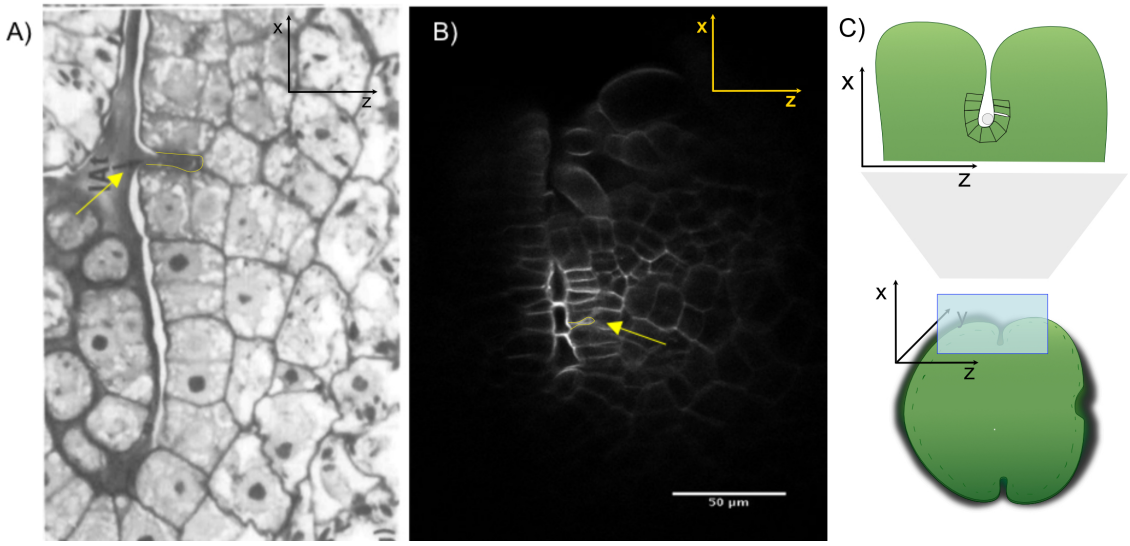


Figure 4.2 **The intercellular apertures initially reach the surface of the thallus.** A) Image from Apostolakos et *al* [93]. The xz plane of the thallus apex shows that the first intercellular apertures reach the surface. B) Optical section of two day old gemma fixed, cleared with ClearSee and stained with Calcofluor White M2R 0.01% w/v (see Chapter 2 section 2.3.2). The yellow arrow shows an intercellular aperture. After 48h from deposition on gel-based growth media intercellular apertures start to appear. C) Schematic representation of the plane shown in A) and B). The plane is parallel to the gemma epidermis in the notch area.

It was postulated that the mutation of the *Marchantia NOPPERABO1* gene stops air chambers development at the very initial stages. This mutant offers the chance to compare its development with the one of wild type plants to understand which step of air chamber development is missing and prevents their appearance. Since the appearance of intercellular apertures at the junction of apical cell segments derivatives is thought to be the initial step of air chamber morphogenesis, the notch area of wild type (WT) gemmae was compared to the ones of Mp*nop1* gemmae. WT and mutant d5 gemmae were fixed overnight in 4% w/v paraformaldehyde, cleared for three days in ClearSee solution, and the cell outline stained overnight with 0.01% w/v Calcofluor White M2R (see Chapter 2, section 2.3.2). The notches of both WT and Mp*nop1* gemmae showed the presence of junctions, (Figure 4.3 B). Looking at the subepidermal layer of the notch, it appeared that intercellular apertures were present only in WT gemmae and they were separated by at least one cell (yellow asterisks in Figure 4.3 XY plane). The notch of Mp*nop1* mutants,

despite the presence of apical cell segments' derivatives junctions, did not show the presence of intracellular apertures or air spaces below the dorsal layer.

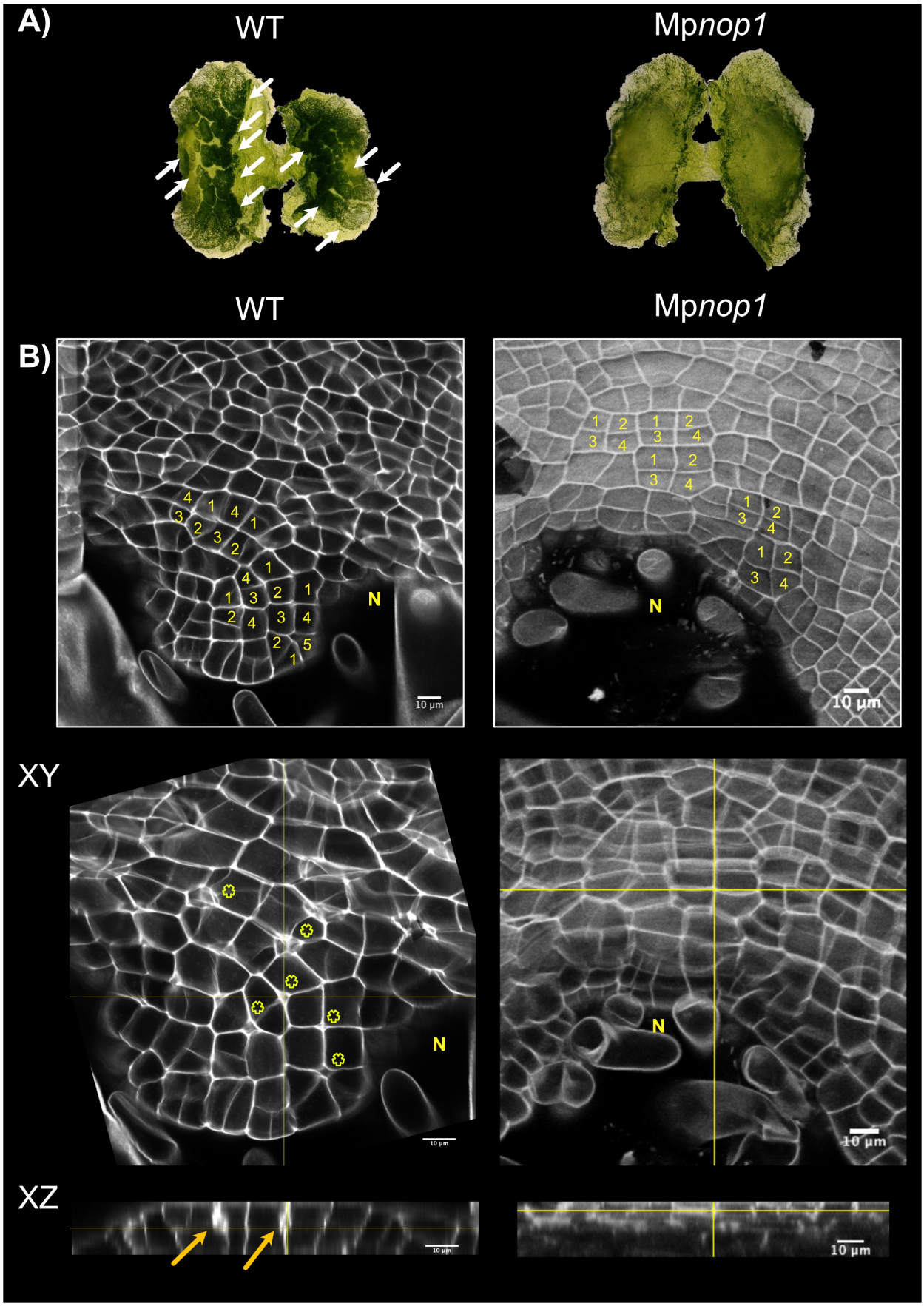

A)
WT
Mpnop1
B)
WT
Mpnop1
N
10 µm
XY
XZ

Figure 4.3 **The initial separation of cell walls and formation of a intercellular apertures is essential for air chambers morphogenesis.**
A) Bright field microscopy images of Cam-1 WT and Mp*nop1* mutants in Cam-1 background gemmalings. WT plants have air chambers (arrows) on the dorsal epidermis, while the Mp*nop1* mutant shows none. B) Z-stack projection and orthogonal views of samples cleared with ClearSee and stained with Calcofluor White M2R 0.01% w/v (see Chapter 2 section 2.3.2) showing WT and Mp*nop1* gemmalings notches. Both WT and Mp*nop1* show the presence of 4 and 5 cell junctions (cells surrounding the junction are shown with yellow numbers), but only WT plants form air spaces below the epidermal layer in between these 4 or 5 cells (indicated by yellow asterisks). The last row shows the XZ planes of the stacks and the arrows indicate the air space in the sub-epidermal layer. N indicates the notch.

4.2.2 Pore diameter increases with distance from the apical notch

The formation of intercellular apertures is followed by an oblique division in the cells surrounding the air space to form a radially symmetric structure [93], from here on called pore complex, which will develop into the final 16 cells/four layer thick air pore of the chamber [93]. These divisions define the first four cells which will form the air pore of the chamber (Figure 4.4).

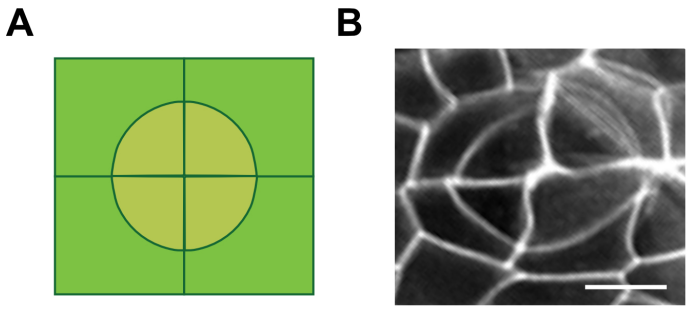


Figure 4.4 **Oblique divisions in the cells surrounding the intercellular apertures generate a radially symmetric structure called the pore complex.**
In A) a schematic version of the oblique divisions taking place at the cell junctions where intercellular apertures appear. B) Confocal image of a pore precursor in a d5 gemma cleared with ClearSee and stained with Calcofluor White M2R 0.01% w/v. Scale bar = 10 μm

The pore complex provides an easy to score and quantifiable surface landmark for air chamber development. The diameter size of the pore complex appeared to enlarge with increased distance from the notch, suggesting that as the intercellular apertures expand to form the subepidermal chamber and the floor and roof cells divide, the pore complex is also increasing in size. In order to test this hypothesis, the diameter of 105 air chamber pore complexes from 25 plants was measured from d5 WT gemmae, as well as the distance of the pore complex from a fixed point located in the gemmaling notch (Figure 4.5 A) and defined as a single point at the centre of the gemmalings notch (see Chapter 2 section 2.3.7 for details). A linear correlation between the air pore diameter and the distance from the notch was found (Pearson's correlation coefficient = 0.86, p-value <2.2e-16), demonstrating that the pore diameter can be used as a quantifiable trait to stage air chamber development (Figure 4.5 B).

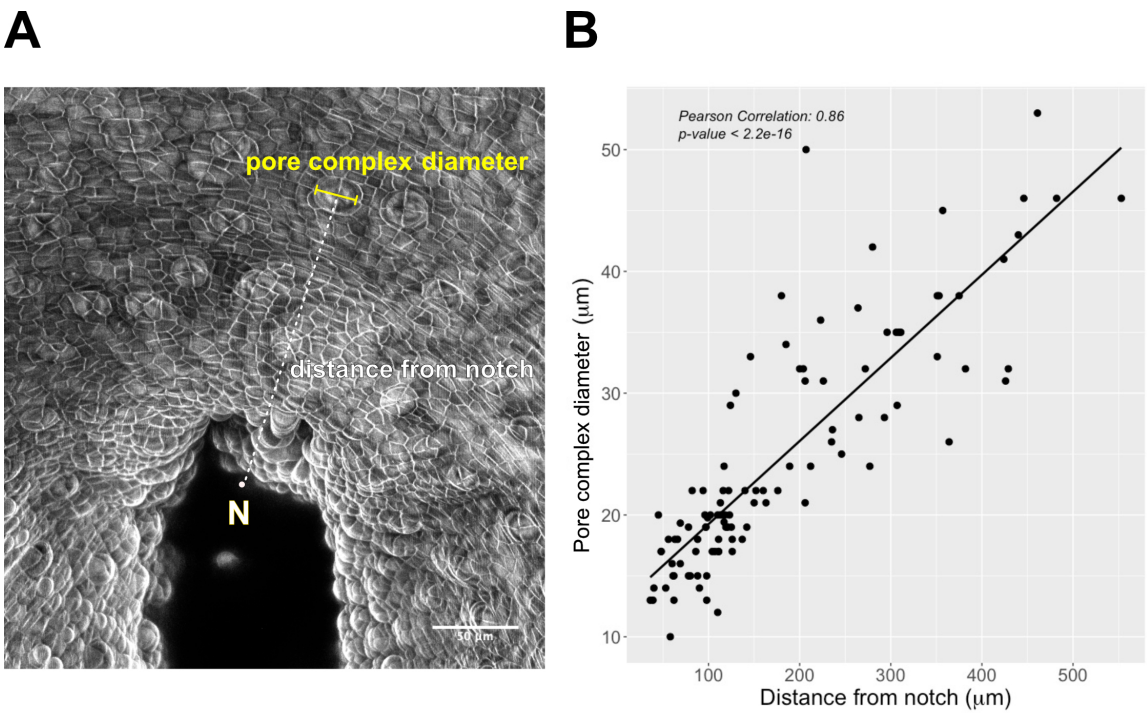


Figure 4.5 **The pore complex diameter increases with distance from the notch as air chambers develop.**
A) Five days gemmalings were fixed and cleared following the ClearSee protocol (Chapter 2, section 2.3.2) and stained with 0.01% Calcofluor White MR2. Samples were imaged under the confocal microscope and z-stacks were projected to measure the diameter of pore complexes using the ImageJ measuring tool. The distance was calculated from a fixed point located at the notch of thallus (N in the figure). B) There is a positive correlation between air pore diameter and distance from the notch. Pearson's correlation coefficient: 0.86, p-value<2.2e-16. Number of samples = 105.

4.2.3 The sequential appearance of photosynthetic filaments and pore opening define three stages of air chamber development

Air chambers are characterised by the presence of elongated cells that protrude from the floor cells of the air chamber. These cells do not form as soon as the intercellular aperture and the pore precursor are formed. Using sectioned tissue, Apostolakos was able to visualise air chambers not containing photosynthetic filaments in close proximity to the notch [93]. My use of cleared samples allowed the observation of the sub-epidermal layers of air chambers to verify the presence of photosynthetic filaments and correlate this to the pore complex size. Photosynthetic filaments started to appear in chambers with a pore complex with a diameter larger than 20 μm (Figure 4.6 A), filling in the chamber underneath the pore. It was also noticed that as the sub-epidermal chamber enlarged, the number of filaments inside increased (data not shown).

As stated in Section 4.2.1, the initial intercellular aperture closes due to thickening in the cell walls facing the air space. Therefore, at early stages of air chamber development, the pore complex is formed above a closed sub-epidermal chamber and it opens only at later times. From the obtained air chambers' images, the opening of the pore complex was monitored and compared with the size of its diameter. The pore complexes appeared to open to the exterior when their diameter was larger than 28 μm (Figure 4.6 B).
Looking at the graphs in Figure 4.6 three different stages of air chambers' morphogenesis could be identified:

1. Air chambers where the pore complex is closed and the subepidermal chamber does not have any photosynthetic filaments.

2. Air chambers with closed pore complexes that start to develop photosynthetic filaments.

3. Air chambers with open pore complexes and filled with photosynthetic filaments.

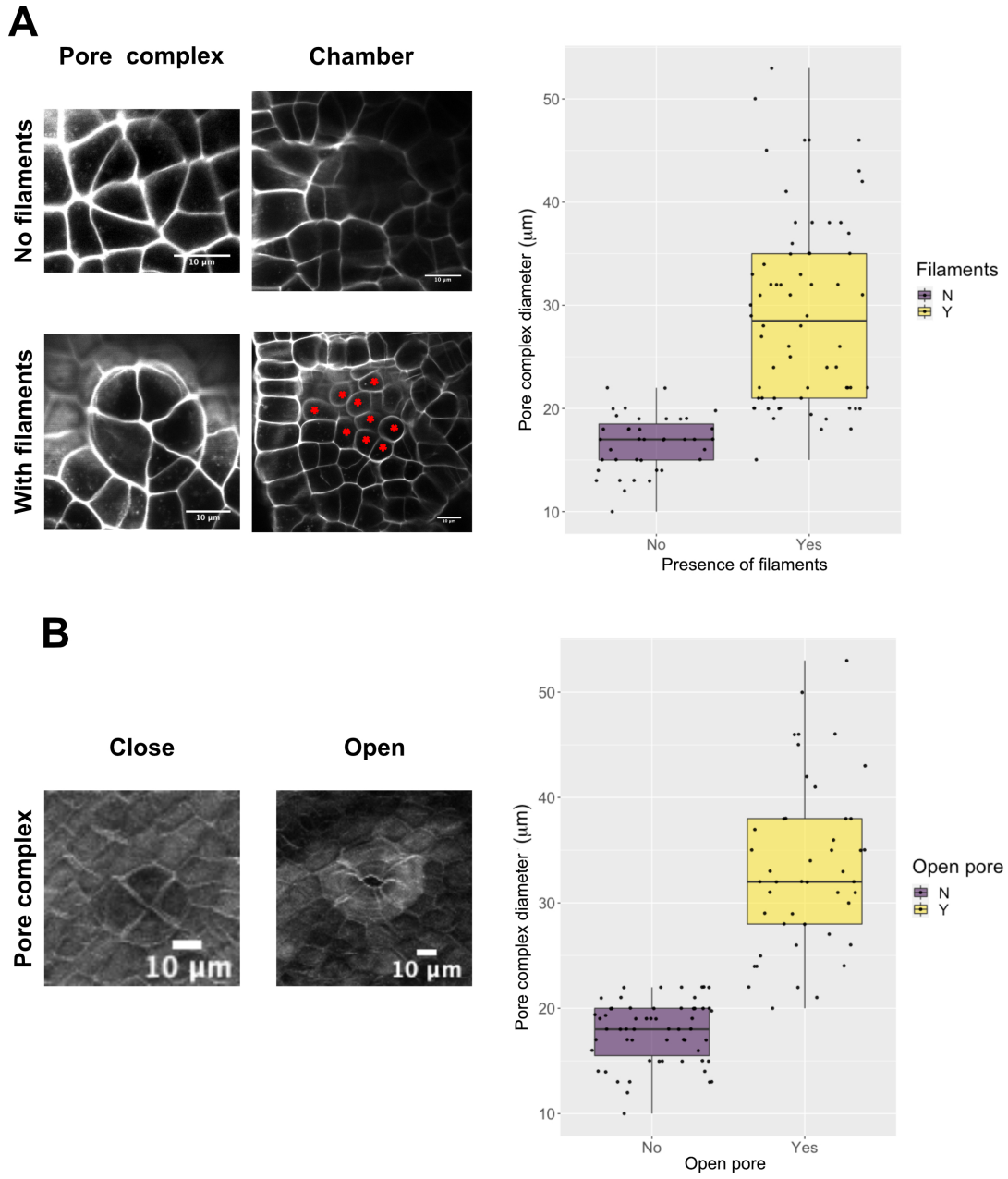


Figure 4.6 **Photosynthetic filaments appear in chambers with pore complexes larger than 20 μm and pores open when the complex is larger than 28 μm.**
For each of the measured pore complexes (n=105), the presence of photosynthetic filaments, as well as the one of an opening of the pore to the environment, was verified. This classification allowed the identification of three stages of air pore development relating the pore complex diameter size to the presence or absence of photosynthetic filaments and the opening of the pore into the environment. A) On the left, a confocal image showing a surface view of the pore complex and the underneath chamber showing absence or presence of photosynthetic filaments (red asterisks). On the right, a boxplot showing the correlation between the pore complex diameter and the presence of photosynthetic filaments. B) On the left a surface view of a close and open pore. On the left, the boxplot shows that pores over 28 μm start to open.

4.2.4 Air chambers walls are only a single cell layer thick

Air chambers are separated from each other by walls which enclose the sub-epidermal chamber containing photosynthetic filaments. Air chambers in different stages of development (mostly in stage 2 and 3 - see Section 4.2.5 - in Figure 4.7) can be observed in a developmental gradient distal to the meristematic notch. The walls of these chambers are shared amongst neighbours and they are one cell layer thick. In Figure 4.7, an XY optical section of the interior of air chambers (circled in red) highlights the single cell layer walls. Air chambers are regularly distributed on the epidermal surface of *Marchantia* thalli and walls ensure the correct spacing between them from the early stages of their development.

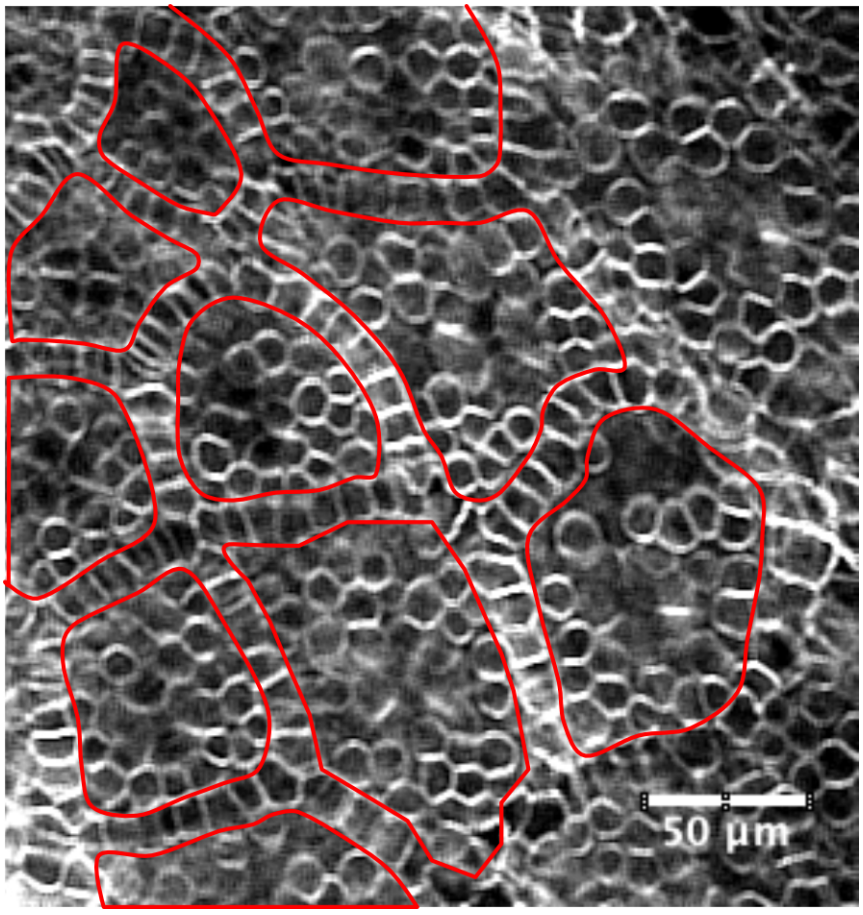


Figure 4.7 **One cell thick walls separate neighbouring air chambers.** XY plane of the interior of air chambers from 5 days old gemmalings show that walls separating the chambers are only one cell thick.

4.2.5 Air chamber morphogenesis takes place in four stages

Based on the observation and the measurements done in the previous sections, four stages could be identified in air chamber development.

- **Stage 0:** in close proximity of the apical cell, 4-5 cells, at the apical cell derivatives' junctions, intercellular apertures form in the sub-epidermal layer.
- **Stage 1:** four radial divisions form the pore complex around the intercellular aperture and close the dorsal layer of the chamber. This structure has a diameter of less than 20 μm and it will later become the air chamber pore. The sub-epidermal chamber begins to enlarge, but it does not contain any photosynthetic filaments. Air chambers at this stage can be found within a distance of <150 μm from the apical notch (Figure 4.5).
- **Stage 2:** the pore complex diameter is between 20 and 28 μm. The pore is not open yet, but the chamber in the sub-epidermal layer starts to enlarge. Photosynthetic filaments start to emerge from the chamber floor. Air chambers at this stage can be found typically at 200-300 μm from the apical notch (Figure 4.5).
- **Stage 3:** the pore complex diameter is larger than 28 μm, the pore is open connecting the internal part of the chamber to the outside. The sub-epidermal space shows protruding photosynthetic filaments. Air chambers in this stage are found 300 μm away from the notch (Figure 4.5).

In Figure 4.8, stages of air chamber morphogenesis are shown in *Marchantia* 5 days old gemmalings. They include a view of the surface of the thallus (XY surface), an image of the sub-epidermal layer to illustrate the interior of the air chamber (XY inner layer) and a cross section (XZ section) to highlight the different degrees of the opening of the pores and the appearance of filaments inside the chamber.

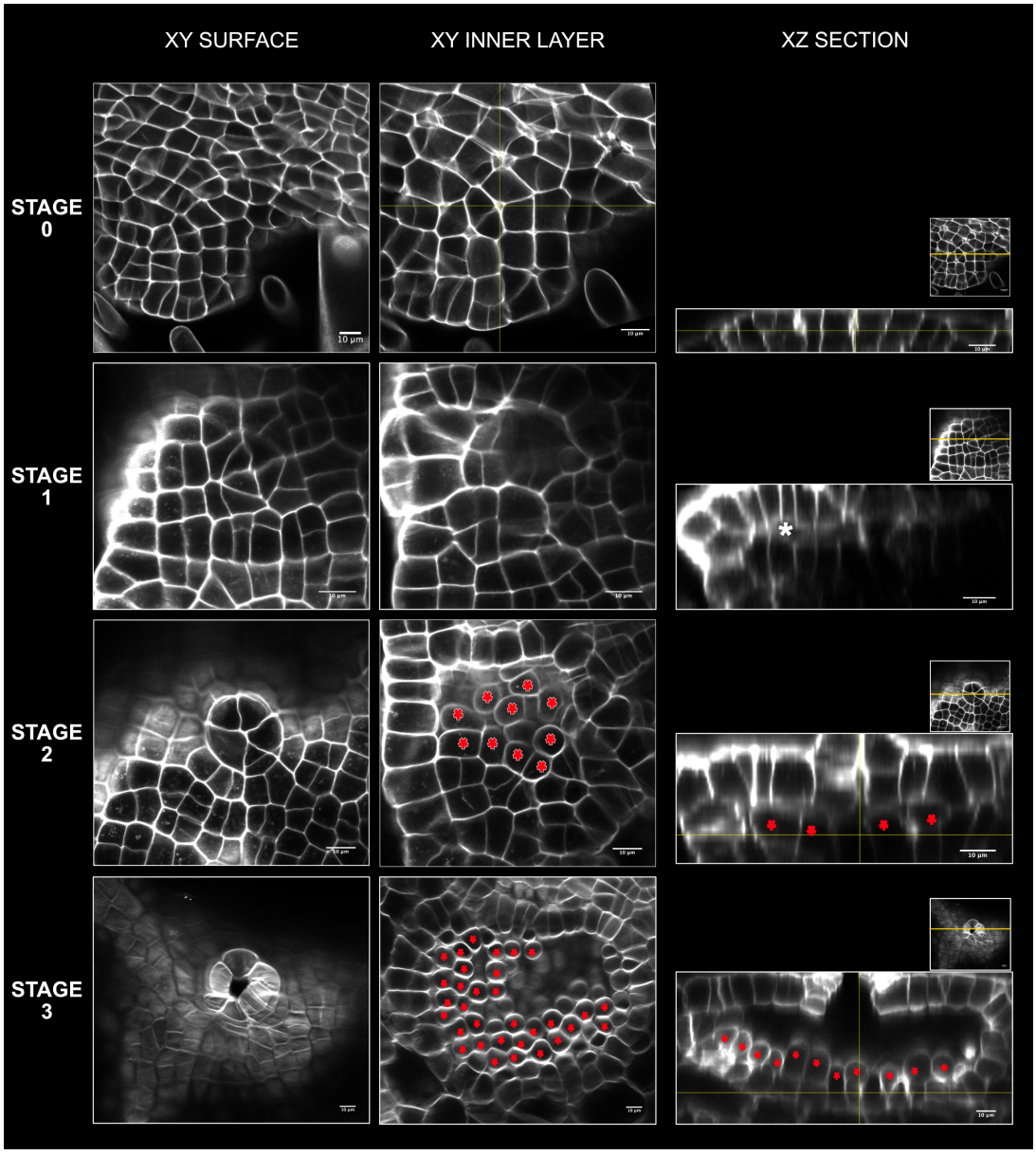


Figure 4.8 **Air chamber morphogenesis can be divided in 4 stages.** Air chambers of 5 day old gemmalings cleared and stained with Calcofluor White M2R, as described in Chapter 3. For each stage, paradermal XY section view of the surface and paradermal section of the sub-epidermal layer are shown. The XZ planes show the inside of the chamber. In the XY planes on top of the sections, the yellow line shows where the xz section was positioned. **Stage 0**: appearance of intercellular apertures at apical cell derivatives' junction, in close proximity of the notch (4-5 cells). **Stage 1**: appearance of the pore precursor and enlargement of the sub-epidermal chamber (white asterisk), before photosynthetic filaments appear. **Stage 2**: Photosynthetic filaments appear inside the chamber (red asterisks) and the sub-epidermal layers of the pore complex cells start to separate. **Stage 3**: the pore shows all the 4 cell layers and photosynthetic filaments are formed inside the chamber (red asterisks). All scale bars are 10 μm.

4.3 Discussion

4.3.1 A new model for air chamber development

Thanks to its open development, *Marchantia* gemmalings provide a highly accessible system to observe the morphogenesis of air chambers, which develop on the dorsal surface. The application of clearing and staining techniques presented in Chapter 3, has allowed the visualisation of air chamber development at the cellular level, including the observation of the sub-epidermal processes. High resolution optical sections were obtained using laser scanning confocal microscopy. This approach allowed the imaging of chambers in their entirety, while maintaining the spatial relationship and three dimensional appearance of the cellular components.

The results presented in this chapter led to the formulation of a revised model for air chamber development, which included and extended earlier former observations from Apostolakos and Ishizaki [93, 98] (Figure 4.9). The new model correlates the development of dorsal features, the air pore, with the sub-epidermal structures, the emergence of photosynthetic filaments, providing a link between the events taking place in the dorsal and sub-epidermal layer. It brings together the observations of Apostolakos and Ishizaki, although it demonstrates that the initial intercellular aperture closes as the sub-epidermal chamber grows, contrary to what was shown by Ishizaki(Figure 4.9 A).
The examination of air chambers in stage 0 has confirmed that the absence of intercellular apertures, as observed in the Mp*nop1* mutant, prevents the formation of air chambers, as previously suggested by Ishizaki [98].

The stage of pore complex development is a marker that is easy to quantify and score and it has allowed me to define three additional stages of air chamber morphogenesis. The size of the pore complex precursor, in fact, correlates with the appearance of photosynthetic filaments in the subepidermal layer and it also correlates with the opening of the pore. In Stage 1, the pore complex is closed and no photosynthetic filaments are present inside the chamber. Stage 2 includes all chamber with a closed pore complex and photosynthetic filaments starting to appear in the subepidermal layer and, finally, Stage 3 comprises air chambers with an open pore in the pore complex and photosynthetic filaments. As the pore complex enlarges on the dorsal surface, it also divides periclinally to form a four cell layer air pore [93]. Further

observations of this process in the pore complex could help define new stages of its development and improve the accuracy of the model.

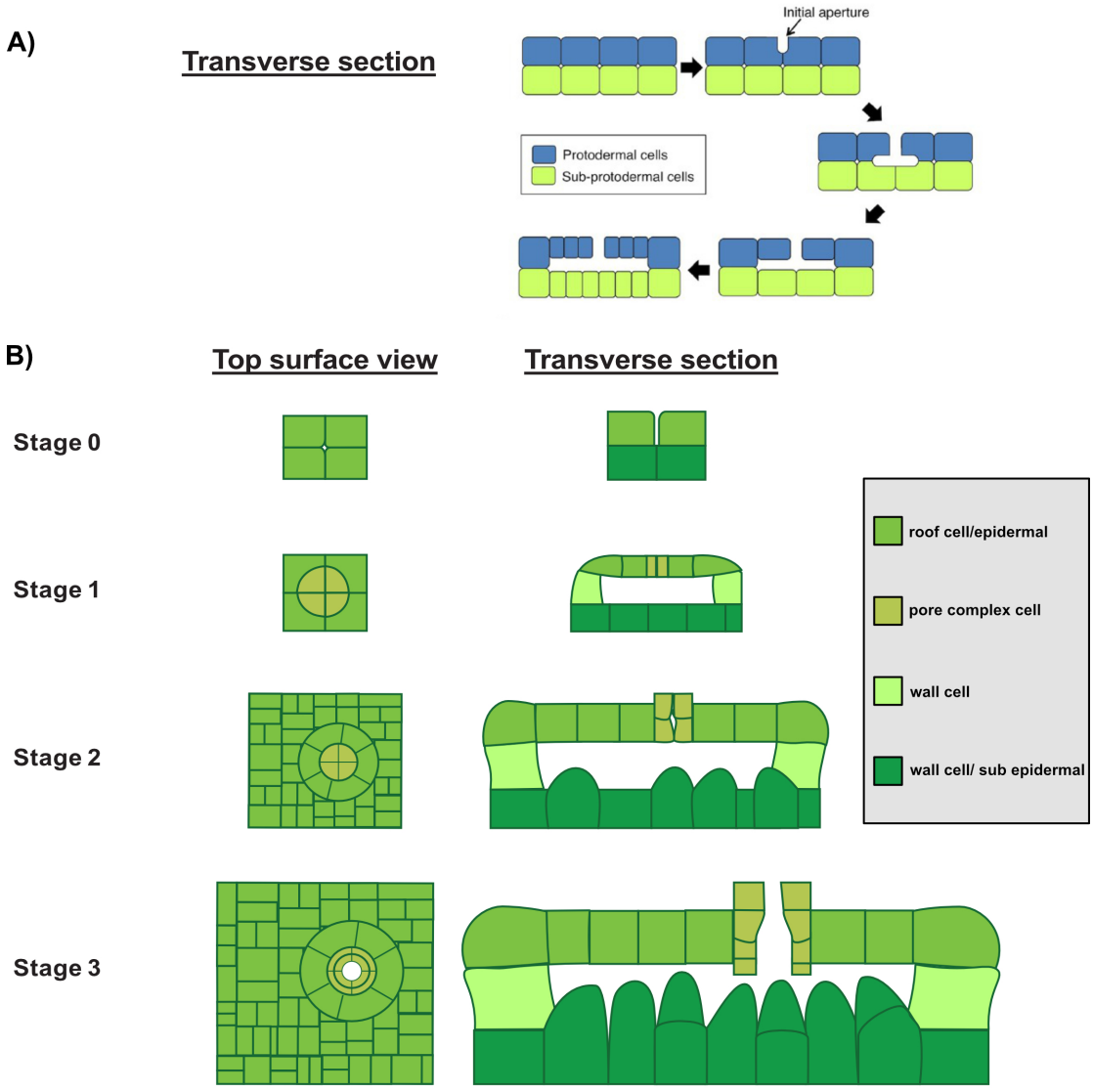


Figure 4.9 **A new proposed model for air chamber morphogenesis.**
A) The model proposed by Ishizaki [98] only describes the formation of the schizogenous space and its enlargement to form the sub-epidermal chamber.
B) The four stages identified in Section 4.2.4 are represented in this diagram. The epidermal structures are shown in bright green, the sub-epidermal layer in dark green, the pore precursor in the pastel green and the wall cells in lime green. The white circle in stage 3 represents the opening of the pore and the protruded cells in Stage 2 and 3 are the photosynthetic filaments.

4.3.2 The floor cells protrude into the chamber to form photosynthetic filaments

Based on the observed samples, air chambers in Stage 2 show the first photosynthetic filaments protruding from the chamber floor. As the chambers enlarge to reach Stage 3 the size of the air space increases as well as the number of photosynthetic filaments. This observation suggests that the cells on the sub-epidermal layer enclosed inside the chamber walls continue to divide and develop into photosynthetic filaments, filling the floor of the chamber.

4.3.3 Air chamber spacing

The walls of air chambers have been shown to be composed of a single cell layer and are shared between neighbouring air chambers. As air chambers increase in size, the wall cells should also divide, but their division plane must then be anticlinal to the perimeter of the chamber, allowing the chamber to enlarge, while the walls maintain the same thickness. This strict pattern might suggest that some kind of lateral inhibition or cell-cell communication is taking place, to restrict the cell division pattern of walls and pores, while allowing the floor and roof of the chamber to divide to expand the surface area of the air chamber. The pore precursors, instead, need to divide periclinally with respect to the dorsal epidermis to form the four layer pore found in Stage 3 air chambers.

The restrictive pattern of divisions taking place in walls and pores suggests that positional cues might be in place to ensure the correct patterning and spacing of air chambers. Positional cues play a key role in the patterning of epidermal structures, like root hairs, trichomes and stomata in *Arabidopsis* [49, 53, 52, 60, 55, 163]. It is possible that similar mechanisms could be in place in *Marchantia*, as already suggested by Ishizaki in relation to the initiation of schizogenous spaces [98].

The Mp*nop1* mutant is not sufficient to understand the mechanisms underlying air chamber patterning: it is possible that the *nop1* mutation is only affecting the mechanism controlling the separation of cell walls, while the distribution and reception of positional cues is still taking place. In order to identify and characterise the presence of underlying genetic networks, the generation of mutants (using UV-induced r T-DNA insertion mutagenesis) with abnormal patterning of air chambers would be

required. These studies could shed light on the nature (e.g. hormones, peptides, environmental conditions etc) of the signals affecting the patterning of air chambers.

Furthermore, the use of cell markers with specific expression in the pore complex cells or photosynthetic filaments could help understand the dynamics of cell divisions and differentiation that lead to the formation of a fully developed air chamber. These type of experiments would help to refine and improve the current model, providing a better reference for the characterisation of mutant lines showing defects in air chamber morphogenesis.

Chapter 5

Identification of candidate genes for cell-type specific markers through transcriptome comparison of wild type Cam-1 and Mp*nop1* mutant plants

5.1 Introduction

Transcriptomic datasets have been a useful resource for scientists to identify key genetic regulators, when comparing transcribed genes from organisms at different developmental stages, with distinct genotypes. Several transcriptomic datasets for *Marchantia polymorpha* have been generated in the past five years. The publicly released data mainly focus on the comparison between gametophyte and sporophyte derived tissues, hormone treatments and related mutants, as well as spore development (see Appendix A.1 for a list of the published and open access transcriptomic datasets). However, none of the available datasets captures systematically the early stages of gemma development, when air chambers start to develop, and no transcriptomic data from air chamber mutants have been published.

The *Marchantia nopperabo1* is the only available mutant that shows an extreme phenotype where no air chamber can be observed on the surface of the thallus. It was first obtained through a T-DNA insertion in the exon 2 of the *NOPPERABO 1*

gene of Tak plants [85]. The Cam-1 Mp*NOP1* gene sequence is similar to the one annotated in v5.1 genome of the Tak 1 ecotype, showing only two transitions (C→T) in the coding sequence, which both lead to silent substitutions not modifying the translated sequence (See Appendix C). No other gene regulating air chamber formation has been identified to date. Comparison of the transcriptome of Mp*nop1* mutants with that of wild type plants could be useful for the identification of new candidate transcription factors involved in air chamber development.

In this chapter, I describe how I produced a transcriptomic dataset capturing the early stages of air chamber development in young gemmae collected from wild type (WT) Cam-1 plants and Mp*nop1* mutants obtained in the Cam ecotype. I collected samples of gemmae grown on half strength Gamborg B5 media for 24, 96 and 192 hours after their removal from gemma cups, to capture key stages of air chamber development. The data produced were mapped to the Tak v 5.1 reference genome and quantified using a pipeline developed by Marius Rebmann. Data was analysed in R using the tximport package [122] and differentially expressed genes were identified and analysed using the R package DESeq2 [123]. The analysis presented in this chapter focuses on the identification of differentially expressed transcription factors between wild type (WT) and Mp*nop1* plants, which have been shown to regulate epidermal patterning in vascular plants, including members of the basic helix loop helix and the WRKY gene families. As the Mp*nop1* mutant lacks air chamber and, to our knowledge, no other gemmaling's structure, the identified gene candidates could be used as cellular markers to track air chamber development.

5.2 Results

5.2.1 Physiological differences between WT and Mp*nop1* mutant plants

Marchantia sporelings of the Cam ecotype constitutively expressing the plant codon optimised version of the endonuclease CRISPR (Clustered Regularly Interspaced Short Palindromic Repeat) associated protein 9, Cas9, [164] under the control of the *ELONGATION FACTOR 1α* promoter [165], were transformed by Mr Owen Male and Ms Linda Silvestri with the pBsgRNA3_nop1 plasmid designed by Dr Bernardo Pollak [166], which contains a 19 bp guide RNA (gRNA) targeting exon 3 of *Marchantia NOPPERABO1* gene (Mp8g11800). A successful transformant exhibiting the Mp*nop1* mutant phenotype (later identified as a male plant) was genotyped and a 5 bp deletion causing a frameshift mutation (Figure 5.1) was identified. This line was selected to conduct the experiments described in this chapter.

Marchantia Cam and Tak ecotypes have very conserved genomes, but small mismatches in their gene coding sequences can be identified. In addition, not all the annotated genes on the Tak genome can, at the moment, be retrieved from the available Cam genome assembly [167]. Although a Mp*nop1* mutant line is already available in the Tak-1 background, for a reliable outcome, it was necessary to generate the mutant in a Cam-1 background. In this way, the mutant was obtained in an isogenic background, avoiding potential differences in the number of annotated genes between WT and mutant plants.

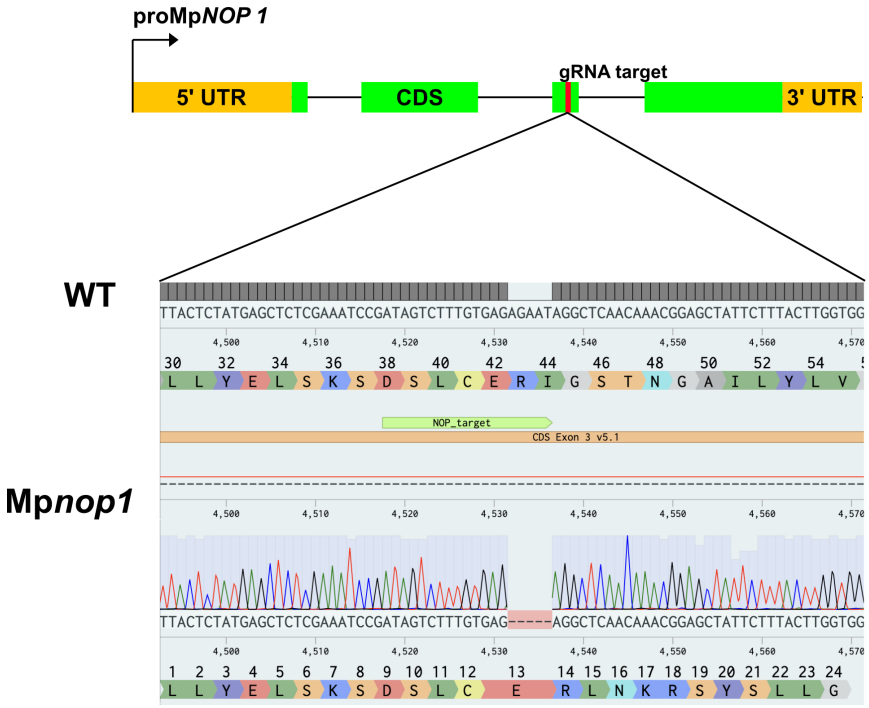
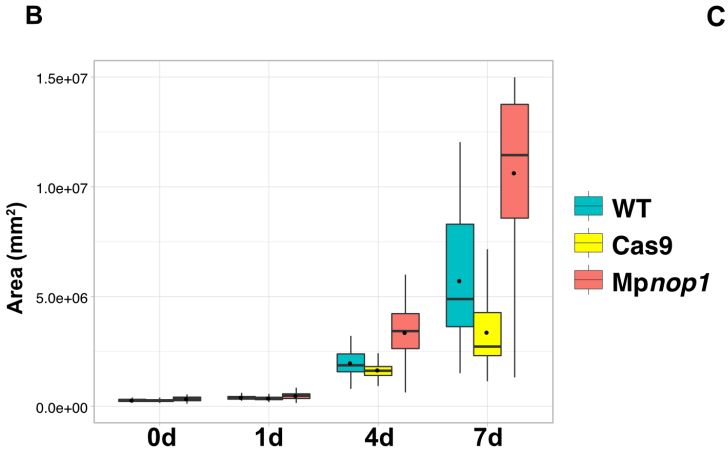
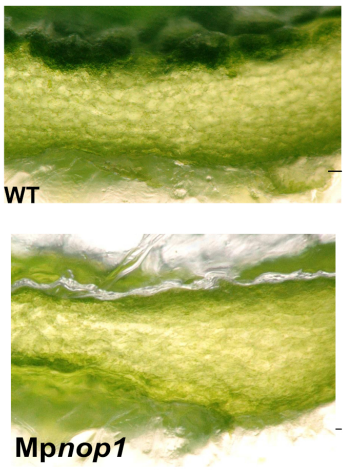


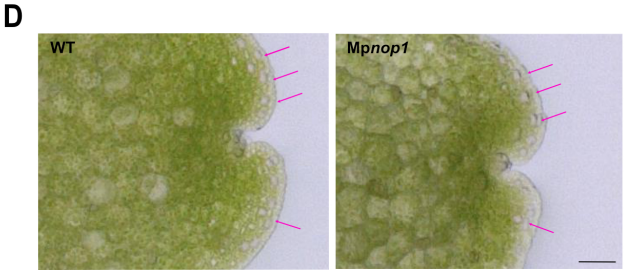
Figure 5.1 **Genotype of Mp*nop1* mutant plants.**
At the top, a diagram of the gene structure with a red rectangle indicating the sequence on exon3 targeted with the gRNA. Orange rectangles represent the 5' and 3' untranslated regions of the gene, while the green boxes show the coding sequence of the protein. The top row shows the WT sequence of the Mp*NOP1* gene and the encoded amino acids (arrows below the sequence). The gRNA target sequence is highlighted with a green arrow. In the bottom row, the CDS sequence of the Mp*nop1* mutant showing the 5 bp deletion that causes a change in the amino acid sequence at the 3' end of the deletion.

The phenotype of Mp*nop1* plants was evaluated monitoring its growth habits and its ability to produce all types of differentiated tissue, other than air chambers, that can be observed in WT thalli (Figure 5.2 C): oil bodies (Figure 5.2 D), rhizoids (Figure 5.2) and gemma cups (data not shown). Mutant Mp*nop1* thalli were able to enter both vegetative and sexual reproduction. To assess differences in growth habits between WT, Cas9 expressing plants and Mp*nop1*, the surface area (mm^2) of 50 WT (blue boxplot in Figure 5.2 B), 50 gemmalings constitutively expressing the Cas9 endonuclease (yellow boxplot in Figure 5.2 B) and 50 Mp*nop1* gemmalings (pink boxplot in 5.2 B) was measured at 0, 1, 4 and 7 days after removal from the gemma cups. Mp*nop1* mutant gemmalings showed a surface area which was on average two times larger than WT or Cas9 expressing gemmalings at 4 and 7 days.


A
1 day
4 day
8 day
notch
WT
nop1


B
1.5e+07
1.0e+07
5.0e+06
0.0e+00
Area (mm²)
0d
1d
4d
7d
WT
Cas9
Mpnop1


C
WT
Mpnop1


D
WT
Mpnop1

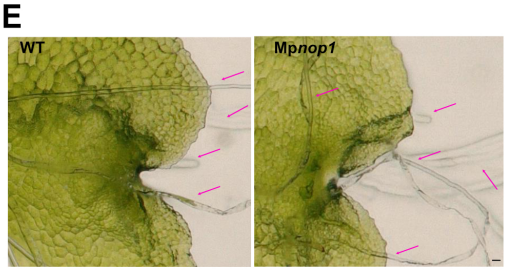

E
WT
Mpnop1

Figure 5.2 **Comparison of WT and Mp*nop1* mutant gemmae development in Cam-1 genetic background.** A) Images of WT and Mp*nop1* gemmae at 1, 4 and 8 days after removal from the gemma cups and grown on half strength Gamborg B5 media plates with 1.2% agar. All scale bars are 100 μm. Air chambers start to form in proximity of the notch right after germination, but they become visible after five days from gemmae germination. The notch column highlights the absence of air chambers in Mp*nop1* gemmalings. Scale bars are 50 μm. B) Thallus growth was calculated measuring thalli surface area at 0, 1, 4 and 7 days after removal from the gemma cups. Wild-type plants are shown in blue (n= 50), plants constitutively expressing Cas 9 in yellow (n= 50) and Mp*nop1* mutants in pink (n= 50). Dots represent the mean and horizontal lines show the median value. C) Sections of WT and Mp*nop1* thalli showing the absence of air chambers in Mp*nop1* gemmae. Scale bars 100 μm. D) Wild type (WT) and Mp*nop1* gemma showing oil bodies (magenta arrows) at the margin of the thallus. Scale bar 50 μm. E) Wild type (WT) and Mp*nop1* gemmalings showing rhizoids, indicated by the magenta arrows. Scale bar 50 μm.

5.2.2 Sample collection, RNA extraction and sequencing

The presence of phenolic compounds and polysaccharides in *Marchantia* tissue represents a challenge for the extraction of genetic material. A fast and easy protocol was developed combining Trizol extraction with column purification to extract large amounts of high quality RNA. *Marchantia* gemmae were collected from WT plants and Mp*nop1* mutant plants. They were grown on 25 ml of half strength Gamborg B5 media with 1.2% w/v agar in 9 cm Petri dishes for 1 day (d1), 4 days (d4) or 8 days (d8) (Figure 5.2 A). For 1 day samples, approximately a hundred gemmae were removed from gemma cups and spread onto a Petri dish and grown for 24 h. Twenty gemmae per plate were grown on each Petri dish for 4 day samples and 10 gemmae per plate for 8 day samples. Plants were placed in a tissue culture room and grown under continuous light (150 μmol/m2/s) at 21℃. 100 milligrams of tissue were collected in triplicates for each time point and tissue was ground using a Tissue Lyser II at 60Hz for 30 s. RNA was extracted using a combination of Trizol and QIAGEN RNeasy columns (Section 2.2.4). All RNA samples were treated with Turbo DNase right after extraction to remove any trace of residual DNA. The RNA integrity was verified using the Agilent 2100 Bioanalyzer equipped with the RNA 6000 Nano Kit and samples with RNA integrity number (RIN)$>$ 6 were selected and sent to Genewiz®. An Illumina Poly A selection library was prepared at the Genewiz® sequencing facility, where samples were single-index sequenced on an Illumina HiSeq platform producing 150 bp paired end reads. 18 libraries were generated with an average of 20 million reads per sample (5.1) . After mapping against

the v5.1 of the *Marchantia* genome and filtering for genes with no counts equal to zero, 16891 out of 19287 transcripts were identified.

Table 5.1 Size of the generated libraries. Numbers represent million (M) reads per library.

Sample	**Rep1 reads (M)**	**Rep2 reads (M)**	**Rep3 reads (M)**
d1_WT	24.2	18.2	21.1
d1_Mp*nop1*	17.5	27.3	25.9
d4_WT	23.7	21.9	23.6
d4_Mp*nop1*	22.6	22.9	19.6
d8_WT	21.8	19.3	18.8
d8_Mp*nop1*	18.5	17.7	18.1

5.2.3 Differential expression analysis

The obtained raw reads were mapped onto the v5.1 of the *Marchantia* genome [99], transcripts were quantified using Kallisto [121] and then differential expression analysis was performed using the DESeq2 package [123]. Principal component analysis (PCA) was performed on gene counts normalised against the library sizes (see Section 2.4.1) to identify the main factors causing sample variability. The PCA plot in Figure 5.3 shows that gemmae age (filled circle: 1 day, filled triangle: 4 days and filled square: 8 days) has a greater effect on transcripts' abundance than the plant genetic background (WT is shown in blue and Mp*nop1* mutant in pink).

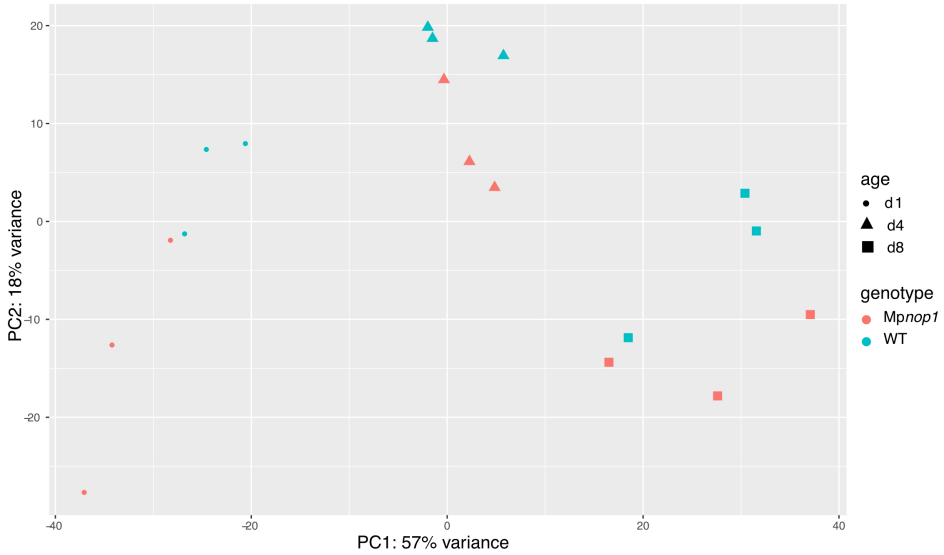


Figure 5.3 **Principal component analysis of sequenced libraries.**
The graph shows that 57% of the difference between the samples is due to the age of the gemmalings (x axis), and only 18% due to different genotypes (y axis). Shapes indicate the different age of the samples: circles, 1 day, triangles 4 day, squares 8 days. The blue dots correspond to WT plants, while pink dots are Mp*nop1* mutant samples.

Differentially expressed genes (DEGs) were identified between the six conditions present in the experiment. The different comparisons are shown in Figure 5.4. Each comparison is shown as a volcano plot, where the x axis shows the expression log2 fold change, while the y axis displays the negative logarithm in base 10 of the adjusted p-value. The adjusted p-value (padj) is an estimate of the statistically significant results, as defined by the Wald test performed by the DEseq function. Hence, the padj column represents the false discovery rate (FDR), i.e., the percentage of false positive results for each statistically significant gene. In Figure 5.4, FDR is set at 0.01, meaning that 1% of the statistically significant genes would be false positives. DEGs with an FDR$<$ 0.01 are shown in red, while the others are shown in grey. It needs to be noted that the differences in the transcript abundance between the samples could be due to differences in transcript levels as well as by different number of cells between WT and Mp*nop1* mutants, as well as different developmental stages.

The comparisons amongst different developmental stages of the same genotype (WT in the second row of Figure 5.4 and Mp*nop1* in the third row), revealed more than 2500 DEGs in the 1 vs 4 day comparison in both WT and Mp*nop1*

gemmalings, 1389 DEGs for WT and 2008 for Mp*nop1* in the 4 vs 8 day comparison and, finally, 4186 DEGs were identified in WT samples and 6115 in Mp*nop1* in the 1 vs 8 day comparison. When the comparison was made between the two genotypes, WT and Mp*nop1*, at each time point, only 203 DEGs were identified in 1 day old gemmalings, 106 in 4 days old gemmalings and 1013 in 8 days old samples. The number of DEGs confirms what the PCA showed: greater differences are found throughout the different developmental stages, rather than between the two genotypes. WT and Mp*nop1* mutant plants have very similar transcriptomes up until 4 days, but an increase in the number of DEGs can be seen at day 8, when the WT dorsal surface is largely covered with air chambers.

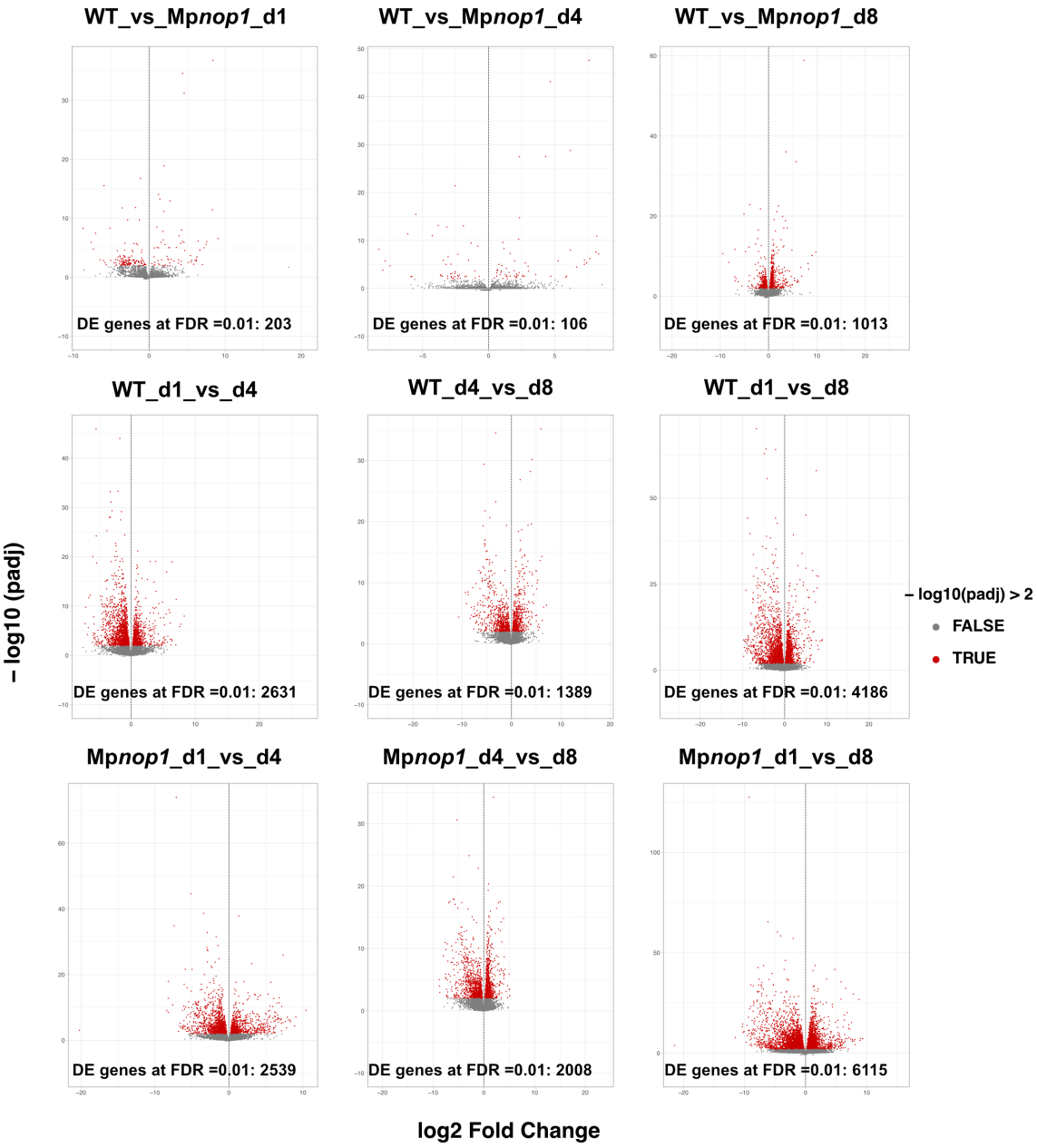


Figure 5.4 **DEGs in all comparisons amongst samples.**
Volcano plots showing differentially expressed genes in 9 of the different comparisons. Red dots indicate genes with an FDR < 0.01, while the grey ones have an FDR >0.01. The black dotted line indicates the 0, separating the upregulated genes on the right and the downregulated genes on the left. The total number of DEGs according to the threshold is indicated at the bottom of the graph.

For the generated lists of DEGs in the WT vs Mp*nop1* mutant comparisons, the gene ontology (GO) terms for the biological processes they were involved in, were identified. Figure 5.5 shows only terms represented by at least five genes GO. DEGs in one day samples appear to be involved in transmembrane transport, oxidation-reduction processes and carbohydrate metabolic processes. In d4 samples DEGs appeared to be involved in oxidative stress responses other than in more general oxidation-reduction processes. The 8 day comparison (d8) showed that DEGs are involved in 37 biological processes. Amongst these, cell-wall related terms (cell macromolecule remodelling process) were identified as well as terms related to photosynthesis (photosynthesis, photosynthesis light harvesting, chlorophyll catabolic process, porphyrin-containing compound biosynthetic process).

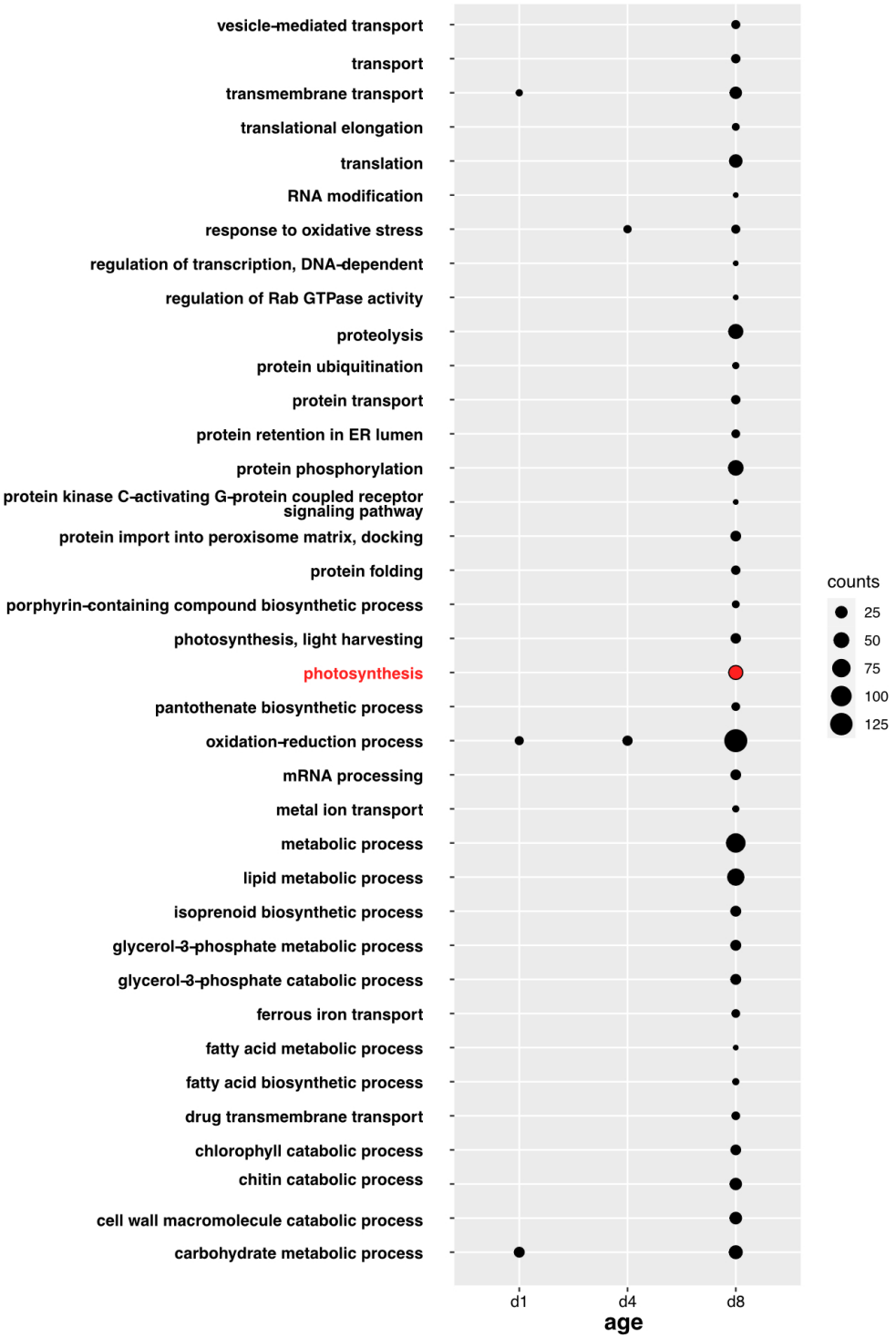


Figure 5.5 **GO (gene ontology) term enrichment of WT vs Mp*nop1* DEGs.** Differentially expressed genes with FDR>0.01 were identified for each time point in the WT vs Mp*nop1* comparison were grouped according to their GO terms corresponding to biological processes. Terms with at least 5 genes are shown and the size of the dots represents the number of genes described by a given term in each comparison.

5.2.4 Identification of DEGs

DEGs between WT and Mp*nop1* mutant plants were identified at each time point. Genes were filtered according to a false discovery rate (FDR or padj) of 0.01, which allowed to identify 203 DEGs in d1, 106 in d4 samples and 1013 in d8 ones. Given the small number of statistically significant DEGs, no additional filtering was applied to the dataset. The number of upregulated and downregulated DEGs identified in all the analysed comparisons can be seen in Figure 5.6.

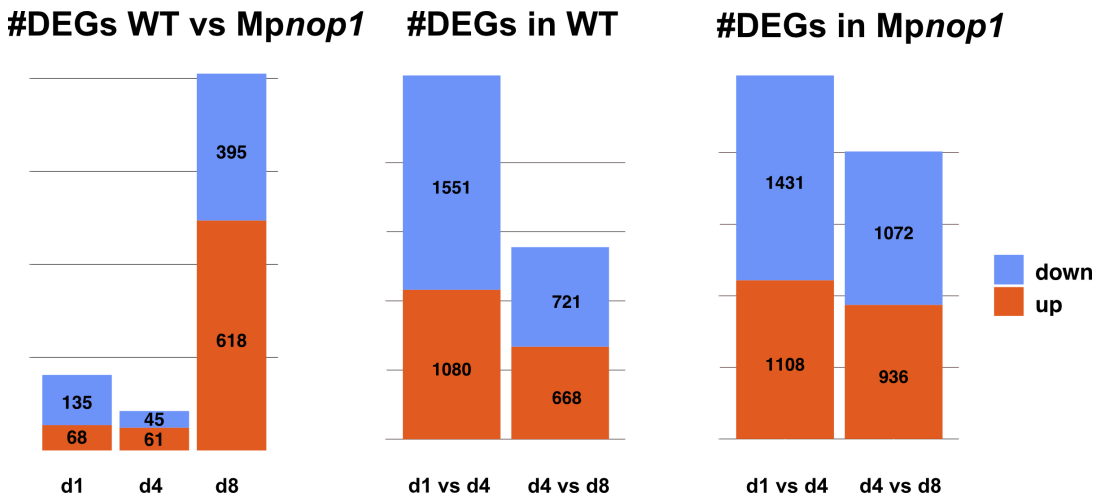


Figure 5.6 **DEGs in all the analysed comparisons.**
Bar charts representing the number of up regulated (orange) and down regulated (blue) genes in each comparison.

DEGs were also identified between 1 and 4 days and 4 and 8 days for each genotype, (Figure 5.6). A GO enrichment analysis was run on these comparisons to identify the biological processes DEGs were involved in and highlight any differences between the three different time points in the two genotypes. DEGs involved in biological processes were identified in WT and Mp*nop1* comparison. GO terms relating to DNA metabolism, transport and protein modification were enriched in both WT and Mp*nop1* samples, while photosynthesis related genes were only enriched in WT samples in the 1 vs 4 day comparison (Figure 5.7).

Figure 5.7 **Gene ontology analysis of DEGs during WT and Mp*nop1* gemmae development.**
Number of enriched genes for the different biological processes in 1 vs 4 day and 4 vs 8 days comparison in WT (blue) and Mp*nop1* plants (pink). The size of the dots represents the number of genes described by a given term in each comparison.

5.2.5 Identification of DEGs involved in cell wall remodelling

Nop1 mutant plants (see Section 5.2.1) seem to have completely lost the ability to form air chambers, including the formation of the initial cell-wall separation at the four way cell junctions, which are thought to initiate the schizogenous process required for the formation of intercellular spaces [98]. In the WT vs Mp*nop1* comparison, 53 genes were differentially expressed at all three time points (Figure 5.8 A).

Amongst these, 3 genes were involved in protein phosphorylation and cell wall modification were identified. A pectinesterase (Mp1g00190), an enzyme modifying pectin residues in the cell wall, shows higher expression in WT plants starting from day 1 and its expression peaks at day 4, while it is lowly expressed in Mp*nop1* samples (Figure 5.8 C). Two serine threonine kinases, Mp3g01920 and Mp3g01930, were found to be highly expressed in WT plants at 4 days, while they have low expression in Mp*nop1* samples. Pectin has been shown to play a key role as a cell wall component in vascular plants, as its chemical modification can lead to cell separation [168]. Moreover, it has been suggested that pectin fragments could bind cell-wall calcium dependent associated kinases during development, antagonising auxin function in modifying the cell wall composition to favour cell elongation [168]. Mp3g01920 and Mp3g01930 are both kinases containing a calcium binding domain [168]. This finding could support the theory that the Mp*NOP1* gene might be involved in a kinase mediated signalling cascade that leads to cell wall modifications to promote the appearance of the initial intercellular apertures [98] (Figure 5.8 B).

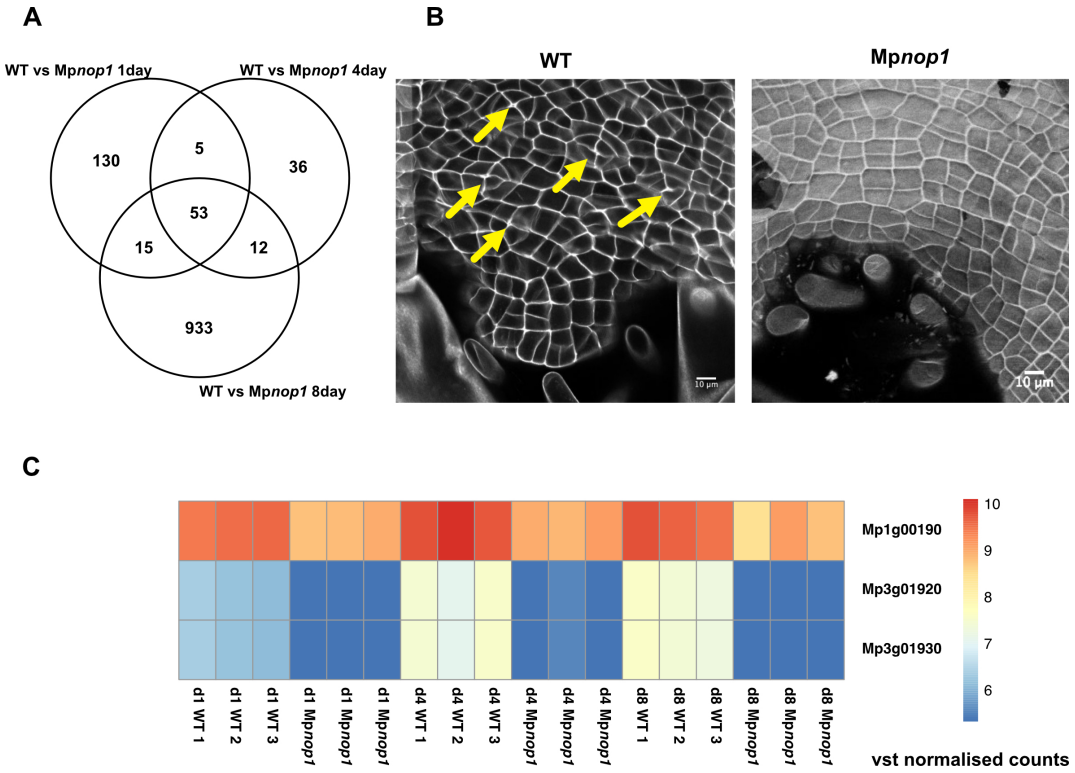


Figure 5.8 **Genes involved in cell wall remodelling which are differentially expressed in all WT vs Mp*nop1* comparisons.**
A) Venn diagram representing the number of DEGs shared in the WT vs Mp*nop1* comparisons amongst the three time points. B) Confocal microscopy images of the notch area of WT and Mp*nop1* 4 days old thalli fixed, cleared and stained with 0.01% w/v Calcofluor White M2R. Yellow arrows indicate the intercellular openings (at four ways cell junctions), which initiate air chambers' development in WT. C) Heatmap showing the vst transformed counts normalised to the library size of *Marchantia* pectinesterase gene Mp1g00190 and kinases Mp3g01920 and Mp3g01930 genes at 1 day, 4 days and 8 days. The vst transformation normalises the counts to the library size and removes dependence of the variance from the mean. This allows to diminish the effect of extreme values, but differences in expression level can still be detected. The blue-red gradient shows low to high values of normalised counts for each gene.

5.2.6 Analysis of differentially expressed transcription factors

In order to identify candidate transcription factors (TFs) that could play a role in air chamber development, differentially expressed TFs in each WT vs Mp*nop1* comparison were retrieved, as well as the ones, which were uniquely expressed during WT plants' development. The low numbers of statistically significant differentially expressed TFs, 18 amongst the three comparisons, could be due to

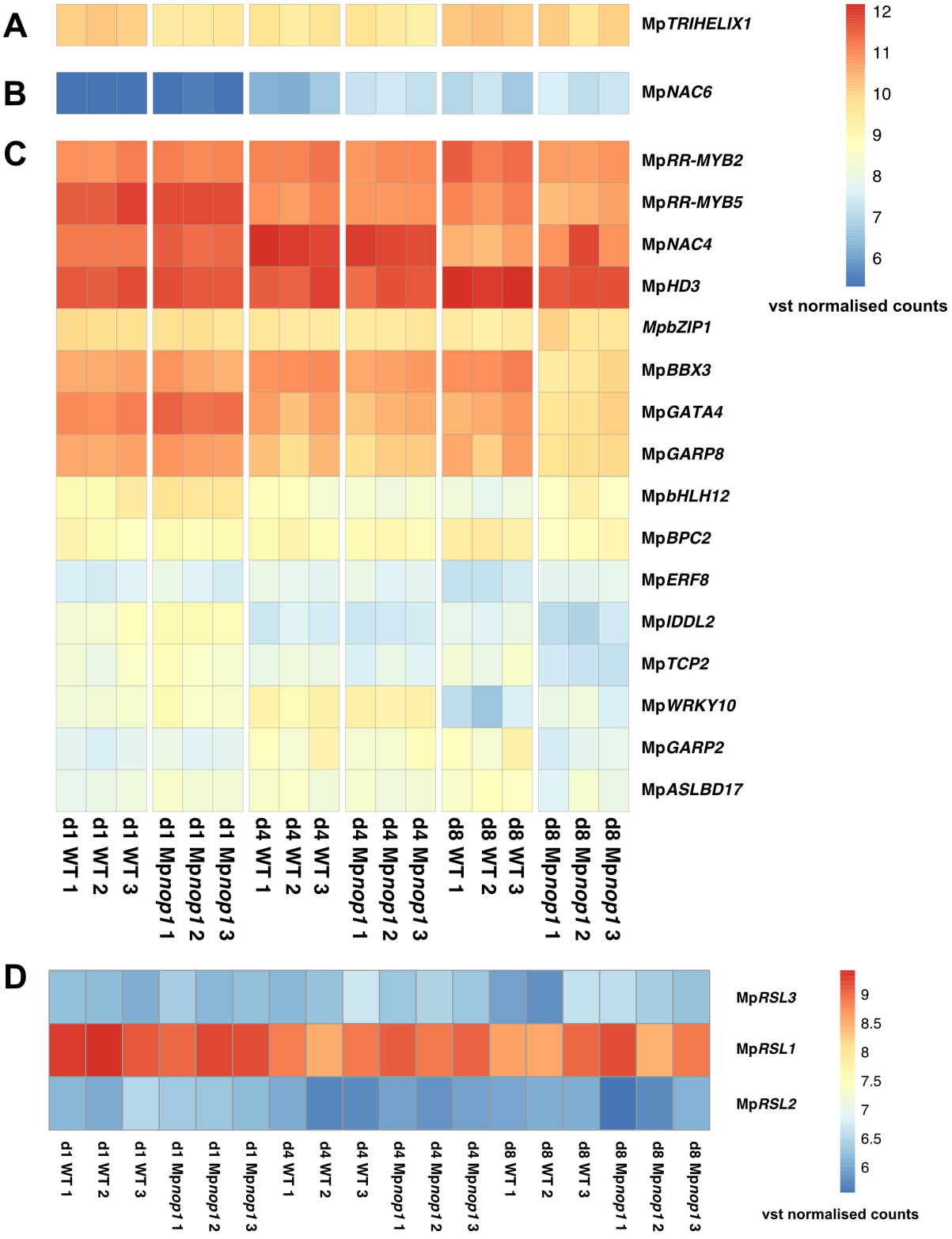
a lower overall expression level of regulatory genes compared to structural genes. Another reason could be that the structures studied in this dissertation, account only for a small portion of the plant tissue, thus differences in expression levels could be obscured by a higher number of cells belonging to different tissues.

In Figure 5.9, a heat map shows the level of gene expression indicated as vst transformed and normalised counts for each of the 18 TFs identified in the WT vs Mp*nop1* comparison. The vst function normalises the counts to the size of the library and removes variance dependence from the mean. Genes that show more counts across all samples are shown in red, while the ones in blue are less expressed. In order to validate that the identified genes were differentially expressed, vst normalised counts of the Mp*RSL1*, Mp*RSL2* and Mp*RSL3* genes, which do not appear to be differentially expressed in WT to Mp*nop1* comparisons, were included in the heat map. Mp*RSL* genes showed similar expression levels in both WT and Mp*nop1* samples, across the three different time points (Figure 5.9 D). In the d1 comparison, the Mp1g16790 trihelix gene was the only differentially expressed TF, showing a higher expression in WT plants (Figure 5.9 A). In *Arabidopsis*, trihelix genes are known to regulate responses related to light and stress, embryo and flower development as well as the development of epidermal structures such as stomata and trichomes [169].

Mp*NAC6* (Mp5g01530) was the only differentially expressed TF in d4 samples (Figure 5.9 B), being highly expressed in Mp*nop1* mutant plants. The NAM (no apical meristem) - ATAF (Arabidopsis transcription activation factor) - CUC (cup shaped cotyledon), NAC, family has been widely investigated in *Arabidopsis* where it has shown to be involved not only in the development of the shoot apical meristem [36], but also in secondary cell-wall biosynthesis [170] and biotic and abiotic stress [171].

Sixteen different TFs were found to be differentially expressed in the WT vs Mp*nop1* comparison at 8 days (Figure 5.9 C). These included the *Marchantia bHLH12* gene (Mp2g05070) and *Marchantia* Mp*WRKY10* gene (Mp7g06550). Mp*bHLH12* belongs to the IIIf subfamily, which regulates rhizoid and trichome patterning in vascular plants [172]. In this process, bHLH genes interact with WD40 and R2R3MYB proteins (MYB-bHLH-WD40 protein complex - MBW) to positively regulate the expression of WRKY TFs, which act upstream of the homeodomain transcription

factors [173], *GLABRA2* (*GL2*), involved in the initiation of the root hair and trichome developmental programs [173]. Hence, Mp*bHLH12* and Mp*WRKY10* represent excellent candidates as markers for epidermal patterning.


A
B
C
D
MpTRIHELIX1
MpNAC6
MpRR-MYB2
MpRR-MYB5
MpNAC4
MpHD3
MpbZIP1
MpBBX3
MpGATA4
MpGARP8
MpbHLH12
MpBPC2
MpERF8
MpIDDL2
MpTCP2
MpWRKY10
MpGARP2
MpASLBD17
12
11
10
9
8
7
6
vst normalised counts
d1 WT 1
d1 WT 2
d1 WT 3
d1 Mpnop1 1
d1 Mpnop1 2
d1 Mpnop1 3
d4 WT 1
d4 WT 2
d4 WT 3
d4 Mpnop1 1
d4 Mpnop1 2
d4 Mpnop1 3
d8 WT 1
d8 WT 2
d8 WT 3
d8 Mpnop1 1
d8 Mpnop1 2
d8 Mpnop1 3
MpRSL3
MpRSL1
MpRSL2
9
8.5
8
7.5
7
6.5
6
vst normalised counts

E **WT vs Nop log2 fold change**

	day 1	day 4	day 8
Mp*RSL3*	–0.3475503	0.09252986	–0.5408187
Mp*IDDL2*	–0.543667	0.09616354	1.08653245
Mp*TRIHELIX1*	0.67993103	0.15042954	0.28470248
MpbZIP1	0.09136737	–0.0565649	–0.4192325
Mp*TCP2*	–0.3768078	0.41102908	1.46272901
Mp*bHLH12*	–0.5795497	0.40529432	–0.9509467
Mp*RR-MYB2*	–0.0788191	0.21178668	0.61487381
Mp*RSL1*	0.15015125	–0.2851805	–0.1773743
Mp*BBX3*	–0.0765982	0.27087348	1.44907818
Mp*GARP2*	–0.124776	0.23192516	1.40758914
Mp*NAC4*	–0.2515103	0.09955657	–0.7866098
Mp*NAC6*	–0.7666754	–1.7848789	–0.5703833
Mp*RR-MYB5*	–0.0980634	0.11263746	0.55000991
Mp*ERF8*	–0.2276973	0.06863225	–0.8134273
Mp*RSL2*	–0.0797004	–0.2824525	0.80677171
Mp*GATA4*	–0.3728519	0.16661748	0.82704404
Mp*WRKY10*	–0.4833496	–0.0825284	–1.1727657
Mp*GARP8*	–0.1714003	0.15994826	0.72703944
Mp*HD3*	–0.01745	0.07839222	0.41448239
Mp*ASLBD17*	–0.5164391	0.15810764	0.78835887
Mp*BPC2*	0.01684169	0.01718597	0.58118485

Figure 5.9 **Heatmap of vst normalised counts of DE TFs in WT vs Mp*nop1* comparison for each time point.**
The heatmap shows transcript counts normalised with the vst function from each TF in all the generated libraries. Blue squares indicate a lower number of counts, while red squares indicate higher expression levels. A) Mp*TRIHELIX1* is the only TF DE in d1 samples between WT and Mp*nop1*. It appears to be more abundant in WT plants. B) Mp*NAC6* transcript abundance increase from d1 to d4, where they are more abundant in WT samples. C) List of the 16 differentially expressed TFs on the d8 comparison. These included different families of TFs. **ARR-B**: Arabidopsis Response Regulators/ B-motif (GARP-like motif) binding (Mp*GARP8*); **BBR-BPC:** Barley B Recombinant (BBR) - BASIC PENTACYSTEINE1 (BPC1), (Mp*BPC2*); **bHLH**: Basic helix loop helix (Mp*bHLH12*); **bZIP**: Basic leucine zipper (Mp*bZIP1*); **C2H2**: CCHH (Zn), (Mp*IDDL2*); **CO-like**: CONSTANS-like (Mp*BBX3*); **ERF**: ethylene response factor (Mp*ERF8*); **G2-like**: Golden2 (G2)-like (Mp*GARP2*); **LBD**: ASYMMETRIC LEAVES2/LATERAL ORGAN BOUNDARIES (Mp*ASLBD17*); **MYB**: Myb proto-oncogene protein (Mp*RR-MYB2*, Mp*RR-MYB5*); **NAC**: NAM, ATAF1,2 and CUC2 (Mp*NAC4*, Mp*NAC6*); **TCP**: TEOSINTE-LIKE1, CYCLOIDEA, and PROLIFERATING CELL FACTOR1 (Mp*TCP2*); **ZF-HD**: Zinc finger homeodomain protein (Mp*HD3*). D) Vst normalised counts for the three bHLH TFs involved in rhizoids development: Mp*RSL1*, Mp*RSL2* and Mp*RSL3*. These genes are not differentially expressed in the WT vs Mp*nop1* comparison, and the transcript abundance has similar levels in WT and Mp*nop1* samples and across time points. E) The table contains the expression log2 fold change for all the genes presented in the heatmap.

5.2.7 Investigation of the potential role of bHLH Ib, IIIb subfamily and WRKY in air chamber development

As described in Sections 1.3.1 and 1.3.2, different bHLH genes are involved in patterning in vascular plants. *Marchantia* homologues of bHLH subfamilies involved in patterning and morphogenesis were identified based on the phylogenetic analysis performed by Catarino et *al* [174], and the number of genes in each subfamily is definitely smaller than in *Arabidopsis* (Figure 5.10).

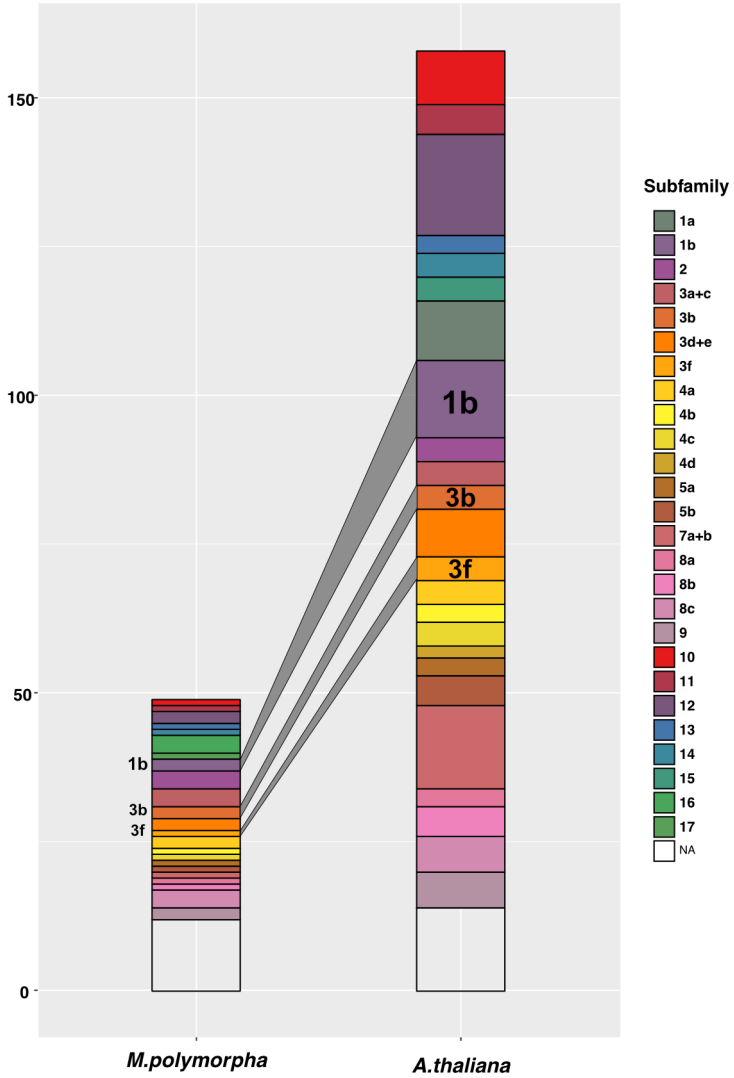


Figure 5.10 **Number of genes in each bHLH subfamily in *Marchantia* and *Arabidopsis*.**
The classification follows the phylogenetic analysis published by Catarino et al [174]. The different colours indicate different families. In bold are the names of the bHLH subfamilies involved in epidermal patterning in *Arabidopsis thaliana*.

The Ib subfamily is involved in the regulation of embryonic development of

epidermis, controlling cell adhesion and cuticle formation and *Marchantia* has two genes belonging to this subfamily: the Mp*bHLH38* and Mp*bHLH24* genes. The IIIb subfamily, which in vascular plants controls stomata patterning has two homologues in *Marchantia* Mp*bHLH16* and *MpbHLH17*. Finally, in the IIIf family includes genes involved in the formation of the MBW (MYB-bHLH-WRKY) protein complex (see section 1.3.1), which controls the development of root hair and trichomes in vascular plants, *Marchantia* has a single homologue Mp*bHLH12* (Mp2g05070). None of these genes, with the exception of Mp*bHLH12* passed the filtering threshold of a padj<0.01, but considering the small proportion of cells compared to the entirety of the plant tissue, the expression of these genes was explored in the different samples, calculating the transcript abundance as number of transcripts per million (TPM) for each time point and each genotype. (Fig 5.11).

Mp*bHLH12* showed a decrease in abundance throughout WT development, while Mp*nop1* plants increased the gene expression level at 8 day. In the Ib subfamily, Mp*bHLH24* transcript abundance does not differ between WT and Mp*nop1* mutant samples, until 8 day, where an increase in WT samples can be observed. Mp*ZOU1* is more abundant in Mp*nop1* plants than WT, but in both samples a decrease can be identified at 4 days. Mp*ICE1* and Mp*ICE2* were both showing a consistent increase in their transcript abundance in Mp*nop1* samples, while in WT plants they drop at 4 days and increase again at 8 days.

Furthermore, amongst DEGs, a WRKY TF, Mp*WRKY10* showed differential expression in the 8 days comparison between WT and Mp*nop1* gemmalings. Mp*WRKY10* (Mp7g06550) transcripts increased at 4 day and decreased again at 8 day, being more abundant in Mp *nop1* compared to WT plants (Figure 5.11). Mp*WRKY10* shows high homology to WRKY genes in higher plants, which include genes regulated by the MBW complex, making it an interesting target for downstream analysis. The presence of differences in transcript abundance of bHLH and a WRKY TFs between WT and Mp*nop1* plants at key stages of air chamber formation, makes these gene families an interesting target for further investigations.

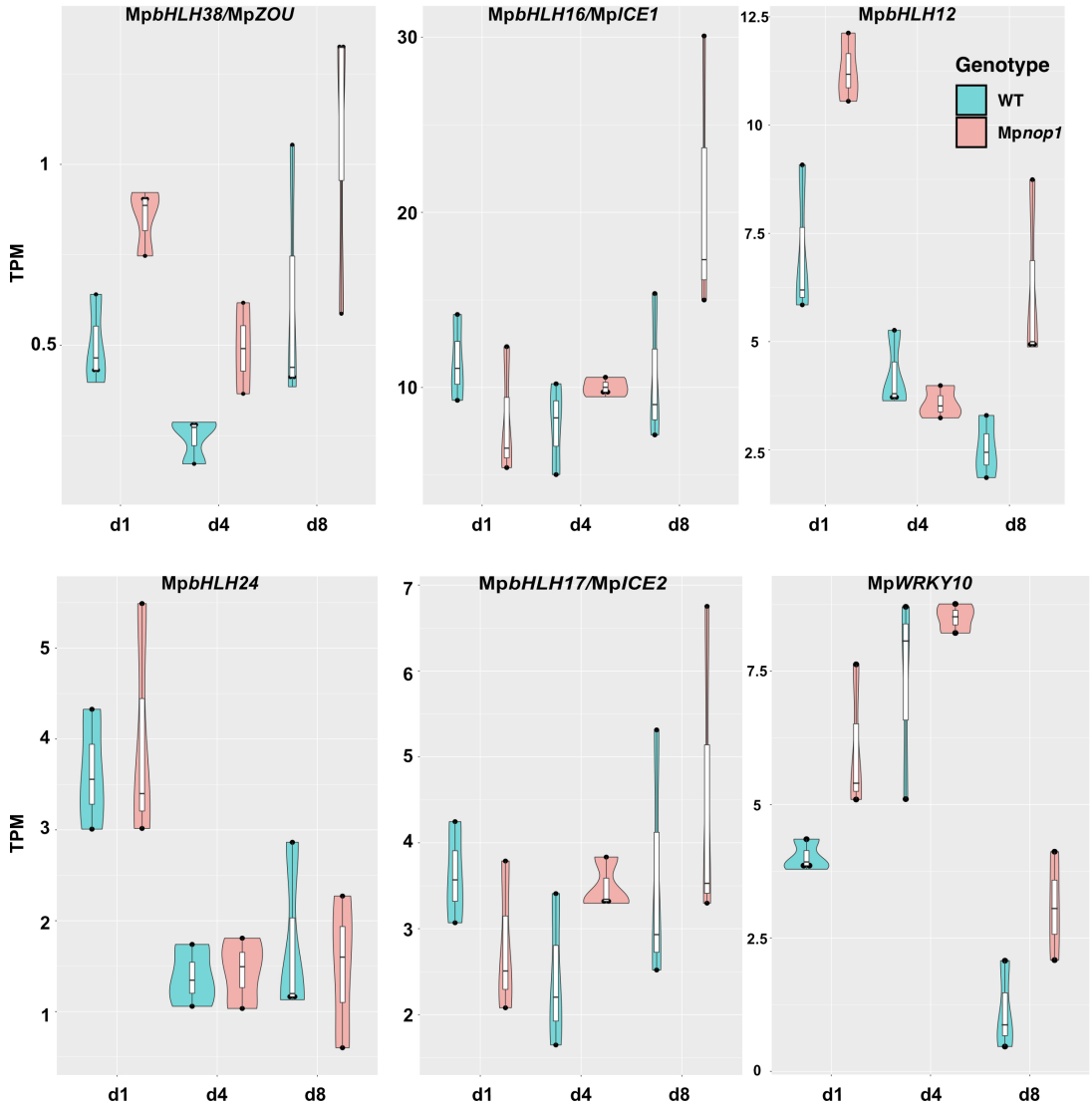


Figure 5.11 **Transcript per million (TPM) of *Marchantia* bHLH and WRKY gene candidates in the control air chamber development.**
The violin plot with the dots shows the transcript per million (TPM) value for each genotype (WT in blue and Mp*nop1* in pink), the boxplot in the middle shows the mean and standard deviation. Mp*bHLH24* shows a similar expression in both genotypes at 1 and 4 days, but it appears to diverge at 8 days. Mp*ICE1* and Mp*ICE2* show a higher abundance in Mp*nop1* at each time point. Mp*bHLH12* shows a decrease in WT plants, while its expression increases again in Mp*nop1* at 8 days. The expression of Mp*WRKY10* follows the same trend in both genotypes, but it appears to be more abundant in Mp*nop1* plants.

5.3 Discussion

5.3.1 Cellular differences between WT and Mp*nop1* plants represent a resource for the identification of cellular markers

The Mp*nop1* mutant represents a unique resource for the study of air chamber development, as it possesses all the cell types that are present in WT plants, but on the dorsal surface there are no air chambers, hence no air pore and no photosynthetic filaments (Figure 5.2). The complete lack of these cell types could translate into a difference in expressed genes, especially those controlling air pore or photosynthetic filament development. This difference could be captured in the transcriptomic analysis of the two different genotypes at key developmental stages of air chamber development.

Transcriptomic analysis has often been used as a starting point for functional studies of genes [175, 176, 177]. Given the different cellular composition, the observed differences in transcript abundance between the WT and Mp*nop1* genotypes might be due to differences in expression levels as well as different numbers of cells between WT and Mp*nop1* plants and the three observed time points. Nevertheless, genes with different number of transcripts represent excellent candidates as cellular markers to track the cellular process of air chambers' formation. The comparison of transcriptomic data between WT and Mp*nop1* mutant showed that the two genotypes are closely related as the retrieved DEGs are involved in the same biological processes (DNA metabolism, transport and protein modification), with the exception of photosynthesis in d8 samples (Figure 5.7). This might be attributed to the fact that WT plants possess photosynthetic filaments, which are instead absent in Mp*nop1* plants. Nevertheless, differentially expressed genes involved in cell wall remodelling and transcription factors known to be involved in cell patterning in higher plants (bHLH and WRKY) were identified when comparing the transcriptome of Mp*nop1* plants to WT ones.

5.3.2 DEGs involved in cell wall remodelling support the proposed model for air chamber origin

The presence of differentially regulated genes involved in cell wall remodelling in the WT vs Mp*nop1* mutant comparison, is consistent with the model for air chamber origin proposed by Ishizaki *et al.* [98]. Wall associated kinases (WAKs) have been shown to be involved in plant development, where they antagonise auxin effects and inhibit cell elongation in stems or prevent rhizogenesis [178, 179], other than in pathogen responses [168]. WAKs are able to tightly bind products produced by the degradation of pectin, one of the key components of cell wall deposition and cell growth. Plant cells change their cell wall composition in response to mechanical stimuli generated within the cells or by neighbouring cells. Plant receptors involved in cell wall remodelling, like kinases belonging to the WAK group, have been identified both in *Physcomitrella* and *Marchantia*, but they are not present in charophytes, indicating a potential adaptation for life on land [180]. The formation of intercellular space triggers the formation of air chambers and requires cell wall modifications to take place in order for the cells to detach and later open a pore. The high level of expression of the WAK proteins (Mp3g01920 and Mp3g019301 genes) in WT plants, which is instead quite low in the mutant, could imply a signalling cascade triggered by cell modifications which are not occuring in the mutant plants.

5.3.3 bHLH and WRKY transcription factors are excellent candidates as cellular markers for air chamber development

DE analysis led to the identification of a short list of TFs, which could have a role in air chamber development. Four bHLH subfamilies (Ib, IIIb, IIIf and a group sister to IIIf) and one WRKY gene were identified as potential transcriptional regulators of air chambers' formation. It has already been shown that in *Marchantia* homologs of the bHLH VIIIc (*RSL* genes) subfamily, which controls root hair in *Arabidopsis*, regulate rhizoids development [11]. Similarly, the bHLH IIIb subfamily is required for the coordination of stomata development in *Arabidopsis* [181], and controls stomata formation in *Physcomitrella* sporophytes [10]. The conservation of these genetic circuits in the control of structures with similar morphologies but different physiological roles, makes it reasonable to assume that conserved circuits, which could be conserved in vascular plants, might control the development of air

chambers.

One of the candidate genes identified belongs to the IIIf bHLH subfamily, which regulates trichomes development in *Arabidopsis* [172]. *Marchantia* has a single gene belonging to this subfamily and its overexpression leads to a change in size and frequency of gemma cups emerging on the thallus [182]. Although no knock out mutants were produced for the IIIf mutants to confirm this hypothesis, this gene does not appear so far to be involved in air chamber development. Two other genes identified in the DE analysis, cluster in a sister group to the one of IIIf genes and, therefore, they could be interesting targets for genetic manipulation.

The second family of TFs identified in the study of WT and Mp*nop1* transcriptomes is the WRKY family. *WRKY* genes are present in all land plants and green algae, and their numbers have expanded during plant evolution: from a single copy in *Chlamydomonas reinhardtii*, 14 in *Marchantia*, 37 in *Physcomitrella*, to 74 in *Arabidopsis* and more than 100 in rice, *Oryza sativa* [183]. The WRKY families include genes controlling plant developmental processes, such as seed development and trichome development, as well as responses to biotic stresses. In *Arabidopsis*, a WRKY TF is also involved in trichome and root hair development. This protein is regulated by the MBW complex, which induces specific cell fate in epidermal cells: it either triggers trichome development in leaves or root hairs in the root [173]. WRKY TFs appear to be differentially expressed in *Marchantia* during pathogen infection [184], but no literature is available on the role of *WRKY* genes in liverworts' development.

Studying the function of the identified Mp*WRKY10* gene and testing its relationship with the Ib, IIIb and IIIf bHLH families could shed light on possible conserved mechanisms, involving bHLH and WRKY controlling epidermal patterning in early land plants. The results produced provide a new dataset, which captures genes potentially involved in air chamber morphogenesis, which can be used as cellular markers for the process.

Chapter 6

Characterisation of cell specific markers for air chamber development

Part of the content of this chapter has been published in:
Susanna Sauret-Güeto, Eftychios Frangedakis, Linda Silvestri, Marius Rebmann, Marta Tomaselli, Kasey Markel, Mihails Delmans, Anthony West, Nicola J. Patron, and Jim Haseloff, "Systematic Tools for Reprogramming Plant Gene Expression in a Simple Model, *Marchantia polymorpha*" [3].

6.1 Introduction

Large scale reporter gene fusions have been exploited to study gene expression in the plant model *Arabidopsis thaliana* [185, 186, 187, 188]. In this kind of screens, reporter genes, typically the luciferase (LUC), the β-glucuronidase (GUS) or the green fluorescent proteins (GFP), are fused to the promoter of a gene of interest to study its temporal and spatial expression pattern during plant development [189]. Numerous studies published at the end of the 1980s and the beginning of the 1990s took advantage of this experimental technique to identify key regulators involved in a variety of processes such as the circadian cycle, hormone signalling pathways and responses to stress, which would have been difficult to identify with conventional phenotypic screens [190, 191, 192]. Epidermal patterning takes place in the outer layer of the plant, hence it is more accessible for visualisation. In the plant epidermis, the use of reporter genes, allows the direct observation of correlations between gene expression and cellular processes, without the need of laborious sample preparation. Stomata [54], trichomes [193] and root hair development have rigorously been documented through the visualisation of reporter gene fusions. The study of their expression pattern [194, 109, 108] allowed the identification of transcription factors (TFs) regulating the cellular changes, which lead to the correct patterning of epidermal structures.

Synthetic biology approaches have been applied in plant systems allowing for fast and efficient assembly of DNA constructs. Most of these DNA assembly protocols are based on Type IIS endonucleases, which cut at a fixed distance from their DNA binding sequence, leaving non-complementary 'sticky overhangs' that can be customised. DNA sequences can be digested and ligated in a one-pot reaction, making this type of assembly highly efficient. Type IIS DNA assembly methods are often combined with the use of standardised DNA parts. These are DNA sequences corresponding to genetic functions (e.g, promoters, protein coding sequences, etc.), containing specific four nucleotides long 5' end and 3' end fusion sites [109]. When the parts are used in combination with a Type IIS DNA assembly method, they generate the correct overhang sequences to be assembled into transcriptional units (TUs) in a one pot reaction. Golden Gate Cloning based toolkits have been developed and tested in vascular plants [195, 194] and a common syntax, including a set of unique overhang sequences for the ordered assembly of plant DNA parts [109], has now been implemented to facilitate the exchange of sequences. Moreover, a new simple, recursive Type IIS method, called Loop assembly [108], has further facilitated the design and building of DNA circuits. This novel technique is based on the recursive use of two Type IIS endonucleases: *BsaI* and *SapI*. Standardised DNA parts, called Level 0 (L0) can be assembled into TUs, named Level 1 (L1) parts, using *BsaI* and DNA ligase. Four TUs can be combined together, digesting them with *SapI* (*LguI*) and joining them with DNA ligase, to produce Level 2 (L2) parts. Thanks to the recursive nature of the assembly, L2 parts can be digested with *BsaI* and ligated in Level 3 constructs containing up to 16 TUs, which can in turn be digested with SapI and ligated into a Level 4 (L4) construct with up to 64 TUs.

These technical advances can be used to design large scale reporter gene screens in *Marchantia polymorpha*. As mentioned in Section 1.6, *Marchantia* has a small sized genome with low gene redundancy, which contains a total of only 398 TFs [100]: a small number when compared to the one of vascular plants. Thus, it was possible to start building a complete collection of TF promoter - fluorescent protein fusion constructs, using the Loop assembly, which can then be used to transform *Marchantia* plants [3]. Transgenic *Marchantia* plants expressing those constructs can then be used to study the expression pattern of the TF promoter - fluorescent protein fusion. This type of screen could help to identify genes with conserved or divergent functions in plant epidermal patterning.

In this chapter, I will describe how I designed and performed a reporter gene screen where the promoter sequences of TFs potentially involved in epidermal patterning were fused

to fluorescent proteins and transformed into *Marchantia*. Successfully transformed plants were selected and their gemmae were screened with confocal microscopy to visualise the expression pattern of fluorescent proteins under the tested promoters. My aim was to identify gene promoters leading to a cell type specific expression of fluorescent proteins in air chambers. Moreover, to screen large numbers of gemmae a high-throughput set up for imaging gemmae was developed. The screen led to the identification of potential markers of air chamber development.

6.2 Results

6.2.1 Synthetic promoter design and synthesis

The sequences of *Marchantia* TFs promoters and 5' untranslated regions (5'UTR) were extracted from the *Marchantia* Cam-1 genome assembly [167] by Dr Mihails Delmans. Synthetic promoters for each TF were designed as 1.8 kb sequences upstream of the annotated 5'UTR of the gene, with the aim of including the core promoter region as well as some of the gene specific cis-regulatory regions to observe cell-type specific expression [196, 197, 198]. The 1.8 kb size of the parts was determined by the sequence size limit for synthesised fragments at the time of the order.

Promoters and 5'UTRs constituted two separate DNA parts, which were domesticated to eliminate any *SapI* or *BsaI* recognition sites (Section 2.1.4) and sent for synthesis or amplified from Cam-1 genomic DNA by PCR. The generated DNA parts were cloned into the new universal acceptor vector, pUAP4, derived from the published pUAP1 [8]. The pUAP4 vector substitutes the pUAP1 *BbsI* sites in the cloning region, with *SapI* restriction sites to make the vector compatible with Loop Assembly [108]. Moreover, the pUAP1 red fluorescent protein expression cassette has been replaced with one expressing the α subunit of the β-glucuronidase to allow for blue-white screening in transformed bacteria. Out of 398 TF, 363 were generated as L0 parts to be used as DNA parts for Loop Assembly. The promoters presented in this chapter were selected based on candidate genes identified in the transcriptomic data analysis described in Chapter 5 or homology with TFs regulating epidermal patterning in vascular plants. The 5'UTRs of the genes presented in this chapter did not contain introns. Since the aim was to test the endogenous activity of these promoters, no intron was added to their sequence to enhance the expression level [199]. Figure 6.1 shows the size of the synthetic promoter (green rectangles) and 5'UTR (orange rectangles) of each gene, which were ordered as L0 parts. It also includes for clarity the gene coding sequence (CDS in cyan), and the genes 3'untranslated regions (3'UTR) shown in yellow.

6.2.2 Plasmid design

Loop Assembly allows positional cloning of DNA parts using a set of specific ordered overhangs. There are two sets of acceptors called pCk (pCk-1, pCk-2, pCk-3 and pCk-4) used in odd level reactions(L1, L3, ...) and pCs (pCsA, pCsB, pCsC, pCsE) used in even reactions (L2, L4, ...) which allows the ordered assembly of TUs.

To screen for the expression pattern of fluorescent proteins fused to the above mentioned

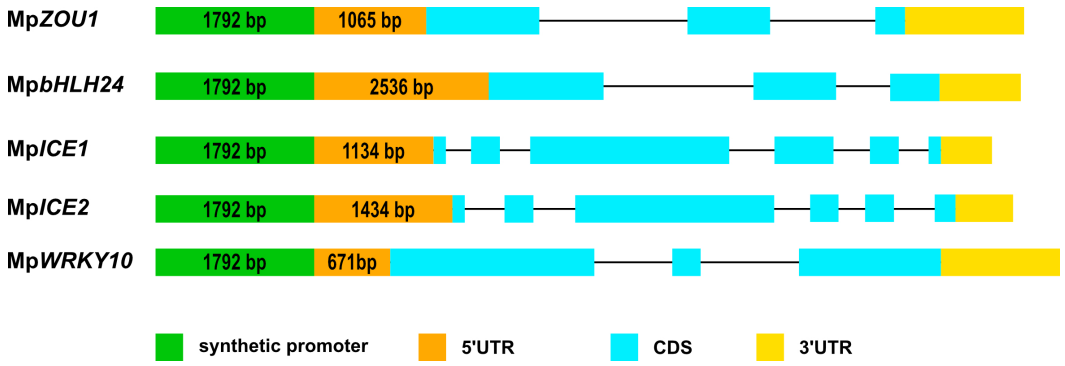


Figure 6.1 **Diagram of the structures of the genes selected for the reporter genes screening.**
A representation of the Mp*ZOU1*, Mp*bHLH24*, Mp*ICE1*, Mp*ICE2* and Mp*WRKY10* gene structures is shown. The synthetic promoter and 5'UTR generated as L0 parts are shown in green and orange respectively. The numbers indicate the length of each DNA fragment. The gene coding sequence (CDS) is shown in cyan and the 3' untranslated region (3'UTR) in yellow. Introns are shown with a black solid lines.

promoters, L2 plasmids were designed to be transformed in *Marchantia*. These were assembled using the following L1 parts: an antibiotic resistance TU for the selection of successfully transformed plants, two constitutively and ubiquitously expressed fluorescent marker TUs, one localised to the plasma membrane to identify the cell outline and one localised to the nucleus to compare with the expression pattern of the fluorescent proteins driven by the studied promoters. The hygromycin phosphotransferase II gene (*HPTII*) selection TU was assembled in pCK-1 to occupy the first position in the L2 construct. The nuclear and plasma membrane fluorescent markers were assigned to positions 3 (pCk-3) and 4 (pCk4). Initially, two constitutive promoters, the *Marchantia Tubulin* α gene promoter (proMp*Tub*α) and the Cauliflower Mosaic Virus 35S promoter (CaMV35s), were used to drive the expression of mTurquoise2 fused to the Arabidopsis N7 peptide [144] to target the protein to the cell nucleus, and the enhanced GFP (eGFP) expression targeted to the cell plasma membrane with the *Arabidopsis* Lti6b peptide [144], respectively.

During the screening, two new improved and highly expressed constitutive and ubiquitous *Marchantia* promoters were identified: the large subunit ribosomal protein (Mp*L40e*) gene promoter and the ubiquitin-conjugating enzyme E2 gene, Mp*UBE2*, promoter. These two promoters were used to drive the expression of the nuclear and membrane marker respectively, in the most recently assembled L2 constructs. The promoter under investigation and the 5'UTR sequences were used to drive the expression of mVenus in the cell nucleus using the N7 targeting peptide and they were assigned to position 2 (assembled in pCk-2 backbones).

A graphical summary of the constructs' design is shown in Figure 6.2 and a complete list of the L1 assembled plasmids used in this chapter is shown in Table 6.1.

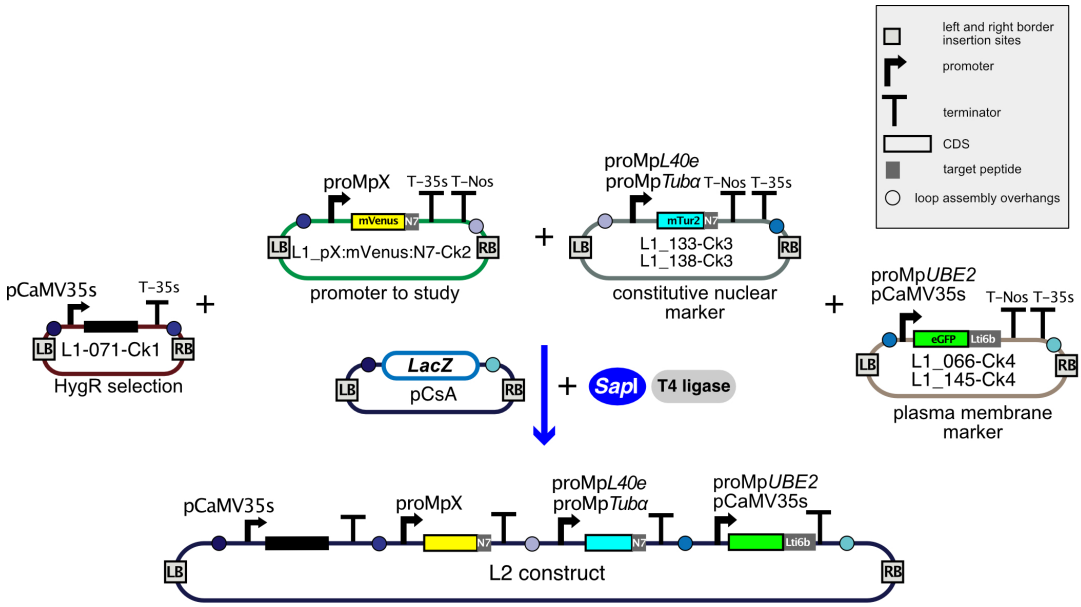


Figure 6.2 **Diagram of level 1 (L1) and level 2 (L2) plasmids used to assess promoter expression.**
Level 1 plasmids are shown at the top. The different coloured ellipses represent the 4 different backbone vectors. Filled circles inside the right and left border indicate the Loop assembly overhangs, which ensure the ordered assembly of the plasmids. Arrows indicate promoters and "T" terminator sequences: T-35s from the Cauliflower Mosaic Virus (CaMV) and T-Nos from *Agrobacterium* nopaline synthase gene. Fluorescent proteins are shown as coloured rectangles and peptides ensuring cellular localisation are shown with the grey squares positioned at the protein C-terminal end. Left border (LB) and Right border (RB) indicate the recombination sites for plant genomic integration. LacZ is the α subunit of the β-glucoronidase expression cassette, which allows for blue-white screening of assembled plasmid. The Ck1/2/3/4 at the end of L1 parts indicates the position occupied by the TUs in the final L2 construct. Level 1 parts were assembled in the L2 acceptor vector pCsA using *SapI* and T4 ligase (Chapter 2 section 2.1.4).

Table 6.1 List of assembled L1 and L2 plasmids.

Plasmid name	**Description**
L1_071-Ck1	pCaMV35s:*HPTII*:35sTerm
L1_1025-Ck2	proMp*ZOU1*:mVenus:N7:T35S:TNos
L1_1026-Ck2	proMp*ICE1*:mVenus:N7:T35S:TNos
L1_1027-Ck2	proMp*ICE2*:mVenus:N7:T35S:TNos
L1_1073-Ck2	proMp*bHLH24*:mVenus:N7:T35S:TNos
L1_1063-Ck2	proMp*WRKY10*:mVenus:N7:T35S:TNos
L1_133-Ck3	proMp*Tub*α:mTurquoise2:N7:TNos:T35S
L1_138-Ck3	proMp*L40e*:mTurquoise2:N7:TNos:T35S
L1_066-Ck4	pCamV35s:eGFP:Lti6b:TNos:T35S
L1_145-Ck4	proMp*UBE2*:eGFP:Lti6b:TNos:T35S
L2_1025-CsA	L1_071-Ck1, L1_1025-Ck2, L1_138-Ck3, L1_145-Ck4
L2_1026-CsA	L1_071-Ck1, L1_1026-Ck2, L1_133-Ck3, L1_066-Ck4
L2_1027-CsA	L1_071-Ck1, L1_1027-Ck2, L1_138-Ck3, L1_145-Ck4
L2_1073-CsA	L1_071-Ck1, L1_1073-Ck2, L1_138-Ck3, L1_145-Ck4
L2_1063-CsA	L1_071-Ck1, L1_1063-Ck2, L1_133-Ck3, L1_066-Ck4

6.2.3 Spore transformation and selection

The spore transformation protocol used in the laboratory was optimised to scale up the number of transformation carried out at the same time and ensure high efficiency [3]. *Marchantia* sporelings were transformed with the L2 constructs described on Table 6.1 using *Agrobacterium tumefaciens* mediated transformation (section 2.2.8) and plated on half strength Gamborg B5 media containing 1.2% (w/v) agar, hygromycin and cefotaxime with a concentration of 20 μg/ml and 100 μg/ml respectively. Successful transformants were screened under an epifluorescent microscope to assess the expression of the different fluorescent markers.

After two weeks, sporelings that survived the antibiotic selection and showed the expression of mTurquoise2 and eGFP were moved onto new selection plates containing 20 μg/ml of hygromycin and 100 μg/ml of cefotaxime. Since transformed sporelings could generate chimeric individuals, plants were grown until gemma cups were produced in transformed regions of the thallus (regions showing expression of mTurquoise and eGFP fluorescent markers). Gemmae, the clonal propagules produced inside gemma cups, were removed from the cups and moved onto new selection plates (half strength Gamborg B5, 1.2% w/v agar with hygromycin 20 μg/ml and cefotaxime 100 μg/m) to produce an isogenic mutant line, called G1. Gemmae produced by G1 plants (called G2) were then used for screening under the confocal microscope to assess the expression of mVenus fused

to the studied promoters.

A total of 77 TFs promoters have been so far fused to fluorescent protein reporter genes, transformed in *Marchantia* sporelings and screened for cell-type specific expression. 49 reporter genes were successfully expressed in gemmalings obtained from transformed plants: most of the promoters generated ubiquitous expression of the fluorescent reporter, but several lines with reporter genes specifically expressed in the notch, oil bodies, mucilage cells, rhizoids or pore cells were also identified.

6.2.4 Developing a procedure for high throughput screening of *Marchantia* gemmae

For each construct, 3 lines resulting from independent transformation events were selected and 6 gemmae from each line were screened in a 7 day long time course. In order to automate the screening of G2 gemmae, a new imaging setup was developed to allow for systematic imaging of large numbers of samples and also ensure a sterile growing environment for the gemmae for the entire duration of the time course.

Transparent 384 well plates with a clear base were filled with half strength gamborg B5 with 1.2% (w/v) agar and they were used to image up to 106 gemmae at the same time (see Chapter2, Section 2.3.5 for plate set up protocol). The clear bottom ensured that plants receive light from all sides, recreating a growing environment similar to the one of a Petri dish that are usually used for gemmae in laboratory conditions. Additional rows (and columns) of wells were filled with media around those containing the gemmae to image. This precaution slowed the evaporation of nutrient media from wells containing the samples (Figure 6.3 A).

As the wells were filled up to the rim with nutrient media, additional space was created by positioning a Gene Frame® on top of the plate to create room for the plant to grow during the 7 day time course. A cover slip was positioned on top of the Gene Frame® after being treated with the B-Fog anti fogging spray (Bollé safety) (Figure 6.3 B-D). This step prevents the formation of condensation on the cover slip due to plant transpiration. The plate set up was then recorded with the 'Mark and Find' option of the Leica LasX software, which stores the x,y, z coordinates of each well and z-stacks were recorded for each well at 0, 1, 4 and 7 day. All imaging parameters are described in Section 2.3.4 of Chapter 2, using the second set of sequential scanning settings.

The collected images were then processed using a macro in ImageJ [118] to produce a

z-plane projection of the acquired images for each individual channel. The maximum intensity option was used for each channel and then they were merged together to create a composite image (Figure 6.4). Combining the channels together, allowed for the identification of cell-type specific expression patterns and compared the level of expression of the different fluorescent markers. A full description of the macro commands can be found in Chapter 2 Section 2.3.6. This imaging setup is optimal for the identification of markers expressed on the dorsal epidermis. Thus, it facilitated the initial selection of markers potentially involved in air chambers' development.

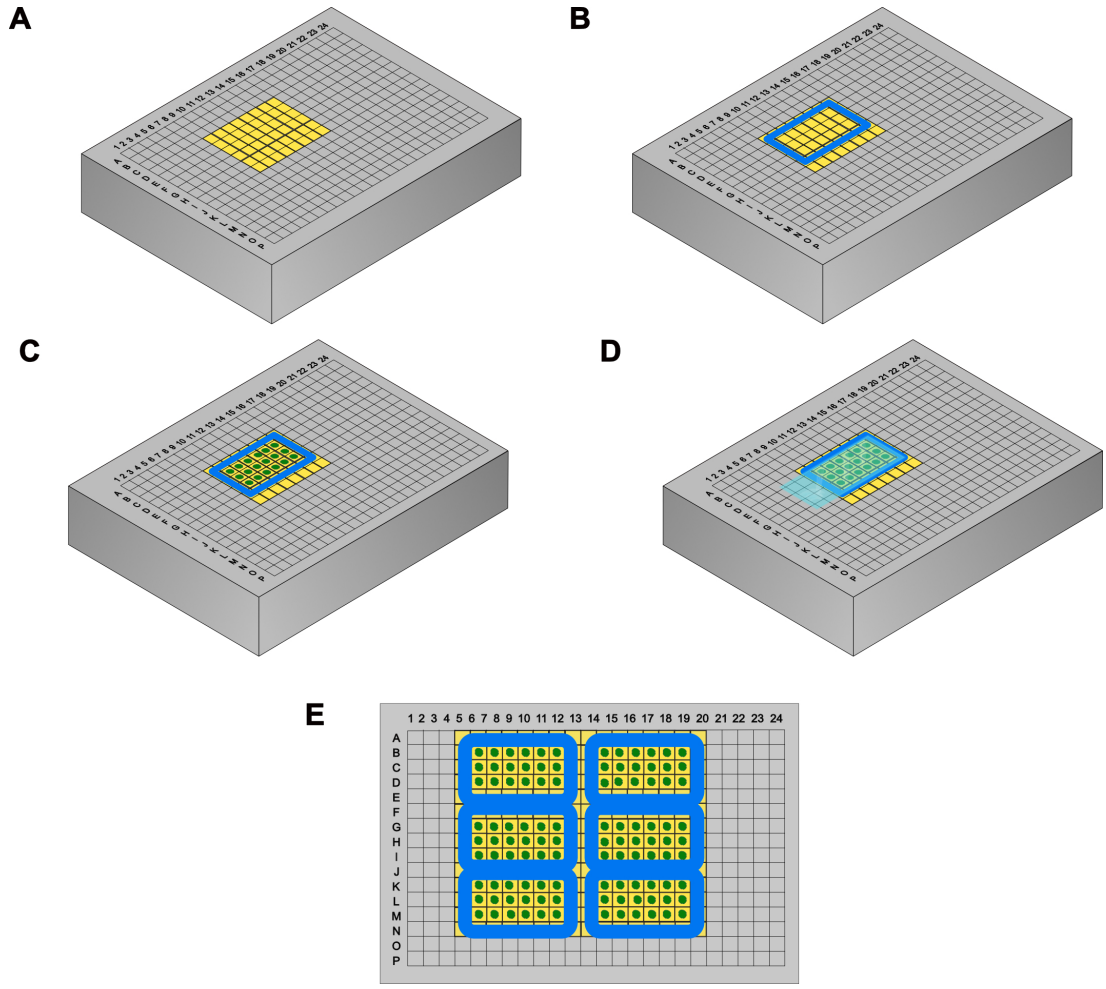


Figure 6.3 **384 multi-well plate set up for high throughput imaging.** Wells filled with media are shown in yellow. A) The wells are filled to the top with half strength Gamborg B5 1.2% (w/v) agar B) A gene frame is positioned around the filled wells making sure that there are extra rows and columns filled with media around it. C) Gemmae are placed in the centre of each well. D) A cover slip sprayed with an anti-fogging agent is positioned on top of the gene frame. E) View from the top of a plate filled with 106 samples.

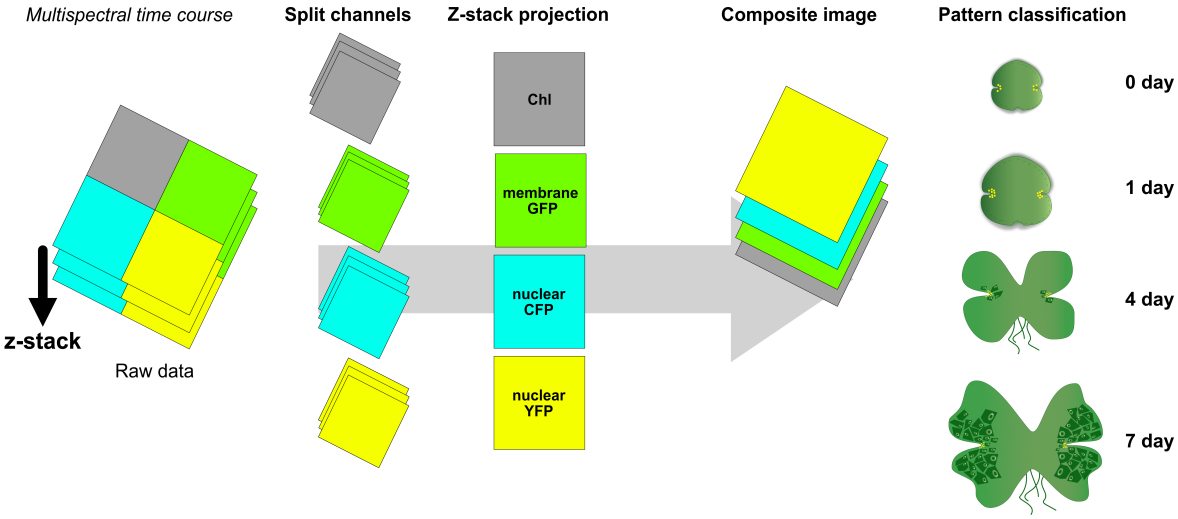


Figure 6.4 **ImageJ Macro diagram for fast processing of imaging data.** Multispectral z-stacks are generated for each imaged gemma on the multiwell plate. The macro (a set of ImageJ software instructions, which are performed repetitively on the input files) takes the raw data for each imaged gemma, it splits the z-stacks for each channel and it projects images on the z-plane using the max intensity function. The obtained projected images are then merged into a single image file, which allows the user to localise fluorescent proteins in different cell types at distinct stages of gemmalings development.

6.2.5 Analysis of the expression of mVenus under the control of bHLH transcription factor promoters involved in epidermal patterning

As discussed in the Introduction, certain subfamilies of the basic loop helix (bHLH) TFs regulate a variety of epidermal patterning processes in vascular plants, which include stomata and trichome and root hairs development. More precisely, the IIIb subfamily regulates stomata guard cell development, while the Ib subfamily regulates epidermal development of the embryo in *Arabidopsis*.

Marchantia has two genes in each of the Ib (Mp*ZOU1* and Mp*bHLH24*) and IIIb (Mp*ICE1* and Mp*ICE2*) bHLH subfamilies and from the transcriptomic analysis presented in Section 5.2.7 of Chapter 5, they show differences in transcript abundance during the development of WT and Mp*nop1* mutant plants. While Mp*ZOU1* appeared to be more highly expressed in Mp*nop1* plants at each of the time points examined, Mp*bHLH24* showed similar expression levels in both WT and mutant plants, with an increase in the number of transcripts at 8 days in Mp*nop1* samples. Mp*ICE1* and Mp*ICE2* showed higher numbers of transcripts in Mp*nop1* plants on day 4 and 8. Therefore, gemmalings transformed with construct containing the promoter of the gene of interest driving mVenus expression in the

cell nucleus, were screened for cell-specific expression of mVenus during a 7 day long time course. This time frame was able to capture the early stages of air chamber appearance [8].

The bHLH IIIb subfamily includes the two *Marchantia* homologs, Mp*ICE1* and Mp*ICE2*. Plants were transformed with the L2_1026-CsA and L2_1027-CsA constructs, which contained the following expression cassettes: proMp*ICE1*:mVenus:N7:T35S:TNos and proMp*ICE2*:mVenus:N7:T35S:TNos, respectively. The expression of mVenus under the control of the Mp*ICE1* gene promoter was restricted to few cells at the centre of the gemma in 0 day samples (Figure 6.5 b). On day 4, the expression of mVenus was further localised to cells in the central zone of the gemmalings (Figure 6.5 e) and it showed a similar pattern in 7 day samples (data not shown).

Gemmalings transformed with L2_1027-CsA, which contains proMp*ICE2*:mVenus:N7:T35S:TNos expression cassette, did not show any signal in the latter time points examined (Figure 6.5 h,k), even after increasing laser power and the detector's gain. The weak signal in the mVenus column is likely to be due to spillover signal from the eGFP protein, whose emission spectra partially overlaps with the ones of the eGFP expressed in the membrane.

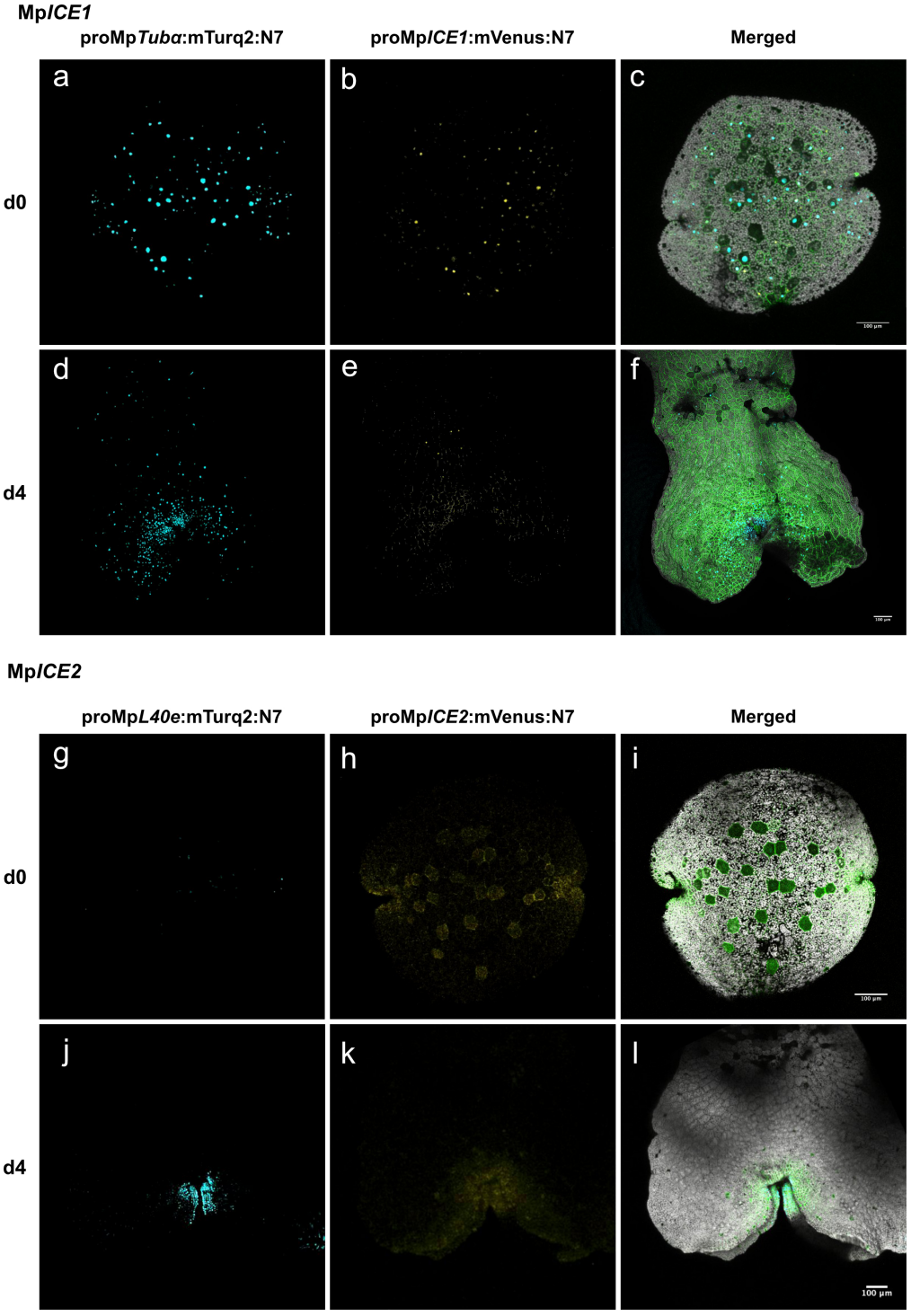

Mp*ICE1*
proMp*Tuba*:mTurq2:N7
proMp*ICE1*:mVenus:N7
Merged
a
b
c
d0
d
e
f
d4
Mp*ICE2*
proMp*L40e*:mTurq2:N7
proMp*ICE2*:mVenus:N7
Merged
g
h
i
d0
j
k
l
d4
100 µm

Figure 6.5 **The expression pattern of the proMp*ICE1*:mVenus:N7:T35S:TNos and proMp*ICE2*:mVenus:N7:T35S:TNos cassettes is not specific to air chamber related cell types.**
The figure shows the gemmae at 0 and 4 days after removal from the gemma cups. **A to f** show gemmae where mVenus expression is under the control of the Mp*ICE1* gene promoter at 0 day (**a-c)** and 4 day (**d-f**). **G to l** show gemmae where mVenus expression is under the control of the Mp*ICE2* gene promoter at 0 day (**g-i**) and 4 day (**j-l**).**A, d, g, j:** the expression of the constitutive nuclear fluorescent markers pro*MpL40e*:mTurq2:N7 and proMp*Tubα*:mTurq2:N7, is shown. **B, e, h, k**: the expression pattern of mVenus driven by the Mp*ICE1* and Mp*ICE2* genes' promoters can be observed. **C, f, i, l:** merged images including the plasma membrane marker in green (proMp*UBE2*:eGFP:Lti6b paired with proMp*ICE1* or CaMV35s:eGFP:Lti6b paired with proMp*ICE2*) and chlorophyll in grey.

Gemmalings transformed with the L2_1025-CsA plasmid, carrying the proMp*ZOU1*:mVenus:N7:T35S:TNos cassette, showed mVenus expression in all the cells of 0 day samples (Figure 6.6 a-c). The expression was detected in most of the dorsal cells of the thallus at day 4 (Figure 6.6 d-f) and 7 (data not shown). In 4 day gemmalings mVenus expression was observed to be scattered on the developing thallus, with stronger signals in cells localised on the margins of the thallus.

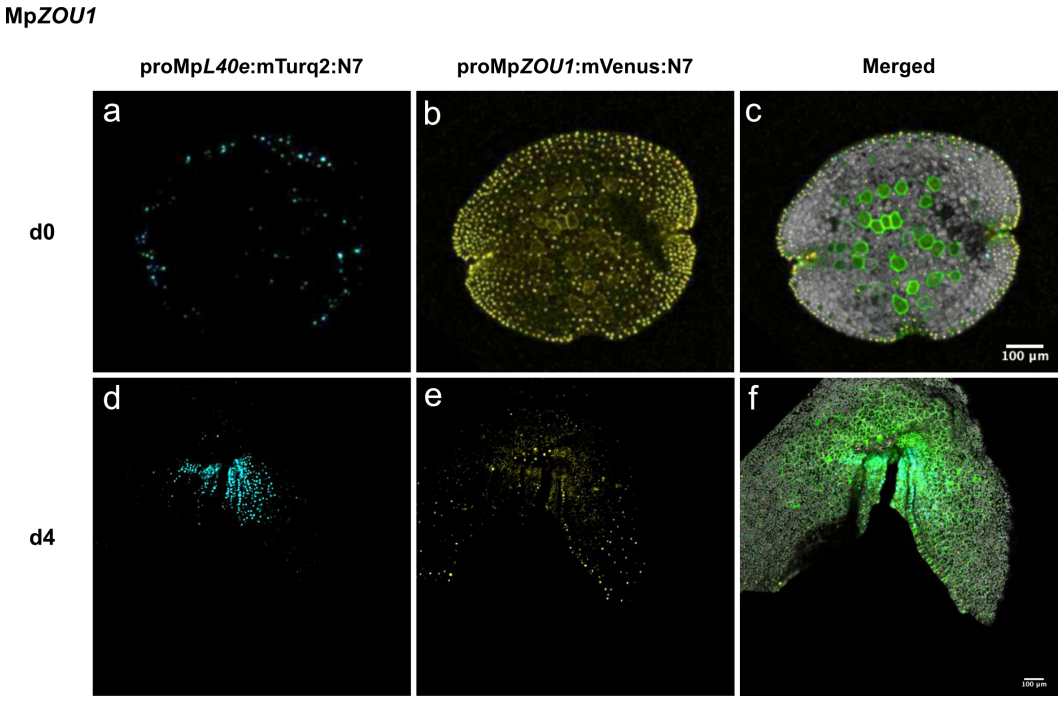


Figure 6.6 **The expression pattern of proMp*ZOU1*:mVenus:T35S:TNos cassette localises to marginal cells and epidermal cells.**
The figure shows the gemmae at 0 and 4 days after removal from the gemma cups. A to f show a gemma where mVenus expression is under the control of the Mp*ZOU1* gene promoter at 0 day (**a-c**) and 4 day (**d-f**). A and d: the expression of the constitutive nuclear fluorescent markers pro*MpL40e*:mTurq2:N7 is shown. **B and e**: the expression pattern of mVenus driven by the Mp*ZOU1* promoter can be observed. **C and f**: merged images including the plasma membrane marker in green (proMp*UBE2*:eGFP:Lti6b) and chlorophyll in grey.

The second member of the bHLH Ib subfamily is the Mp*bHLH24* gene. Gemmaling carrying the L2_1073-CsA plasmid containing the proMp*bHLH24*:mVenus:N7:T35S:TNos expression cassette, showed mVenus expression in most of the cells of a 0 day gemmae, with stronger signal in the cells on the gemma outline. The mVenus expression domain expanded to the centre of the thallus in 4 day samples and finally in day 7 gemmalings it showed strong expression on the pores of air chambers farthest from the notch area (Figure 6.7 a-i). When looking at air chambers, it can be observed that Mp*bHLH24* promoter drives the transcription of mVenus also in other dorsal cells, but it appears to be slightly brighter in the cells forming the air chamber pore (Figure 6.7 j-l) compared to the expression in the nucleus of the dorsal epidermis cells.

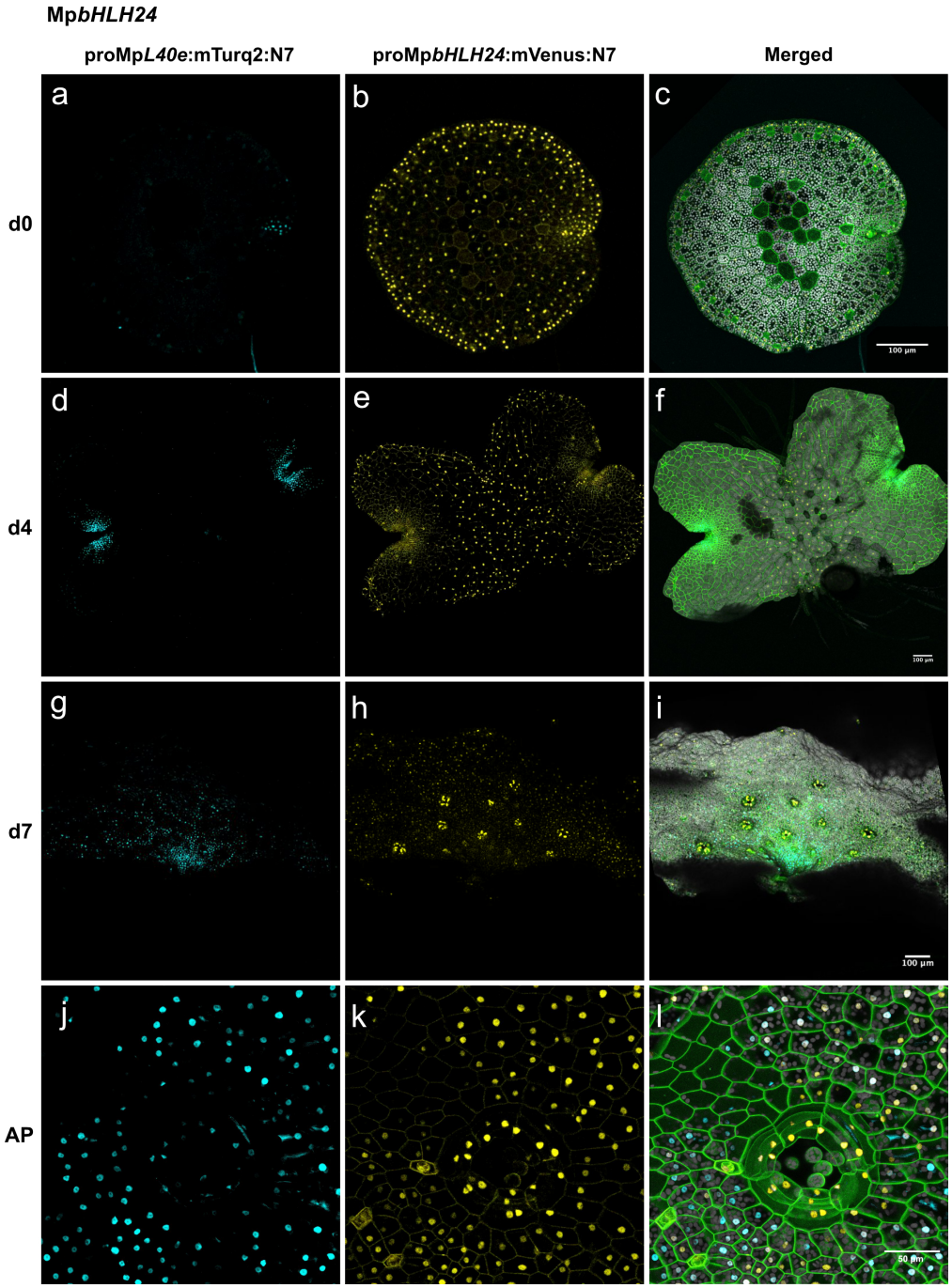


Figure 6.7 **The expression pattern of proMp*bHLH24*:mVenus:N7:T35S:TNos cassette localises to marginal cells in 0 day gemmae and it expands to the air pore in 7 days old gemmalings.**
The figure shows gemmae at 0, 4 and 7 days after removal from the gemma cups. A to i show a gemma where mVenus expression is under the control of the Mp*bHLH24* gene promoter at 0 day (**a-c**), 4 day (**d-f**) and 7 day (**g-i**). A focus on the air chamber pore is shown in **j to l**. **A, d, g, j**: the expression of the constitutive nuclear fluorescent markers proMp*L40e*:mTurq2:N7 is shown. **B, e, h, k**: the expression pattern of mVenus driven by the Mp*bHLH24* promoter can be observed. *C, f, i, l*: merged images including the plasma membrane marker in green (proMp*UBE2*:eGFP:Lti6b) and chlorophyll in grey. Scale bars of *a-i* are 100 micrometres, while *j-l* scale bars are 50 micrometres.

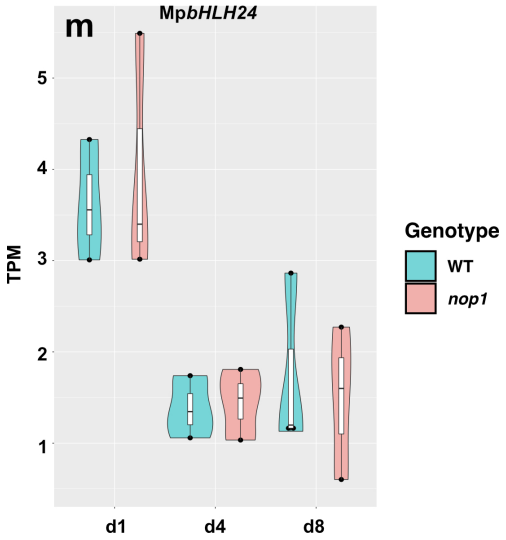


Figure 6.7 **m** Transcript per million (TPM) of Mp*bHLH24* in WT (blue) and Mp*nop1* (pink) at d1, d4 and d8.

6.2.6 Analysis of the expression of mVenus under the control of proMp *WRKY10* in epidermal cells

In the comparison of WT plants and Mp*nop1* mutant transcriptomes (see Chapter 5 section 5.2.7), Mp *WRKY10* was identified as a potential candidate marker for air chamber patterning. WRKY transcription factors are known to play a role in epidermal patterning in Arabidopsis, where they regulate root hair and trichome development through the formation of a regulatory complex with bHLH and MYB TFs [200]. Based on some preliminary imaging results, the Mp *WRKY10* promoter showed to drive mVenus expression only in oil bodies in 0 day gemmae (Figure 6.8 a-c). This pattern of expression was maintained in 1 and 4 days old gemmalings (Figure 6.8 d-f), but in 7 day samples it showed a specific expression in the air pore cell, with no expression in any other dorsal epidermis cells (Figure 6.7 g-l).

In the notch area, where air chambers are continuously produced, it can be observed that the Mp *WRKY10* promoter drives expression of mVenus in the notch area from the very early stages of gemma development. The signal coming from the notch cells is dimmer than the one captured from oil bodies (indicated by red asterisks in Figure 6.9), but it persists through gemma development. In d4 gemmalings, the first pores can be visualised (white arrows) and mVenus signal becomes localised to air pores as air chambers reach stage 3 of their development and move away from the notch, indicated with the letter N in Figure 6.9.

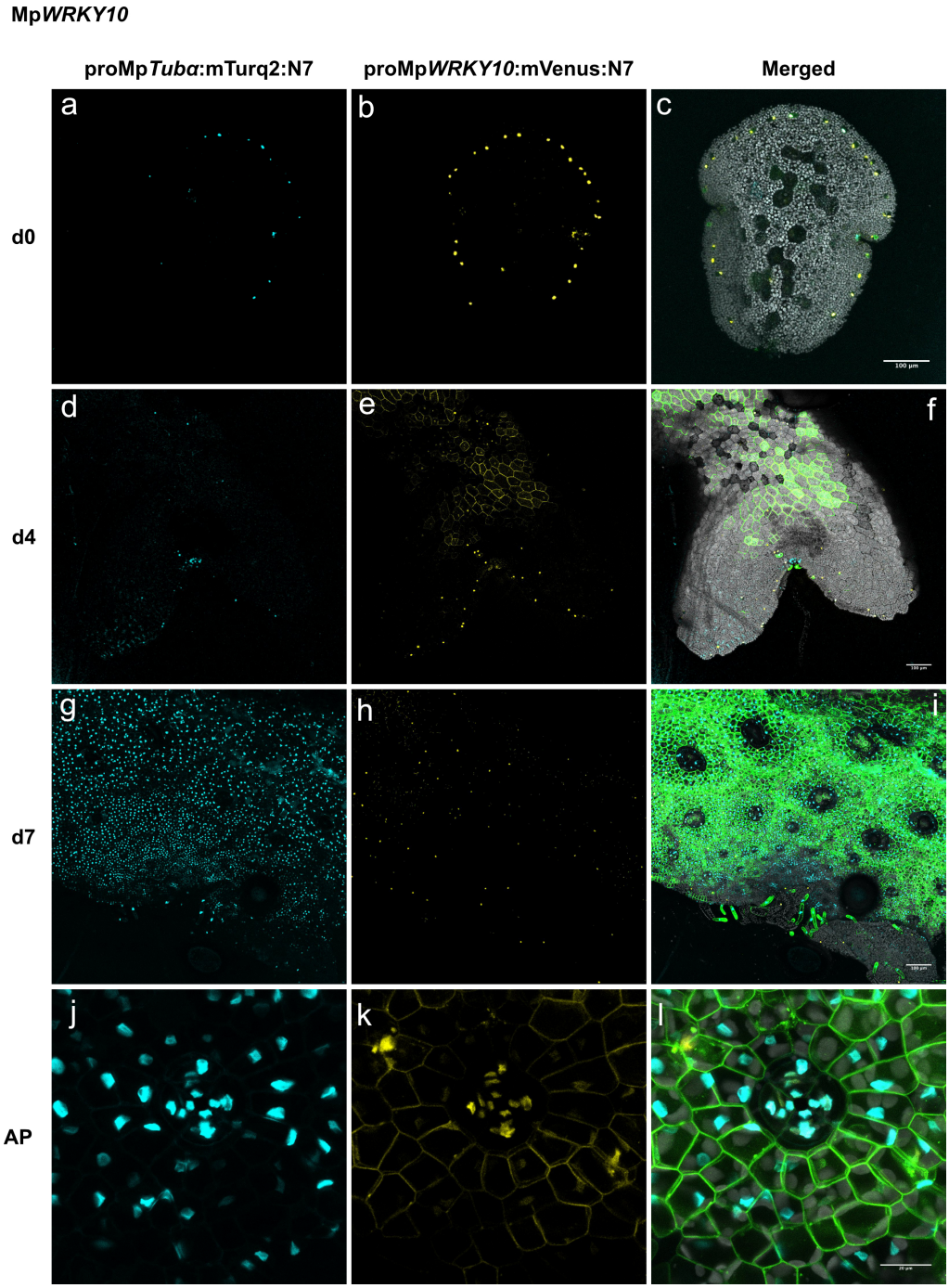


Figure 6.8 **Expression pattern of proMpWRKY10:mVenus:N7:T35S:TNos cassette in 0, 4 and 7 days old gemmalings.**
The figure shows the gemmae at 0, 4 and 7 days after removal from the gemma cups. A to i show a gemma where mVenus expression is under the control of the Mp*WRKY10* gene promoter at 0 day (**a-c**), 4 day (**d-f**) and 7 day (**g-i**). A focus on the air chamber pore is shown in j to l. **A, d, g, j**: the expression of the constitutive nuclear fluorescent markers proMp*Tubα*:mTurq2:N7 is shown. **B, e, h, k**: the expression pattern of mVenus driven by the Mp*WRKY10* promoter can be observed. **C, f, i, l**: merged images including the plasma membrane marker in green (CaMV35s:eGFP:Lti6b) and chlorophyll in grey. Scale bars of **a-i** are 100 micrometres, while **j-l** scale bars are 50 micrometres.

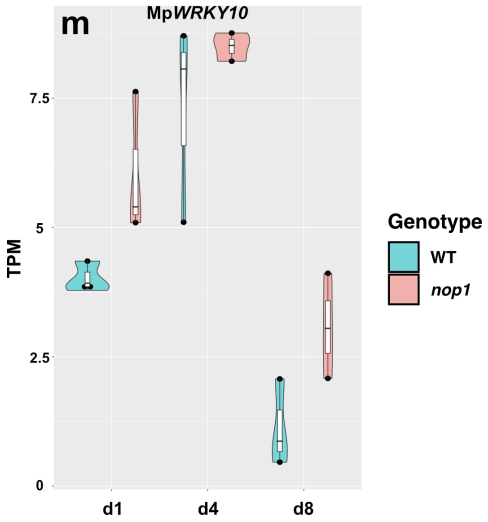


Figure 6.8 **m** Transcript per million (TPM) of Mp*WRKY10* in WT (blue) and Mp*nop1* (pink) at d1, d4 and d8.

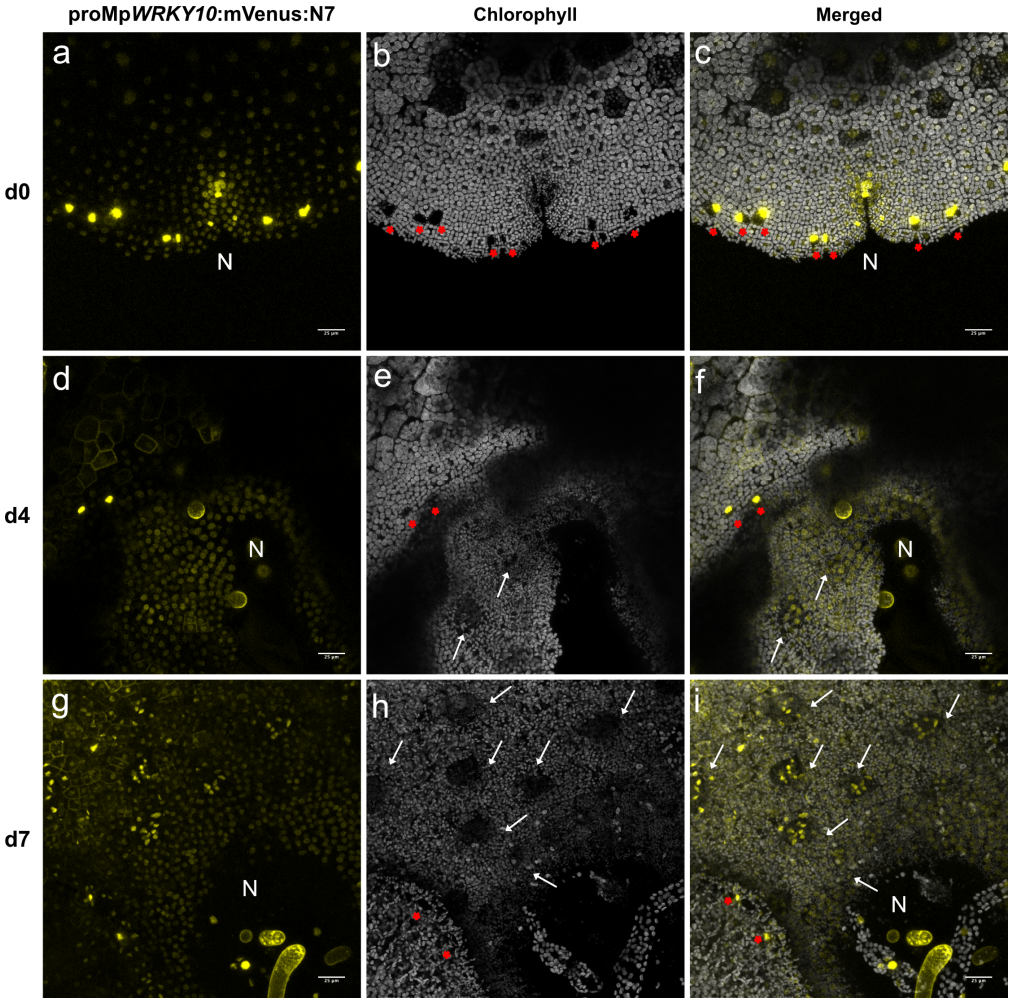


Figure 6.9 **Mp*WRKY10* controls mVenus expression in the notch of gemmalings and air pore specific expression starts in d4 gemmae.**
Cellular localisation of mVenus expression driven by pro*MpWRKY10* was visualised imaging gemmalings' notches with high magnification objectives (40x water immersion). A to i show the notch area of 0 day (**a-c**), 4 day (**d-f**) and 7 day (**g-i**) gemmalings. **A, d, g** expression pattern of mVenus driven by the Mp*WRKY10* promoter can be observed. The letter N indicates the notch. **B, e, g**: gemmalings chlorophyll is shown in gray. White arrows indicate air pores, while red asterisks highlight oil bodies. **C, f, i**: merged images. Scale bars are 25 micrometres.

6.3 Discussion

6.3.1 Reporter gene fusions as a resource for the identification of cellular markers

The collection of *Marchantia* synthetic promoters represents a valuable resource for the identification of key plant developmental genes. A series of vectors were produced to test the cell-specific expression of fluorescent proteins fused to a set of promoters of interest identified in the transcriptome comparison of WT and Mp*nop1* mutant plants, described in Chapter 5. The expression patterns of the fluorescent proteins transcribed from the promoters of interest were monitored in gemmalings in a week-long time course to capture any pattern correlating with the emergence of air chambers from the very early stages. Although the length of the designed synthetic promoters might not cover the complete size of the gene promoter, they still represent a valuable tool for the study of cellular development in *Marchantia*.

6.3.2 Mp*WRKY10* and Mp*bHLH24* promoter expression pattern correlate with air pore emergence

Out of the five genes' promoters that were selected for the screening, proMp*ICE2* did not show any mVenus expression, while two other promoters (proMp*ICE1* and proMp*ZOU1*) showed expression of the reporter in cell types that do not necessarily correlate with air chamber development (dorsal epidermis and central region of the thallus).

The complete lack of expression of mVenus from the proMp*ICE2* might be due to an insufficient length of the promoter, or the lack of important regulatory regions in the cloned promoter fragment. The Mp*bHLH24* gene promoter was able to induce broad expression of mVenus on the surface of the thallus, with higher expression in the cells on the margin of the gemmalings and in the pore cells of air chambers, although expression could also be visualised in the cells belonging to the dorsal epidermis of the thallus. Similarly, the expression of mVenus driven by the Mp*WRKY10* gene promoter showed strong expression in oil bodies in 0 day gemmae and dimmer expression in the cells around the notch. As air chambers reached their final stages of development at 7 day, mVenus expression, under the control of proMp*WRKY10*, localised in the pore cells specifically.

In order to achieve cellular resolution to monitor the localisation of mVenus under the control of the proMp*WRKY10*, samples should be cleared and stained with cell-wall markers to visualise the different cell types. Alternatively, plants should be transformed with a

plasmid containing the proMp*UBE2*:eGFP:Lti6b:Nos35STerm membrane marker instead of pCamV35S:eGFP:Lti6b:Nos35STerm, as it has been shown to have a more ubiquitous and constitutive expression in gemmalings, including in the notch area [3].

6.3.3 Reporter gene fusions suggest that Mp*WRKY10* and Mp*bHLH24* could have a function in air chamber development

Both the Mp*WRKY10* and Mp*bHLH24* genes' promoters showed an air pore specific expression of mVenus in gemmalings.

The expression patterns visualised through the fusion of a promoter of interest to a fluorescent protein might be determined by the length of the regulatory region cloned in the vector. Some regulator binding sites might have been excluded from the selected sequence, preventing enhancer or inhibitor proteins to bind to the DNA sequence and regulate gene expression. The insertion site of the T-DNA region might also influence gene expression, but this scenario was excluded as gemmae produced by independent transformation events were screened and the observed expression pattern was consistent amongst them.

Nevertheless, the visualised patterns demonstrated the power of promoter fusions to reporter genes as a tool for identification of cellular markers, which can subsequently functionally characterised. The pore specific expression of Mp*bHLH24* and Mp*WRKY10* genes promoters makes them excellent candidate genes for functional analysis, in an attempt to identify new genes regulating air chamber developments. Therefore, in order to test if Mp*bHLH24* and Mp*WRKY10* have a role in this process, the production of mutants would be necessary to screen for aberrant phenotypes in air chamber patterning and/or development.

Chapter 7

Conclusions

7.1 Summary

This dissertation describes the development of cellular markers used to study the patterning of the dorsal epidermis of *Marchantia polymorpha*, more specifically the morphogenesis of air chambers. In my project, I developed faster, more effective and safer tissue clearing techniques to observe tissue in three dimensions. These were applied to produce a map of air chamber development, which included all specific cell types present in this structure. I identified candidate genes as cellular markers for air chamber morphogenesis by comparing the transcriptome of wild type plants to the transcriptome of a mutant lacking air chambers. I used the candidate genes identified through transcriptomic analysis to create a collection of reporter genes' fusions to be used as cellular marker for air chamber morphogenesis. Their expression patterns were visualised and characterised *in planta*. Two genes, Mp*WRKY10* and Mp*bHLH24* were selected for further functional characterisation, based on the localisation of their expression in the air chamber's pore.

7.2 Simple, fast and safe clearing technique for high-resolution microscopy

Previously used microscopy techniques for the visualisation of *Marchantia* gemmae in three dimensions made use of potentially hazardous chemicals with sample preparation processes taking up to a week [201].

Clearing techniques have been widely used in mammalian research to study whole organs, especially brain tissue [138, 139, 137, 127], to avoid the use of expensive instruments, such as computed tomography [202] or magnetic resonance imaging [203]. In mammalian cells

clearing solutions needs to penetrate through cell membranes, while plant cells present an additional barrier: the cell wall, which causes spherical aberration and light scattering [143]. Damaging the wall structure can lead to a loss of the cell-cell relationships in the plant organism, hence clearing techniques should be gentle enough to preserve the wall structure, while being able to remove or mitigate the effect of all those components causing light absorption and scattering [127], including large molecules such as chlorophylls and carotenoids amongst others.
Three dimensional imaging of entire organs in plants has been used for a long time on those structures which are already transparent, such as the root tip and the meristems [204]. Nowadays, leaves, stems and shoot apices can be visualised in three dimension without the need of physical sectioning [142, 116]. The most recent clearing techniques had been tested in *Arabidopsis* tissues, including ClearSee, which was also used to clear moss gametophytes [116]. None of them were tested in liverworts.
Hence, I developed and tested a new protocol using the ClearSee clearing solution, which makes use of safer chemicals compared to other clearing techniques, takes only three days to clear the sample and is able to preserve fluorescent proteins localised in the nucleus.

Additionally, this technique allows visualisation of the fluorescent markers beyond the outer epidermal layers of the plant and, for the first time, it was paired with chemical stains, showing to be compatible with Calcofluor White M2R and SCRI Renaissance 2200 [205] for cell outlines visualisation. Finally, the ClearSee protocol uses the same solution for clearing and mounting, allowing the maintenance of a high refractive index, throughout the tissue and the mountant.

7.3 A cellular description of air chamber morphogenesis

The ease of preparation of cleared tissue has allowed me to develop a model for air chamber development, which confirms a number of the observations made by Apostolakos [93]. I was able to identify four stages of air chamber development: the initial formation of intercellular spaces, which initiate air chamber morphogenesis in close proximity of the notch (Stage 0); the appearance of radial division in the cells surrounding the initial aperture leading to the formation of the pore complex (Stage 1); the emergence of photosynthetic filaments in the sub-epidermal chamber (Stage 2) and the final opening of the pore in the pore complex (Stage 3).

Optical clearing allows to preserve samples integrity and it avoids tissue destruction to

access internal layers, which can be visualised by optical sections. The entire air chamber arrangement could be visualised in three dimensions after the appearance of the initial intercellular apertures without the need for physical sections and maintaining the spatial relationship between the cell types found in the epidermal and sub-epidermal layers. I was able to correlate the development of the pore complex size with other cellular events such as the appearance of photosynthetic filaments in the sub-epidermal cell layer, and the opening of an aperture at the centre of the pore complex to connect the sub-epidermal chamber to the outer environment.

This work allows correlation of cellular events taking place in the dorsal layer of the thallus with events happening in the sub-epidermal chamber. I was also able to visualise the air chamber walls, which appeared to be composed of a single cell layer. This suggests that these cells might acquire their fate at early stages in air chamber development, to ensure the correct spacing between the emerging air chambers. Further investigation of the series of cell divisions leading to the formation of the air chamber complex might be required to generate a more accurate model, but the description shown here represents a useful reference for mapping cell marker expression and to correlate them with the stages of air chambers' development.

7.3.1 Next steps in the identification of air chambers patterning and differentiation processes

The formation of air chambers on the epidermal surface of *Marchantia* might be regulated by positional cues, which could trigger cell differentiation. Ishizaki [98], through the observation of the Mp*nop1* mutant phenotype, has suggested the existence of positional cues that allow cell separation to take place. This speculation is further supported by the observations by Apostolakos [94], who documented that this event only takes place at a defined distance (4-5 cells) from the apical cell. Those positional cues could be hormones, peptides or genetic material (such as mRNA or miRNA) moving between neighbouring cells.
For example, the application of exogenous auxin (and auxin inhibitors) to *Marchantia* thalli causes abnormalities (distortions) in air chambers shape [206, 207].

The*Marchantia* Mp*arf3* mutant, which is not able to respond to auxin signalling, shows air chambers on the dorsal side, but it loses the ability to produce gemmae inside its gemma cups [208]. The authors report that the loss of function mutant shows areas without air chambers, but this is often observed while recovering early transformants and these regions tend to disappear with time (data not shown). They also observe a difference in the air

pore size between WT and mutant thalli, but no quantitative data is provided to support this statement. Enhanced and overexpression of the Mp*ARF3* gene results in the reversion to undifferentiated cell state of the thallus [208].
It is reasonable to think that auxin acts as a morphogen in *Marchantia*, but further clarification on auxin distribution and perception in air chambers is required to understand how it affects their patterning and differentiation. To achieve this goal, there is a need for more sensitive auxin reporter system for *Marchantia*. Other positional cues though might be in place to ensure that air chambers are regularly spaced, separated by single cell-layer walls and making sure that each chamber has a single pore.

Early experiments treating *Marchantia* WT gemmae for 12 days with exogenous gibberellic acid (GA 4/7 50 μM) and its inhibitor Paclobutrazol (50 μM) show respectively a wider spacing of air chambers, which are no longer next to each other (Figure 7.1) in GA treated samples, or the formation of clusters of undifferentiated cells in the Paclobutrazol treated ones (Figure 7.1).

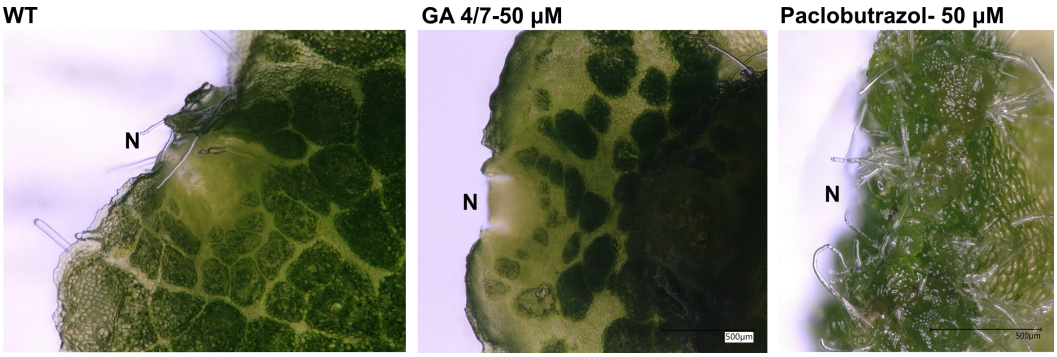


Figure 7.1 **Gibberellic acid and Paclobutrazol treatment on *Marchantia* WT gemmae.**
Marchantia gemmae were grown for 12 days were grown on half-Gamborg B5 media with 1.2% agar and GA4/7 50 μM or Paclobutrazol 50 μM. The spacing of air chambers is altered in the gemmalings grown on GA 4/7, while Paclobutrazol induces the formation of clusters of undifferentiated cells.

Gibberellins have been shown to be involved in the initiation and morphogenesis of epidermal structures in *Arabidopsis* [209]. *Marchantia* has a DELLA homologue but no GID1 receptor [210, 211, 100], hence it is still not clear how GA perception and signalling takes place in *Marchantia*.
Sequencing the transcriptome of plants with altered distribution of air chambers, such as the ones treated with gibberellic acid could help to discover additional positional cues involved in the patterning process. The identification of such cues would help to integrate the Ishizaki model [98] of spatial patterning into the the one presented in this manuscript.

7.4 Transcriptomic analysis for the identification of candidate marker genes

The Mp*nop1* mutant was obtained by knocking out a E3 ubiquitin ligase protein, but no other members of the regulatory network it belongs, have been identified. The absence of air chambers was the only phenotype observed in Mp*nop1* plants, which were still able to develop other epidermal structures, such as rhizoids and gemma cups. This well-defined phenotype made Mp*nop1* mutants an excellent system to identify transcription factors with specific expression in cells belonging to air chambers. Transcriptomic analysis was able to reveal the presence of differences in gene expression between the Mp*nop1* mutant and wild type plants.

A series of differential expressed transcription factors, which could be involved in air chamber patterning, differentiation, metabolism or related processes were identified, such as members of the Ib, IIIb and IIIf subfamily of bHLH and WRKY genes. The small number of cells forming air chambers, compared to the total number of thallus cells, complicates the enrichment of transcripts, which are specifically expressed in air chamber cell types, but further sequencing experiments could be performed on samples enriched in air chamber cells. Nevertheless, the analysis presented in this dissertation was able to provide two candidate genes, one bHLH and one WRKY transcription factors, which showed specific expression in the air chambers' air pore.

7.5 Creation of a collection of cell specific markers

Marchantia has only 398 transcription factors [100] and thanks to the low redundancy in each family, represents an excellent system to build and test a collection of reporter gene fusions to transcription factor promoters, which can be used as cellular markers.

A library of potential cellular markers was generated and the constructs were transformed in plants and screened with the use of confocal microscopy to characterise the promoter:fluorescent protein expression pattern at different developmental stages of gemma development. For the aim of this dissertation, I specifically looked for markers expressed in air chamber related cell types, choosing gene promoters based on candidate genes identified during the transcriptomic analysis and possesing homology with genes involved in epidermal patterning in higher plants.

I tested the promoters of all the *Marchantia* TF members belonging to the Ib and IIIb subfamily of bHLHs as well as the promoter of Mp*WRKY10*. Amongst them, Mp*bHLH24*

and Mp*WRKY10* promoters showed air pore specific expression. Moreover, markers for oil bodies, rhizoids and mucilage cells were also identified during the screening process. This collection represents a valuable resource for the study of developmental processes as well as cellular dynamics.

7.6 Identification of potential regulators of air chamber development

The identification of air pore specific expression of mVenus controlled by the Mp*WRKY10* gene promoter, provided an excellent candidate gene for air pore development.
The development of other epidermal structures of *Marchantia* have been shown to be regulated by gene families, which are also involved in epidermal cell differentiation in late divergent plants. Rhizoids development is controlled by the bHLH VIII subfamily [11], which also controls the formation of new gemmae (but not gemma cups) on the dorsal side of the thallus. The same gene family controls root hair development in *Arabidopsis*. Similarly, gemma cup differentiation is controlled by an R2R3 MYB transcription factor (subfamily 14) and its orthologue in angiosperms controls the formation of axillary meristems in late divergent plants [80]. Nevertheless, none of these genes seem to affect air chamber development.
Mp*WIP* expression is not strictly localised to the air pore, but it is also expressed in other epidermal cells in the dorsal and ventral part of the plant. Similarly, the two markers investigated in this dissertation, Mp*WRKY 10* and Mp*bHLH24* are expressed both in the air pore and the thallus epidermis. Their expression on the ventral part of the thallus or in the internal tissue could not be verified during the preparation of this manuscript.
The wide expression domain could suggest that these promoters might be active in a variety of tissues, where they exert different functions depending on their regulation and interaction with other TFs. This would show that a simple genetic toolkit can be used to regulate multiple developmental processes, as it already happens for trichomes and root hairs in *Arabidopsis*.

7.7 Future experiments

The work presented in this dissertation provides a series of tools which can be used to characterise and monitor cell specific markers in *Marchantia* gemmalings.

7.7.1 Using cell markers to identify patterns and create cell type specific transcriptomic data

Cell-type specific fluorescent markers also represent a resource for the generation of a more accurate map of *Marchantia* development, where the map of gene expression can be combined with transcriptomic information. The high-throughput imaging system developed allows for the collection of high-quality and standardised images that can be used to train artificial intelligence (AI) algorithms for automatic detection of the expression pattern of fluorescent proteins. This will further streamline the work of researchers, who could select suitable cell specific fluorescent markers based on their expression patterns.

Cell specific expression of fluorescent markers can be used to sequence samples enriched in the desired cell type thanks to fluorescence-assisted cell sorting (FACS). This could lead, for example, to the generation of a transcriptomic data set specific to air chamber cell types, such as pores and wall cells and photosynthetic filaments, identifying the genes required for the differentiationof these cell types. This might also help to understand the physiological function of air chambers.

7.7.2 Understanding the role of Mp*WRKY 10* and Mp*bHLH24* in air chamber development

In order to understand the role of Mp*WRKY10* and Mp*bHLH24* in air chamber development, further validation is required. Gene editing experiments were started during the writing of this manuscript, but could not be finalised due to the current COVID-19 pandemic.

The generation of knock-out and overexpression mutants, in both the WT and Mp*nop1* mutant genetic background, would reveal if Mp*WRKY10* and Mp*bHLH24* play a role in air chamber development. Plasmids containing two guide RNAs and the Cas9 endonucleases (CRISPR/Cas9 system) to generate large deletions in exon 1 of the Mp*WRKY10* gene have been transformed in *Marchantia* sporelings and they are currently being screened. Some of the early recovered sporelings show patches with no air chambers or photosynthetic filaments directly protruding from the dorsal epidermis.

Similarly a plasmid targeting the first exon of Mp*bHLH24* gene has been prepared with a single gRNA. This choice was made to avoid off-target events due to the high degree of sequence conservation of the bHLH TF family. The observation of mutant phenotypes induced by the gene loss of function, will reveal if these transcription factor play a role in air pores patterning and/or differentiation. Furthermore, the expression of the Mp*WRKY10* and Mp*bHLH24* promoters fused to mVenus in Mp*nop1* mutant plants, could also provide insights into which cell types are absent (if any) in the mutant and how the expression pattern of these constructs diverges from the one observed in WT plants.

7.7.3 Identifying genes with epidermis specific expression

The ATML1 homeobox protein defines the epidermal tissue in the *Arabidopsis* and regulates the transcription of all the epidermis specific genes [39, 43]. This transcription factor binds to a specific promoter region called L1 box [39], which is present in the promoter of many genes with epidermal specific expression [43], including *ATML1* itself.
Marchantia has a single homologue of this gene: Mp7g09710. It would be interesting to investigate if the *Marchantia* homologue is able to bind to the same L1 box sequence or has an analogous one to induce epidermal specific expression. A genome wide analysis of Mp7g09710 binding sites in *Marchantia* genes could reveal those that might be involved in epidermal tissue differentiation.

Thanks to the new techniques and knowledge produced in this work (Figure 7.2), the investigation of the genetic and cellular mechanisms controlling the patterning and development of air chamber on the dorsal epidermis of *Marchantia* can be carried out with straightforward protocols and a reference model.

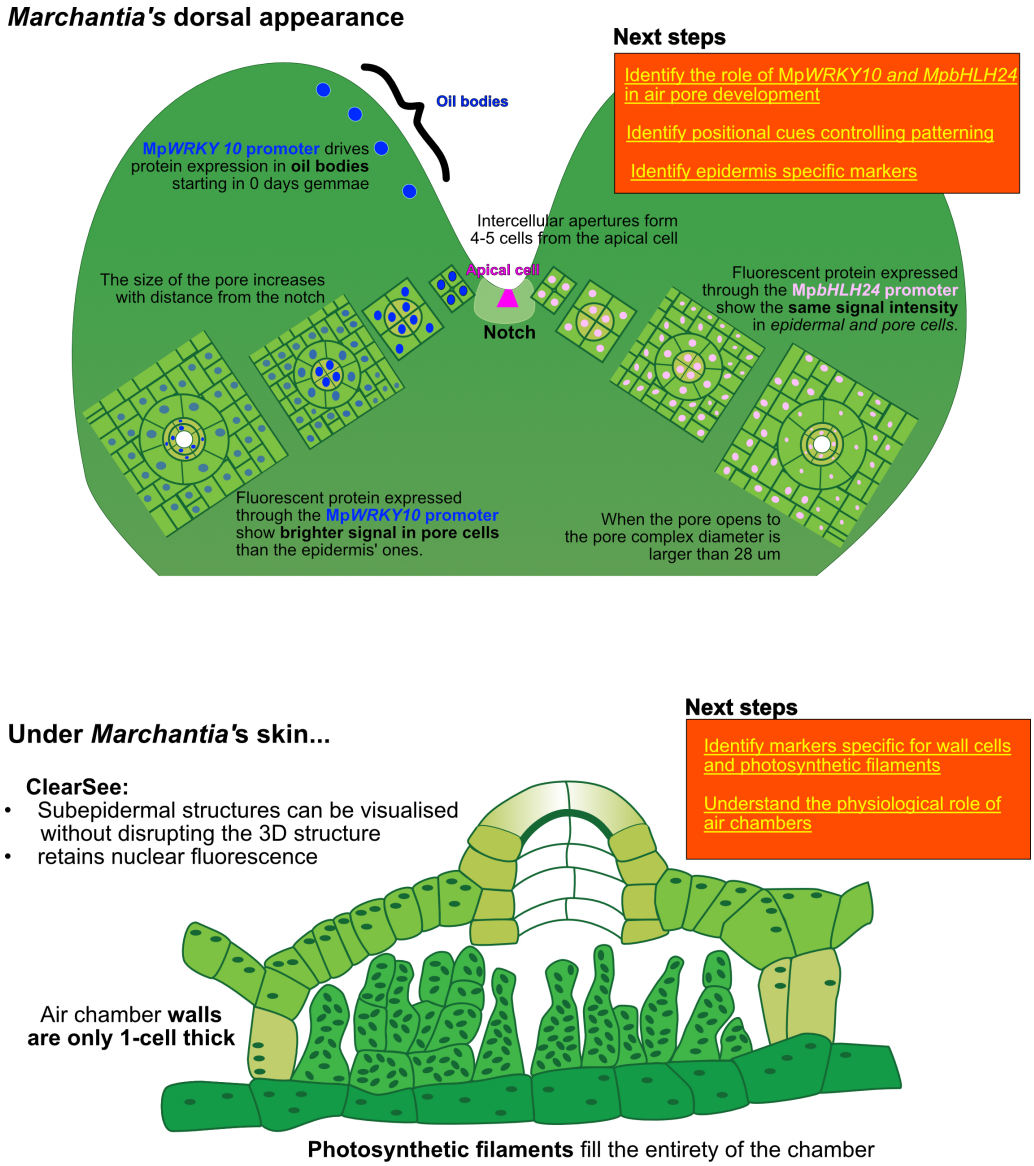


Figure 7.2 **Visual summary.**
The figure illustrates the main findings reported in this manuscript, indicating a list of next steps to undertake to investigate the patterning and differentiation of air chambers.

Appendices

.1 Sequencing primers

Primers used to validate inserts in L1 and L2 plasmids with pCambia backbone.

pC_F: 5' GCAACGCTCTGTCATCGTTAC 3'
pC_R: 5' GTAACTTAGGACTTGTGCGACATGTC 3'

Primer used to validate the assembly of reporter gene fusions in L2 plasmids:

67_F: 5' TACTCGCCGATAGTGGAAACC 3'

Primer used to validate the insertion of sgRNA target sequences in L1_LacZgRNA-Ck2 and L2_lacZgRNA-Cas9-CsA:

EF021_F: 5' CATATCTCTGCAAGAGCTTATTC 3'

Primers designed for Mp*WRKY10* mutants genotyping:

MT216_F: 5' TGCAGGAATGTCGAGAAGAC 3'
MT217_R: 5' CGAGGACATGGAGGAATTCG 3'

Appendix A

Commands for RNAseq data analysis

```
library(tximport)
library(DESeq2)
library(dplyr)
library(tidyr)
library(ggplot2)
library(RColorBrewer)
library(pheatmap)
library(gridExtra)

|#set working directory
setwd("/Users/Marta/Desktop/RNAseq_CamvsNop/Analysis  Genom
v5.1/")

# Optional step to automatise the rest of the code for different datasets
study <- c("MT_CAM_NOP")

# load metadata file
meta <- read.table(paste("",study,".txt", sep=""), header=T, sep =
"\t")
```

```
# rename table rows and reshape table columns
rownames(meta) <- meta$run_accession
meta <- separate(meta, condition, c("genotype", "strain", "tissue",
"age"), sep = "_", remove = F)

# defining interactions between genotype and tissue
meta$genotype.tissue <- interaction(meta$strain, meta$tissue,
sep=".")
# defining interactions between condition and study
meta$condition.study <- interaction(meta$condition,
meta$study_accession, sep=".")

# defining interactions between condition and technical replicate
meta$condition.rep <- interaction(meta$condition, meta$rep, sep=".")

#importing kallisto output using tximport

# create file path
dir<-"/Users/Marta/Desktop/RNAseq_CamvsNop/Analysis Genom
v5.1/Kallisto"

#imports all tsv files in the directory defined in dir files <- file.path(dir,
meta$run_accession, "abundance.tsv")

# renaming tsv files accordingly to the run-accession names names(files)
<- meta$run_accession

# building gene list from file. Can be
any table containing genes ID. tx2gene <-
```

```
read.table("/Users/Marta/Desktop/RNAseq_CamvsNop/Analysis
Genom v5.1/abundance_copy.tsv", header=T)
tx2gene <- tibble(tx2gene$target_id, gsub("
..*", "", tx2gene$target_id))

# renaming columns of tx2gene table
colnames(tx2gene) <- c("TXNAME", "GENEID")

# Pooling results from each sequenced sample into one file
txi <- tximport(files, type = "kallisto", tx2gene = tx2gene)

# saving table containing counts or abundance from pooled studies
dir.create(study)
write.csv(txi$counts, paste(study,"/",study,"_counts.csv", sep=""))
write.csv(txi$abundance,paste(study,"/",study,"_abundance.csv",
sep="") )

# creating a DESeq object from txi file
de <- DESeqDataSetFromTximport(txi, meta, ~condition)

# Checking that all genes are included in the analysis. The number is
the product of tot number of genes * number of tested conditions

# filtering out genes with < 5 reads in total
keep <- rowSums(counts(de)) > 5
de <- de[keep,]

# checking how many genes have been discarded
length(counts(de))
```

```
# filtering out genes with no expression in any condition
keep2 <- apply(counts(de), 1, function(row) all(row !=0 ))
de <- de[keep2,]

# checking how many genes have been discarded
length(counts(de))

#normalising data from DESSeq object for downstream analysis, with variance stabilising transformation
vsd <- vst(de)

#transforming the obtained object in dataframe to use it for diagrams and saving it
All_gene_vsd<-as.data.frame(assay(vsd)) %>%
mutate(id = rownames(vsd))
write.csv(All_gene_vsd, 'All_gene_vsd.csv')

# Identifiying log2fold change, p-value, padj, etc for each of the transcript in the de object
deGenes <- DESeq(de)
```

A.1 List of publicly available transcriptomic datasets of *Marchantia*

Study	Ecotype	Genotype	Tissue	Age	Ref
PRJDB4420	Tak-1	WT	thallus	9 day	[212]
	Tak-1		antheridiophore	NA	
	Tak-1		antheridia	NA	
	Tak-2		archegoniophore	NA	
PRJDB5890	Tak-1	WT	thallus	42 day	[213]
		Mp*PAP1-OE*	thallus	42 day	
		Mp*PAP2-OE*	thallus	42 day	
PRJDB6579	Tak-1	WT	thallus	7 day	[214]
	BC3-83	WT	thallus	7 day	
	Tak-1	*hpt2040*	thallus	7day	
	Tak-2	*hpt2040*	thallus	7 day	
PRJDB7023	Tak-1	WT	thallus	7 day	[100]
	Tak-1		gemma cup	NA	
	Tak-1		gemma	1 day	
	Tak-1		no meristem	NA	
	BC4		archegoniophore	NA	
	Tak-2		archegoniophore	NA	
PRJNA218052	Tak-1	WT	thallus	NA	[215]
	Tak-2		thallus	NA	
	Tak-1		antheriodiophore	NA	
	Tak-2		immature antheriodiophore	NA	
	Tak-1		immature antheridiophore	NA	

	Tak-2		archegoniophore	NA	
PRJNA265205	Tak	WT	thallus apical cell sporophyte	9 day 9 day NA	NA
PRJNA397394	Tak-1	WT WT+auxin	thallus	10 day	NA
PRJNA433456	Tak-1	WT mir160 Mp*arf3*	thallus	30 day	NA
PRJNA251267	Tak-1 Tak-1 BC4 Tak-1 Tak-1 Tak-1 Tak-1	WT WT WT heat shock regen 24 regen 0 WT	sporophyte spore archegoniophore thallus basal fragment basal fragment antheridiophore	NA 5 day NA 7 day 7 day 7 day NA	NA
PRJDB5847	Tak-1	WT WT+ABA Mp*PLY1*-OE Mp*PLY1*-OE+ABA Mp*ply1* Mp*ply1*+ABA	thallus	4 day	NA
PRJDB8103	Tak-1	WT Mp*met-3* Mp*met-4*	thallus	NA	[216]
PRJNA488718	Tak-1	WT Mp*CL3*-OE-2 Mp*CL3*-OE-3	thallus	21 day	[182]

PRJNA350270	Cam-1	WT	spores	0 day 1 day 2 day 3 day 4 day	[100]

Appendix B

Commands for image processing

```
//ImageJ macro to generate from raw LIF files, maxima projection for each channel and composite images in tiff format

macro'convert LIF tiff'{
run("Bio-Format Macro Extensions");
dir1= getDirectory("/Users/YourFileFolder");
list = getFileList(dir1);

setBatchMode(true);

// create folder for output tiff files
dir1parent = File.getParent(dir1);
dir1name = File.getName(dir1);
dir2 = dir1parent+File.separator+dir1name+"–Output"; if (File.exists(dir2)==false ) {
File.makeDirectory(dir2);
}
//create a list of files for( i=0; i<list.length; i++){
showProgress(i+1, list.length);
print("processing..."+i+1+"/"+list.length+"\n "+list[i]);
```

```
path=dir1+list[i];

//counting number of series in LIF file Ext.setId(path);
Ext.getSeriesCount(seriesCount);

for (j=1; j<=seriesCount; j++) { run("Bio-Formats", "open=path autoscale color_mode=Custom, view=Hyperstack stack_order=XYCZT series_"+j);
name=File.nameWithoutExtension;

//retrieve name of the series from metadata

text=getMetadata("Info");
n1=indexOf(text," Name = ")+8;

n2=indexOf(text,"SizeC = ");
seriesname=substring(text, n1, n2-2);
seriesname=replace(seriesname,"/","-");
rename(seriesname);

//define colours for each channel

if (nSlices>1){
Stack.setChannel(1)
run("Cyan");
Stack.setChannel(3)
run("Green");
Stack.setChannel(4)
run("Grey");
```

```
Stack.setChannel(5)
run("Yellow");
//create z-stack maximum intensity projection run("Z Project...",
"projection=[Max Intensity]);

//save images from individual channels
Stack.setChannel(1)
run("Scale Bar...", "width=100 height=4 font=14 color=White back-
ground=None location=[Lower Right] overlay");
saveAs("Jpeg",dir2+File.separator+name+"_"+seriesname+"_CFP.jpg");
Stack.setChannel(3)
run("Scale Bar...", "width=100 height=4 font=14 color=White back-
ground=None location=[Lower Right] overlay");
saveAs("Jpeg",dir2+File.separator+name+"_"+seriesname+"_GFP.jpg");
Stack.setChannel(4)
run("Scale Bar...", "width=100 height=4 font=14 color=White back-
ground=None location=[Lower Right] overlay");
saveAs("Jpeg",dir2+File.separator+name+"_"+seriesname+"_Chl.jpg");
Stack.setChannel(5)
run("Scale Bar...", "width=100 height=4 font=14 color=White back-
ground=None location=[Lower Right] overlay");
saveAs("Jpeg",dir2+File.separator+name+"_"+seriesname+"_YFP.jpg");

make composite image
run("Arrange Channels...", "new=1345");
run("Make Composite");
run("Scale Bar...", "width=100 height=4 font=14 color=White back-
ground=None location=[Lower Right] overlay");
```

```
saveAs("Jpeg",dir2+File.separator+name+"_"+seriesname+"_composite.jpg");
}
}
}
showMessage("– finished –");
run("Close All");
setBatchMode(false);
}
```

Appendix C

Comparison of MpNOP1 and MpWRKY10 amino acid sequence between Cam-1 and Tak-1

Figure C.1 **MpNOP1 protein alignment.**

MpWRKY*10*

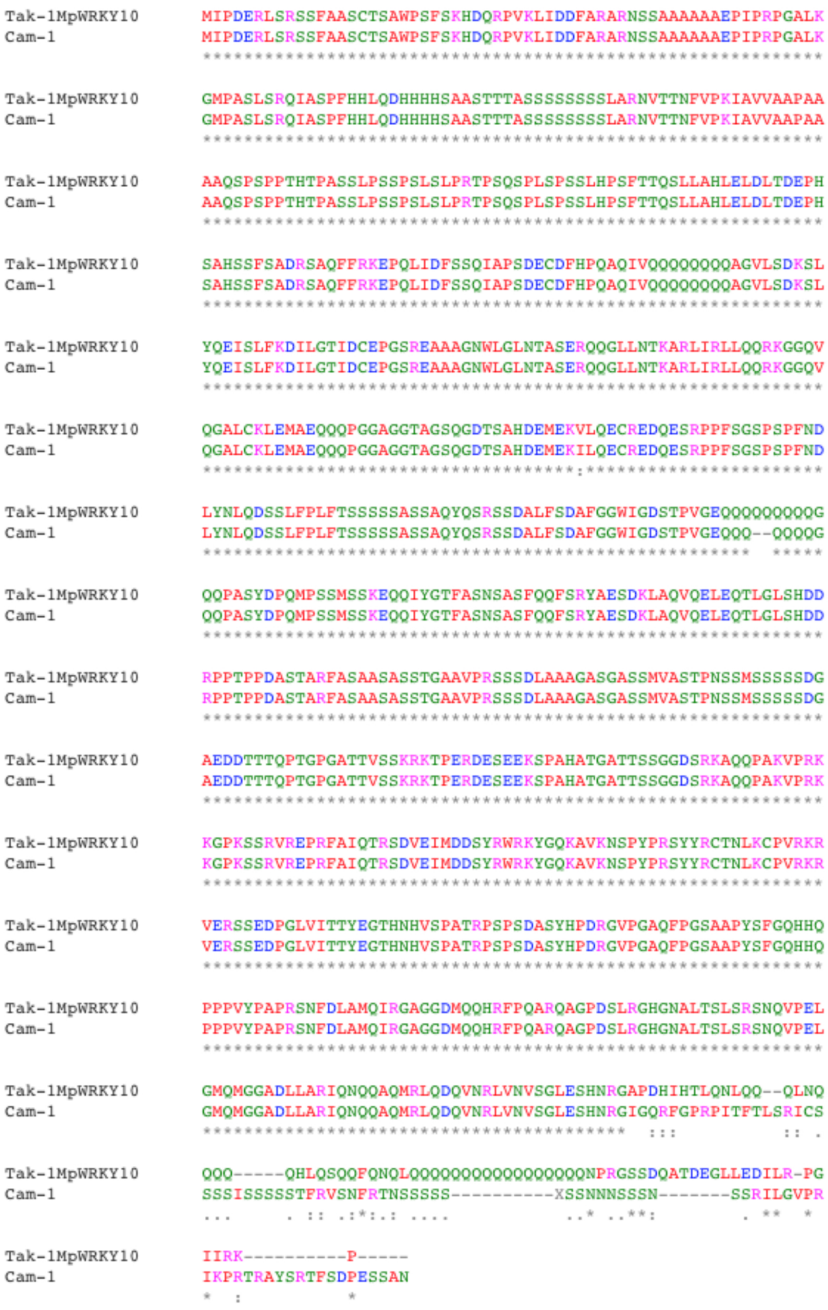


Figure C.2 **MpWRKY10 protein alignment.**

Bibliography

[1] Linda E Graham et al. *Origin of land plants.* John Wiley & Sons, Inc., 1993.

[2] Charles H Wellman, Peter L Osterloff, and Uzma Mohiuddin. Fragments of the earliest land plants. *Nature*, 425(6955):282–285, 2003.

[3] Susanna Sauret-Gueto, Eftychios Frangedakis, Linda Silvestri, Marius Rebmann, Marta Tomaselli, Kasey Markel, Mihails Delmans, Anthony West, Nicola J Patron, and Jim Haseloff. Systematic tools for reprogramming plant gene expression in a simple model, marchantia polymorpha. *ACS Synthetic Biology*, 9(4):864–882, 2020.

[4] Elias J Durand. The development of the sexual organs and sporogonium of *Marchantia polymorpha*. *Bulletin of the Torrey Botanical Club*, 35(7):321–335, 1908.

[5] Charles R Barnes and William Jesse Goad Land. Bryological papers. II. The origin of the cupule of *Marchantia*. *Botanical Gazette*, 46(6):401–409, 1908.

[6] Niina Tarén. Factors regulating the initial development of gemmae in *Marchantia polymorpha*. *The Bryologist*, 61(3):191–204, 1958.

[7] Mary Ellen O'Hanlon. Germination of spores and early stages in development of gametophyte of *Marchantia polymorpha*. *Botanical Gazette*, 82(2):215–222, 1926.

[8] Jeremy E Solly, Nik J Cunniffe, and C Jill Harrison. Regional growth rate differences specified by apical notch activities regulate liverwort thallus shape. *Current Biology*, 27(1):16–26, 2017.

[9] Caspar C Chater, Robert S Caine, Marta Tomek, Simon Wallace, Yasuko Kamisugi, Andrew C Cuming, Daniel Lang, Cora A MacAlister, Stuart Casson, Dominique C Bergmann, et al. Origin and function of stomata in the moss *Physcomitrella patens*. *Nature Plants*, 2(12):1–7, 2016.

[10] Caspar CC Chater, Robert S Caine, Andrew J Fleming, and Julie E Gray. Origins and evolution of stomatal development. *Plant Physiology*, 174(2):624–638, 2017.

[11] Hélène Proust, Suvi Honkanen, Victor AS Jones, Giulia Morieri, Helen Prescott, Steve Kelly, Kimitsune Ishizaki, Takayuki Kohchi, and Liam Dolan. RSL class I genes controlled the development of epidermal structures in the common ancestor of land plants. *Current Biology*, 26(1):93–99, 2016.

[12] Brogan J Harris, C Jill Harrison, Alistair M Hetherington, and Tom A Williams. Phylogenomic Evidence for the Monophyly of Bryophytes and the Reductive Evolution of Stomata. *Current Biology*, 2020.

[13] Karen Sue Renzaglia, R Joel Duff, Daniel L Nickrent, and David J Garbary. Vegetative and reproductive innovations of early land plants: implications for a unified phylogeny. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 355(1398):769–793, 2000.

[14] KAREN S RENZAGLIA, JUAN CARLOS VILLARREAL AGUILAR, and DAVID J GARBARY. Morphology supports the setaphyte hypothesis: mosses plus liverworts form a natural group. *Bryophyte Diversity and Evolution*, 40(2):11–17, 2018.

[15] Jennifer L Morris, Mark N Puttick, James W Clark, Dianne Edwards, Paul Kenrick, Silvia Pressel, Charles H Wellman, Ziheng Yang, Harald Schneider, and Philip CJ Donoghue. The timescale of early land plant evolution. *Proceedings of the National Academy of Sciences*, 115(10):E2274–E2283, 2018.

[16] Eftychios Frangedakis, Masaki Shimamura, Juan Carlos Villarreal, Fay-Wei Li, Marta Tomaselli, Manuel Waller, Keiko Sakakibara, Karen S Renzaglia, and Péter Szövényi. The hornworts: Morphology, evolution and development. *New Phytologist*, 229(2):735–754, 2021.

[17] Yin-Long Qiu, Libo Li, Bin Wang, Zhiduan Chen, Volker Knoop, Milena Groth-Malonek, Olena Dombrovska, Jungho Lee, Livija Kent, Joshua Rest, et al. The deepest divergences in land plants inferred from phylogenomic evidence. *Proceedings of the National Academy of Sciences*, 103(42):15511–15516, 2006.

[18] James H Leebens-Mack, Michael S Barker, Eric J Carpenter, Michael K Deyholos, Matthew A Gitzendanner, Sean W Graham, Ivo Grosse, Zheng Li, Michael Melkonian, Siavash Mirarab, et al. One thousand plant transcriptomes and the phylogenomics of green plants. 2019.

[19] Fay-Wei Li, Tomoaki Nishiyama, Manuel Waller, Eftychios Frangedakis, Jean Keller, Zheng Li, Noe Fernandez-Pozo, Michael S Barker, Tom Bennett, Miguel A Blázquez, et al. Anthoceros genomes illuminate the origin of land plants and the unique biology of hornworts. *Nature plants*, 6(3):259–272, 2020.

[20] Jian Zhang, Xin-Xing Fu, Rui-Qi Li, Xiang Zhao, Yang Liu, Ming-He Li, Arthur Zwaenepoel, Hong Ma, Bernard Goffinet, Yan-Long Guan, et al. The hornwort genome and early land plant evolution. *Nature plants*, 6(2):107–118, 2020.

[21] Olivier Hamant, Jan Traas, and Arezki Boudaoud. Regulation of shape and patterning in plant development. *Current Opinion in Genetics & Development*, 20(4): 454–459, 2010.

[22] Evan Heller and Elaine Fuchs. Tissue patterning and cellular mechanics. *Journal of Cell Biology*, 211(2):219–231, 2015.

[23] Erin Sparks, Guy Wachsman, and Philip N Benfey. Spatiotemporal signalling in plant development. *Nature Reviews Genetics*, 14(9):631–644, 2013.

[24] Richard Flavell. Role of model plant species. In *Plant Genomics*, pages 1–18. Springer, 2009.

[25] A Colette, Kuan-Ju Lu, and Dolf Weijers. Building a plant: cell fate specification in the early *Arabidopsis* embryo. *Development*, 142(3):420–430, 2015.

[26] G Venugopala Reddy, Marcus G Heisler, David W Ehrhardt, and Elliot M Meyerowitz. Real-time lineage analysis reveals oriented cell divisions associated with morphogenesis at the shoot apex of *Arabidopsis thaliana*. *Development*, 131(17): 4225–4237, 2004.

[27] Erika E Kuchen, Samantha Fox, Pierre Barbier De Reuille, Richard Kennaway, Sandra Bensmihen, Jerome Avondo, Grant M Calder, Paul Southam, Sarah Robinson, Andrew Bangham, et al. Generation of leaf shape through early patterns of growth and tissue polarity. *Science*, 335(6072):1092–1096, 2012.

[28] Susana Ubeda-Tomás, Fernán Federici, Ilda Casimiro, Gerrit TS Beemster, Rishikesh Bhalerao, Ranjan Swarup, Peter Doerner, Jim Haseloff, and Malcolm J Bennett. Gibberellin signaling in the endodermis controls *Arabidopsis* root meristem size. *Current Biology*, 19(14):1194–1199, 2009.

[29] Justyna Wiśniewska, Jian Xu, Daniela Seifertová, Philip B Brewer, Kamil Růžička, Ikram Blilou, David Rouquié, Eva Benková, Ben Scheres, and Jiří Friml. Polar PIN localization directs auxin flow in plants. *Science*, 312(5775):883–883, 2006.

[30] Holger Breuninger, Enno Rikirsch, Marita Hermann, Minako Ueda, and Thomas Laux. Differential expression of WOX genes mediates apical-basal axis formation in the *Arabidopsis* embryo. *Developmental Cell*, 14(6):867–876, 2008.

[31] Thomas Berleth and G Jurgens. The role of the monopteros gene in organising the basal body region of the *Arabidopsis* embryo. *Development*, 118(2):575–587, 1993.

[32] Mitsuhiro Aida, Dimitris Beis, Renze Heidstra, Viola Willemsen, Ikram Blilou, Carla Galinha, Laurent Nussaume, Yoo-Sun Noh, Richard Amasino, and Ben Scheres. The PLETHORA genes mediate patterning of the *Arabidopsis* root stem cell niche. *Cell*, 119(1):109–120, 2004.

[33] John F Emery, Sandra K Floyd, John Alvarez, Yuval Eshed, Nathaniel P Hawker, Anat Izhaki, Stuart F Baum, and John L Bowman. Radial patterning of *Arabidopsis* shoots by class III HD-ZIP and KANADI genes. *Current Biology*, 13(20):1768–1774, 2003.

[34] Klaus FX Mayer, Heiko Schoof, Achim Haecker, Michael Lenhard, Gerd Jürgens, and Thomas Laux. Role of WUSCHEL in regulating stem cell fate in the *Arabidopsis* shoot meristem. *Cell*, 95(6):805–815, 1998.

[35] Michael J Prigge, Denichiro Otsuga, José M Alonso, Joseph R Ecker, Gary N Drews, and Steven E Clark. Class III homeodomain-leucine zipper gene family members have overlapping, antagonistic, and distinct roles in *Arabidopsis* development. *The Plant Cell*, 17(1):61–76, 2005.

[36] Mitsuhiro Aida, Tetsuya Ishida, Hidehiro Fukaki, Hisao Fujisawa, and Masao Tasaka. Genes involved in organ separation in *Arabidopsis*: an analysis of the cup-shaped cotyledon mutant. *The Plant Cell*, 9(6):841–857, 1997.

[37] Jeff A Long, Erich I Moan, June I Medford, and M Kathryn Barton. A member of the KNOTTED class of homeodomain proteins encoded by the STM gene of *Arabidopsis*. *Nature*, 379(6560):66–69, 1996.

[38] John L Bowman and Yuval Eshed. Formation and maintenance of the shoot apical meristem. *Trends in Plant Science*, 5(3):110–115, 2000.

[39] Mitsutomo Abe, Hiroshi Katsumata, Yoshibumi Komeda, and Taku Takahashi. Regulation of shoot epidermal cell differentiation by a pair of homeodomain proteins in *Arabidopsis*. *Development*, 130(4):635–643, 2003.

[40] Rita San-Bento, Etienne Farcot, Roberta Galletti, Audrey Creff, and Gwyneth Ingram. Epidermal identity is maintained by cell–cell communication via a universally active feedback loop in *Arabidopsis thaliana*. *The Plant Journal*, 77(1):46–58, 2014.

[41] Mitsutomo Abe, Taku Takahashi, and Yoshibumi Komeda. Identification of a cis-regulatory element for L1 layer-specific gene expression, which is targeted by an L1-specific homeodomain protein. *The Plant Journal*, 26(5):487–494, 2001.

[42] Miyuki Nakamura, Hiroshi Katsumata, Mitsutomo Abe, Naoto Yabe, Yoshibumi Komeda, Kotaro T Yamamoto, and Taku Takahashi. Characterization of the class IV homeodomain-leucine zipper gene family in *Arabidopsis*. *Plant Physiology*, 141 (4):1363–1375, 2006.

[43] Shinobu Takada, Nozomi Takada, and Ayaka Yoshida. ATML1 promotes epidermal cell differentiation in *Arabidopsis* shoots. *Development*, 140(9):1919–1923, 2013.

[44] Martin Hülskamp. Plant trichomes: a model for cell differentiation. *Nature Reviews Molecular Cell Biology*, 5(6):471–480, 2004.

[45] Daniel B Szymanski, Alan M Lloyd, and M David Marks. Progress in the molecular genetic analysis of trichome initiation and morphogenesis in *Arabidopsis*. *Trends in Plant Science*, 5(5):214–219, 2000.

[46] Jeffrey J Esch, Margaret Chen, Mark Sanders, Matthew Hillestad, Sampson Ndkium, Brian Idelkope, James Neizer, and M David Marks. A contradictory GLABRA3 allele helps define gene interactions controlling trichome development in *Arabidopsis*. *Development*, 130(24):5885–5894, 2003.

[47] Tetsuya Ishida, Tetsuya Kurata, Kiyotaka Okada, and Takuji Wada. A genetic regulatory network in the development of trichomes and root hairs. *Annu. Rev. Plant Biol.*, 59:365–386, 2008.

[48] Victor Kirik, Marissa Simon, Martin Huelskamp, and John Schiefelbein. 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(2):506–513, 2004.

[49] Su-Hwan Kwak, Ronglai Shen, and John Schiefelbein. Positional signaling mediated by a receptor-like kinase in *Arabidopsis*. *Science*, 307(5712):1111–1113, 2005.

[50] Matt Geisler, Jeanette Nadeau, and Fred D Sack. Oriented asymmetric divisions that generate the stomatal spacing pattern in *Arabidopsis* are disrupted by the too many mouths mutation. *The Plant Cell*, 12(11):2075–2086, 2000.

[51] Lynn Jo Pillitteri and Keiko U Torii. Mechanisms of stomatal development. *Annual Review of Plant Biology*, 63:591–614, 2012.

[52] Kenta Hara, Toshiya Yokoo, Ryoko Kajita, Takaaki Onishi, Saiko Yahata, Kylee M Peterson, Keiko U Torii, and Tatsuo Kakimoto. Epidermal cell density is autoregulated via a secretory peptide, EPIDERMAL PATTERNING FACTOR 2 in *Arabidopsis* leaves. *Plant and Cell Physiology*, 50(6):1019–1031, 2009.

[53] Kenta Hara, Ryoko Kajita, Keiko U Torii, Dominique C Bergmann, and Tatsuo Kakimoto. The secretory peptide gene EPF1 enforces the stomatal one-cell-spacing rule. *Genes & Development*, 21(14):1720–1725, 2007.

[54] Lee Hunt and Julie E Gray. The signaling peptide EPF2 controls asymmetric cell divisions during stomatal development. *Current Biology*, 19(10):864–869, 2009.

[55] Elena D Shpak, Jessica Messmer McAbee, Lynn Jo Pillitteri, and Keiko U Torii. Stomatal patterning and differentiation by synergistic interactions of receptor kinases. *Science*, 309(5732):290–293, 2005.

[56] Cora A MacAlister, Kyoko Ohashi-Ito, and Dominique C Bergmann. Transcription factor control of asymmetric cell divisions that establish the stomatal lineage. *Nature*, 445(7127):537–540, 2007.

[57] Lynn Jo Pillitteri, Daniel B Sloan, Naomi L Bogenschutz, and Keiko U Torii. Termination of asymmetric cell division and differentiation of stomata. *Nature*, 445(7127): 501–505, 2007.

[58] Kyoko Ohashi-Ito and Dominique C Bergmann. *Arabidopsis* FAMA controls the final proliferation/differentiation switch during stomatal development. *The Plant Cell*, 18 (10):2493–2505, 2006.

[59] Lien B Lai, Jeanette A Nadeau, Jessica Lucas, Eun-Kyoung Lee, Tsuyoshi Nakagawa, Liming Zhao, Matt Geisler, and Fred D Sack. The *Arabidopsis* R2R3 MYB proteins FOUR LIPS and MYB88 restrict divisions late in the stomatal cell lineage. *The Plant Cell*, 17(10):2754–2767, 2005.

[60] Emily B Abrash and Dominique C Bergmann. Regional specification of stomatal production by the putative ligand CHALLAH. *Development*, 137(3):447–455, 2010.

[61] Tatsuhiko Kondo, Ryoko Kajita, Aya Miyazaki, Mayumi Hokoyama, Touko Nakamura-Miura, Satoko Mizuno, Yuichi Masuda, Kazuhiro Irie, Yuki Tanaka, Shinobu Takada, et al. Stomatal density is controlled by a mesophyll-derived signaling molecule. *Plant and Cell Physiology*, 51(1):1–8, 2010.

[62] Shigeo S Sugano, Tomoo Shimada, Yu Imai, Katsuya Okawa, Atsushi Tamai, Masashi Mori, and Ikuko Hara-Nishimura. Stomagen positively regulates stomatal density in *Arabidopsis*. *Nature*, 463(7278):241–244, 2010.

[63] Juan Dong and Dominique C Bergmann. Stomatal patterning and development. In *Current topics in Developmental Biology*, volume 91, pages 267–297. Elsevier, 2010.

[64] Jeanette A Nadeau and Fred D Sack. Control of stomatal distribution on the *Arabidopsis* leaf surface. *Science*, 296(5573):1697–1700, 2002.

[65] Juergen A Knoblich. Mechanisms of asymmetric stem cell division. *Cell*, 132(4): 583–597, 2008.

[66] Heather N Cartwright, John A Humphries, and Laurie G Smith. A receptor-like protein that promotes polarization of an asymmetric cell division in maize. *Science*, 323(5914):649–651, 2009.

[67] Juan Dong, Cora A MacAlister, and Dominique C Bergmann. BASL controls asymmetric cell division in *Arabidopsis*. *Cell*, 137(7):1320–1330, 2009.

[68] Lynn Jo Pillitteri, Kylee M Peterson, Robin J Horst, and Keiko U Torii. Molecular profiling of stomatal meristemoids reveals new component of asymmetric cell division and commonalities among stem cell populations in *Arabidopsis*. *The Plant Cell*, 23 (9):3260–3275, 2011.

[69] Matthew Rowe, Juan Dong, Annika Weimer, and Dominique Bergmann. A plant-specific polarity module establishes cell fate asymmetry in the *Arabidopsis* stomatal lineage. *bioRxiv*, page 614636, 2019.

[70] Ying Zhang, Dominique C Bergmann, and Juan Dong. Fine-scale dissection of the subdomains of polarity protein basl in stomatal asymmetric cell division. *Journal of experimental botany*, 67(17):5093–5103, 2016.

[71] Dave Jacobs, Danielle Glossip, Heming Xing, Anthony J Muslin, and Kerry Kornfeld. Multiple docking sites on substrate proteins form a modular system that mediates recognition by erk map kinase. *Genes & development*, 13(2):163–175, 1999.

[72] Motohiro Tani and Yusuf A Hannun. Neutral sphingomyelinase 2 is palmitoylated on multiple cysteine residues: role of palmitoylation in subcellular localization. *Journal of Biological Chemistry*, 282(13):10047–10056, 2007.

[73] Hayley A Mattison, Takashi Hayashi, and Andres Barria. Palmitoylation at two cysteine clusters on the C-terminus of GluN2A and GluN2B differentially control synaptic targeting of nmda receptors. *PLoS One*, 7(11):e49089, 2012.

[74] Céline F Mouchel, Georgette C Briggs, and Christian S Hardtke. Natural genetic variation in arabidopsis identifies BREVIS RADIX, a novel regulator of cell proliferation and elongation in the root. *Genes & development*, 18(6):700–714, 2004.

[75] brx.

[76] Masahiro M Kanaoka, Lynn Jo Pillitteri, Hiroaki Fujii, Yuki Yoshida, Naomi L Bogenschutz, Junji Takabayashi, Jian-Kang Zhu, and Keiko U Torii. SCREAM/ICE1 and SCREAM2 specify three cell-state transitional steps leading to *Arabidopsis* stomatal differentiation. *The Plant Cell*, 20(7):1775–1785, 2008.

[77] Jin Suk Lee, Takeshi Kuroha, Marketa Hnilova, Dmitriy Khatayevich, Masahiro M Kanaoka, Jessica M McAbee, Mehmet Sarikaya, Candan Tamerler, and Keiko U Torii. Direct interaction of ligand–receptor pairs specifying stomatal patterning. *Genes & Development*, 26(2):126–136, 2012.

[78] Benoît Menand, Keke Yi, Stefan Jouannic, Laurent Hoffmann, Eoin Ryan, Paul Linstead, Didier G Schaefer, and Liam Dolan. An ancient mechanism controls the development of cells with a rooting function in land plants. *Science*, 316(5830): 1477–1480, 2007.

[79] James D Masucci and John W Schiefelbein. The *rhd6* mutation of *Arabidopsis thaliana* alters root-hair initiation through an auxin-and ethylene-associated process. *Plant Physiology*, 106(4):1335–1346, 1994.

[80] Yukiko Yasui, Shigeyuki Tsukamoto, Tomomi Sugaya, Ryuichi Nishihama, Quan Wang, Hirotaka Kato, Katsuyuki T Yamato, Hidehiro Fukaki, Tetsuro Mimura, Hiroyoshi Kubo, et al. GEMMA CUP-ASSOCIATED MYB1, an ortholog of axillary meristem regulators, is essential in vegetative reproduction in *Marchantia polymorpha*. *Current Biology*, 29(23):3987–3995, 2019.

[81] Takuma Hiwatashi, Honzhen Goh, Yukiko Yasui, Li Quan Koh, Hideyuki Takami, Masataka Kajikawa, Hiroyuki Kirita, Takehiko Kanazawa, Naoki Minamino, Taisuke Togawa, et al. The RopGEF KARAPPO is essential for the initiation of vegetative reproduction in *Marchantia polymorpha*. *Current Biology*, 29(20):3525–3531, 2019.

[82] Qiaohong Duan, Daniel Kita, Chao Li, Alice Y Cheung, and Hen-Ming Wu. FERONIA receptor-like kinase regulates RHO GTPase signaling of root hair development. *Proceedings of the National Academy of Sciences*, 107(41):17821–17826, 2010.

[83] John A Humphries, Zuzana Vejlupkova, Anding Luo, Robert B Meeley, Anne W Sylvester, John E Fowler, and Laurie G Smith. ROP GTPases act with the receptor-

like protein PAN1 to polarize asymmetric cell division in maize. *The Plant Cell*, 23 (6):2273–2284, 2011.

[84] Michelle R Facette, Yeri Park, Dena Sutimantanapi, Anding Luo, Heather N Cartwright, Bing Yang, Eric J Bennett, Anne W Sylvester, and Laurie G Smith. The SCAR/WAVE complex polarizes PAN receptors and promotes division asymmetry in maize. *Nature Plants*, 1(2):1–8, 2015.

[85] Kimitsune Ishizaki, Miya Mizutani, Masaki Shimamura, Akihide Masuda, Ryuichi Nishihama, and Takayuki Kohchi. Essential role of the E3 ubiquitin ligase nopperabo1 in schizogenous intercellular space formation in the liverwort *Marchantia polymorpha*. *The Plant Cell*, 25(10):4075–4084, 2013.

[86] Victor AS Jones and Liam Dolan. Mp*WIP* regulates air pore complex development in the liverwort *Marchantia polymorpha*. *Development*, 144(8):1472–1476, 2017.

[87] Masaki Shimamura. *Marchantia polymorpha*: taxonomy, phylogeny and morphology of a model system. *Plant and Cell Physiology*, 57(2):230–256, 2016.

[88] Charles R Barnes and William Jesse Goad Land. Bryological papers. I. The origin of air chambers. *Botanical Gazette*, 44(3):197–213, 1907.

[89] Karen S Renzaglia, Juan C Villarreal, and R Joel Duff. New insights into morphology, anatomy, and systematics of hornworts. *Bryophyte biology*, 2:139–171, 2009.

[90] Philip Carella, Anna Gogleva, Marta Tomaselli, Carolin Alfs, and Sebastian Schornack. Phytophthora palmivora establishes tissue-specific intracellular infection structures in the earliest divergent land plant lineage. *Proceedings of the National Academy of Sciences*, 115(16):E3846–E3855, 2018.

[91] Linda E Graham, Martha E Cook, David T Hanson, Kathleen B Pigg, and James M Graham. Structural, physiological, and stable carbon isotopic evidence that the enigmatic paleozoic fossil prototaxites formed from rolled liverwort mats. *American Journal of Botany*, 97(2):268–275, 2010.

[92] Hubert Leitgeb. *Untersuchungen über die Lebermoose: Die Marchantieen und allgemeine Bemerkungen über Lebermoose. VI.(Schluss)-Heft.* Leuschner & Lubensky, 1881.

[93] P Apostolakos, B Galatis, and K Mitrakos. Studies on the development of the air pores and air chambers of *Marchantia paleacea*: 1. Light Microscopy. *Annals of Botany*, 49(3):377–396, 1982.

[94] P Apostolakos and B Galatis. Studies on the development of the air pores and air chambers of *Marchantia paleacea*. II. Ultrastructure of the initial aperture formation with particular reference to cortical microtubule organizing centres. *Canadian Journal of Botany*, 63(4):744–756, 1985.

[95] Hidemasa Suzuki, C Jill Harrison, Masaki Shimamura, Takayuki Kohchi, and Ryuichi Nishihama. Positional cues regulate dorsal organ formation in the liverwort *Marchantia polymorpha*. *Journal of Plant Research*, pages 1–11, 2020.

[96] Karen S Renzaglia, Juan Carlos Villarreal, Bryan T Piatkowski, Jessica R Lucas, and Amelia Merced. Hornwort stomata: architecture and fate shared with 400-million-year-old fossil plants without leaves. *Plant Physiology*, 174(2):788–797, 2017.

[97] Cora A MacAlister and Dominique C Bergmann. Sequence and function of basic helix–loop–helix proteins required for stomatal development in *Arabidopsis* are deeply conserved in land plants. *Evolution & Development*, 13(2):182–192, 2011.

[98] Kimitsune Ishizaki. Development of schizogenous intercellular spaces in plants. *Frontiers in Plant Science*, 6:497, 2015.

[99] Sean A Montgomery, Yasuhiro Tanizawa, Bence Galik, Nan Wang, Tasuku Ito, Takako Mochizuki, Svetlana Akimcheva, John L Bowman, Valérie Cognat, Laurence Maréchal-Drouard, et al. Chromatin organization in early land plants reveals an ancestral association between H3K27me3, transposons, and constitutive heterochromatin. *Current Biology*, 30(4):573–588, 2020.

[100] John L Bowman, Takayuki Kohchi, Katsuyuki T Yamato, Jerry Jenkins, Shengqiang Shu, Kimitsune Ishizaki, Shohei Yamaoka, Ryuichi Nishihama, Yasukazu Nakamura, Frédéric Berger, et al. Insights into land plant evolution garnered from the *Marchantia polymorpha* genome. *Cell*, 171(2):287–304, 2017.

[101] Kanji Ohyama, Hideya Fukuzawa, Takayuki Kohchi, Hiromasa Shirai, Tohru Sano, Satoshi Sano, Kazuhiko Umesono, Yasuhiko Shiki, Masayuki Takeuchi, Zhen Chang, et al. Chloroplast gene organization deduced from complete sequence of liverwort *Marchantia polymorpha* chloroplast DNA. *Nature*, 322(6079):572–574, 1986.

[102] Kenji Oda, Katsuyuki Yamato, Eiji Ohta, Yasukazu Nakamura, Miho Takemura, Naoko Nozato, Kinya Akashi, Takeshi Kanegae, Yutaka Ogura, Takayuki Kohchi, et al. Gene organization deduced from the complete sequence of liverwort *Marchantia polymorpha* mitochondrial DNA: a primitive form of plant mitochondrial genome. *Journal of Molecular Biology*, 223(1):1–7, 1992.

[103] Kimitsune Ishizaki, Shota Chiyoda, Katsuyuki T Yamato, and Takayuki Kohchi. *Agrobacterium*-mediated transformation of the haploid liverwort *Marchantia polymorpha L.*, an emerging model for plant biology. *Plant and Cell Physiology*, 49(7): 1084–1091, 2008.

[104] Shota Chiyoda, Kimitsune Ishizaki, Hideo Kataoka, Katsuyuki T Yamato, and Takayuki Kohchi. Direct transformation of the liverwort *Marchantia polymorpha L.* by particle bombardment using immature thalli developing from spores. *Plant cell Reports*, 27(9):1467, 2008.

[105] Shota Chiyoda, Philip J Linley, Katsuyuki T Yamato, Hideya Fukuzawa, Akiho Yokota, and Takayuki Kohchi. Simple and efficient plastid transformation system for the liverwort *Marchantia polymorpha L.* suspension-culture cells. *Transgenic Research*, 16(1):41–49, 2007.

[106] Ryuichi Nishihama, Sakiko Ishida, Hiroko Urawa, Yasuhiro Kamei, and Takayuki Kohchi. CRISPR/Cas9-mediated targeted mutagenesis in the liverwort *Marchantia polymorpha* using its own heat-shock promoter and Cre/lox P-mediated site-specific recombination. *Plant and Cell Physiology*, 57(2):271–280, 2016.

[107] Ryuichi Nishihama, Sakiko Ishida, Hiroko Urawa, Yasuhiro Kamei, and Takayuki Kohchi. Conditional gene expression/deletion systems for *Marchantia polymorpha* using its own heat-shock promoter and Cre/lox P-mediated site-specific recombination. *Plant and Cell Physiology*, 57(2):271–280, 2016.

[108] Bernardo Pollak, Ariel Cerda, Mihails Delmans, Simón Álamos, Tomás Moyano, Anthony West, Rodrigo A Gutiérrez, Nicola J Patron, Fernán Federici, and Jim Haseloff. Loop assembly: a simple and open system for recursive fabrication of DNA circuits. *New Phytologist*, 222(1):628–640, 2019.

[109] Nicola J Patron, Diego Orzaez, Sylvestre Marillonnet, Heribert Warzecha, Colette Matthewman, Mark Youles, Oleg Raitskin, Aymeric Leveau, Gemma Farré, Christian Rogers, et al. Standards for plant synthetic biology: a common syntax for exchange of DNA parts. *New Phytologist*, 208(1):13–19, 2015.

[110] Peter Hajdukiewicz, Zora Svab, and Pal Maliga. The small, versatile pPZP family of *Agrobacterium* binary vectors for plant transformation. *Plant Molecular Biology*, 25(6):989–994, 1994.

[111] Maximilian Haeussler, Kai Schönig, Hélène Eckert, Alexis Eschstruth, Joffrey Mianné, Jean-Baptiste Renaud, Sylvie Schneider-Maunoury, Alena Shkumatava, Lydia Teboul, Jim Kent, et al. Evaluation of off-target and on-target scoring algorithms

and integration into the guide RNA selection tool CRISPOR. *Genome Biology*, 17 (1):148, 2016.

[112] Jeffrey Doyle. DNA protocols for plants. In *Molecular techniques in taxonomy*, pages 283–293. Springer, 1991.

[113] Sue Porebski, L Grant Bailey, and Bernard R Baum. Modification of a CTAB DNA extraction protocol for plants containing high polysaccharide and polyphenol components. *Plant Molecular Biology Reporter*, 15(1):8–15, 1997.

[114] Rolf Deblaere, Benny Bytebier, Henri De Greve, Francine Deboeck, Jeff Schell, Marc Van Montagu, and J Leemans. Efficient octopine Ti plasmid-derived vectors for *Agrobacterium*-mediated gene transfer to plants. *Nucleic acids research*, 13(13):4777–4788, 1985.

[115] Roger P Hellens, E Anne Edwards, Nicola R Leyland, Samantha Bean, and Philip M Mullineaux. pGreen: a versatile and flexible binary Ti vector for *Agrobacterium*-mediated plant transformation. *Plant Molecular Biology*, 42(6):819–832, 2000.

[116] Daisuke Kurihara, Yoko Mizuta, Yoshikatsu Sato, and Tetsuya Higashiyama. ClearSee: a rapid optical clearing reagent for whole-plant fluorescence imaging. *Development*, 142(23):4168–4179, 2015.

[117] Johannes Schindelin, Ignacio Arganda-Carreras, Erwin Frise, Verena Kaynig, Mark Longair, Tobias Pietzsch, Stephan Preibisch, Curtis Rueden, Stephan Saalfeld, Benjamin Schmid, et al. Fiji: an open-source platform for biological-image analysis. *Nature Methods*, 9(7):676–682, 2012.

[118] Caroline A Schneider, Wayne S Rasband, and Kevin W Eliceiri. NIH image to ImageJ: 25 years of image analysis. *Nature Methods*, 9(7):671–675, 2012.

[119] Simon Andrews, Felix Krueger, Anne Segonds-Pichon, Laura Biggins, Christel Krueger, and Steven Wingett. FastQC. Babraham Institute, January 2012.

[120] Felix Krueger. Trim galore. *A wrapper tool around Cutadapt and FastQC to consistently apply quality and adapter trimming to FastQ files*, 516:517, 2015.

[121] Nicolas L Bray, Harold Pimentel, Páll Melsted, and Lior Pachter. Near-optimal probabilistic RNA-seq quantification. *Nature Biotechnology*, 34(5):525–527, 2016.

[122] Charlotte Soneson, Michael I Love, and Mark D Robinson. Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. *F1000Research*, 4, 2015.

[123] Michael I Love, Wolfgang Huber, and Simon Anders. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, 15(12):550, 2014.

[124] Yoav Benjamini and Yosef Hochberg. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, 57(1):289–300, 1995.

[125] Hadley Wickham. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York, 2016. ISBN 978-3-319-24277-4. URL `http://ggplot2.org`.

[126] Raivo Kolde. Pheatmap: pretty heatmaps. *R package version*, 1(2), 2012.

[127] Douglas S Richardson and Jeff W Lichtman. Clarifying tissue clearing. *Cell*, 162(2): 246–257, 2015.

[128] Arthur W Toga, Amir Goldkorn, Karen Ambach, Kuo Chao, Bruce C Quinn, and Patrick Yao. Postmortem cryosectioning as an anatomic reference for human brain mapping. *Computerized Medical Imaging and Graphics*, 21(2):131–141, 1997.

[129] Arthur W Toga, Karen L Ambach, and Stefanie Schluender. High-resolution anatomy from in situ human brain. *Neuroimage*, 1(4):334–344, 1994.

[130] Marvin Minsky. Memoir on inventing the confocal scanning microscope. *Scanning*, 10(4):128–138, 1988.

[131] José-Angel Conchello and Jeff W Lichtman. Optical sectioning microscopy. *Nature Methods*, 2(12):920–931, 2005.

[132] Jerome Mertz. Optical sectioning microscopy with planar or structured illumination. *Nature Methods*, 8(10):811, 2011.

[133] Emmanuel G Reynaud, Uroš Kržič, Klaus Greger, and Ernst HK Stelzer. Light sheet-based fluorescence microscopy: more dimensions, more photons, and less photodamage. *HFSP journal*, 2(5):266–275, 2008.

[134] B Clancy and LJ Cauller. Reduction of background autofluorescence in brain sections following immersion in sodium borohydride. *Journal of Neuroscience Methods*, 83 (2):97–102, 1998.

[135] T Cowen, AJ Haven, and G Burnstock. Pontamine sky blue: a counterstain for background autofluorescence in fluorescence and immunofluorescence histochemistry. *Histochemistry*, 82(3):205–208, 1985.

[136] Michael Neumann and Detlef Gabel. Simple method for reduction of autofluorescence in fluorescence microscopy. *Journal of Histochemistry & Cytochemistry*, 50(3):437–439, 2002.

[137] Hiroshi Hama, Hiroshi Kurokawa, Hiroyuki Kawano, Ryoko Ando, Tomomi Shimogori, Hisayori Noda, Kiyoko Fukami, Asako Sakaue-Sawano, and Atsushi Miyawaki. Scale: a chemical approach for fluorescence imaging and reconstruction of transparent mouse brain. *Nature Neuroscience*, 14(11):1481, 2011.

[138] Etsuo A Susaki, Kazuki Tainaka, Dimitri Perrin, Fumiaki Kishino, Takehiro Tawara, Tomonobu M Watanabe, Chihiro Yokoyama, Hirotaka Onoe, Megumi Eguchi, Shun Yamaguchi, et al. Whole-brain imaging with single-cell resolution using chemical cocktails and computational analysis. *Cell*, 157(3):726–739, 2014.

[139] Meng-Tsen Ke, Satoshi Fujimoto, and Takeshi Imai. SeeDB: a simple and morphology-preserving optical clearing agent for neuronal circuit reconstruction. *Nature Neuroscience*, 16(8):1154, 2013.

[140] Kwanghun Chung and Karl Deisseroth. CLARITY for mapping the nervous system. *Nature Methods*, 10(6):508, 2013.

[141] William M Palmer, Antony P Martin, Jamie R Flynn, Stephanie L Reed, Rosemary G White, Robert T Furbank, and Christopher PL Grof. PEA-CLARITY: 3D molecular imaging of whole plant organs. *Scientific Reports*, 5(1):1–6, 2015.

[142] Thomas J Musielak, Daniel Slane, Christian Liebig, and Martin Bayer. A versatile optical clearing protocol for deep tissue imaging of fluorescent proteins in *Arabidopsis thaliana*. *PLoS One*, 11(8):e0161107, 2016.

[143] Jim Haseloff. Old botanical techniques for new microscopes. *Biotechniques*, 34(6): 1174–1182, 2003.

[144] Sean R Cutler, David W Ehrhardt, Joel S Griffitts, and Chris R Somerville. Random GFP cDNA fusions enable visualization of subcellular structures in cells of *Arabidopsis* at a high frequency. *Proceedings of the National Academy of Sciences*, 97(7):3718–3723, 2000.

[145] Felix Althoff, Sarah Kopischke, Oliver Zobell, Kentaro Ide, Kimitsune Ishizaki, Takayuki Kohchi, and Sabine Zachgo. Comparison of the MpEF1α and CaMV35 promoters for application in *Marchantia polymorpha* overexpression studies. *Transgenic research*, 23(2):235–244, 2014.

[146] Jinyoung Seo, Minjin Choe, and Sung-Yon Kim. Clearing and labeling techniques for large-scale biological tissues. *Molecules and Cells*, 39(6):439, 2016.

[147] Liam Dolan, Kees Janmaat, Viola Willemsen, Paul Linstead, Scott Poethig, Keith Roberts, and Ben Scheres. Cellular organisation of the *Arabidopsis thaliana* root. *Development*, 119(1):71–84, 1993.

[148] SG Mansfield and LG Briarty. Early embryogenesis in *Arabidopsis thaliana*. II. the developing embryo. *Canadian Journal of Botany*, 69(3):461–476, 1991.

[149] Seung Y Rhee, Kenneth D Birnbaum, and David W Ehrhardt. Towards building a plant cell atlas. *Trends in Plant Science*, 24(4):303–310, 2019.

[150] Leopold Kny. *Botanische Wandtafeln: mit erläuterndem Text; Abth. 1-3*. Wiegandt, Hempel & Parey, 1874.

[151] Jean-Jules Beauverie. Etudes des modifications morphologiques et anatomiques de thalles de Marchantia et de Lunularia obtenues expérimentalement. *Publications de la Société Linnéenne de Lyon*, 44(1):57–69, 1898.

[152] Gerhard Lorbeer. Untersuchungen über Reduktionsteilung und Geschlechtsbestimmung bei Lebermoosen. *Zeitschrift für Induktive Abstammungs-und Vererbungslehre*, 44(1):1–109, 1927.

[153] M Mirbel. Researches anatomiques et physiologiques sur le *Marchantia polymorpha*. *Mém Acad Roy Sc Inst France*, 13:337–436, 1835.

[154] John L Bowman. A brief history of *Marchantia* from Greece to genomics. *Plant and Cell Physiology*, 57(2):210–229, 2016.

[155] Tiesheng Gu, Maria Mazzurco, Waheeda Sulaman, Dinah D Matias, and Daphne R Goring. Binding of an arm repeat protein to the kinase domain of the S-locus receptor kinase. *Proceedings of the National Academy of Sciences*, 95(1):382–387, 1998.

[156] Moonil Kim, Hye Sun Cho, Do-Myung Kim, Jeong Hee Lee, and Hyun-Sook Pai. CHRK1, a chitinase-related receptor-like kinase, interacts with NtPUB4, an armadillo repeat protein, in tobacco. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*, 1651(1-2):50–59, 2003.

[157] Marcus A Samuel, Jennifer N Salt, Shin-Han Shiu, and Daphne R Goring. Multifunctional arm repeat domains in plants. *International Review of Cytology*, 253:1–26, 2006.

[158] Marcus A Samuel, Yashwanti Mudgil, Jennifer N Salt, Frédéric Delmas, Shaliny Ramachandran, Andrea Chilelli, and Daphne R Goring. Interactions between the S-domain receptor kinases and AtPUB-ARM E3 ubiquitin ligases suggest a conserved signaling pathway in *Arabidopsis*. *Plant Physiology*, 147(4):2084–2095, 2008.

[159] Malick Mbengue, Sylvie Camut, Fernanda de Carvalho-Niebel, Laurent Deslandes, Solene Froidure, Dörte Klaus-Heisen, Sandra Moreau, Susana Rivas, Ton Timmers, Christine Hervé, et al. The *Medicago truncatula* E3 ubiquitin ligase PUB1 interacts with the LYK3 symbiotic receptor and negatively regulates infection and nodulation. *The Plant Cell*, 22(10):3474–3488, 2010.

[160] Dongping Lu, Wenwei Lin, Xiquan Gao, Shujing Wu, Cheng Cheng, Julian Avila, Antje Heese, Timothy P Devarenne, Ping He, and Libo Shan. Direct ubiquitination of pattern recognition receptor FLS2 attenuates plant innate immunity. *Science*, 332 (6036):1439–1442, 2011.

[161] Nuri Purswani Ramchandani. *Introducing the gemma of the liverwort Marchantia polymorpha L. as a simple morphogenetic system*. PhD thesis, University of Cambridge, 2015.

[162] P Apostolakos and B Galatis. Studies on the development of the air pores and air chambers of *Marchantia paleacea*. IV: Cell plate arrangement in initial aperture cells. *Protoplasma*, 128(2-3):136–146, 1985.

[163] John C Larkin, Nevin Young, Michael Prigge, and M David Marks. The control of trichome spacing and number in *Arabidopsis*. *Development*, 122(3):997–1005, 1996.

[164] Jian-Feng Li, Julie E Norville, John Aach, Matthew McCormack, Dandan Zhang, Jenifer Bush, George M Church, and Jen Sheen. Multiplex and homologous recombination–mediated genome editing in *Arabidopsis* and *Nicotiana benthamiana* using guide RNA and Cas9. *Nature Biotechnology*, 31(8):688–691, 2013.

[165] Shingo Nagaya, Miho Takemura, and Kanji Ohyama. Endogenous promoter, 5UTR and transcriptional terminator enhance transient gene expression in a liverwort, *Marchantia polymorpha L.* *Plant Biotechnology*, pages 1112260060–1112260060, 2011.

[166] Bernardo Pollak Williamson. *Frameworks for reprogramming early diverging land plants*. PhD thesis, University of Cambridge, 2018.

[167] Mihails Delmans, Bernardo Pollak, and Jim Haseloff. MarpoDB: an open registry for *Marchantia polymorpha* genetic parts. *Plant and Cell Physiology*, 58(1):e5–e5, 2017.

[168] Sebastian Wolf. Plant cell wall signalling and receptor-like kinases. *Biochemical Journal*, 474(4):471–492, 2017.

[169] Ruth N Kaplan-Levy, Philip B Brewer, Tezz Quon, and David R Smyth. The trihelix family of transcription factors–light, stress and development. *Trends in Plant Science*, 17(3):163–171, 2012.

[170] Yoshimi Nakano, Masatoshi Yamaguchi, Hitoshi Endo, Nur Ardiyana Rejab, and Misato Ohtani. NAC-MYB-based transcriptional regulation of secondary cell wall biosynthesis in land plants. *Frontiers in Plant Science*, 6:288, 2015.

[171] Mohammed Nuruzzaman, Akhter Most Sharoni, and Shoshi Kikuchi. Roles of NAC transcription factors in the regulation of biotic and abiotic stress responses in plants. *Frontiers in Microbiology*, 4:248, 2013.

[172] C Thomas Payne, Fan Zhang, and Alan M Lloyd. GL3 encodes a bHLH protein that regulates trichome development in *Arabidopsis* through interaction with GL1 and TTG1. *Genetics*, 156(3):1349–1362, 2000.

[173] Cameron S Johnson, Ben Kolevski, and David R Smyth. TRANSPARENT TESTA GLABRA2, a trichome and seed coat development gene of *Arabidopsis*, encodes a WRKY transcription factor. *The Plant Cell*, 14(6):1359–1375, 2002.

[174] Bruno Catarino, Alexander J Hetherington, David M Emms, Steven Kelly, and Liam Dolan. The stepwise increase in the number of transcription factor families in the Precambrian predated the diversification of plants on land. *Molecular Biology and Evolution*, 33(11):2815–2819, 2016.

[175] Matthew WB Spencer, Stuart A Casson, and Keith Lindsey. Transcriptional profiling of the *Arabidopsis* embryo. *Plant Physiology*, 143(2):924–940, 2007.

[176] Markus Schmid, Timothy S Davison, Stefan R Henz, Utz J Pape, Monika Demar, Martin Vingron, Bernhard Schölkopf, Detlef Weigel, and Jan U Lohmann. A gene expression map of *Arabidopsis thaliana* development. *Nature Genetics*, 37(5):501–506, 2005.

[177] Joakim Palovaara, Shunsuke Saiga, and Dolf Weijers. Transcriptomics approaches in the early *Arabidopsis* embryo. *Trends in Plant Science*, 18(9):514–521, 2013.

[178] C Branca, G De Lorenzo, and F Cervone. Competitive inhibition of the auxin-induced elongation by α-d-oligogalacturonides in pea stem segments. *Physiologia Plantarum*, 72(3):499–504, 1988.

[179] Daniela Bellincampi, Maura Cardarelli, Daniela Zaghi, Giovanna Serino, Giovanni Salvi, Christiane Gatz, Felice Cervone, Maria Maddalena Altamura, Paolo Costantino, and Giulia De Lorenzo. Oligogalacturonides prevent rhizogenesis in rolb-transformed tobacco explants by inhibiting auxin-induced expression of the rolb gene. *The Plant Cell*, 8(3):477–487, 1996.

[180] Anne Dievart, Céline Gottin, Christophe Périn, Vincent Ranwez, and Nathalie Chantret. Origin and diversity of plant receptor-like kinases. *Annual Review of Plant Biology*, 71, 2020.

[181] Dominique C Bergmann and Fred D Sack. Stomatal development. *Annu. Rev. Plant Biol.*, 58:163–181, 2007.

[182] Haruka Arai, Kazuya Yanagiura, Yuko Toyama, and Kengo Morohashi. Genome-wide analysis of MpBHLH12, a IIIf basic helix-loop-helix transcription factor of *Marchantia polymorpha*. *Journal of Plant Research*, 132(2):197–209, 2019.

[183] Paul J Rushton, Imre E Somssich, Patricia Ringler, and Qingxi J Shen. WRKY transcription factors. *Trends in Plant Science*, 15(5):247–258, 2010.

[184] Philip Carella, Anna Gogleva, David John Hoey, Anthony John Bridgen, Sara Christina Stolze, Hirofumi Nakagami, and Sebastian Schornack. Conserved Biochemical Defenses Underpin Host Responses to Oomycete Infection in an Early-Divergent Land Plant Lineage. *Current Biology*, 29(14):2282–2294, 2019.

[185] Damian R Page and Ueli Grossniklaus. The art and design of genetic screens: *Arabidopsis thaliana*. *Nature Reviews Genetics*, 3(2):124–136, 2002.

[186] Viswanathan Chinnusamy, Becky Stevenson, Byeong-ha Lee, and Jian-Kang Zhu. Screening for gene regulation mutants by bioluminescence imaging. *Science's STKE*, 2002(140):pl10–pl10, 2002.

[187] Kiyoshi Onai, Kazuhisa Okamoto, Harumi Nishimoto, Chisato Morioka, Minako Hirano, Nobunori Kami-ike, and Masahiro Ishiura. Large-scale screening of *Arabidopsis* circadian clock mutants by a high-throughput real-time bioluminescence monitoring system. *The Plant Journal*, 40(1):1–11, 2004.

[188] Megan M Southern and Andrew J Millar. Circadian genetics in the model higher plant, *Arabidopsis thaliana*. In *Methods in Enzymology*, volume 393, pages 23–35. Elsevier, 2005.

[189] Jim Haseloff, Emma-Louise Dormand, and Andrea H Brand. Live imaging with green fluorescent protein. In *Confocal microscopy Methods and Protocols*, pages 241–259. Springer, 1999.

[190] Andrew J Millar, Isabelle A Carre, Carl A Strayer, Nam-Hai Chua, and Steve A Kay. Circadian clock mutants in *Arabidopsis* identified by luciferase imaging. *Science*, 267 (5201):1161–1163, 1995.

[191] Carsten Meier, Thomas Bouquin, Mads Eggert Nielsen, Dora Raventos, Ole Mattsson, Anne Rocher, Fritz Schomburg, Richard M Amasino, and John Mundy. Gibberellin response mutants identified by luciferase imaging. *The Plant Journal*, 25(5): 509–519, 2001.

[192] Liming Xiong, Manabu Ishitani, Hojoung Lee, and Jian-Kang Zhu. The *Arabidopsis* LOS5/ABA3 locus encodes a molybdenum cofactor sulfurase and modulates cold stress–and osmotic stress–responsive gene expression. *The Plant Cell*, 13(9):2063–2083, 2001.

[193] Daniel B Szymanski, Ross A Jilk, Susan M Pollock, and M David Marks. Control of GL2 expression in *Arabidopsis* leaves and trichomes. *Development*, 125(7):1161–1171, 1998.

[194] Athanasios Lampropoulos, Zoran Sutikovic, Christian Wenzl, Ira Maegele, Jan U Lohmann, and Joachim Forner. GreenGate-a novel, versatile, and efficient cloning system for plant transgenesis. *PLOS ONE*, 8(12):e83043, 2013.

[195] Carola Engler, Mark Youles, Ramona Gruetzner, Tim-Martin Ehnert, Stefan Werner, Jonathan DG Jones, Nicola J Patron, and Sylvestre Marillonnet. A golden gate modular cloning toolbox for plants. *ACS Synthetic Biology*, 3(11):839–843, 2014.

[196] Nrisingha Dey, Shayan Sarkar, Sefali Acharya, and Indu B Maiti. Synthetic promoters *in planta*. *Planta*, 242(5):1077–1094, 2015.

[197] Sajid Ali and Won-Chan Kim. A fruitful decade using synthetic promoters in the improvement of transgenic plants. *Frontiers in Plant Science*, 10, 2019.

[198] Yaomin Cai, Kalyani Kallam, Henry Tidd, Giovanni Gendarini, Amanda Salzman, and Nicola J Patron. Rational Design of Minimal Synthetic Promoters for Plants. *bioRxiv*, 2020.

[199] Betty YW Chung, Cas Simons, Andrew E Firth, Chris M Brown, and Roger P Hellens. Effect of 5'UTR introns on gene expression in *Arabidopsis thaliana*. *BMC genomics*, 7(1):1–13, 2006.

[200] Sitakanta Pattanaik, Barunava Patra, Sanjay Kumar Singh, and Ling Yuan. An overview of the gene regulatory network controlling trichome development in the model plant, *Arabidopsis*. *Frontiers in Plant Science*, 5:259, 2014.

[201] Elisabeth Truernit, Hélène Bauby, Bertrand Dubreucq, Olivier Grandjean, John Runions, Julien Barthélémy, and Jean-Christophe Palauqui. High-resolution whole-mount imaging of three-dimensional tissue organization and gene expression enables the study of phloem development and structure in *Arabidopsis*. *The Plant Cell*, 20 (6):1494–1503, 2008.

[202] AM Cormack. Reconstruction of densities from their projections, with applications in radiological physics. *Physics in Medicine & Biology*, 18(2):195, 1973.

[203] Paul C Lauterbur. Image formation by induced local interactions: examples employing nuclear magnetic resonance. *nature*, 242(5394):190–191, 1973.

[204] Elisabeth Truernit. Phloem imaging. *Journal of experimental botany*, 65(7):1681–1688, 2014.

[205] Thomas J Musielak, Laura Schenkel, Martina Kolb, Agnes Henschen, and Martin Bayer. A simple and versatile cell wall staining protocol to study plant reproduction. *Plant Reproduction*, 28(3-4):161–169, 2015.

[206] Eduardo Flores-Sandoval, D Magnus Eklund, and John L Bowman. A simple auxin transcriptional response system regulates multiple morphogenetic processes in the liverwort *Marchantia polymorpha*. *PLoS Genet*, 11(5):e1005207, 2015.

[207] Hirotaka Kato, Kimitsune Ishizaki, Masaru Kouno, Makoto Shirakawa, John L Bowman, Ryuichi Nishihama, and Takayuki Kohchi. Auxin-mediated transcriptional system with a minimal set of components is critical for morphogenesis through the life cycle in *Marchantia polymorpha*. *PLoS Genet*, 11(5):e1005084, 2015.

[208] Eduardo Flores-Sandoval, D Magnus Eklund, Syuan-Fei Hong, John P Alvarez, Tom J Fisher, Edwin R Lampugnani, John F Golz, Alejandra Vázquez-Lobo, Tom Dierschke, Shih-Shun Lin, et al. Class C ARFs evolved before the origin of land plants and antagonize differentiation and developmental transitions in *Marchantia polymorpha*. *New Phytologist*, 218(4):1612–1630, 2018.

[209] Marie-Theres Hauser. Molecular basis of natural variation and environmental control of trichome patterning. *Frontiers in Plant Science*, 5:320, 2014.

[210] Jorge Hernández-García, Asier Briones-Moreno, Renaud Dumas, and Miguel A Blázquez. Origin of Gibberellin-Dependent Transcriptional Regulation by Molecular

Exploitation of a Transactivation Domain in DELLA Proteins. *Molecular Biology and Evolution*, 36(5):908–918, 01 2019. ISSN 0737-4038. doi: 10.1093/molbev/msz009.

[211] Pengfei Wang, Qianqian Zhang, Yingchun Chen, Yanxia Zhao, Fengshan Ren, Hongmei Shi, and Xinying Wu. Comprehensive identification and analysis of della genes throughout the plant kingdom. *BMC Plant Biology*, 20(1):1–12, 2020.

[212] Asuka Higo, Masaki Niwa, Katsuyuki T Yamato, Lixy Yamada, Hitoshi Sawada, Tomoaki Sakamoto, Tetsuya Kurata, Makoto Shirakawa, Motomu Endo, Shuji Shigenobu, et al. Transcriptional framework of male gametogenesis in the liverwort *Marchantia polymorpha L. Plant and Cell Physiology*, 57(2):325–338, 2016.

[213] Hiroyoshi Kubo, Shunsuke Nozawa, Takuma Hiwatashi, Youichi Kondou, Ryo Nakabayashi, Tetsuya Mori, Kazuki Saito, Kojiro Takanashi, Takayuki Kohchi, and Kimitsune Ishizaki. Biosynthesis of riccionidins and marchantins is regulated by R2R3-MYB transcription factors in *Marchantia polymorpha*. *Journal of Plant Research*, 131(5):849–864, 2018.

[214] Shohei Yamaoka, Ryuichi Nishihama, Yoshihiro Yoshitake, Sakiko Ishida, Keisuke Inoue, Misaki Saito, Keitaro Okahashi, Haonan Bao, Hiroyuki Nishida, Katsushi Yamaguchi, et al. Generative cell specification requires transcription factors evolutionarily conserved in land plants. *Current Biology*, 28(3):479–486, 2018.

[215] Niharika Sharma, Prem L Bhalla, and Mohan B Singh. Transcriptome-wide profiling and expression analysis of transcription factor families in a liverwort, *Marchantia polymorpha*. *BMC Genomics*, 14(1):915, 2013.

[216] Yoko Ikeda, Ryuichi Nishihama, Shohei Yamaoka, Mario A Arteaga-Vazquez, Adolfo Aguilar-Cruz, Daniel Grimanelli, Romain Pogorelcnik, Robert A Martienssen, Katsuyuki T Yamato, Takayuki Kohchi, et al. Loss of GC methylation in *Marchantia polymorpha* causes disorganization of cell division and reveals unique DNA methylation regulatory mechanisms of non-GC methylation. *Plant and Cell Physiology*, 59 (12):2421–2431, 2018.