

The 2-dimensional to 3-dimensional growth transition in *Physcomitrium patens*



Zoe Weeks
Balliol College
University of Oxford
A thesis submitted for the degree of
Doctor of Philosophy
Hilary Term 2024

Abstract

Plants colonised the land around 470 million years ago. This was facilitated by the acquisition of apical cells capable of dividing in three dimensions, which allowed the development of complex 3-dimensional (3D) morphologies. This is in contrast to the charophyte green algae, from which land plants evolved, which have apical cells that can divide in only 1 or 2 dimensions. Therefore, 3D growth is a pivotal and unifying feature of all land plants, but we know little about how it is regulated. The model moss *Physcomitrium patens* has been developed as a model system in which to dissect the genetic basis of 3D growth, which also involves complex auxin-cytokinin crosstalk. Previously, a 3D-defective mutant named *Ppnog1R* was generated in a forward genetics screen. *Ppnog1R* exhibits cell division orientation defects that result in early developmental arrest of the 3D “shoots” and consequently a complete failure to establish 3D growth. Therefore, 3D growth in *P. patens* depends on the *PpNOG1* gene, which encodes a protein with a ubiquitin-associated domain. In this work, the function of *PpNOG1* was further clarified by genetic manipulation of the *Ppnog1R* mutant and analysis of the resulting phenotypes. To reveal genetic interactors of *PpNOG1*, a suppressor screen was performed using an independent *nog1* mutant, which yielded two *suppressor of nog1* mutants (*snog1a* and *snog1b*) with restored capacity for 3D growth. The causative mutation in *snog1a* was mapped and the *PpSNOG1A* gene was identified as an inhibitor of 3D growth and genetic antagonist of *PpNOG1*. *PpSNOG1A* also encodes a protein with a ubiquitin-associated domain. Finally, several RNA-seq experiments were performed, which allowed examination of the transcriptomic landscape of *Ppnog1R*, another 3D-defective mutant, *Ppnog2R*, and the *snog1a* suppressor mutant. These analyses revealed processes and pathways that have been impacted in these mutants.

Declaration

I declare that this thesis was composed by myself and that the work contained herein is my own except where explicitly stated in the text. The work has not been submitted for any degree or professional qualification except as specified.

Zoe Weeks

Hilary Term, 2024

Dedication

This work is dedicated to my parents, for their provision of love, support and good food during this endeavour.

Acknowledgements

First and foremost, I would like to thank my supervisor Dr Laura Moody for being a wonderful mentor and giving me so much support, guidance and encouragement. I would also like to thank Professor Nick Harberd for his ideas and advice. This work could not have been completely without the generous funding of the Biotechnology and Biological Sciences Research Council and the Royal Society. My thanks also to the Oxford Interdisciplinary Doctoral Training Partnership for giving me this opportunity. I would like to express my appreciation for everyone in my lab group, past and present, for their help, advice and camaraderie. In particular, a special thanks to Emily day and Gargi Chaturvedi for their contributions to my work, but also Katy Mather, Rency Raquid, Professor Josefina Rodriguez-Enriquez, Richard Jaeger, Siobhan Bridson, Dr Aishah Binti Mohd Kamal and Abir Patwary. I am hugely grateful for the assistance of the technical staff, in particular Sarah Rodgers, Caroline O'Brien, Patricia Bowman and John Baker. I am very thankful to Professor Steven Kelly for his invaluable contribution to this work. Additionally, I would like to recognise my previous assessors Dr Reka Toth, Professor Francesco Licausi and Professor Gail Preston, who have given me excellent guidance and advice along the way. Many thanks to all of my colleagues who have shared their expertise, ideas, advice and materials over the last few years; Professor Hugh Dickinson, Professor Mark Fricker, Dr Maddy Seale, Dr Charlotte Kirchelle, Dr Liam Elliot, Dr Xinyu Zhang, Dr Jennifer Holter, Dr Nattapong Sanguankiattichai, Dr Henry Hung, Dr Dana Vlad, Dr Pierre Buscaill, Dr Zhe Ji, Dr Hanwei Yan and Professor Sid Shaw. I am also thankful for all the wonderful people who I have shared an office with, who have given me moral support, let me use their equipment and borrow their reagents and made my working environment so positive and collaborative; Morgan Wade,

Floren Scrafton, Ashleigh MacDonald, Brett Wadley, Athira Menon, Allison Ho, William Kay, Rose Bourdon, Dave Aquino, Carrol Fazolato, Lauren Chappell and all the wonderful guests and students we have had. A huge thanks to Professor Seb Shimeld for his guidance during my time as an undergraduate and for encouraging me to apply for a DPhil. Finally, I could never have done this without the love and support of my amazing family and friends. My Mum, Dad, siblings and grandmother have done so much to help me during this time and before. My wonderful friends have cheered me up when things were tough and cheered me on when they were going well. In particular I want to thank Ana, Hannah, Ellen, Bailey and Emily who shared this journey with me and made it really special.

Contents

1	Chapter 1: Literature review	12
1.1	The colonisation of land coincided with the acquisition of 3D growth	12
1.2	3D shoot growth in land plants is organised by meristems	13
1.3	The alternation of generations	16
1.4	<i>Physcomitrium patens</i> as a model plant for studying the transition from 2D to 3D growth	18
1.5	Hormonal Regulation of the 2D-to-3D growth transition	20
1.5.1	Cytokinin	22
1.5.2	Auxin	24
1.6	The role of the cytoskeleton in the 2D to 3D growth transition	26
1.7	Genetic regulation of the 2D-to-3D growth transition	28
1.7.1	Cell signalling	29
1.7.2	Transcriptional regulation	30
1.7.3	Epigenetic regulation	36
1.7.4	Cell trafficking	36
1.7.5	Metabolite biosynthesis and the moss cuticle	37
1.7.6	Protein degradation	40
1.8	Summary	50
1.9	Project aims and objectives	53
2	Chapter 2: Materials and methods	55
2.1	<i>P. patens</i> lines and growth conditions	55
2.2	<i>P. patens</i> transformation	56
2.3	<i>E. coli</i> transformation	58

2.4	Construct assembly	58
2.4.1	Generation of pHSP::NOG1-GFP/GUS constructs	58
2.4.2	Generation of the <i>PpNOG1</i> deletion construct	59
2.4.3	Generation of <i>PpNOG1</i> rescue constructs	60
2.4.4	Generation of the <i>PpDEK1</i> deletion construct	60
2.4.5	Generation of the <i>PpHEC1a/b</i> , <i>PpCCS52A</i> and <i>PpMOC1a/b</i> overexpression constructs	61
2.4.6	Generation of the <i>PpNOG2</i> overexpression construct	61
2.4.7	Generation of Yeast two-hybrid (Y2H) constructs	62
2.4.8	Generation of the <i>Pp3c8-19720</i> constructs	62
2.5	Yeast two-hybrid	63
2.6	Phenotypic characterisation	65
2.6.1	Imaging	65
2.6.2	Gametophore assays	65
2.6.3	Hormone treatments	66
2.7	GUS Staining	66
2.8	UV mutagenesis	66
2.9	Plant crossing and phenotyping of segregating progeny	67
2.10	Genomic DNA extraction	68
2.11	Southern Blotting	69
2.12	Whole Genome Sequencing and bulk segregant analysis	72
2.12.1	Whole Genome Sequencing by Novogene	72
2.12.2	Read processing and variant calling	72
2.12.3	Bulk segregant analysis	73
2.13	Phylogenetics	73
2.14	RNA extraction	75
2.15	cDNA synthesis	75
2.16	Quantitative RT-PCR	76
2.17	Transcriptome analysis	76
2.17.1	RNA-seq by Novogene	76

2.17.2	Read alignment, quantification and DEG analysis	76
2.18	Protein extraction	77
2.19	SDS-PAGE	78
2.20	Western blotting	78
2.21	Heat shock and proteasome inhibitor treatments	79
3	Chapter 3: Genetic manipulation of the <i>Ppnog1R</i> mutant allows further characterisation of the function of <i>PpNOG1</i>.	80
3.1	Introduction	80
3.2	Results	82
3.2.1	<i>NOG1</i> -like sequences are found in the green algae	82
3.2.2	PpNOG1-GFP but not PpNOG1-GUS can rescue the <i>Ppnog1R</i> phenotype	82
3.2.3	PpNOG1 has nuclear and cytosolic localisation	86
3.2.4	PpNOG1 is degraded by the proteasome	86
3.2.5	Complementation of <i>Ppnog1R</i> with NOG1 variants results in some partial rescues	88
3.2.6	PpNOG1 may interact with polyubiquitin.	97
3.2.7	Knockdown of <i>PpDEK1</i> expression in <i>Ppnog1R</i> produces an overbudding phenotype	99
3.3	Discussion	103
4	Chapter 4: A suppressor screen reveals a new regulator of the 2D-to-3D growth transition.	109
4.1	Introduction	109
4.2	Results	110
4.2.1	The <i>nog1dis</i> mutant resembles <i>Ppnog1R</i>	110
4.2.2	UV mutagenesis produced two <i>suppressor of nog1</i> mutants in which 3D growth is restored	112
4.2.3	The mutated <i>snog1a</i> locus was mapped to a region on chromosome 8	119

4.2.4	Point mutations in the coding sequence of <i>Pp3c8_19720</i> resulted in the introduction of two premature STOP codons	122
4.2.5	Overexpression of <i>Pp3c8_19720</i> in <i>snog1a</i> complements the <i>snog1a</i> mutant phenotype	123
4.2.6	Deletion of <i>Pp3c8_19720</i> in <i>nog1dis</i> recapitulates the <i>snog1a</i> phenotype	126
4.2.7	SUPPRESSOR OF NOG1 A is an uncharacterised protein that shares some structural similarities with PpNOG1	128
4.2.8	PpSNOG1A protein homologues are found across land plants and green algae	131
4.3	Discussion	134
5	Chapter 5: RNA-seq experiments yield insights into gene expression in 3D-defective mutants.	141
5.1	Introduction	141
5.2	Results	142
5.2.1	The transcriptional response to cytokinin was severely perturbed in <i>Ppnog1R</i>	142
5.2.2	Many known 3D regulators were differentially expressed in <i>Ppnog1R</i>	150
5.2.3	The overexpression of the transcription factor HECATE1 in <i>Ppnog1R</i> resulted in partial rescue of the <i>Ppnog1R</i> phenotype	158
5.2.4	Fewer genes were differentially expressed in <i>Ppnog2R</i> , than in <i>Ppnog1R</i>	168
5.2.5	<i>PpNOG2</i> controls the expression of genes associated with cellular transport	171
5.2.6	Many known 3D regulators were differentially expressed in <i>Ppnog2R</i> and <i>PpNOG2OE</i>	175
5.2.7	Fewer genes were differentially expressed in <i>nog1dis</i> compared to <i>Ppnog1R</i>	180

5.2.8	A few known 3D regulators exhibit partial restoration of cytokinin responsiveness in <i>snog1a</i>	186
5.3	Discussion	189
6	Chapter 6: General discussion	200
7	Bibliography	209
8	Appendix	249

Chapter 1: Literature review

1.1 The colonisation of land coincided with the acquisition of 3D growth

3-dimensional (3D) growth is a unifying feature of all land plants (embryophytes) that is absent from their aquatic progenitors, the charophyte green algae. The evolution of 3D growth in the early ancestors of land plants likely facilitated colonisation of the land, which occurred approximately 470 million years ago (Delwiche and Cooper, 2015). Along with the acquisition of desiccation tolerance and UV radiation tolerance, 3D growth enabled plants to adapt to the harsher terrestrial environment. As land plants evolved, novel 3D structures emerged, such as vasculature, roots, leaves, seeds and flowers, to name a few (Harrison, 2017; Moody, 2020).

The ability to grow in 3D is conferred by the possession of apical cells (plant stem cells) that can divide in three or more planes. This simple feature is responsible for the diverse 3D morphologies seen in land plants (Moody, 2020). In contrast, multicellular green algae have apical cells with only one or two division planes and so are only capable of 1-dimensional (1D) or 2-dimensional (2D) growth. This constraint limits morphologies to filamentous (1D unbranched or 2D branched) or planar disc structures (Kenrick and Crane, 1997; Wickett et al., 2014; Harrison, 2017). Understanding the mechanisms that control 3D growth is of great importance as it is key to understanding the most fundamental principles underlying the development of all plant tissues and lends insight about the genetic toolkit that

had to evolve to bring about one of the most momentous evolutionary transitions in the history of this planet.

1.2 3D shoot growth in land plants is organised by meristems

Land plant morphology is organised by meristem tissues. The structure of plant meristems varies in size and complexity across clades and between different tissue types. Crucially, meristems all contain undifferentiated stem cells, that can divide in three dimensions to maintain themselves and generate initial cells that will differentiate to form all the tissues and structures required to build a plant (Harrison, 2017).

The shoot apical meristem (SAM) in angiosperms determines the architecture of the aerial structures of the plant. The SAM is a multicellular, dome-shaped tissue at the apex of the plant that consists of several different populations of cells, which are organised into distinct zones and layers (Figure 1.1). The central zone, at the very tip of the meristem, is where a population of totipotent stem cells is maintained. The central zone also generates cells to replenish the surrounding zones, where differentiation occurs. From the central zone, cells are either recruited to the peripheral zone, which encircles the central zone, or they are recruited to the rib zone, below the central zone. Cells in the peripheral zone differentiate to become lateral organs, whereas cells in the rib zone become the stem. The boundaries between the different zones or layers, and the number and identity of cells within those areas, have to be tightly maintained to preserve the integrity of the meristem and allow continual indeterminate growth. Initiating and maintaining this meristem organisation requires a complex interplay between several regulatory genes. For example, a feedback system between the homeobox gene *WUSCHEL* and the *CLAVATA* signalling pathway is required to maintain the identity and size of the central zone (Bowman and Eshed, 2000; Traas and Vernoux, 2002; Somssich et al., 2016).

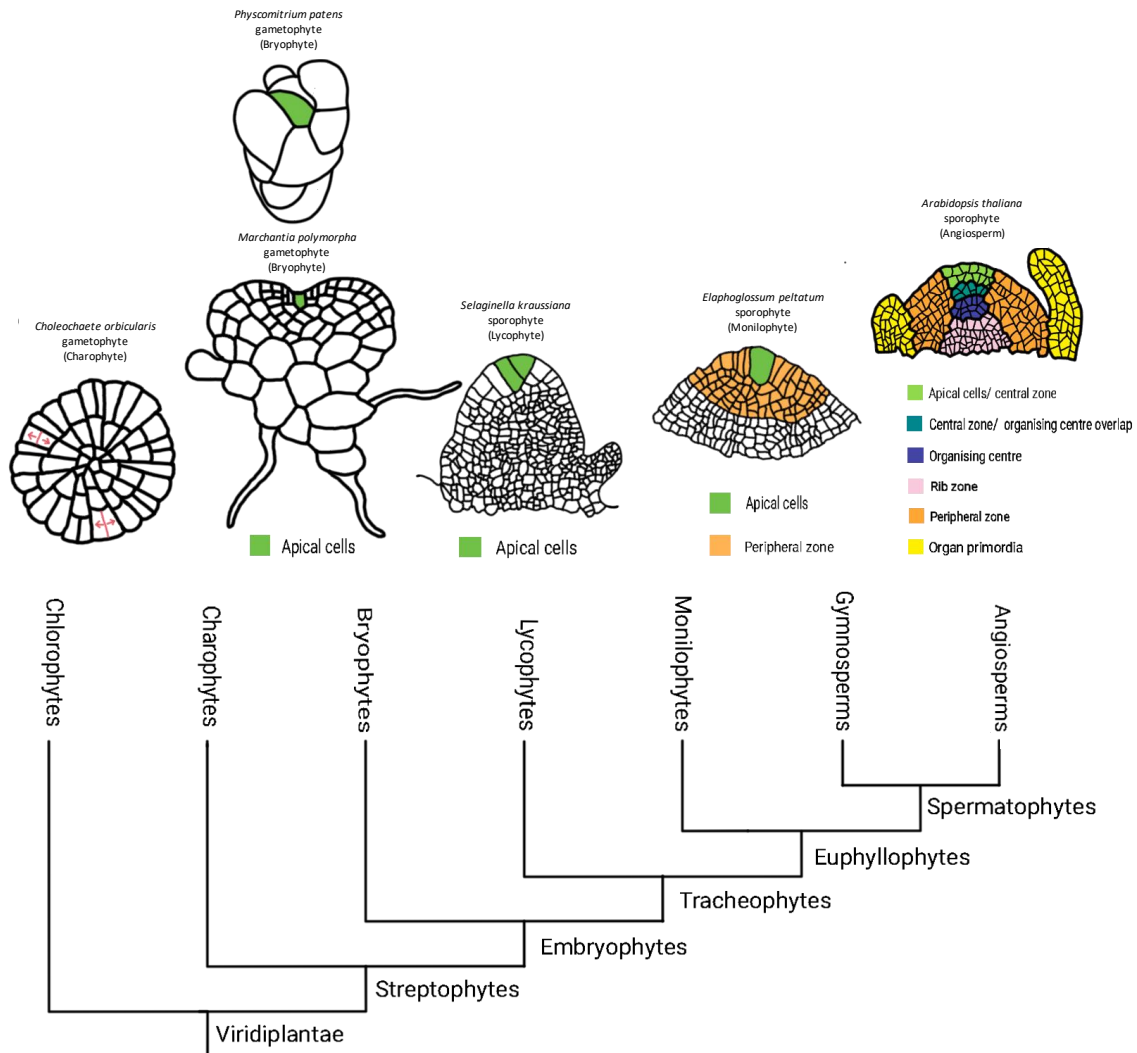


Figure 1.1: Representative diagrams of meristems from various green plant lineages; charophytes, bryophytes, lycophytes, monilophytes and angiosperms. For *Choleochaete orbicularis*, red lines and arrows indicate the two cell division planes in which the peripheral apical cells divide to generate a planar mat. Image adapted from Graham et al. (2015). *Physcomitrium patens* image adapted from Harrison et al. (2009). *Marchantia polymorpha* image adapted from Kny (1890). *Selaginella kraussiana* image adapted from Harrison et al. (2007). *Elaphoglossum peltatum* image adapted from Ambrose and Vasco (2016). *Arabidopsis thaliana* image adapted from Fuchs and Lohmann (2020). Cells highlighted in green are apical cells.

In contrast to the complex, multicellular meristems seen in angiosperms and gymnosperms, non-seed plants often possess simplex meristems consisting of one or two apical cells (Figure 1.1). Bryophytes (mosses, liverworts and hornworts) possess unicellular meristems, which consist of a single apical cell. For example, in *Marchantia polymorpha*, a model liverwort, the gametophyte initially grows as a unistratose thallus, which has a lenticular-shaped apical cell that can divide in two planes to produce lateral expansion of the thallus. The lenticular-shaped apical cell then converts into a wedge-shaped apical cell that can also divide dorsally and ventrally and so has a total of four cutting faces. This wedge-shaped apical cell resides in clefts called apical notches that form around the edge of the thallus. This allows the formation of a dichotomously branching, layered thallus (Shimamura, 2016). Similarly, the gametophyte of the hornwort *Anthoceros agrestis* grows as a thallus and has a wedge-shaped cell with four cutting faces housed in an apical notch (Frangedakis et al., 2021). In the moss *Physcomitrium patens*, the gametophyte produces many small shoots called gametophores and at the apex of each is a single apical cell with a characteristically tetrahedral shape that resembles an inverted pyramid. These unicellular meristems in bryophytes are sufficient to generate the morphologically simple architecture of these plants (Harrison, 2017; Moody, 2020).

In lycophytes, shoot growth is driven by meristems containing one or two apical cells. For example, in *Selaginella kraussiana*, sporophyte shoot growth is organised by two transient apical cells. However, *S. kraussiana* branches by bifurcation of the apical meristem. Before this happens, the number of apical cells increases to around four, so that when the apical meristem splits to form two branches, there are around two apical cells at each of the new apices. Therefore the number of apical cells is more variable in *S. kraussiana* (Harrison et al., 2007).

Monilophyte (fern) sporophytes have a more complex multicellular meristem than lycophytes and bryophytes. The fern meristem has a peripheral zone, but at the centre of the meristem is a single, large apical cell, rather than a population of apical cells, as in seed plants (Figure 1.1; Ambrose and Vasco, 2016).

1.3 The alternation of generations

Land plants evolved from the charophycean green algae; a group of freshwater aquatic algae, which include the extant representatives *Chara braunii*, *Choleochaete orbicularis* and *Klebsormidium nitens* (Qiu et al., 2006). Charophytes display a haplobiontic life cycle whereby the only multicellular stage is the haploid gametophyte, which produces gametes by mitosis. After fertilisation and the formation of unicellular diploid zygote, meiosis occurs immediately to generate haploid spores that become new gametophytes. In contrast, the diplobiontic land plant life cycle consists of two alternating multicellular generations with different ploidy. As in algae, the haploid gametophytes of land plants produces gametes by mitosis. However, after gamete fusion takes place, the diploid zygote divides by mitosis and produces another multicellular stage called the sporophyte, which eventually generates haploid spores via meiosis that restart the cycle. The emergence of the multicellular sporophyte was one of the key evolutionary transitions that occurred during plant terrestrialisation (Kenrick, 1994).

In each land plant lineage, the gametophyte and sporophyte have distinct body plans, with one generation being the dominant stage and the other being smaller and more transient. In some cases, both stages have meristems capable of 3D growth, but in others 3D growth only occurs in one of the generations (Harrison, 2017). In the bryophytes, which are the sister lineage to all other land plants, the gametophyte is the dominant, indeterminate, free-living stage with 3D-capable apical cells. The sporophyte in bryophytes is a small, transient and determinate structure that emerges from, and is nutritionally dependent on, the gametophyte. The sporophytes of mosses do possess an apical cell, but the sporophytes of hornworts and liverworts seem to develop and form sporangia without generating an apical cell (Prigge and Bezanilla, 2010; Kofuji and Hasebe, 2014; Harrison, 2017). Hornwort sporophytes instead form a basal meristem that divides to produce cells for upward growth (Frangedakis et al., 2021). Liverwort sporophyte growth seems to be derived from two separate cells. The hypobasal cell forms the

lower tissues of the sporophyte and the epibasal cell produces the spore capsule (Shimamura, 2016). Over the course of land plant evolution, there has been an evolutionary shift towards a more dominant sporophyte. All tracheophytes (vascular plants; lycophytes, monilophytes, gymnosperms and angiosperms) primarily exist as a sporophyte, with the gametophyte becoming increasingly more reduced in more derived lineages (Plackett et al., 2015). In lycophytes, the sporophyte is the dominant stage and exhibits 3D growth. The gametophyte grows independently of the sporophyte but remains encapsulated within the spore coat and produces archegonia and antheridia, the gamete producing organs, without generating an apical cell capable of 3D growth (Spencer et al., 2021). In ferns, an apical cell is present in both the large sporophyte, which bears the fronds, and the gametophyte, which is much smaller than the sporophyte, but grows independently (Harrison, 2017). Gymnosperms and angiosperms, the seed plants or spermatophytes, are the most derived of the plant lineages and represent the majority of extant plants. As with other vascular plants, the sporophyte dominates the life cycle, but the gametophyte is extremely reduced and consists of only a few cells in the pollen grain or the embryo sac, which develop within the parent plant, without generating any apical cell. Therefore, 3D growth only occurs during the development of the sporophyte in seed plants (Harrison, 2017).

It is important to be cognizant of these differences in generation dominance and morphology when studying 3D development in land plants. The developmental programs that control the development of the sporophyte versus the gametophyte, although encoded by the same genome, are known to be distinct, sharing some regulatory components but also employing some that are generation specific (Friedman, 2013; Pandey et al., 2022).

1.4 *Physcomitrium patens* as a model plant for studying the transition from 2D to 3D growth

In the angiosperm sporophyte, the transition to 3D growth occurs early during embryogenesis. By the ‘heart stage’ in the embryo of *Arabidopsis thaliana* (Arabidopsis), the root and shoot apical meristems (and therefore 3D growth) have been fully established (Jürgens, 2001). Consequently, it is difficult to investigate the regulation of 3D growth in flowering plant models like Arabidopsis, as disrupting 3D growth is likely to cause early embryo lethality. Therefore, an alternative model is required to effectively examine these fundamental genetic pathways.

In the gametophyte of the moss *Physcomitrium patens* (previously *Physcomitrella patens*), a prolonged 2D filamentous growth phase precedes the transition to 3D growth. This 2D phase can be maintained indefinitely by vegetative propagation, allowing for the generation of 3D-defective mutants that are still viable. Furthermore, genetic studies are facilitated by the fact that the gametophyte is haploid, so mutagenesis approaches can be simplified, circumventing the need for backcrossings (Moody et al., 2018; Rensing et al., 2020). *P. patens* is also unusual among plants in that it has a very high rate of homologous recombination, similar to the rate seen in yeast, and thus targeted genome editing is simple and effective (Kamisugi et al., 2006). The *P. patens* genome was one of the first plant genomes to be published and was the first available bryophyte genome. *P. patens* is phylogenetically significant, as the bryophytes are the sister lineage to all other land plants, with a divergence date more than 400 million years ago, and therefore it is useful for comparative studies to investigate plant evolution (Rensing et al., 2008; Kenrick and Crane, 1997; Prigge et al., 2010). Thus, *P. patens* is a useful model system for studying the regulatory mechanisms that govern the transition from 2D to 3D growth.

During the life cycle of *P. patens*, the gametophyte phase begins with the germination of a haploid spore. From the spore emerges a chloronemal initial

cell which divides to produce a chloronemal filament that extends by tip growth. Chloronemal cells are very chloroplast dense and their cell walls are perpendicular to the axis of the filament. Chloronemal cells can differentiate to produce a second filamentous cell type called the caulonema. Caulonemal cells contain fewer chloroplasts and their cell walls form at an oblique angle (Figure 1.2). Chloronemal and caulonemal filaments can grow in one dimension, or form branches to grow in two dimensions (2D growth). This branching filamentous network consisting of chloronema and caulonema is collectively known as the protonema (Cove and Knight, 1993; Cove, 2005). 3D ‘leafy shoots’ called gametophores arise from caulonemal cells but not chloronemal cells. When a caulonemal cell divides laterally to produce a side branch initial cell, that cell will either develop into a secondary filament (95%), or a gametophore (5%) (Harrison et al., 2009; Aoyama et al., 2012). The decision for a side branch initial to acquire either 2D or 3D fate results from signalling within the parental caulonemal cell, from which it is derived (Tang et al., 2019). This signalling is very localised as secondary protonema and gametophore buds can occur in very close proximity, even emerging from the same caulonemal cell (Harrison et al., 2009). It is possible to distinguish between a protonemal initial and a gametophore initial before they divide, due to subtle morphological differences. The angle of the division at the base of the initial, where it meets the parental caulonemal cell, is greater in gametophore initials than in protonemal initials. Additionally, gametophore initials are wider than protonemal initials as they begin to swell and expand diffusely instead of extending by tip growth like a filament (Harrison et al., 2009; Tang et al., 2019).

The first division of a protonemal initial will be parallel to the axis of the parental caulonemal filament, from which it is derived, whereas the gametophore initial will divide obliquely (Figure 1.2). The two subsequent divisions in the gametophore bud are perpendicular to the first division. This is followed by two more rotating cell divisions, which establishes a tetrahedral apical cell, with three cutting faces (Figure 1.2). After formation of the gametophore initial, but before the establishment of the tetrahedral apical cell, this structure will be referred to as a gametophore

bud. It is the establishment and maintenance of the tetrahedral apical cell that represents the transition to 3D growth in the *P. patens* gametophyte. The apical cell divides asymmetrically, in a rotating manner, to maintain itself and generate organ primordia (phyllid initials) that become the leaf-like phyllids, arranged in spiral phyllotaxy around the central stem (Harrison et al., 2009). At the apex of the gametophores, gametangia (antheridia and archegonia) form and gametes are produced by mitosis. Fertilisation requires the presence of water, which allows the flagellated sperm cells, which are released from the antheridia, to swim to the archegonia and fertilise the eggs contained within. The resulting zygote develops into a sporophyte, which grows from the gametophore apex and produces spores via meiosis, which are then released from the spore capsule to restart the life cycle (Figure 1.2; Cove, 2005).

1.5 Hormonal Regulation of the 2D-to-3D growth transition

Genetic regulation of gametophore induction is integrated with phytohormone signalling. The hormones that are usually regulators of plant development are auxin, cytokinin, brassinosteroids and gibberellin (Wolters and Jürgens, 2009). Gibberellin has not been detected in *P. patens* and the genes for gibberellin biosynthesis and signalling seem not to have been functionally conserved, although some precursors of gibberellin have been found (Hirano et al., 2007; Guillory and Bonhomme, 2021). There is still some debate over whether brassinosteroids are present in *P. patens* (Guillory and Bonhomme, 2021). Another phytohormone, abscisic acid (ABA), often regulates biotic and abiotic stress responses, but is also known to have some developmental roles (Chen et al., 2020). In *P. patens* it has been shown that ABA does regulate stress responses as it does in other plants (Sakata et al., 2009; Stevenson et al., 2016; Xiao et al., 2018). There is some evidence that ABA also has some role in regulating the 2D-to-3D transition. For example, in another moss, *Funaria hygrometrica*, ABA treatment resulted in inhibition of bud

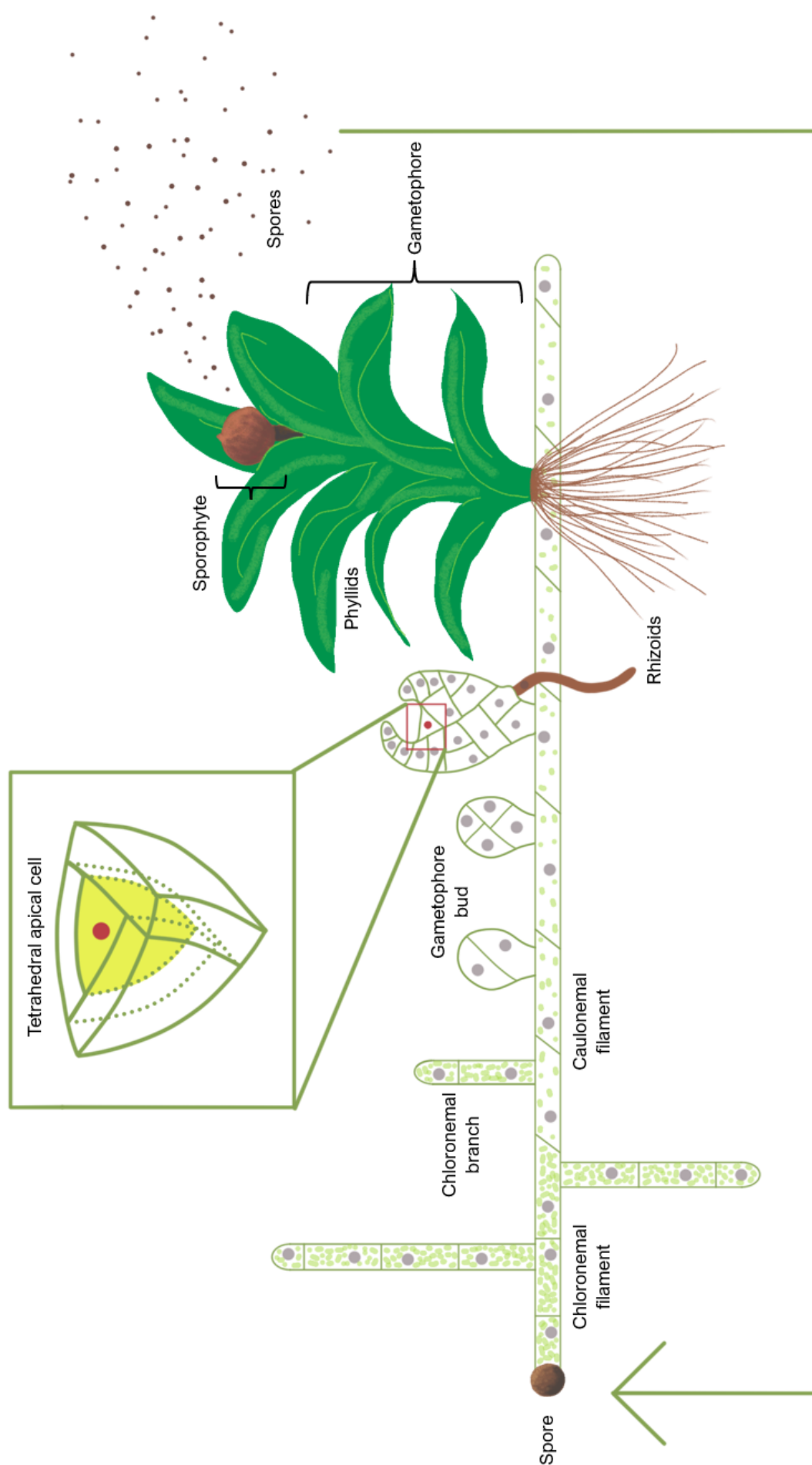


Figure 1.2: Lifecycle of *Physcomitrium patens*, showing the establishment of the tetrahedral apical cell with three cutting faces (red circle). Apical cell diagram adapted from Hata and Kyoizuka (2021).

initiation through antagonism of cytokinin-mediated bud initiation (Christianson, 2000). However not much else is known about the significance of ABA in this context. The importance of cytokinin and auxin and the crosstalk between them during the transition to 3D growth is well-established.

1.5.1 Cytokinin

It has long been known that cytokinin induces the formation of gametophore initial cells. Treatment with cytokinin results in a higher frequency of buds, in a dose-dependent manner, and sustained exposure to cytokinin is required for buds to commit to gametophore fate, as its removal at the early stages of bud development results in reversion to protonemal identity (Brandes and Kende, 1968). However, persistent treatment with high levels of cytokinin after the establishment of the gametophore will not maintain 3D growth as it leads to the formation of callus-like buds due to cell proliferation without differentiation (Ashton et al., 1979). Treatment with lower levels of cytokinin results in small gametophores with broad, short phyllids and the formation of ectopic apical cells on the gametophores (Cammarata et al., 2023). Mutants thought to be impaired in their ability to synthesise cytokinin are unable to produce gametophores, indicating its necessity for the 2D-to-3D transition (Ashton et al., 1979).

The importance of cytokinin perception for bud initiation was demonstrated by knocking out the classical cytokinin receptors, the CHASE domain-containing histidine kinases *PpCHK1,2* and *3*. Triple knockout mutants ($\Delta chk1,2,3$) had reduced bud frequency. Additionally, gametophores were much smaller than in wild type and appeared later (von Schwartzenberg et al., 2016). *CYTOKININ OXIDASE/DEHYDROGENASE* (*CKX*) genes regulate cytokinin concentrations by degrading cytokinin in a cell-specific manner. In *P. patens*, *CKX1* was found to be specifically expressed in stem cells, where active division is occurring, including the tetrahedral apical cell of the gametophore. Constitutive overexpression of *PpCKX1* resulted in reduced levels of cytokinin that led to delayed bud initiation (Hyoung et al., 2020).

The essential role of cytokinin in maintaining apical cell identity by promoting cell division and inhibiting differentiation is well documented in studies of the *Arabidopsis* shoot apical meristem (SAM; Wybouw and De Rybel, 2019). When cytokinin levels are decreased or cytokinin signalling suppressed, the size of the SAM is reduced (Werner et al., 2003; Ishida et al., 2008). It has been shown that cytokinin directly induces the expression of *WUSCHEL* (*WUS*), which is necessary for maintaining apical cell identity within the SAM. Cytokinin accumulates at the SAM where it is perceived by ARABIDOPSIS HISTIDINE KINASE 2-4, triggering a phosphorelay that results in the phosphorylation and therefore activation of type-B ARABIDOPSIS RESPONSE REGULATORS (ARRs; Figure 1.3). The activated type-B ARRs bind directly to the *WUS* promoter, inducing *WUS* expression (Meng et al., 2017; Zürcher et al., 2013; Wybouw and De Rybel, 2019).

The evidence from *P. patens* indicates that cytokinin performs the same role in the establishment of the simplex meristem in the gametophore; the tetrahedral apical cell. However it does so through a different mechanism than in *Arabidopsis* (Cammarata et al., 2019). *WUS* is a member of the *WUSCHEL*-related *HOMEODOMAIN* (*WOX*) family, representatives of which are found in all land plants. However, the clade of *WOX* genes that contains *WUS* arose after the divergence of the bryophytes and the vascular plants (Deveaux et al., 2008). Three *WOX* family genes were identified in the *P. patens* genome; *WUSCHEL*-related *HOMEODOMAIN 13-like A* (*WOX13LA*), *B* (*WOX13LB*) and *C* (*WOX13LC*). Expression of *WOX13LC* has not been detected in *P. patens* tissues, indicating that it may be a pseudogene (Deveaux et al., 2008; Sakakibara et al., 2014). It was found that *WOX13LA* and *WOX13LB* do not regulate the apical cell of *P. patens* gametophores as *wox13lab* double disruption mutants were not defective in gametophore development. However, *WOX13LA* and *WOX13LB* have an important role in the sporophyte as *wox13lab* mutants had severely impaired zygote development, resulting in the formation of much fewer sporangia, all of which were malformed (Sakakibara et al., 2014). As the *P. patens* *WOX* genes do not function to maintain the apical cell in the gametophore, cytokinin must be

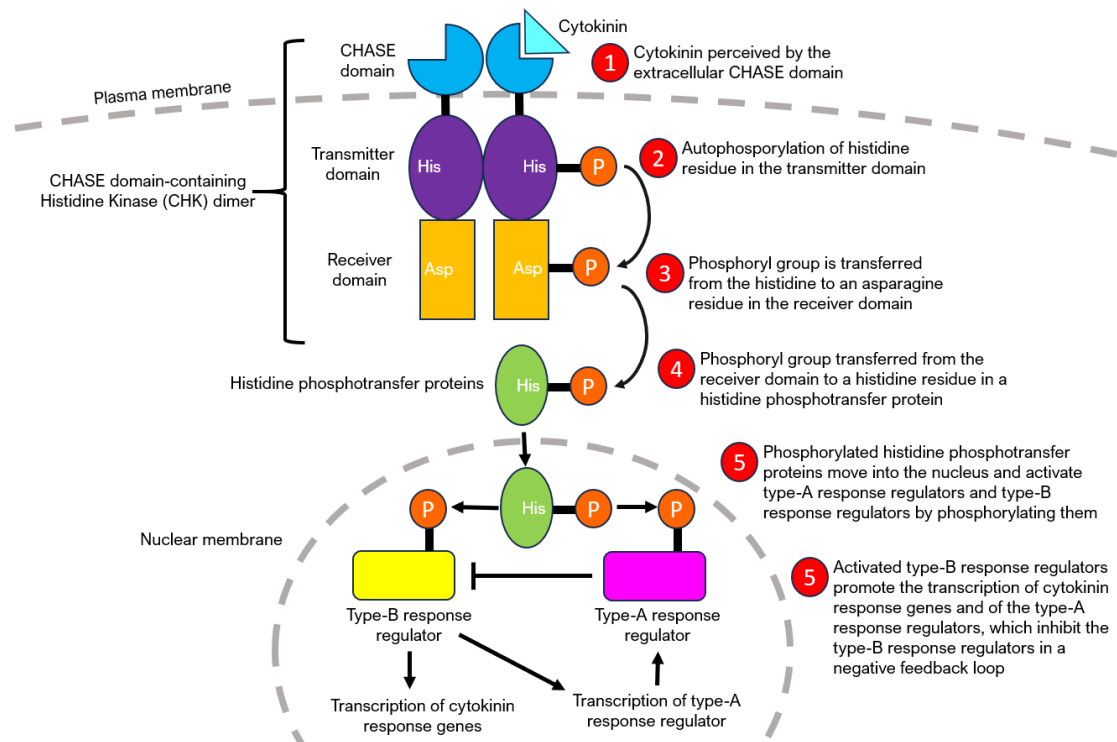


Figure 1.3: The cytokinin signalling cascade in plants, a two-component phosphorelay system. Based on a review by Keshishian and Rashotte (2015).

interacting with alternative, but currently unknown factors to promote gametophore development in *P. patens*. As is the case in the SAM of Arabidopsis, cytokinin sensing is high in the gametophore apical cell and low in the surrounding tissue (Hata et al., 2023). The expression of *PpCKX1* in the apical cell, and close neighboring cells, indicates that cytokinin levels are being precisely regulated in the apex of developing gametophores (Hyoung et al., 2020).

1.5.2 Auxin

Auxin is required for initiation and maintenance of 3D growth as cytokinin-induced bud formation is dependent on the presence of auxin. It is thought that auxin acts first, by priming the target caulonemal cells to respond to cytokinin and produce a bud initial (Lehnert and Bopp, 1983). However, auxin also antagonises cytokinin, as application of exogenous auxin can counteract the effect of cytokinin treatment and reverse bud initiation (Brandes and Kende, 1968).

Auxin also affects gametophore development in a dose-dependent manner. Exogenous application of the auxin analogue α -naphthaleneacetic acid (NAA) at

low levels (<500 nM) has very little effect on 3D growth. However, at moderate to high levels (>2.5 μ M), NAA leads to gametophores with highly stunted growth, large buds that cannot produce phyllids or a complete inhibition of bud initiation (Ashton et al., 1979). In *P. patens*, as in angiosperms, it is believed that auxin promotes cell differentiation and antagonises cytokinin, which maintains the apical cell niche (Hata and Kyoizuka, 2021). This suggestion is supported by the fact that auxin levels in the developing gametophore decreases upon establishment of the tetrahedral apical cell. A ratiometric auxin sensing reporter shows that during the first few divisions of the gametophore bud, auxin levels are high in all cells. Upon the formation of the apical cell, auxin levels become cell-type specific, with a minima in the apical cell. This is consistent with exclusion of auxin from the apical cell in order to maintain stem cell identity and prevent differentiation. On the other hand, auxin levels are high in fully differentiated tissues. In small, newly-formed phyllids, auxin levels are generally low. As phyllids mature, an area of high auxin appears at the tip and spreads proximally. At the base of the phyllid, where active cell division is occurring, a zone of low auxin remains. The appearance of higher auxin levels in more distal cells precedes a period of cell expansion before growth arrest (Thelander et al., 2019). Thus, precise auxin distribution is critical for apical cell maintenance and to control organ initiation and differentiation.

The components of auxin signalling are largely conserved between angiosperms and bryophytes. PIN-FORMED (PIN) proteins responsible for polar auxin transport have been identified and characterised in *P. patens* and disrupting *PIN* genes perturbs phyllid development and the chloronemal to caulonemal transition (Bennett et al., 2014; Viaene et al., 2014). Thus, maintaining the proper distribution of auxin and cytokinin is vital to 3D development. Importantly, it seems to be the ratio of auxin to cytokinin in different tissues and in individual cells that is the most important for organising development. For example, if auxin synthesis is inhibited with the application of L-kynurenine, the phenotype resembles plants treated with low concentrations of cytokinin; the formation of small gametophores with ectopic apical cells (Cammarata et al., 2023). The crosstalk between these two hormones

and the genes under their regulation is complex and still being uncovered.

1.6 The role of the cytoskeleton in the 2D to 3D growth transition

The correct orientation of cell divisions during the transition to 3D growth is essential for the successful establishment of the apical cell and subsequent development. Misorientation of the early cell divisions within the gametophore bud often results in a failure to generate a tetrahedral apical cell, which normally leads to developmental arrest (Causier et al., 2012; Perroud et al., 2014; Moody et al., 2018, 2021).

At the mechanistic level, the cell division plane is determined by the cytoskeleton. The sister chromatids separate and migrate to form new nuclei in the right regions of the cell via the mitotic spindle. The phragmoplast is a cytoskeletal scaffolding that forms during cytokinesis to guide the formation of the cell plate and the deposition of materials to form the new primary cell wall (Li et al., 2015). Another cytoskeletal structure that usually guides division orientation in plants is the preprophase band (PPB), a microtubule array that encircles the cell before the onset of mitosis, marking exactly where the division plane will be positioned (Mineyuki, 1999). However, evidence suggests that the PPB is not the absolute determinant of division plane orientation, but rather is responsible for fine-tuning plane positioning, as mutants lacking the PPB often have fairly normal overall morphology (Schaefer et al., 2017). Interestingly, in the gametophore bud, the first three asymmetric cell divisions are thought to occur in a PPB-independent manner as the PPB has not been observed during these divisions, but it does appear during the later development of the gametophore. Recently, a novel microtubule organising centre, known as the ‘gametosome’, was identified in *P. patens* and shown to guide the orientation of the spindle during these initial asymmetric divisions, therefore influencing the orientation of the division plane. During the first three divisions of the bud, the gametosome emerges at prophase

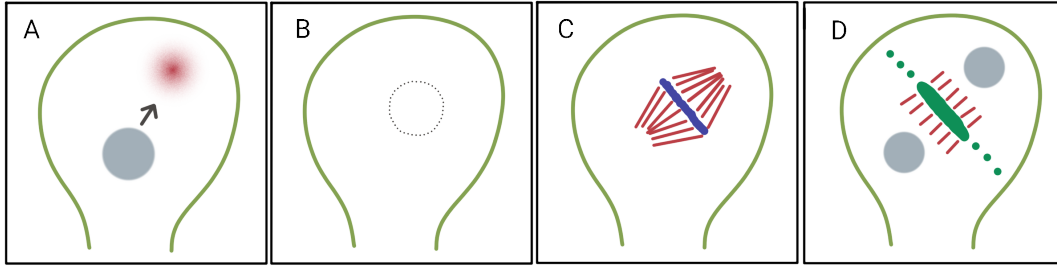


Figure 1.4: The gametosome plays a role in orientating the asymmetric divisions of the gametophore bud in a PPB-independent manner. A) During prophase the gametosome (red) appears and the nucleus (grey) migrates towards it. B) At nuclear envelope breakdown, the gametosome dissipates to merge with the forming spindle. C) The spindle (red) forms at an angle that is influenced by the gametosome and the sister chromatids are separated. D) Daughter nuclei (grey) reform and the phragmoplast (red) forms to guide the expansion of the cell plate (dark green). The new cell wall is deposited at the cell plate.

as an amorphous cloud of microtubules. The nucleus migrates towards the gametosome, and at nuclear envelope breakdown, the gametosome merges with the spindle. Transient microtubule destabilisation by treatment with oryzalin during gametosome formation, followed by removal of the oryzalin before the spindle forms, results in defective spindle orientation (Kosetsu et al., 2017). A gene called *Sad1/UNC-84 DOMAIN PROTEIN 2* (*SUN2*) is important for the association of the gametosome with the nuclear envelope. In *sun2* knockout mutants the gametosome is mislocalised and the spindle does not align correctly, resulting in misoriented divisions in the bud (Yoshida et al., 2023). During cytokinesis, the initial angle of the phragmoplast is determined by the spindle orientation. However, the developing cell plate undergoes a further rotation which determines the final plane of division. This final rotation of the phragmoplast and cell plate is not dependent on the presence of the gametosome, which indicates that additional mechanisms are operating in concert with the gametosome to control the orientation of these critical divisions (Kosetsu et al., 2017).

The gametosome plays a role in determining the position of the mitotic spindle during the critical asymmetric cell divisions of the bud, but other factors are also involved in this process. A study in *P. patens* of the gene *targeting protein for Xklp2*

(*TPX2*), which is known to be important for microtubule nucleation, revealed that it is required to maintain the position of the spindle during these asymmetric cell divisions. There are five *TPX2* genes in *P. patens* that seem to act redundantly. A quadruple knockout mutant, lacking four of these genes (*TPX2 1-4*) appeared similar to wild type. It was not possible to generate a quintuple mutant, but a hypomorphic mutant was created in $\Delta tpx2\ 1-4$, whereby the expression of *TPX2-5* was greatly reduced. This hypomorphic mutant displayed delayed mitosis and produced dwarfed gametophores. Also, during the first oblique division of the gametophore initial, it was observed that the cell plate was shifted towards the base of the gametophore initial, which altered the apical/basal cell size ratio from what is seen during normal wild type development. Live cell imaging showed that the misplaced cell plate position was caused by spindle motility after nuclear envelope breakdown, which does not occur in normal circumstances. Interestingly, the basal movement of the spindle seen in the hypomorphic mutant can be completely suppressed by treatment with latrunculin A, which depolymerises actin filaments, indicating that the actin cytoskeleton is driving this basal spindle migration (Kozgunova et al., 2022). Thus, both actin filaments and microtubules play important roles in determining cell division plane during the establishment of 3D growth in gametophores.

1.7 Genetic regulation of the 2D-to-3D growth transition

Some of the genetic regulators of the transition to 3D growth have been discerned. Several genes that are known to be involved in meristem regulation in angiosperm plant models have been investigated in *P. patens* and some have then been shown to have an analagous role in gametophore initiation or development. Other important regulators have been found using a forward genetics approach which has revealed previously undescribed developmental functions in several genes.

1.7.1 Cell signalling

The *CLAVATA* signalling pathway

The *CLAVATA* signalling (*CLV*) pathway arose in the land plants, and is completely absent in charophyte green algae (Whitewoods et al., 2018). In *Arabidopsis*, the *CLV* pathway is required for apical meristem maintenance. *CLV* genes are part of a feedback loop with the transcription factor *WUSCHEL* (*WUS*). *WUS* expression is induced in the SAM by cytokinin. *WUS* promotes stem cell identity, but also induces the expression of *CLV3*, a small mobile peptide that negatively regulates the expression of *WUS* via leucine-rich repeat receptor-like protein kinases (*CLAVATA1*, *CLAVATA2*, *BARELY ANY MERISTEM* and *CORYNE*). This provides a negative feedback mechanism that regulates the size of the meristem by limiting the region in which *WUS* is expressed (Clark et al., 1996; Fletcher et al., 1999; Brand et al., 2000; Schoof et al., 2000; Rojo et al., 2002; Leibfried et al., 2005; Muller et al., 2008; Somssich et al., 2016). Interestingly, even though *WUSCHEL*-related genes are not involved in apical cell maintenance in *P. patens* gametophores, the function of the *CLAVATA* signalling pathway in regulating apical cells is conserved between vascular plants and bryophytes (Whitewoods et al., 2018; Cammarata et al., 2019).

P. patens has twelve *CLV* genes: nine *CLAVATA3/EMBRYO SURROUNDING REGION* (*CLE*) ligands (*PpCLE1-9*), two homologues of the *CLV1/BAM* receptors (*PpCLV1a* and *PpCLV1b*) and one homologue of the *RPK2* receptor (*PpRPK2*) (Nemec-Venza et al., 2022). Several *clv* mutants have been created in which various *CLV* genes have been disrupted. In all cases, gametophore development was defective, with incorrectly oriented divisions in the bud, but despite these early defects, these mutants were able to produce stunted gametophores (Whitewoods et al., 2018).

There is evidence that auxin and cytokinin crosstalk is perturbed in the *Ppclv* mutants. For example, aspects of development that are controlled by auxin and cytokinin, such as cell proliferation and cell fate, are disrupted in the *Ppclv* mutants (Whitewoods et al., 2018). A *Ppclv1a Ppclv1b* double mutant and a

Pprpk2 mutant both have gametophores that produce ectopic apical cells on their stems. This phenotype resembles plants that have been treated with low concentrations of exogenous cytokinin (10 nM - 100 nM of the cytokinin analogue 6-benzylaminopurine), indicating an antagonism between *CLV* and cytokinin signalling (Cammarata et al., 2019, 2022). *CLV* genes have also been shown to promote auxin synthesis and transport (Nemec-Venza et al., 2022). These data signify that *CLV* signalling is an important mediator of auxin-cytokinin crosstalk during the transition from 2D to 3D growth.

1.7.2 Transcriptional regulation

AP2-type transcription factors

A family of AP2-type transcription factors are necessary and sufficient for the induction of gametophore initial cells. The four *P. patens* genes in this family are abbreviated as *PpAPB1-4* as they are orthologues of the genes *AINTEGUMENTA*, *PLETHORA* and *BABYBOOM* from *Arabidopsis* (Aoyama et al., 2012). These genes contribute to aspects of angiosperm sporophyte development including embryogenesis, stem cell niche formation, cell proliferation and organ growth (Boutilier et al., 2002; Hu et al., 2003; Aida et al., 2004).

The *PpAPBs* have a striking expression pattern during the transition from 2D to 3D growth. *PpAPB1*, *3* and *4* are all expressed in the parental caulonemal cell. When a side branch initial is formed, those that become gametophore initials retain *PpAPB* expression, but expression is suppressed in protonemal initials. *PpAPB2* is only expressed in side branch initials and similarly persists only in those that are fated for 3D growth. A quadruple knockout of all four genes completely abolishes gametophore induction, even in the presence of cytokinin. Interestingly, single, double or triple knockouts of *PpAPB1,2,3* and *4* either have no effect or only cause partial suppression of gametophore formation, indicating that the *PpAPB* genes likely act redundantly. A single mutant lacking *PpAPB4* has the most striking bud suppression phenotype of the single knockout lines and induction of

PpAPB4 overexpression is sufficient to drive an increase in gametophore induction. Auxin positively regulates the expression of the *PpAPBs*. It is hypothesised that the *PpAPBs* could be triggering the expression of cytokinin signalling genes, or those involved in cytokinin perception, to provide a mechanism for regulating the frequency of bud induction (Aoyama et al., 2012).

TOPLESS

The *TOPLESS* (*TPL*) transcriptional corepressors are a plant-specific gene family that are known to be involved in many processes including, but not limited to, embryo patterning, meristem maintenance, auxin signalling, organ growth, circadian clock and biotic/abiotic stress response (Plant et al., 2021). *TOPLESS* is so named because the most severe phenotype seen in a dominant negative Arabidopsis *tpl* mutant is a conversion of the embryonic shoot into another root. This mutant produces a range of phenotypes that are associated with disruption of apical embryonic patterning (Szemenyei et al., 2008; Smith and Long, 2010). TPL proteins act as corepressors and so function by binding to transcription factors via a motif called the repression domain, rather than binding directly to DNA (Causier et al., 2012).

In *P. patens* there are 2 *TPL* genes. Single knockout lines of each have no discernible phenotype indicating functional redundancy. It has not been possible to recover double mutants and thus it has been assumed that *TPL* is essential for *P. patens* viability. A dominant negative *Pptpl2* mutant was generated, which initiates much fewer gametophores than wild type. Furthermore, developing gametophores in this mutant undergo premature arrest because early division planes are misoriented. Transcriptomic analyses revealed that several genes known to regulate the transition to 3D growth are downregulated in the *tpl2* dominant negative mutant. These include *PpAPB1-4*, *NO GAMETOPHORES 2*, *CLE3*, *4* and *6* and *PpCLV1a*. *CLE8* was upregulated in *tpl2*. Other important genes had unaffected transcription; *DEFECTIVE KERNEL 1*, *NO GAMETOPHORES 1*, *PpCLV1b*, *PpRPK2*, *PpCLE 1,2,5* and *7* (Causier et al., 2023). Therefore, *PpTPL* genes seem to be early

regulators of the transition to 3D growth that operate upstream of the *PpAPBs*.

ANGUSTIFOLIA 3

ANGUSTIFOLIA 3 (*AN3*), also known as *GRF-INTERACTING FACTOR 1* (*GIF1*) is a transcriptional coactivator. *AN3* genes possess a highly conserved synovial sarcoma translocation protein (SYT) N-terminal homology (SNH) domain, which is required for direct interactions between AN3 proteins and the closely related GROWTH REGULATING FACTOR (GRF) transcription factors as well as the SWI/SNF chromatin remodelling complex. By binding to GRF transcription factors or the SWI/SNF chromatin remodelling complex, AN3 controls the transcription of target genes (Vercruyssen et al., 2014; Kim and Tsukaya, 2015). In Arabidopsis *Atan3* mutants, leaf size is decreased and the SAM is reduced. This was shown to be due to lower cell proliferation in the meristematic tissues and leaf primordia. *AtAN3/AtGIF1* is closely related to *AtGIF2* and *AtGIF3*. Double and triple mutants of these genes have increasingly severe phenotypes where cell proliferation is perturbed. Thus, together *AtAN3/AtGIF1*, *AtGIF2* and *AtGIF3* act semi-redundantly to promote cell proliferation in the SAM and lateral organs. Expression of cell cycle regulators is suppressed in triple *gif* mutants, indicating that the *GIF* genes promote cell proliferation by inducing the expression of cell cycle regulators (Lee et al., 2009).

There are four *PpAN3* genes in *P. patens* known as *PpAN3-1,2,3* and *4*. The generation of single, double, triple and quadruple mutants results in the formation of stunted gametophores. The severity of the phenotype is proportionate to the number of genes disrupted, indicating a gene dosage effect. Fewer phyllids are formed in higher order mutants indicating an inability of the apical cell to maintain adequate proliferation for the formation of phyllid initials. Application of the cytokinin analogue 6-benzylaminopurine (BAP) to *Ppan3* mutants and wild type resulted in similar levels of bud initiation. Therefore, *PpAN3* is not required for initiation, but is needed to maintain apical cell proliferation (Kawade et al., 2020).

Transcriptomic analyses of *Ppan3* triple and quadruple knockout mutants

revealed that the set of downregulated genes was enriched in genes with metabolic functions. When this was probed further, it was determined that disruption of *PpAN3* lead to accumulation of arginine in the gametophores. Furthermore, the phenotype of *Ppan3* mutants can be recreated by growing wild type plants on medium containing high concentrations of arginine. Therefore *PpAN3* regulates arginine metabolism, which in turn regulates gametophore development, although the underlying mechanism for this remains unclear (Kawade et al., 2020).

PpGRAS12

PpGRAS12 is a member of the plant-specific GRAS family of transcription factors (Beheshti et al., 2021). *PpGRAS12* and its homologues in other plants are known to form a regulatory module with the microRNA miR171. In Arabidopsis, there are three homologues, known either as *HAIRY MERISTEM1,2* and *3*, *LOST MERISTEM1,2* and *3* or *SCL6-II*, *III* and *IV* and the spatiotemporal distribution of their transcripts is negatively regulated by miR171. In Arabidopsis, this GRAS/miR171 module has been found to be important for maintaining organisation in shoot apical and axillary meristems by promoting cell differentiation in the peripheral zone (Schulze et al., 2010; Engstrom et al., 2011; Ma et al., 2014; Beheshti et al., 2021). In *P. patens*, a *Ppgras12* knockout mutant showed no phenotypic differences at the protonemal or early gametophore stage. An inducible *PpGRAS12* overexpression line was generated, in which the miR171 binding site had been mutated to prevent miR171-mediated degradation of the *PpGRAS12* transcript. Upon induction of *PpGRAS12* overexpression, many new ectopic apical cells formed at the top of gametophores. These ectopic apical cells developed normally to form new gametophores, growing from the top of pre-existing gametophores, even once induction of *PpGRAS12* expression had ceased (Beheshti et al., 2021).

As this phenotype resembles that of *clavata* mutants, expression of the CLAVATA genes was investigated after induced *PpGRAS12* overexpression. It was found that all the *PpCLE* genes were upregulated, whereas *PpCLV1a*, *PpCLV1b* and *PpRPK2* were downregulated. Other known 3D regulators, including *PpAPB1*

and *PpAPB4*, as well as *NO GAMETOPHORES 1* and *DEFECTIVE KERNEL 1* were also shown to be downregulated (Beheshti et al., 2021). The downregulation of *PpCLV1a/b* and *PpRPK2* may explain the multiple ectopic apical cells as knockout mutants in those genes present a similar phenotype (Beheshti et al., 2021; Cammarata et al., 2022).

***SQUAMOSA PROMOTER BINDING PROTEIN* and miR156**

SQUAMOSA PROMOTER BINDING PROTEIN-LIKE (*SPL*) genes are a plant-specific family of transcription factors and are so named as SPL proteins were discovered to bind to the promoter of the gene *SQUAMOSA*, a regulator of floral meristem identity in *Antirrhinum majus*. SPL proteins contain a conserved DNA-binding domain called the SBP-box, which allows them to bind to the *SQUAMOSA* promoter (Cardon et al., 1999; Poethig, 2009). *SPLs* are regulators of plant maturation as they control the switch from juvenile to adult morphology and the transition to floral development (Poethig, 2009; Xu et al., 2016). The *SPL* genes form a regulatory module with miR156 which is known to regulate the transition to flowering in angiosperms. The levels of miR156 are high in plant embryos and decline as the plant gets older. Conversely, the expression levels of the 10 *SPL* genes that are targeted by miR156 increase with age (Wu and Poethig, 2006; Xu et al., 2016).

In the *P. patens* gametophyte, rather than regulating the transition from vegetative growth to flowering, the *SPL*/miR156 module is involved in the protonema to gametophore transition. Interestingly, while miR156 inhibits maturation in angiosperms, the reverse seems to be the case in *P. patens*, where miR156 promotes the transition to 3D growth. As in angiosperms, miR156 is temporally regulated. Levels of miR156 are low in protonemal tissue but increase to a peak during bud initiation and then decrease again as the gametophores mature. In *P. patens*, the targets of miR156 are *SQUAMOSA PROMOTER BINDING PROTEIN 3,6* and *13* (*SBP3,6* and *13*). Expression of *SBP3,6* and *13* is low in protonema and during bud initiation, despite miR156 levels also being low

initially. Expression of *SBP3,6* and *13* then starts to increase in more mature tissues with gametophores, after bud initiation, when miR156 has reached its peak and is declining again (Cho et al., 2012).

In *P. patens*, the overexpression of a miR156 target mimic, which reduces miR156 function, results in delayed bud initiation and fewer buds being produced. The reduction in miR156 function results in increased expression of *SBP3*, *SBP6* and *SBP13* during the bud initiation period. In wild type plants *SBP3* has the highest relative expression so it is thought to have the dominant function in regulating this growth transition, rather than *SBP6* and *SBP13*. An *sbp3* deletion mutant exhibits the opposite phenotype to the reduced miR156 function line, with increased bud initiation. Therefore, miR156 promotes bud initiation by repressing the expression of *SBP3*, a negative regulator of bud initiation. In addition, miR156 suppresses the accumulation of certain trans-acting small interfering RNAs (tasiRNAs). These tasiRNAs negatively regulate the expression of *APETALA2* (*AP2*) transcription factors and *AUXIN RESPONSE FACTOR* (*ARF*) genes. As yet, it is unclear if the targets of tasiAP2 have a role in the transition to 3D growth but the targets of tasiARF, the *AUXIN RESPONSE FACTOR* genes, are integral to auxin signalling pathways and therefore important for this transition. The production of these tasiRNAs is dependent on another microRNA, miR390, that displays an inverse expression pattern to miR156 during development, with a minimum expression level during bud initiation. However, miR156 does not regulate the expression of miR390 and *vice versa*. An miR390 overexpression mutant also exhibited a delay in the formation of gametophores. miR156 levels were unaffected in the miR390 overexpression mutant, but there was an overaccumulation of tasiAP2 and tasiARF, with a corresponding reduction in their targets. Cytokinin, which promotes bud initiation, does not affect the expression of the miRNAs or tasiRNAs. This indicates that cytokinin regulates bud initiation through a distinct pathway to these miRNAs and tasiRNAs (Cho et al., 2012). It seems that this network of regulatory small RNAs is important for fine-tuning the timing of the developmental transition to 3D growth through temporal regulation of the levels of their targets; *SBP* mRNAs and

ARF mRNAs.

1.7.3 Epigenetic regulation

The Polycomb Repressive Complex 2 (PRC2) is an epigenetic modifier that mediates tri-methylation of lysine 27 of histone 3 (H3K27me3), which is associated with suppression of gene expression. The PRC2 is made up of four core subunits. Three of these four subunits have multiple homologues in plants, so different combinations of the subunit versions create different complexes (Mozgova et al., 2015). Loss of any individual subunit of the PRC2 reduces the level of H3K27me3 in genes targeted by the PRC2, often resulting in substantial phenotypic defects (Ohad et al., 1999; Yadegari et al., 2000; Köhler et al., 2003; Okano et al., 2009; Mosquna et al., 2009; Godwin and Farrona, 2022; Raquid et al., 2023).

In *P. patens*, two components of the PRC2, *CURLY LEAF* (*CLF*) and *FERTILISATION INDEPENDENT ENDOSPERM* (*FIE*), are required for the repression of the sporophyte body plan within the gametophyte generation. Mutants lacking either *PpCLF* or *PpFIE* produce side branch initials that form sporophyte-like structures in place of filaments and gametophores. Unlike sporophytes however, these structures exhibit indeterminate growth (Okano et al., 2009; Mosquna et al., 2009). In a *clf* mutant, the phenotype can be partially suppressed by exogenous auxin treatment, but is unaffected by cytokinin treatment. This suggests that the balance between auxin and cytokinin is perturbed in the *clf* mutant (Raquid et al., 2023).

1.7.4 Cell trafficking

The exocyst is a vesicle tethering complex that is crucial for spatial control of exocytosis. The exocyst is an octamer comprising SEC3, SEC5, SEC6, SEC8, SEC10, SEC15, EXO70 and EXO84 subunits (Heider and Munson, 2012).

In *P. patens*, two exocyst genes have been found to be involved with regulation of the 2D to 3D growth transition; *PpSEC6* and *PpEXO70.3d* (Rawat et al., 2017; Brejšková et al., 2021). In *Ppsec6* mutants, gametophore buds form but they arrest

at an early stage. Examination of the buds in detail showed that division planes are disorganised resulting in developmental arrest because the apical cell cannot be established. Further analysis revealed incomplete cell walls indicating defective cytokinesis in the *Ppsec6* mutants (Brejšková et al., 2021). In *Ppexo70.3d* mutants, the transition from chloronema to caulonema is impaired, resulting in protonema consisting of mostly chloronemal filaments and some filaments with confused chloronemal/caulonemal cell identity. Furthermore, *Ppexo70.3d* gametophores can initiate from chloronemal filaments. The division planes of the early bud are misoriented but despite that, they still manage to form stunted gametophores, although some do abort development. There were also some cases where buds dedifferentiated back to filaments, indicating an inability to maintain 3D fate in those buds. Defective cytokinesis was observed in the protonema but cytokinesis during bud development was not examined. PpEXO70.3d was also shown to function in the deposition of the moss cuticle, another process that is known to have significance for the transition to 3D growth. (Rawat et al., 2017; Moody et al., 2021; Zhang et al., 2023).

1.7.5 Metabolite biosynthesis and the moss cuticle

NO GAMETOPHORES 2* and *CYP98

The *NO GAMETOPHORES 2* (*PpNOG2*) gene encodes a shikimate O-hydroxycinnamoyltransferase that contains two transferase domains. *PpNOG2* is required for the transition to 3D growth (Moody et al., 2021; Kriegshauser et al., 2021). The *no gametophores 2-Reference* (*Ppnog2R*) mutant, which was generated by UV mutagenesis, produces excessive numbers of gametophore initials. However, all of these gametophore initial cells fail to become mature gametophores due to developmental defects. The plane of the first oblique division in the gametophore initial is correct in *Ppnog2R*, as is the second, but subsequently defects occur that invariably result in developmental arrest. Consequently, it was concluded that the *PpNOG2* gene negatively regulates gametophore initiation but positively regulates

gametophore development (Moody et al., 2021).

The *Ppnog2R* mutation is a premature stop codon in the coding sequence. However, *PpNOG2* transcripts are not truncated, but display splicing defects that cause the loss of one of both of the transferase domains. This is likely the result of nonsense-mediated alternative splicing, which suppresses the expression of potentially deleterious truncated proteins by upregulating the expression of alternative splice variants (Wang et al., 2002; Moody et al., 2021). The PpNOG2 protein has three predicted transmembrane domains, and has been shown to localise to the plasma membrane (Moody et al., 2021).

The Arabidopsis orthologue of *PpNOG2* is *hydroxycinnamoyl-coenzyme A: shikimate hydroxycinnamoyl transferase (HCT)*, a component of the phenylpropanoid biosynthesis pathway, which leads to lignin production (Hoffmann et al., 2004; Vanholme et al., 2019). Lignin is not found in *P. patens*, but PpNOG2 functions instead in the ascorbic acid pathway, which culminates in the production of moss cuticle biopolymers (Kriegshauser et al., 2021). Therefore, genes from the ascorbic acid pathway in *P. patens*, such as *PpNOG2* and *PpCYP98*, the enzyme that catalyses a subsequent step in the pathway, appear to have been co-opted from cuticle biosynthesis to lignin biosynthesis in vascular plants (Renault et al., 2017). The defects in gametophore development exhibited by *Ppnog2R* mutants may therefore be partly caused by disruption of cuticle biosynthesis. In support of this, *Ppcyp98* mutants fail to maintain 3D growth. Gametophores are initiated but development is aborted due to organ fusion (Renault et al., 2017). The cuticle is present in gametophores but absent in protonemal tissues (Lee et al., 2020). The question still remains as to exactly when the cuticle begins to be deposited during the transition to 3D growth and whether 3D growth is dependent on the presence of the cuticle. Cuticle deposition is known to be important for morphogenesis as correctly deposited cuticle is vital for controlling anisotropic cell expansion (Zhang et al., 2023; Bridson et al., unpublished).

The expression of *PpNOG2* is induced by application of the cytokinin analogue BAP, but not the auxin analogue NAA. While auxin treatment does not affect

PpNOG2 expression, auxin homeostasis is disrupted in *Ppnog2R*. Application of NAA to *Ppnog2R* suppresses the overbudding phenotype, implying that auxin levels in *Ppnog2R* are not high enough to negatively regulate bud formation. Also, in *Ppnog2R*, the expression of the auxin responsive genes *PpDOX*, *PpRSL* and *PpCBS* are altered in a manner that is consistent with lower levels of auxin in *Ppnog2R* than in wild type (Moody et al., 2021). A possible explanation can be found by looking at the substrate of *PpNOG2*, *p*-coumaroyl-CoA. In Arabidopsis, *p*-coumaroyl-CoA sits at the metabolic junction between lignin and flavonoid biosynthesis. When *HCT* is knocked out, all substrate is diverted into the flavonoid biosynthesis pathway via the enzyme chalcone synthase (CHS), resulting in flavonoid overaccumulation. This is thought to retard plant growth, as flavonoids are inhibitors of polar auxin transport (Brown et al., 2001). The inhibitory effect of the flavonoid naringenin on auxin transport has also been demonstrated in *P. patens* and the presence of functional chalcone synthases ascertained (Jiang et al., 2006; Bennett et al., 2014). Therefore, it is hypothesised that the *Ppnog2R* phenotype is partially the result of perturbed auxin homeostasis caused by blocking the ascorbic acid pathway and the resulting increased flux into flavonoid biosynthesis (Moody et al., 2021; Kriegshauser et al., 2021).

Glycerol-3-phosphate acyltransferase

A study of the cuticle in *P. patens*, found that the gene *Glycerol-3-phosphate acyltransferase 2* (*PpGPAT2*) is required for cuticle formation. Content analysis of the components of the *P. patens* cuticle revealed that the cuticle that forms on gametophores is similar to the cuticle layer of more derived land plants. Therefore the *P. patens* genome was scrutinised for homologues of Arabidopsis cuticle biosynthesis genes. From this search, the *GPAT* genes stood out because homologues were found in land plant representatives but not in the algal species that were examined; *Chlamydomonas reinhardtii*, *Klebsormidium nitens* and *Chara braunii*. The *GPAT* gene family therefore seems to be restricted to land plants (Lee et al., 2020). In Arabidopsis the *GPAT* enzymes are involved in the synthesis

of the cuticle component cutin (Li-Beisson et al., 2009; Yang et al., 2010, 2012). In *P. patens*, there are 7 *GPAT* homologues (*PpGPAT1-7*). The expression of *PpGPAT3* could not be detected, but all of the other *PpGPAT* genes were expressed in the gametophores. Only *PpGPAT4* expression was detected in the protonema. *PpGPAT1,2,4,5,6* and 7 all localise to the endoplasmic reticulum (Lee et al., 2020). The Arabidopsis *gpat6-2* mutant is cutin deficient, resulting in increased cuticle permeability in petals and organ fusion in the flowers (Li-Beisson et al., 2009). Interestingly, the phenotype of the *gpat6-2* mutant was rescued by heterologous expression of *PpGPAT2* and *PpGPAT4*, indicating functional conservation across land plants. The other *PpGPAT* genes did not rescue the *gpat6-2* mutant phenotype. This suggested that *PpGPAT2* and *PpGPAT4* are involved in cutin biosynthesis in *P. patens*. Subsequently, $\Delta Ppgpat2$ and $\Delta Ppgpat4$ knockout mutants were generated. Upon examination $\Delta Ppgpat4$ was indistinguishable from wild type. However, the $\Delta Ppgpat2$ mutant produced approximately 50% fewer gametophores than wild type and those gametophores exhibited growth retardation. The presence of cutin monomers was approximately 38-40% lower in $\Delta Ppgpat2$ than in wild type, confirming that this mutant formed a defective cuticle, which was responsible for delayed gametophore development (Lee et al., 2020).

1.7.6 Protein degradation

The levels of proteins within a cell are determined by the relative rates of protein synthesis and protein degradation. Some proteins persist for long periods in the cell, while others are turned over very rapidly, depending on active retention/degradation and the inherent stability of the protein. The regulation of protein degradation is therefore vital for cell function and to allow rapid responses to internal or external cues (Vierstra, 1993). Protein degradation is also important for quality control as it removes misfolded or damaged proteins so that they don't have negative impacts on the cell (Goldberg, 2003). There are two major protein degradation pathways; the ubiquitin-proteasome system (UPS) and the lysosomal proteolysis pathway.

The Ubiquitin-Proteasome System

The 26S proteasome is an ATP-dependent macromolecular protease complex. The catalytically active 20S proteasome or core particle, a barrel shaped structure, is at the centre of the complex. At each end of the 20S proteasome is a cap of regulatory proteins called the 19S regulatory particle (Tanaka, 2009).

The proteasome selectively degrades proteins that have been tagged with polyubiquitin. The process of ubiquitination requires the sequential activity of 3 families of proteins; E1, E2 and E3. Monoubiquitin is activated by an E1 enzyme in an ATP-dependent manner. The E1 enzyme then transfers the activated monoubiquitin to the active site of an E2 ubiquitin-conjugating enzyme. Subsequently, an E3 ubiquitin ligase binds both the target protein and the E2 bearing activated ubiquitin, bringing them in close proximity. This allows the E2 to catalyse the ubiquitination reaction, attaching the ubiquitin to the target protein substrate (Kleiger and Mayor, 2014).

Once a protein has been monoubiquitinated, additional ubiquitin units can be added to generate a polyubiquitin chain, which is recognised by the proteasome (Figure 1.5). The residue at which the ubiquitin-ubiquitin linkage is formed influences the topology of the chain. Ubiquitin can be linked to another ubiquitin at any of 8 different sites (Lys6, Lys11, Lys27, Lys29, Lys33, Lys48, Lys63 or Met1), with Lys48-linkage being the most common in plants and most commonly associated with functioning in the UPS (Kim et al., 2013; Kwon and Ciechanover, 2017).

Ubiquitination has several functions in addition to the UPS, including but not limited to splicing, DNA repair, histone modification, endocytosis/trafficking, transcription, translation, cell signalling and altering protein function (Rotin et al., 2000; Dover et al., 2002; Katzmann et al., 2002; Muratani and Tansey, 2003; Sun and Chen, 2004; Hirano et al., 2006; Chen and Sun, 2009; Rahighi et al., 2009; Bomar et al., 2010; Cano et al., 2015; Bonacci et al., 2017; Dougherty et al., 2020; Zajicek and Yao, 2021).

Polyubiquitin chains can be branched or unbranched and contain one or more types of linkage, creating a diverse array of potential topologies and allowing them

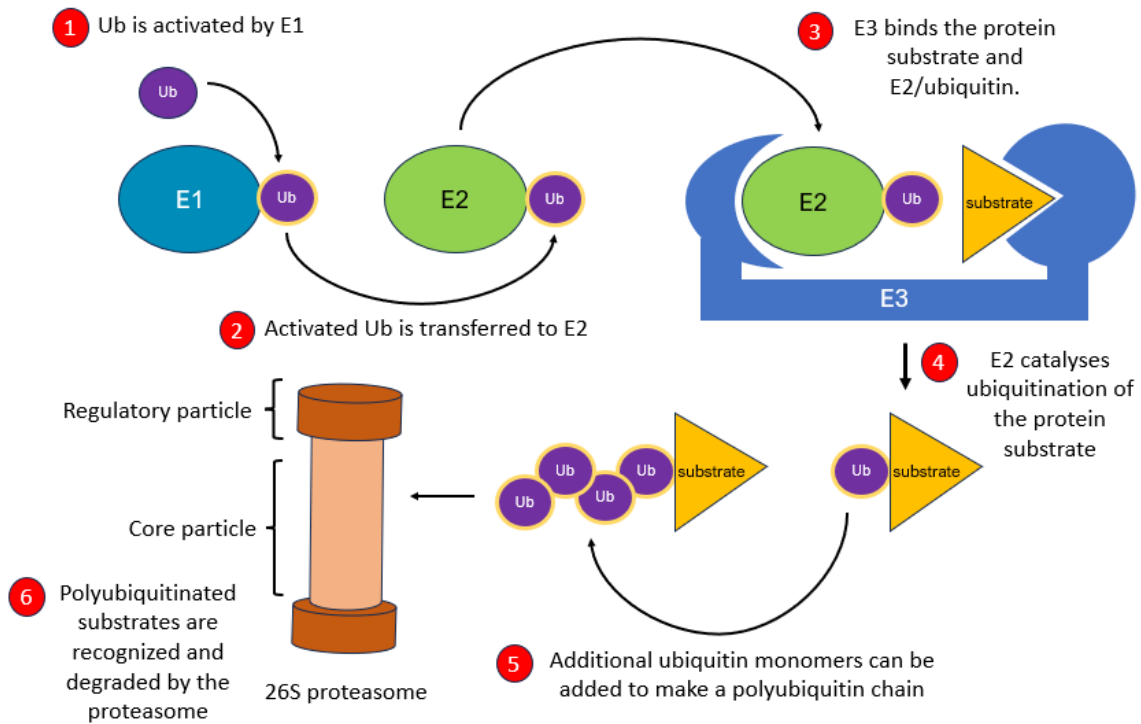


Figure 1.5: The Ubiquitin Proteasome System (UPS); a major protein degradation pathway.

to act as a multifunctional signal and therefore confer a variety of fates to the targeted protein (Pickart and Fushman, 2004; Woelk et al., 2007; Ye et al., 2012; Kwon and Ciechanover, 2017; Orosa-Puente and Spoel, 2022). There are several types of ubiquitin-binding domain which act as receptors for polyubiquitin chains. Some ubiquitin-binding domain-containing proteins function as shuttle proteins, that guide ubiquitinated substrates to the proteasome, whereas others play a role in other ubiquitin-dependent processes. Examples of ubiquitin-binding domains include the Ubiquitin-Interacting Motif (UIM) and the Ubiquitin-Associated (UBA) domain, but many others have been characterised (Hicke et al., 2005; Hurley et al., 2006; Husnjak and Dikic, 2012; Randles and Walters, 2012).

Ubiquitination is carefully regulated and occurs at certain motifs within proteins. Some of these motifs are constitutive and always present in a protein sequence, whereas others are conditional. Two examples of conditional motifs are N-degrons and C-degrons. Cytosolic proteases can cleave proteins to leave exposed motifs at the N-terminus of the C-terminal fragment (N-degron) and/or the C-terminus of the N-terminal fragment (C-degron). These newly created motifs are recognised

by E3 ubiquitin ligases, initiating ubiquitination and subsequent targeting to the proteasome for destruction (Varshavsky, 2019; Timms and Koren, 2020).

Lysosomal proteolysis

Lysosomes are membrane-bound organelles containing a variety of proteases and other digestive enzymes with broad specificity. Proteins are trafficked to the lysosome via vesicles. Extracellular proteins are taken up by endocytosis before being delivered to the lysosome. Intracellular proteins are taken up in autophagosomes, where small volumes of the cytosol are engulfed into vesicles at the endoplasmic reticulum. Vesicles fuse with the lysosome, releasing their contents into the organelle, where the proteins are degraded by proteases (Pillay et al., 2002).

Several of the proteins that are required for the 3D growth transition have been implicated in protein degradation processes.

Arginyl-tRNA protein transferase and the N-degron pathway

Arginyl-tRNA protein transferase (ATE) is a post-translational modifier that adds arginine to the N-terminus of proteins. The N-terminal arginine is recognised by certain E3 ubiquitin ligases (called N-recognins) that are components of the N-degron pathway and the protein is ubiquitinated and subsequently degraded by the proteasome (Kwon et al., 1999; Graciet et al., 2009). In *P. patens* there is a single *PpATE* gene and the protein exhibits a cytosolic localisation. A translational reporter revealed no expression in the protonema but strong expression in the developing bud, at the apex and base of young gametophores and throughout mature gametophores. A *Ppate* knockout mutant displayed severely delayed 3D growth. Although bud-like structures formed after two weeks, gametophores only developed after 4 months and were severely stunted and deformed. In the protonema of *Ppate*, chloronemal cells were both shorter and wider than wild type, whereas caulonemal cells were longer and wider than wild type. Cell divisions along the filaments were disorganised both in angle and spacing. Protonemal cells in the mutant were densely packed with chloroplasts that exhibited altered morphology. Those chloroplasts

were shown to have reduced chlorophyll content due an overaccumulation of starch (Schuessele et al., 2016).

It has been shown that PpATE interacts directly with a class I small heatshock protein; sHSP17.2a, which is not a target for arginylation but rather an interaction partner that facilitates function. This interaction was confirmed *in vivo* using Förster Resonance Energy Transfer (FRET). Four arginylation targets were identified in a pull-down; Acylamino acid-releasing enzyme (PpAARE), an uncharacterised protein (although I note that the Arabidopsis homologue AT3G19990 was more recently annotated as an E3 ubiquitin-protein ligase in Araport11), a putative AAA-type ATPase (PpATAD3.1) and an ABC transporter family protein (PpABCB20) (Hoernstein et al., 2016). PpAARE1 and 2, which are highly conserved proteases, have been shown to regulate ageing in *P. patens*. In double *Ppaare1/2* knockout lines, gametophore formation and senescence occurs earlier than in wild type. Therefore, PpAARE1 and 2 inhibit ageing (Hoernstein et al., 2023). This may explain the delayed development of the *Ppate* mutant in which the PpAARE genes are not being arginylated and degraded by the proteasome and are therefore inhibiting developmental progression. Therefore, *PpAARE* genes regulate the timing of the initiation of the transition to 3D growth. The regulation of PpAARE does not explain why the gametophores in *Ppate* mutants were severely stunted and deformed. One of the other targets of *PpATE* is likely important for correct gametophore development (Schuessele et al., 2016).

Multiubiquitin chain binding proteasome subunit

Multiubiquitin Chain Binding Subunit (MCB1; also known as RPN10) is a component of the 26S proteasome complex. MCB1 is a receptor for polyubiquitin that resides in the 19S regulatory particle and therefore mediates the interaction between ubiquitinated targets and the proteasome, although there are also several alternative ubiquitin receptors (Girod et al., 1999; Finley et al., 1998; Elsasser et al., 2004). Monoubiquitination of the MCB1 protein itself interferes with the ability of the MCB1 Ubiquitin Interacting Motif (UIM) to percieve polyubiquitin. Therefore,

monoubiquitination regulates the receptor activity of MCB1 (Isasa et al., 2010). *Ppmcb1* mutants are unable to initiate gametophores, even after months of growth. In the protonema of *Ppmcb1*, chloronemal cells appear nearly normal, although slightly shorter than wild type. However, caulonemal cells were shorter and had swollen tips. The inability of *Ppmcb1* to produce gametophores was partially rescued by application of both the auxin indoleacetic acid (IAA; 5 nM) and the cytokinin isopentenyladenine (IPA; 50 nM). Hormone treated *Ppmcb1* was able to form normal caulonema and able to initiate buds which developed into gametophores, albeit after some delay. However, the gametophores were severely stunted, failed to reach maturity and only yielded 4-6 phyllids. Treatment with either one of the hormones individually was less effective at rescuing the *Ppmcb1* phenotype. 5 μ M IAA restored caulonemal growth, but not bud initiation. 5 μ M IPA restored bud initiation but there was not much further development of the gametophores. Surprisingly, expression of a version of MCB1 in which a critical hydrophobic patch within the UIM was mutated, or a version where the UIM was completely removed, mostly rescued the *Ppmcb1* mutant. This indicated that PpMCB1 has additional important roles besides the perception of polyubiquitinated substrates (Girod et al., 1999).

DEFECTIVE KERNEL 1

The DEFECTIVE KERNEL 1 (PpDEK1) protein has a 23 transmembrane domain anchor which contains an intracellular loop and at the C-terminus there is a cytosolic calpain cysteine protease domain (Demko et al., 2014). PpDEK1 is the only calpain protein found in plants, although calpain domain-containing proteins are found across all eukaryotes (Lid et al., 2002). PpDEK1 is implicated in perceiving and transmitting positional information to facilitate 3D development in angiosperms. In Arabidopsis, homozygous *Atdek1* mutants are embryo lethal, as cell divisions in the embryo are disordered, causing developmental arrest by the globular stage (Lid et al., 2005). In maize, most *Zmdek1* mutants are also embryo lethal (Lid et al., 2002). *DEK1* also has a more widespread role in plant development as viral-induced silencing of the gene in *Nicotiana benthamiana* caused developmental arrest in organs

due to excessive ectopic cell proliferation (Ahn et al., 2004).

In *P. patens*, during gametophore development, PpDEK1 localises to the plasma membrane between recently divided cells, but expression is abolished in cells that are no longer elongating or dividing (Perroud et al., 2020). *Ppdek1* loss-of-function mutants are unable to produce gametophores, but form excessive buds that undergo early developmental arrest. *PpDEK1* is thus a negative regulator of gametophore induction. In *Ppdek1* buds, the first cell division is unaffected but subsequent cell divisions are incorrectly oriented, resulting in a failure to correctly establish the tetrahedral apical cell that is required for 3D growth. Therefore, *PpDEK1* is also required for correct division plane placement during early gametophore development (Perroud et al., 2014). Interestingly, the *Ppdek1* phenotype can be rescued by expression of the calpain domain alone, although gametophore development is slower and some phyllids are abnormal (Perroud et al., 2014). *Atdek1* mutants can also be complemented by the Arabidopsis DEK1 calpain domain, or indeed the *P. patens* DEK1 calpain domain, which indicates that DEK1 is functionally conserved across land plants (Liang et al., 2013). However, the same is not true in reverse as *Ppdek1* mutants can not be complemented with the Arabidopsis DEK1 calpain domain. This non-reciprocal complementation could indicate that *PpDEK1* retains a broader functionality than *AtDEK1*. It was suggested that because AtDEK1 usually functions in the Arabidopsis sporophyte, perhaps it is lacking a necessary sporophyte-specific component to achieve function when expressed in the moss gametophyte (Perroud et al., 2014). Notably, *Atdek1* cannot be complemented by the calpain domain of a DEK1-like protein from the unicellular charophyte *Mesostigma viride*, so it seems that there is not functional conservation between algae and land plants (Liang et al., 2013). The extracellular loop domain is important for PpDEK1 function, as mutants lacking the loop produce gametophores with aberrant phyllid morphologies due to misregulation of mitosis. In the *Ppdek1* mutant, expression of *PpAPB2* and *3* is elevated, which may explain the excessive budding phenotype, but does not explain why the buds abort (Demko et al., 2014). It has proven difficult to identify the proteolytic targets of the PpDEK1 calpain domain

by direct experimental observation, so a transcriptomics data-based approach was used to examine the role of PpDEK1. The proposed model is that PpDEK1 has a dual role. Firstly, PpDEK1 directly cleaves protein targets to alter the function of, or activate, a protein. Secondly, PpDEK1 indirectly affects protein stability by generating N-degrons through cleavage, resulting in the target being degraded by the proteasome. Given the impact of PpDEK1 on the transcriptome, the targets of PpDEK1 are likely to include several transcription factors, leading to affects on the expression of downstream genes. This hypothesis would explain the severe pleiotropic phenotypes of *Ppdek1* mutants and is supported by transcriptome data. Gene expression is significantly impacted in the *Ppdek1* mutant, with 35% of all genes differentially expressed in the mutant versus wild type. When comparing the transcriptomes of *Ppdek1* and a *PpDEK1* overexpression line, many of the differentially expressed genes showed expression patterns consistent with those genes being the targets of either a transcriptional repressor or transcriptional activator that is cleaved by PpDEK1, and consequently targeted for destruction. Gene Regulatory Network analysis examined these genes that were affected in the transcriptome data and identified predicted upstream transcription factors that may be direct targets of PpDEK1 for cleavage. It was found that the predicted transcription factor targets of PpDEK1 are enriched in calpain cleavage sites that would produce an N-degron which further supports this hypothesis (Demko et al., 2024). This model would explain the significant effect PpDEK1 has on the transcriptome and on *P. patens* development, although it is yet to be validated experimentally.

NO GAMETOPHORES 1

NO GAMETOPHORES 1 (*PpNOG1*) was the first 3D regulator identified from the forward genetics screen carried out by our lab. The *no gametophores 1-Reference* (*Ppnog1R*) mutant produces fewer gametophore initials than wild type, even in the presence of exogenous cytokinin. Those gametophore initials that do form develop abnormally and arrest at an early stage, resulting in a complete failure to generate any gametophores. The typical pattern of cell divisions in the developing bud is

completely defective in this mutant, as the division plane orientation is consistently misaligned. Instead of the first division of the gametophore initial cell being oblique, as it is in the wild type, the division is roughly parallel to the parent caulonemal filament. Subsequent divisions occur randomly and bud development normally arrests by the 5-cell stage (Moody et al., 2018).

Transcript levels of *PpAPB1-4*, positive regulators of gametophore initiation, are suppressed in *Ppnog1R*, providing a potential explanation for the repressed budding phenotype. Several components of the *PpCLV* signalling pathway are affected in *Ppnog1R*; *PpCLE4* and *PpCLE5* are downregulated, whereas *PpCLE7* is upregulated. *PpNOG2* expression is suppressed in the *Ppnog1R* mutant. The *PpCLV* genes that are affected in *Ppnog1R*, show identical perturbations in *Ppnog2R*, indicating that the *CLV* pathway operates downstream of both *PpNOG1* and *PpNOG2* (Moody et al., 2018, 2021).

The causative mutation in *Ppnog1R* was mapped to a locus that contains the gene now known as *NO GAMETOPHORES 1* (*PpNOG1*), which encodes a protein that possesses a C-terminal ubiquitin-associated (UBA) domain. A UV-induced C $\rightarrow$ T point mutation introduced a premature stop codon (R⁷⁶Ter) into the coding sequence, resulting in a truncated protein that lacks the UBA domain. Ectopic expression of *PpNOG1* in the *Ppnog1R* mutant complemented the mutant phenotype and an independent knockout of this gene in wild type recapitulated the mutant phenotype. This confirmed that a mutation in *PpNOG1* was the cause of the *Ppnog1R* phenotype. Thus, *PpNOG1* is a positive regulator of gametophore initiation and development (Moody et al., 2018).

A previous study described the rice orthologue of *PpNOG1*, *HEADING AND GRAIN WEIGHT* (*HGW*), which was identified in a screen for genes that influence rice grain yield. The *hgw* mutant had a later heading (flowering) date, and lower grain weight than wild type. All *hgw* mutants were heterozygous for the mutant *hgw* allele and it was not possible to obtain a homozygous mutant. This suggests that a homozygous mutant would be embryo lethal and indicates that *HGW* has a dosage dependent effect on phenotype. The *HGW* protein tagged with a fluorescent

protein showed cytosolic and nuclear localisation (Li et al., 2012).

The UBA domain can be found across all eukaryotes and was identified as a motif commonly found in proteins associated with the ubiquitination pathway. However, as the UBA domain has no conserved polar residues it is unlikely that it is directly involved in the ubiquitination reaction itself (Hofmann and Bucher, 1996). Rather, it has been shown that this motif is capable of binding non-covalently to ubiquitin. The structure of the domain is a compact three α -helix bundle. Between the first two of these α -helices is a hydrophobic patch which is thought to be the UBA domain-ubiquitin interaction surface, as well as being important for maintaining the structure of the three α -helix bundle. In the yeast protein Ddi1, which contains a C-terminal UBA domain, mutating a particular conserved leucine residue to alanine in this hydrophobic patch results in complete loss of ubiquitin-binding affinity (Bertolaet et al., 2001). Between helix one and two in the bundle resides the highly conserved ‘MGF’ (Met-Gly-Phe) loop, which is important for regulating binding affinity of the UBA domain. Studies of the structure of the UBA from the yeast protein Dsk2 showed that mutating either the glycine or phenylalanine in the MGF loop dramatically decreases the affinity of the UBA for binding ubiquitin (25-fold and 68-fold respectively). The methionine is also crucial for stabilising the interface between the UBA and ubiquitin. Structural analysis of the Dsk2^{UBA}/ubiquitin complex revealed the methionine residue is in direct contact with ubiquitin and the critical hydrophobic patch and also accepts a hydrogen bond from ubiquitin (Ohno et al., 2005).

UBA domains bind to both monoubiquitin and polyubiquitin chains, with some being selective for Lys48-linked polyubiquitin or Lys63-linked polyubiquitin. Some UBA domains do not bind any tested form of ubiquitin (Raasi et al., 2005). The well-studied UBA domain-containing proteins Rad23, Dsk2 and Ddi1 from yeast are known to function in the UPS. All three genes have a C-terminal UBA domain (Rad23 has an additional one midway through the protein) and an N-terminal Ubiquitin-like (UbL) domain. This is a common feature in proteins that interact directly with ubiquitinated substrates and the proteasome (Su and Lau, 2009;

Farmer et al., 2010). For Rad23 and Dsk2 it was shown that the C-terminal UBA domain not only acts as a receptor for polyubiquitinated substrates but also protects the protein from itself being degraded by the proteasome (Heessen et al., 2005; Heinen et al., 2011).

Besides shuttling ubiquitinated substrates to the proteasome, other functions of UBA domains have been described. For example, the UBA domain-containing Edl1 protein from yeast binds monoubiquitinated proteins as part of its function in the endocytic pathway (Swanson et al., 2006). The E2 conjugating enzyme known as Ubc1 in yeast, or Ube2K in mammals, catalyses the formation of Lys48-linked ubiquitin chains. However the protein contains a UBA domain with an affinity for Lys63-linked ubiquitin. It was revealed that the UBA facilitates the formation of branched Lys63/Lys45-linked ubiquitin chains (Pluska et al., 2021). The scaffolding protein p62 sequesters ubiquitinated targets into aggregates that are then deposited into autophagosomes and trafficked to the lysosomal proteolytic pathway. p62 interacts with the ubiquitinated substrates via its UBA domain and aggregate formation occurs by homo-oligomerisation of its PB1 domain. The ability of p62 to sequester ubiquitinated substrates is actually enhanced by ubiquitination of the UBA domain (Lee et al., 2017). Homodimerisation of UBA domains has been shown to inhibit function by blocking the ubiquitin-UBA interface and therefore preventing the binding of ubiquitin to the UBA (Long et al., 2010).

Because of the diverse functions of UBA domains that have already been revealed, the mechanism by which PpNOG1 functions remains an important question, but an interaction with a ubiquitinated target (or targets) and a role in the ubiquitin-proteasome pathway is the most likely scenario.

1.8 Summary

Our current knowledge of the regulatory network that controls the 2D-to-3D transition is summarised in Figure 1.6. Auxin and cytokinin crosstalk are central to the control of 3D initiation and maintenance. Auxin promotes differentiation and

cytokinin promotes cell proliferation so the spatiotemporal regulation of cytokinin and auxin signalling is crucial for coordinating apical cell establishment and maintenance with organogenesis and growth (Hata and Kyozyuka, 2021).

The transition from 2D growth to 3D growth begins with the formation of a gametophore initial from a caulonemal filament. Cytokinin is necessary and sufficient to induce gametophore initiation (Brandes and Kende, 1968; Ashton et al., 1979). Auxin is required for cytokinin-mediated gametophore induction, but treatment with large doses of auxin antagonises the effect of cytokinin and inhibits gametophore initiation (Brandes and Kende, 1968; Lehnert and Bopp, 1983).

The *PpAPB* transcription factors are the master regulators of gametophore initial formation (Aoyama et al., 2012). The *PpTPL* coactivators, *PpNOG1* and *PpMCB1* also promote gametophore initiation, whereas *PpDEK1*, *PpNOG2*, *PpSBP3* and the PRC2 inhibit gametophore initiation (Girod et al., 1999; Okano et al., 2009; Mosquana et al., 2009; Cho et al., 2012; Perroud et al., 2014; Moody et al., 2018, 2021; Causier et al., 2023).

After the formation of the gametophore initial, precisely oriented cell divisions occur resulting in the establishment of the tetrahedral apical cell. This cell is the 3D capable cell that divides to renew itself and generate organ initials for organogenesis (Harrison et al., 2007). Misregulation of cell divisions in the early gametophore can lead to early developmental arrest or other developmental defects such as the formation of ectopic apical cells. Several genes are known to be important for regulating these cell division planes; *CLAVATA* signalling genes, *PpNOG1*, *PpNOG2*, the *PpTPL* coactivators and *PpDEK1* (Perroud et al., 2014; Whitewoods et al., 2018; Moody et al., 2018; Cammarata et al., 2022; Moody et al., 2021; Causier et al., 2023). The cytoskeleton is also critical for controlling the angle of cell divisions (Kosetsu et al., 2017; Kozgunova et al., 2022).

It is becoming increasingly clear that the moss cuticle is critical for 3D development. *PpNOG2*, *PpCYP98* and *PpGPAT2* are all involved in cuticle synthesis and have been shown to be required for 3D growth (Renault et al., 2017; Lee et al., 2020; Moody et al., 2021; Kriegshauser et al., 2021). The exocyst, a

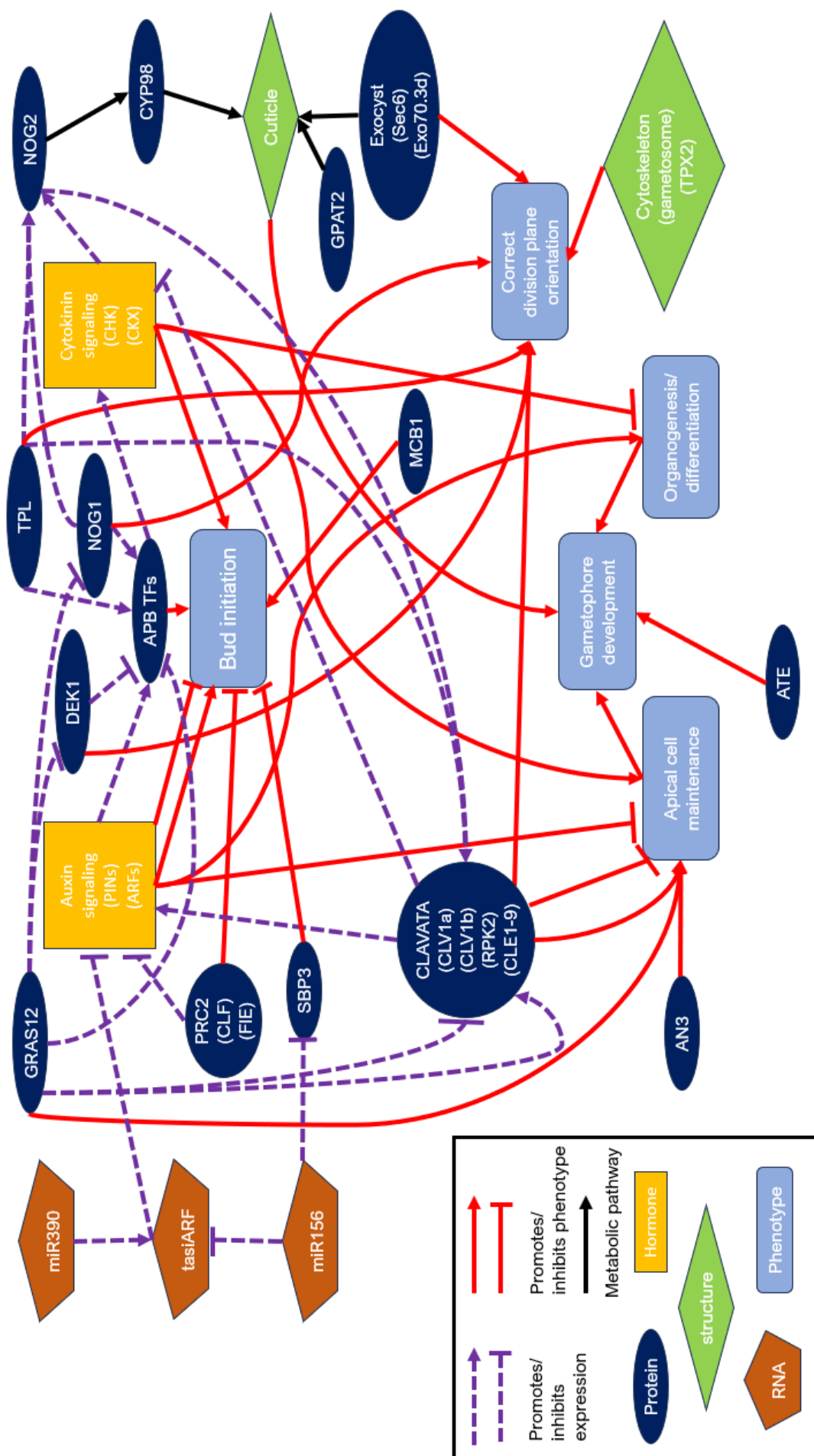


Figure 1.6: Our current understanding of the regulatory network underpinning the transition from 2D to 3D growth in *P. patens*.

complex important for cell trafficking has been demonstrated to be important both for cuticle deposition and for proper cell division (Rawat et al., 2017; Brejšková et al., 2021).

Once the apical cell has been established it must be maintained to allow gametophores to develop to maturity. Cytokinin promotes apical cell identity. Cytokinin is antagonised by the *CLAVATA* signalling pathway, which inhibits apical cell identity (Cammarata et al., 2022). *PpAN3* is required to promote sufficient proliferation of the apical cell (Kawade et al., 2020). PpATE targets proteins to the proteasome by arginylation of the N-terminus, which activates the N-degron pathway (Schuessele et al., 2016). One of its targets, PpAARE controls the rate of ageing, and therefore the timing of gametophore initiation (Hoernstein et al., 2016, 2023). One or several of the other targets of PpATE likely have a role in regulating gametophore development (Schuessele et al., 2016).

The relationships between many of these regulators have been inferred indirectly. Expression level data and hormone treatments can tell us much about where in the network these regulators are acting with relation to each other. However, little is known about the direct functional interactions that are occurring within this network. Consequently, there are gaps in this model and research is needed to elucidate the mechanisms by which these regulators bring about the transition to 3D growth.

1.9 Project aims and objectives

My main aim was to genetically dissect the 2D to 3D growth transition in *P. patens*. My project was primarily focused on *NO GAMETOPHORES 1* as it is an early regulator of the transition to 3D growth, and its direct function is completely uncharacterised. Therefore, I set out to do the following:

- Identify interaction partners of the PpNOG1 protein
- Investigate the functional importance of the ubiquitin-associated domain

- Investigate the genetic interaction between *PpNOG1* and *PpDEK1*
- Identify novel genetic interaction partners of *PpNOG1*
- Perform transcriptomics to identify genes that are differentially expressed in 3D-defective mutants

Chapter 2: Materials and methods

2.1 *P. patens* lines and growth conditions

The species used was *Physcomitrium patens* subsp. *patens*. The lines used in this work are as follows:

<u>Line</u>	<u>Source</u>	<u>Citation</u>
Gransden wild type (GdWt)	Widely available	N/A
Villersexel wild type (VxWt)	Widely available	N/A
Villersexel::mCherry (VxCh)	Pierre-François Perroud	Perroud et al. (2011)
Gransden::GFP (GdGFP)	Pierre-François Perroud	Perroud et al. (2011)
Reute::mCherry (ReCh)	Pierre-François Perroud	Perroud et al. (2019)
<i>no gametophores 1-Reference (Ppnog1R)</i>	Previously generated	Moody et al. (2018)
<i>no gametophores 2-Reference (Ppnog2R)</i>	Previously generated	Moody et al. (2021)

All tissue culture was performed in a laminar flow hood. To encourage bud development, *P. patens* was grown on cellophane-overlaid BCD medium (1 mM MgSO₄, 1.84 mM KH₂PO₄ (pH 6.5), 10 mM KNO₃, 45 μ M FeSO₄; 1 mM CaCl₂; 0.1% Trace Elements Solution: 116 μ M AlK(SO₄)₂, 220 μ M CuSO₄, 10 mM H₃BO₄, 235 μ M KBr, 660 μ M LiCl, 230 μ M CoCl₂, 190 μ M ZnSO₄, 2mM MnCl₂, 170 μ M KI, 124 μ M SnCl₂) containing 0.8% agar. For routine propagation, and to stimulate filamentous growth, *P. patens* was grown on BCD medium supplemented with 5 mM ammonium tartrate (BCDAT). For propagation, tissues were harvested, homogenized in 10 ml sterile water using an Tissue Homogeniser (Omni; TH-03) and then 1-2 ml was pipetted onto cellophane-overlaid BCDAT plates (Cove et al., 2009a). Plates were then placed in a growth cabinet at 24 °C with a 16 h: 8 h light (300 μ mol m⁻² s⁻¹): dark cycle. When protoplasts were created, they were plated onto cellophane-overlaid Protoplast Regeneration Medium (PRMB; BCDAT supplemented with 10 mM CaCl₂, 0.5% glucose and 6% (w/v) D-Mannitol). The

PRMB medium had the same osmolarity as the protoplasts, which prevented them from bursting while their cell walls regenerated (Cove et al., 2009b).

2.2 *P. patens* transformation

Prior to a *P. patens* transformation, large quantities of plasmid DNA were acquired by midiprep using a Plasmid Midi Kit (Qiagen; 12143), according to manufacturer instructions. Approximately 20 μg of plasmid was linearized by restriction digest at 37°C, overnight. For the last hour of the digest, the reaction was treated with either 0.1 volumes of Calf Intestinal Alkaline Phosphatase (Invitrogen; 18009019) or FastAP Thermosensitive Alkaline Phosphatase (Thermo Scientific; EF0651). To precipitate the linearised plasmid, 0.1 volumes of 3 M sodium acetate (pH 5.2) and two volumes of 100% ethanol were added to the digest reaction, incubated at -20 °C for >1 h and centrifuged at full speed for 30 min. The DNA pellet was washed with 70% ethanol, air-dried and resuspended in 20-30 μl sterile water.

Before starting the transformation, all solutions required were made fresh and filter sterilized using a 0.2 μm PES syringe filter (Corning; 431229). The transformation was carried out in a laminar flow hood. 2 g of polyethylene glycol (PEG) 6000 in a flat-bottomed vial, that had been sterilised in the autoclave, was melted in the microwave for approximately 1 min. 5 ml D-Mannitol/ $\text{Ca}(\text{NO}_3)_2$ solution (0.8% D-Mannitol, 0.1 M $\text{Ca}(\text{NO}_3)_2$, 10 mM Tris pH 8.0) was added to the molten PEG 6000, shaken vigorously to completely mix, and allowed to cool for 2-3 h.

Next, *P. patens* tissue was digested to remove cell walls and generate large numbers of protoplasts for the transformation. For each *P. patens* line that was being transformed, a 50 ml Falcon tube with 0.1 g of Driselase (Sigma; D9515) in 10 ml 8% D-Mannitol was prepared. The tube containing the enzyme solution was wrapped in foil and rocked on an orbital shaking platform for approximately 10 min at room temperature. The Driselase/D-Mannitol solution was centrifuged for 3 min at 41 000 $\times g$, and the supernatant was filter sterilised through a 0.2 μm PES syringe filter

(Corning; 431229), into a fresh, sterile 50 ml Falcon tube. For each *P. patens* line that was to be transformed, 1-2 plates of tissue grown on BCDAT were harvested. *P. patens* tissue was placed in the tube with the Driselase solution, wrapped in foil and rocked very gently until the tissue appeared well digested (approximately 40 min).

The digested tissue was passed through a 40 μm nylon cell strainer (Corning; 431750) to separate protoplasts from the remaining tissue debris. The protoplasts were centrifuged for 3 min at no more than 120 xg without braking, the supernatant was removed and the protoplasts were resuspended in 6 ml 8% D-Mannitol. This step was repeated once more. The cell density of the protoplasts was determined using a haemocytometer (Fuchs-Rosenthal). The protoplasts were then centrifuged for 3 min at 120 xg without braking and resuspended in MMM (0.5 M D-Mannitol, 0.15 M MgCl_2 , 0.1% MES pH5.6), to achieve a cell density of approximately 1.5×10^6 cells ml^{-1} . The next steps were carried out without delay as the protoplasts cannot tolerate the MMM solution for very long.

At least 10 μg of the linearised construct was pipetted into the bottom of a sterile, round-bottomed 14 ml tube. To this was added 300 μl of protoplast suspension. 300 μl of the PEG solution was then added in drops and the tube was swirled gently after the addition of each drop. The tube was heat-shocked at 45 °C in a waterbath for 5 min and then incubated at room temperature for a further 5 min.

Next, 300 μl 8% D-Mannitol was added to the tube, 5 times at 3 min intervals and the mixture was gently swirled after each addition. Then, 1 ml 8% D-Mannitol was added to the tube 5 times at 3 min intervals, gently tilting the tube after each addition to mix it. The tube was centrifuged at 120 xg without braking, supernatant removed, and the cells were resuspended in 3 ml 8% D-Mannitol. This suspension was gently pipetted onto three cellophane-overlaid PRMB plates (1 ml per plate). These plates were sealed with micropore tape and wrapped in foil for 24 h to prevent the activation of light-induced DNA repair. Subsequently, they were unwrapped and placed in normal growth conditions for 5-7 days.

After about one week, the cellophanes, along with the regenerating protoplasts,

were transferred to BCDAT medium supplemented with antibiotics (either 40 μgml^{-1} G418 disulphate [Sigma; A1720], 15 μgml^{-1} Hygromycin B [Sigma; H9773] or 100 μgml^{-1} Zeocin [Invitrogen; R25001]). After another week on selective plates, the cellophanes are transferred to BCDAT plates with no selective agent to allow recovery. After another 1-2 weeks, the surviving *P. patens* colonies were individually transferred to selective BCD plates, to select for stably transformed plants. Following another 1 – 2 weeks on selection, plants that survived both rounds of selection were transferred to non-selective BCD plates and allowed to grow and generate tissue for further study.

2.3 *E. coli* transformation

All cloning was carried out using High Efficiency 10-beta Competent *E. coli* (NEB; C3019). For transformation of the *E. coli*, plasmid and bacterial cells were combined in an eppendorf tube (at a 1:10 ratio for purified plasmid and 4:10 ratio for a ligation reaction mixture). The cells were then put on ice for 15-20 min to acclimate, then heat shocked at 42°C for 1 min and returned to ice for 5 min. In the sterile field close to a Bunsen burner flame, 250 μl of S.O.C medium was added to the tube and the recovered cells were incubated at 37°C, for 1 h, shaking at 200 rpm. In the sterile field close to a Bunsen burner flame, the cells were plated onto LB plates containing the appropriate antibiotic (50 mgml^{-1} ampicillin or kanamycin) and incubated overnight at 37°C in a static incubator.

2.4 Construct assembly

All primer sequences can be found in the Appendix Table S1.

2.4.1 Generation of pHSP::NOG1-GFP/GUS constructs

These constructs were assembled by Emily Day. The *PpNOG1* coding sequence (without the STOP codon) was amplified from wild type cDNA

with PpNOG1_F_BamHI/ PpNOG1noSTOP_R_BamHI and ligated into pJET1.2 (Thermo Scientific; K1231) by blunt end ligation. After verification of the amplified sequence by Sanger sequencing (Source Bioscience), it was digested out of pJET1.2 with BamHI-HF (NEB; R3136) and ligated into both pHSP-MCS-GUS-108-35SNPT and pHSP-MCS-GFP-108-35SNPT (kindly provided by Younoussse Saidi), both cut with same enzyme. The resulting pHSP-NOG1-GUS-108-35SNPT (Figure S2) and pHSP-NOG1-GFP-108-35SNPT (Figure S1) constructs were linearised with ApaI (NEB; R0114) and transformed into *Ppnog1R* protoplasts. Stable transformants were selected for using 40 μgml^{-1} G418 disulphate (Sigma; A1720).

2.4.2 Generation of the *PpNOG1* deletion construct

This construct and the resultant *P. patens* line were made by Laura Moody. A genomic DNA fragment from 827bp upstream of the start codon, up to but excluding the start codon of the *PpNOG1* sequence, was PCR-amplified from wild type genomic DNA using NOG1.5FKpnI/ NOG1.5RXhoI primers and ligated into pJET1.2 (Thermo Scientific; K1231) by blunt end ligation. After verification of the amplified sequence by Sanger sequencing (Source Bioscience), it was digested out of pJET1.2 with KpnI-HF (NEB; R3142) and XhoI (NEB; R0146) and ligated into pAHG1 (a kind gift from Yasuko Kamisugi and Andrew Cuming), digested with the same enzymes, to create pAHG1-*NOG1*-5'. A *PpNOG1* genomic DNA fragment including the stop codon, up to 1420bp downstream of the *PpNOG1* stop codon was PCR-amplified using NOG1.3FNotI/ NOG1.3RNotI primers and ligated into pJET1.2 (Thermo Scientific; K1231) by blunt end ligation. After verification of the amplified sequence by Sanger sequencing (Source Bioscience), it was digested out of pJET1.2 with NotI-HF (NEB; R3189) and ligated into pAHG1-*NOG1*-5', also cut with NotI-HF, to create p*NOG1*delH (Figure S3). p*NOG1*delH was linearised using KpnI-HF prior to transformation into *P. patens*. Stable transformants were selected using 15 μgml^{-1} Hygromycin B (Sigma; H9773).

2.4.3 Generation of *PpNOG1* rescue constructs

These constructs and the resultant *P. patens* lines were made by Laura Moody. The *PpNOG1* (without the STOP codon) coding sequence was amplified from *P. patens* wild type cDNA with PpNOG1_F_SalI/ PpNOG1noSTOP_R_HindIII. The *PpNOG1* UBA domain alone (without the STOP codon) was also amplified using PpNOG1UBA_F_SalI/ PpNOG1noSTOP_R_HindIII. The *AtNOG1* coding sequence (without the STOP codon) was amplified from Arabidopsis Col-0 cDNA with AtNOG1_F_SalI/ AtNOG1noSTOP_R_HindIII. Each of these inserts was ligated into pJET1.2 (Thermo Scientific; K1231) by blunt end ligation, verified by Sanger sequencing (Source Bioscience). Additionally, versions of the *PpNOG1* coding sequence were synthesised so that key residues were mutated and SalI and HindIII restriction sites were included at the 5' and 3' end respectively. All inserts were digested out of the vectors with SalI-HF (NEB; R3138) and HindIII-HF (NEB; R3104) and into pZAG (kindly provided by Yuji Hiwatashi), digested with the same enzymes (Figures S4, S5, S6, S7 and S8). For *P. patens* transformation, each plasmid was linearised with SacII (NEB; R0157). Stable transformants were selected for with 100 μgml^{-1} Zeocin (Invitrogen; R25001).

2.4.4 Generation of the *PpDEK1* deletion construct

This construct was assembled by Laura Moody. A genomic DNA fragment from 833 bp upstream of the start codon, to 124 bp into the *PpDEK1* open reading frame, was PCR-amplified from *P. patens* wild type genomic DNA using DEK1.5FNotI/ DEK1.5RXbaI and ligated into pJET1.2 (Thermo Scientific; K1231) by blunt end ligation and verified by Sanger sequencing (Source Bioscience). The *PpDEK1* 5' flanking fragment was digested out of pJET1.2 with NotI-HF (NEB; R3189) and XbaI (NEB; R0145) and ligated into pMBL10A (kindly provided by Andrew Cuming), digested with the same enzymes to create pMBL10A-*DEK*-5'. A genomic DNA fragment from 522 bp before the end of the stop codon to 678 bp after the stop codon was amplified using DEK1.3FEcoRI/ DEK1.3REcoRI and ligated into

pJET1.2 (Thermo Scientific; K1231) by blunt end ligation and verified by Sanger sequencing (Source Bioscience). The *PpDEK1* 3' flanking fragment was digested out of pJET1.2 with EcoRI-HF (NEB; R3101) into pMBL10A-*DEK*-5', also digested with EcoRI-HF to create the *DEK1del* construct (Figure S9). The construct was linearised with NotI-HF and transformed into protoplasts isolated from either GdWt or *Ppnog1R*. Stable transformants were selected for with 40 μgml^{-1} G418 disulphate (Sigma; A1720).

2.4.5 Generation of the *PpHEC1a/b*, *PpCCS52A* and *PpMOC1a/b* overexpression constructs

The coding sequences of the following genes were synthesised by TWIST Bioscience with no stop codon, flanked by 5' BamHI and 3' XmaI sites; *PpCCS52A* (*Pp3c3_3400*), *PpMOC1a* (*Pp3c6_16410*), *PpMOC1b* (*Pp3c5_17790*), *PpHEC1a* (*Pp3c10_19020*) and *PpHEC1b* (*Pp3c3_28970*). The coding sequences were isolated from the commercial vector using BamHI-HF (NEB; R3136) and XmaI (NEB; R0180) and ligated into pUBI-MCS-GFP-108-35SNPT (kindly provided by Younouss Saidi) cut with the same restriction enzymes (Figures S10, S11, S12, S13 and S14). For transformation into *P.patens*, the construct was linearized with NotI-HF (NEB; R3189) and transformed into *Ppnog1R* and wild type protoplasts. Stable transformants were selected for with 40 μgml^{-1} G418 disulphate (Sigma; A1720).

2.4.6 Generation of the *PpNOG2* overexpression construct

This construct and the resultant *P. patens* line were made by Laura Moody. The *PpNOG2* coding sequence was amplified from *P. patens* wild type cDNA with PpNOG2_F_Sall/ PpNOG2noSTOP_R_HindIII and ligated into pJET1.2 (Thermo Scientific; K1231) by blunt end ligation and verified by Sanger sequencing (Source Bioscience). The *PpNOG2* coding sequence was then digested out of pJET1.2 with Sall-HF (NEB; R3138) and HindIII-HF and ligated into pZAG (kindly provided

by Yuji Hiwatashi), digested with the same enzymes (Figure S15). For *P. patens* transformation, the plasmid was linearised with KpnI-HF (NEB; R3142) and transformed into wild type protoplasts. Stable transformants were selected for with 100 μgml^{-1} Zeocin (Invitrogen; R25001).

2.4.7 Generation of Yeast two-hybrid (Y2H) constructs

The vectors for Y2H were the pGADT7 AD ‘prey’ vector (Figure S16; Takara Bio; 630442) and the pGBKT7 ‘bait’ vector (Figure S17; Takara Bio; 630443). The coding sequences of the proteins being tested were cloned from *P. patens* wild type cDNA. The primers that were used for each gene are listed in Table S1. The genes were PCR-amplified and ligated into yeast two-hybrid (Y2H) vectors digested with the appropriate restriction enzymes; either XmaI (NEB; R0180) and BamHI-HF (NEB; R3136) or just XmaI. For the cases when only XmaI restriction enzyme was used, the orientation of the inserts in the vectors was ascertained by Sanger sequencing (Source Bioscience) using the forward primer for the gene.

2.4.8 Generation of the *Pp3c8_19720* constructs

The full-length coding sequence of *Pp3c8_19720* without the STOP codon was synthesised by TWIST Bioscience, flanked by SalI and HindIII sites. Two silent mutations were introduced into the coding sequence to disrupt the SalI and HindIII restriction sites within the sequence. The destination vector (pZAG) has only SalI and HindIII sites at the MCS, making it necessary to edit those sites within the *Pp3c8_19720* sequence. The synthesised insert contained within the commercial vector was digested with SalI-HF (NEB; R3138) and HindIII-HF (NEB; R3104) and ligated into the pZAG vector (kindly provided by Yuji Hiwatashi). The final plasmid (Figure S18) was checked by full plasmid sequencing using the Source Bioscience Nanopore 30 service. For *P. patens* transformation, the plasmid was linearised with SacII (NEB; R0157) and transformed into wild type and *snog1a* protoplasts. Stable transformants were selected using 100 μgml^{-1} Zeocin (Invitrogen; R25001).

In addition, the *Pp3c8_19720* coding sequence without the STOP codon

was amplified from wild type cDNA with the primers Pp3c8_19720_F_BamHI/Pp3c8_19720noSTOP_R_XmaI and ligated into pJET1.2 (Thermo Scientific; K1231) by blunt end ligation and verified by Sanger sequencing (Source Bioscience). The *Pp3c8_19720* coding sequence was then digested out of pJET1.2 with BamHI-HF (NEB; R3136) and XmaI (NEB; R0180) and ligated into pUBI-MCS-GFP-108-35SNPT (kindly provided by Younousse Saidi), digested with the same enzymes (Figure S19). For transformation into *P.patens*, the construct was linearised with NotI-HF (NEB; R3189) and transformed into wild type and *snog1a* protoplasts. Stable transformants were selected for with 40 μgml^{-1} G418 disulphate (Sigma; A1720).

A *Pp3c8_19720* deletion construct (Figure S20) was synthesised by TWIST Bioscience. It contained a 5' flanking sequence from 592 bp upstream of the start codon to 515 bp into the ORF, followed by a G418 resistance cassette and a 3' flanking sequence from 142 bp downstream of the stop codon to 1086 bp downstream of the stop codon. The construct was provided in a generic vector backbone, which was linearised with PvuI-HF (NEB; R3150) for transformation into wild type and *nog1dis* protoplasts. Stable transformants were selected using 40 μgml^{-1} G418 disulphate (Sigma; A1720).

2.5 Yeast two-hybrid

AH109 yeast cells were streaked onto YDPA medium plates (2% bactopectone, 1% yeast extract and 2% ash agar was autoclaved, then when cooler 0.003% adenine hemisulphate and 2% D-glucose, that had been separately autoclaved, was added). Cells were streaked onto plates divided into three sections and incubated at 30°C for approximately 48 h.

For the dual co-transformation of the yeast with yeast two-hybrid constructs, 15 μl sheared salmon sperm DNA (10 mgml^{-1}) per transformation was pre-boiled at 100°C for 20 min and then immediately put on ice. For every four transformations to be carried out, 1 ml of sterile water was pipetted into a 1.5 ml eppendorf tube

and into the water was added 1/3 of a plate of yeast (approx 100 μ l yeast) and mixed by flicking. The suspended cells were centrifuged at maximum speed for 30 seconds and the supernatant removed. The pellet was resuspended in 1 ml sterile 100 mM lithium acetate, and mixed thoroughly by flicking (not by pipetting). The cells were then incubated at 30°C for 10 min. Each tube of cell mixture was then divided between four fresh 1.5 ml eppendorf tubes (250 μ l in each), so that there was one tube of yeast suspension per transformation. The tubes were centrifuged again at maximum speed for 30 s and the supernatant removed. To the pellet was added the following, in this order and without mixing: 240 μ l 50% PEG3350 (filter sterilised), 36 μ l 1M lithium acetate, 10 μ l sheared salmon sperm DNA (pre-boiled), 1 μ g bait plasmid, 1 μ g prey plasmid and sterile water up to 350 μ l. The tubes were vortexed for 1 min and then incubated in a 42°C waterbath for 20 min. The tubes were centrifuged at maximum speed for 10 s, the supernatant removed and the cells resuspended in 200 μ l of sterile water. The contents of each tube were then plated onto SD/-Leu/-Trp dropout medium (26.7 gL⁻¹ minimal SD base [Takara Bio; 630417], 0.6 gL⁻¹ DO supplement -Leu/-Trp [Takara Bio; 630411], 1% ash agar, pH 5.8). The transformed yeast was incubated at 30°C for 2-3 days until colonies grew. Colonies that grew on SD/-Leu/-Trp dropout medium had successfully been co-transformed with the pGADT7 AD and pGBKT7 vectors.

To test for protein interaction between genes of interest, co-transformed yeast were plated onto SD/-Ade/-His/-Leu/-Trp dropout medium (26.7 gL⁻¹ minimal SD base [Takara Bio; 630417], 0.6 gL⁻¹ DO supplement -Ade/-His/-Leu/-Trp [Takara Bio; 630428], 1% ash agar, pH 5.8). The yeast were grown in liquid SD/-Leu/-Trp medium (26.7 gL⁻¹ minimal SD base [Takara Bio; 630417], 0.6 gL⁻¹ DO supplement -Leu/-Trp [Takara Bio; 630411], pH 5.8) for two days, shaking at 30°C. Yeast cell density was determined by measuring absorbance at OD600 on a spectrophotometer. The density of the yeast cultures was normalised by addition of liquid SD/-Leu/-Trp medium in appropriate amounts. 50 μ l of each yeast culture was spotted onto SD/-Ade/-His/-Leu/-Trp plates. The yeast was incubated at 30°C for 5 days.

2.6 Phenotypic characterisation

2.6.1 Imaging

To stain cell walls, *P. patens* tissues were submerged in propidium iodide ($10\ \mu\text{gml}^{-1}$) for around 10 min and then mounted on slides in water. Images were acquired using a Leica SP5 scanning confocal microscope with a 20x and 40x water immersive lens or a Zeiss LSM 880 Airy Scan with a 25x and 40x water immersive lens. A 488 nm argon laser was used to excite the propidium iodide and fluorescence was detected at 600-630 nm. Z-stack images of tissue structures were captured and were processed in ImageJ. For detection of Green Fluorescent Protein (GFP) a 488 nm laser was used to excite the GFP and fluorescence was detected at 500-550 nm. Autofluorescence from chlorophyll was detected at 680-700 nm.

2.6.2 Gametophore assays

To assay the number of gametophores produced by different *P. patens* lines, tissues were grown on cellophane-overlaid BCDAT plates for one week and then the tissues were harvested into separate 50 ml Falcon tubes containing 10 ml sterile water. The tissues were each homogenized with a Tissue Homogeniser (Omni; TH-03) for 20 s. Then 1 ml of homogenate was removed to measure optical density (OD₆₀₀) in a spectrophotometer. The homogenate density was then normalized by adding the required amount of water to each tube. 50 μl of homogenate was pipetted onto BCD medium in small petri dishes, with five repeats for each line. These plates were grown for around two months under standard conditions, so that mature gametophores had enough time to develop. The number of gametophores on each plate were then counted. To measure gametophore height, a representative sample from each line was extracted from the tissue intact, placed in a petri dish and photographed. The gametophore heights were measured in ImageJ, by drawing a line from the apex to the base.

2.6.3 Hormone treatments

Homogenised tissue was pipetted onto cellophane-overlaid BCD plates and allowed to grow for one week. For treatments with NAA (Sigma; N0640), NPA (Sigma; 33371) or BAP (Sigma; B3408), the hormone was added to molten BCD medium at the desired concentration for the experiment. The cellophanes with tissue on top, that had been growing on BCD, were transferred to the BCD + hormone plates (or just BCD plates for the controls) and returned to the growth cabinet for the desired length of treatment time (three days for imaging or one week for RNA-seq experiments).

2.7 GUS Staining

Active GUS staining solution was prepared; 90% GUS Staining Buffer (100 mM Sodium phosphate buffer, 10 mM EDTA, pH7), 10% 5 mgml⁻¹ X-Gluc (Thermo Scientific; R0851) dissolved in DMSO (prepared fresh), 0.1% 5 mM K₃Fe(CN)₆. *P. patens* tissue was submerged in the active GUS staining solution and incubated overnight at 37 °C in the dark (plate wrapped in foil). The next day the tissue was cleared by rocking with multiple washes of 100% ethanol until the chlorophyll had been removed and the blue staining could be seen clearly.

2.8 UV mutagenesis

UV mutagenesis was performed by Emily Day. One-week old protonemata from *nog1dis* was digested for 1 h in 1% Driselase (Sigma; D9515) dissolved in 8% D-Mannitol. The resulting cell suspension was passed through a 40 µm nylon cell strainer (Corning; 431750) and centrifuged for 3 min at 120 xg without braking. Protoplasts were subsequently washed twice in 8% D-Mannitol, with repeated centrifugation steps in between washes. Cells were resuspended in 6 ml 8% D-Mannitol, counted using a haemocytometer (Fuchs-Rosenthal), and then plated at a density of 50,000 cells per plate onto cellophane overlaid PRMB. Protoplasts

were immediately exposed to a 75000 mJ dose of UV light using a Stratagene UV Stratalinker 2400 and this was performed with the petri dish lids removed. Following irradiation, the lids were quickly replaced, and the plates were wrapped with micropore tape. Plates were incubated at 24°C in the dark for 24 h, to prevent photoactivatable DNA damage repair, before being transferred to standard growth conditions for a further 2-3 weeks. When visible, regenerating protoplasts were transferred into individual wells of 24 well plates containing BCD. After a few weeks of growth, the phenotype of each individual mutant line was scrutinised and lines displaying the desired phenotype were propagated for further study.

2.9 Plant crossing and phenotyping of segregating progeny

The *P. patens* lines Reute::mCherry and *snog1a* were grown in magenta pots approximately half-filled with BCD medium, under standard growth conditions (16 h light: 8 h dark, 24°C) for approximately 2-3 months. Once both lines appeared to have a high density of mature gametophores the pots were transferred to sporophyte induction conditions (8 h light: 16 h dark, 16°C). After approximately three weeks, gametophores were examined for the presence of gametangia (archegonia and antheridia), using a light microscope. Once these were identified, the plants were ready for crossing. The male parent (Reute::mCherry) was flooded with sterile water and the pot was agitated to dislodge sperm. The sperm suspension was then transferred to a pot containing *snog1a* and left for several hours or overnight. The sperm suspension was then poured off and the magenta pot returned to sporophyte induction conditions in the growth cabinet. This was repeated weekly. When a sporophyte developed on *snog1a* it was removed with forceps by grasping the seta (stalk), taking care not to burst the sporangium (capsule) containing the spores. The sporophyte was examined under a Leica M165 fluorescence stereomicroscope to verify that it was expressing mCherry. This confirmed that the sporophyte was the result of an outcrossing between *snog1a* and Reute::mCherry, rather than self

fertilisation. The sporophyte was sterilised in 70% ethanol for four min and then washed three times with sterile water. The sporophyte was then kept at 4°C in the dark for one week to promote spore germination. In a laminar flow hood, the sporophyte was transferred to a 15 ml Falcon tube. 10 ml of sterile water was added and the sporangium was ruptured by applying pressure with a pipette tip. The spore suspension was then plated onto ten individual cellophane-overlaid BCDAT plates and placed in standard growth conditions to germinate. Once spores had germinated and were clearly visible, sporelings were transferred to individual wells of a 24-well plate containing BCD medium and allowed to grow for several more weeks. The progeny were phenotyped using a stereomicroscope. Based on absence of gametophores (*nog1*-like ‘wild type pool’) or presence of gametophores (*snog1a*-like ‘mutant pool’), the progeny were sorted into two populations and around 70-100 individuals in each population were selected to be used for bulk segregant analysis. The selected individuals were all homogenised and plated onto cellophane-overlaid BCDAT medium and grown for approximately two weeks to generate tissue for genomic DNA extraction.

2.10 Genomic DNA extraction

7 μ l 2-mercaptoethanol and 10 mg ascorbic acid were added to 10 ml aliquots of extraction buffer (100 mM Tris-HCl pH 8.0, 1.42 M NaCl, 2% CTAB 20 mM EDTA, 2% PVP-40). One aliquot is enough for approximately 13 extractions. The extraction buffer aliquots were pre-warmed at 65°C in a waterbath. One plate of tissue was harvested using a pipette tip, dried using a paper towel to remove as much water as possible, placed in a 2 ml eppendorf tube and frozen in liquid nitrogen. The tip of a plastic pestle was dipped in liquid nitrogen and then used to grind the tissue into powder. 500 μ l of pre-warmed extraction buffer was added to the powder and the tissue was ground further until the mixture was homogenised. Another 200 μ l of extraction buffer was added and the tube was incubated at 65°C for 10 min. 600 μ l of chloroform-isoamyl alcohol (24:1) was added and the tube shaken and then

centrifuged at full speed for 10 min. Then the upper aqueous phase of the mixture was transferred to a fresh tube, 0.7x volume of isopropanol was added, the tube shaken and spun again immediately at full speed for 10 min. Then the pellet was washed in 70% ethanol, allowed to air-dry and then resuspended in 30 μ l dH₂O. A NanoDrop[™] spectrophotometer was used to assess the quantity and quality of the extracted DNA.

2.11 Southern Blotting

Digoxigenin (DIG)-labelled probes were generated by PCR amplification using GoTaq (Promega; M7801). Two reactions were set up; one contained unmodified dNTPs and one contained dNTPs in which dTTP had been partially displaced by DIG-dUTP (stock: 10mM A, C, G, 5 mM T, 10 mM DIG-dUTP (Roche; 11093088910). The primers used were NOG1 3F SphI/ NOG1 3R NotI (Table S1). Both reactions were run on a gel side by side to confirm that DIG-dUTP had been incorporated into the probe, which is evident if there is an increase in size of the PCR product from the reaction containing DIG-dUTP. The DIG-incorporated product was excised from the gel and purified using a QIAquick Gel Extraction Kit (Qiagen; 28706), according to manufacturer instructions.

At least 5 μ g genomic DNA was digested overnight with BamHI-HF (NEB; R3136). The digested DNA was run alongside 5 μ l DNA Molecular-Weight Marker II, DIG-labelled (Roche; 11218590910) on a 1% agarose gel containing SYBR safe (0.1 μ lml⁻¹. Invitrogen; S33102) for 4 h at 50 V. The tank, gel tray and comb were decontaminated using 5% Decon 90 (Fisher Scientific; 10335650), to remove any traces of ethidium bromide. The gel was placed in a dish with the wells facing down and rocked gently in Depurination solution (0.25 M HCl) for 15 min. Then it was rocked in Denaturation solution (0.5 M NaOH, 1.5 M NaCl) for 30 min and then in Neutralising solution (0.5 M Tris pH 7.5, 1.5 M NaCl) for 20 min, twice.

A capillary transfer stack was set up to transfer the DNA from the gel to a Hybond[®]-N+ hybridization membrane (Amersham). A large glass baking dish was

filled with 10X SSC (20X SSC: 3 M NaCl, 300 mM sodium citrate). A smaller upturned dish was placed in the larger dish and a glass plate was placed on top. A strip of Whatman paper was thoroughly wetted with the 10X SSC solution, then placed over the glass plate with both ends in contact with the 10X SSC. The gel was placed wells down on the Whatman paper on top of the glass plate and any air bubbles were removed with a roller. The Hybond[®]-N+ hybridization membrane was cut to size, dipped in 2X SSC and placed on top of the gel, with all air bubbles removed. Two squares of Whatmann paper were cut to the same size as the gel and membrane, dipped in 2X SSC and placed on top of the membrane. Strips of parafilm were placed all around the gel so that capillary flow only occurred through the membrane. A stack of paper towels was placed on top (about half a packet), weighted down and left overnight. The next day the blot was disassembled, the membrane was dipped in 2X SSC and then UV-crosslinked on both sides using a Stratagene UV Stratalinker 2400 (auto crosslink setting).

The membrane was placed in sealable container, DNA-side up, covered with DIG Easy Hyb (Roche; 11603558001), and incubated at 43°C for at least 1 h, shaking gently. In a tube, 4 μ l 10 mgml⁻¹ sheared salmon sperm was combined with 10 - 30 μ l of the probe and made up to 100 μ l with sterile water. The tube was heat shocked at 100 °C for 5 min and then placed on ice for 5 min. The probe mixture was added to 30 ml DIG Easy Hyb in a Falcon tube and mixed well. After incubation, the DIG Easy Hyb was poured off the membrane and replaced with the DIG Easy Hyb plus probe mixture. The membrane was incubated at 52°C overnight (less than 18 h).

The next day the hybridisation solution was replaced with low stringency wash (2X SSC, 0.1% SDS) and rocked for 5 min at room temperature, twice. The low stringency wash was replaced with high stringency wash (0.5 X SSC, 0.1% SDS; pre-heated to 55 °C) and the membrane was rocked for 30 min at 55°C, twice.

The membrane was placed in a glass dish, DNA-side up, covered in 1x Maleic acid buffer (10X Maleic acid buffer: 1 M maleic acid, 1.5 M NaCl) and rocked at room temperature for 3 min. The 1X Maleic acid buffer was then replaced with Blocking buffer (1X Maleic Acid buffer, 1% w/v skimmed milk powder) and rocked

at room temperature for at least 2 h. The antibody used was Anti-Digoxigenin-AP Fab Fragments (Roche; 11093274910). The tube of antibody was centrifuged at maximum speed for 5 min. 5 μ l of antibody was added to 50 ml Blocking buffer. The Blocking buffer was removed from the membrane, replaced with the Blocking buffer and antibody solution and rocked for 30 min at room temperature.

The antibody solution was replaced with 1X Washing buffer (10X Washing buffer: 97% Maleic acid buffer, 3% Tween 20) and rocked for 15 min at room temperature, three times. The Washing buffer was replaced with 1X Detection buffer (10X Detection buffer: 1M Tris-HCl, 1M NaCl, pH to 9.5 by addition of NaOH) and rocked for 3 min at room temperature. 5 μ l CDP-Star reagent (Tropix) was added to 1 ml Detection buffer and vortexed. The Detection buffer was removed and the membrane placed on a large square petri dish, DNA-side up. The 1 ml of CDP-Star and Detection buffer was pipetted over the membrane and spread evenly over the surface by turning over the membrane several times. The membrane was incubated in the presence of CDP-star for 5 min. The membrane was placed in the other half of the petri dish and imaged in an Invitrogen[™] iBright 1500 Imager, using the Chemiblot setting. Different length exposures were used as needed to visualise bands.

After imaging, the membrane was stripped, so that it could be re-probed in future. In a glass dish the membrane was covered with Stripping buffer (0.2 M NaOH, 1% SDS) and rocked for 15 min at 37 °C, twice. The Stripping buffer was replaced with 1X Maleic acid and rocked for 3 min at room temperature. The membrane was then placed in a plastic wallet and stored at 4 °C.

2.12 Whole Genome Sequencing and bulk segregant analysis

2.12.1 Whole Genome Sequencing by Novogene

Four genomic DNA samples were prepared for sequencing; one each for the parental lines (*nog1dis* and Reute::mCherry) and one for each of the two phenotypically distinct progeny populations (wild type pool and mutant pool). These latter two samples were prepared by pooling 1 μ g genomic DNA from each of the individuals in that population.

Genomic sequencing was performed by Novogene. After sample quality checks had been completed, the genomic DNA was fragmented and a library prepared. Sequencing was carried out with the Illumina NovaSeq PE150 platform; 150 bp paired-end reads were generated at 44X coverage.

2.12.2 Read processing and variant calling

The read processing, variant calling and bulk segregant analysis was carried out by Steven Kelly. Methods for these are as previously described in Moody et al. (2018). Raw reads were subject to quality filtering using Trimmomatic to remove low quality bases, read-pairs and contaminating adaptor sequences prior to read mapping and variant calling (Bolger et al., 2014). Sequences were searched for all common Illumina adaptors and the settings used for read processing by Trimmomatic were LEADING:20 TRAILING:20 SLIDINGWINDOW:5:20 MINLEN:50. Each of the trimmed quality filtered paired-end read libraries were mapped using BWA-MEM to the *P. patens* genome (Version Ppatens_318_v3) obtained from Phytozome V11 (Li, 2013). Duplicated mapped reads were removed, mapped reads were realigned around indels, and variants called according to best practice guidelines using GATK v3.6 (McKenna et al., 2010).

2.12.3 Bulk segregant analysis

Comparison of the sequence data from the *nog1dis* and Reute::mCherry parental strains identified a set of single nucleotide variants that distinguished the two strains. To conduct bulk segregant analysis, reads from both wild type and mutant pools were mapped to the *P. patens* reference genome and the allele frequency of the strain-distinguishing variants in both pools was determined. Strain-distinguishing variants were used for analysis if: 1) the coverage depth in both the wild-type and mutant pools was > 0.5 and < 2 times the mean coverage depth and; 2) the variant quality was > 500 . These allele frequencies were used to identify the chromosomal region linked to the mutant allele. SNPs that were present in only the mutant pool and not in either *nog1dis* or Reute::mCherry were identified as UV-induced mutations. Candidates were identified by locating UV-induced SNPs within gene coding sequences in the chromosome region linked to the mutant allele. Mutations were confirmed by PCR-amplifying the region containing the putative mutation in *snog1a* and *nog1dis* and sequencing the amplicon, to confirm presence of the mutation in *snog1a* and absence in *nog1dis*. Sequencing of the PCR product was done by Source Bioscience.

2.13 Phylogenetics

For each of the species included in the analysis, proteome sequences (primary transcript only) were obtained. The data used were: *Physcomitrium patens* v3.3 (Lang et al., 2018), *Marchantia polymorpha* v3.1 (Bowman et al., 2017), *Selaginella moellendorffii* v1.0 (Banks et al., 2011), *Oryza sativa* v7.0 (Ouyang et al., 2007), *Zea mays* PH207 v1.1 (Hirsch et al., 2016), *Sorghum bicolor* v3.1.1 (McCormick et al., 2018), *Brachypodium distachyon* v3.1 (International Brachypodium Initiative, 2010), *Arabidopsis thaliana* Araport11 (Cheng et al., 2017), *Solanum lycopersicum* ITAG2.4 (Sato et al., 2012) *Medicago truncatula* Mt4.0v1 (Tang et al., 2014), *Populus trichocarpa* v3.1 (Tuskan et al., 2006), *Micromonas pusilla* CCMP1545 v3.0 (Worden et al., 2009), *Ostreococcus lucimarinus* v2.0 (Palenik et al., 2007),

Chara braunii Cbr_1.0 (Nishiyama et al., 2018), *Chlamydomonas reinhardtii* v5.5 (Merchant et al., 2007), *Amborella trichopoda* v1.0 (Amborella Genome Project et al., 2013), *Botryococcus braunii* v2.1 (Browne et al., 2017), *Anthoceros agrestis* Oxford (Li et al., 2020), *Azolla filiculoides* v1.1 (Li et al., 2018), *Brassica rapa* FPsc v1.3 (DOE-JGI, <http://phytozome.jgi.doe.gov/>), *Ceratodon purpureus* R40 v1.1 (Carey et al., 2021), *Ceratopteris richardii* v2.1 (Marchant et al., 2022), *Chlorokybus atmophyticus* CCAC 0220 v1.1 (Wang et al., 2020), *Klebsormidium nitens* NIES-2285 v1.1 (Hori et al., 2014), *Mesostigma viride* NIES-296 (Liang et al., 2020), *Spirogloea muscicola* CCAC 0214 (Cheng et al., 2019), *Mesotaenium endlicherianum* SAG 12.97 (Cheng et al., 2019), *Coccomyxa subellipsoidea* C-169 v2.0 (Blanc et al., 2012), *Dunaliella salina* v1.0 (Polle et al., 2017), *Volvox carteri* v2.1 (Prochnik et al., 2010), *Porphyra umbilicalis* v1.5 (Brawley et al., 2017), *Sphagnum fallax* v1.1 (Healey et al., 2023), *Diphasiastrum complanatum* v3.1 (DOE-JGI, <http://phytozome-next.jgi.doe.gov/>), *Ginkgo biloba* v2021 (Liu et al., 2021), *Glycine max* Wm82 ISU-01 v2.1 (Schmutz et al., 2010), *Gossypium raimondii* v2.1 (Paterson et al., 2012), *Musa acuminata* v1 (D’hont et al., 2012), *Panicum hallii* v3.2 (Lovell et al., 2018), *Salvinia cucullata* v1.2 (Li et al., 2018), *Spirodela polyrhiza* v2 (Wang et al., 2014), *Thuja plicata* v3.1 (Shalev et al., 2022), *Vitis vinifera* v2.1 (Jaillon et al., 2007), *Saccharomyces cerevisiae* R64-1-1 (Mewes et al., 1997; Liachko et al., 2013), *Drosophila melanogaster* BDGP6.32 (Adams et al., 2000) and *Homo sapiens* GRCh38 (Hawkins et al., 2001). The proteomes were BLASTed against each other using DIAMOND and OrthoFinder was used to identify orthogroups (Buchfink et al., 2015; Emms and Kelly, 2019). Once the orthogroup of interest was identified from the OrthoFinder output, protein sequences were aligned using MAFFT (L-IN-SI method; Katoh and Standley, 2013). A tree was constructed using IQ-TREE, with automatic model selection (ModelFinder) and ultrafast bootstrapping (UFBoot) with 1000 replicates (Nguyen et al., 2015; Kalyaanamoorthy et al., 2017; Hoang et al., 2018). Trees were rooted and edited in Interactive Tree of Life (iTOL; Letunic and Bork, 2021).

2.14 RNA extraction

P. patens tissues were grown on BCD medium overlaid with cellophanes for two weeks. For experiments where samples were treated with BAP, the tissue was initially grown on BCD for one week and then transferred to plates containing BCD + 1 μ M BAP for another week. Control samples that were not treated with BAP were transferred to a fresh BCD plate containing no hormone at the same stage. For each line and/or treatment there were three biological replicates. Tissues were harvested and frozen in liquid nitrogen, ground with a pestle and the RNA was extracted using a RNeasy[®] Plant Mini Kit (Qiagen; 74905), according to manufacturer instructions. The RNA was treated with DNase from the TURBO DNA-free[™] Kit (Invitrogen; AM1907), according to manufacturer instructions. A NanoDrop[™] spectrophotometer was used to assess the quantity and quality of the extracted RNA. If the RNA samples were to be sent for sequencing, no further steps were required.

2.15 cDNA synthesis

If RNA samples were to be used for quantitative or non-quantitative RT-PCR, then the RNA was converted to cDNA. First strand cDNA synthesis was achieved with SuperScript[™] III Reverse Transcriptase (Invitrogen[™]) with Oligo (dT)₂₀ and 10 mM dNTPs, following the manufacturer's instructions. To check that there was no genomic DNA present in the RNA extracts after treatment with the DNase, an additional minus reverse transcriptase reaction was carried out for each sample, where the SuperScript[™] III Reverse Transcriptase (Invitrogen[™]) was substituted with water. PCR reactions were set up with the minus reverse transcriptase controls and PptubF/ PptubR primers (Table S1). The reactions were then run on an agarose gel with positive controls, to ascertain that there was no PCR product and therefore no genomic DNA present in the minus reverse transcriptase controls.

2.16 Quantitative RT-PCR

Quantitative RT-PCR (RT-qPCR) was used to quantify relative transcript levels. There were three biological replicates for each *P. patens* line and three technical replicates for each cDNA sample and water controls. On a 96-well PCR plate, each well contained 2 μ l of cDNA template, with 6 μ l SYBR[®] Green PCR Master Mix (Applied Biosystems[™]), 0.5 μ l each of the forward and reverse qPCR primers for the target or control genes (Table S1) and 1 μ l water. A StepOnePlus[™] Real-Time PCR System and the accompanying StepOne[™] Software (Applied Biosystems[™]), with a Relative Standard Curve quantitation was used to carry out the RT-qPCR. The cycling conditions were 5 min at 95°C, followed by 40 cycles of 15 s at 95°C and 1 min at 60°C. The raw amplification data was processed with Real-time PCR Miner (Zhao and Fernald, 2005), which calculated Cycle Threshold (CT) values. From the CT values, the expression levels of genes of interest were calculated relative to the control gene using the formula $2^{-\Delta CT}$ as in Moody et al., 2018. An ANOVA test was used to determine if there was statistically significant variation in relative transcript levels between *P. patens* lines.

2.17 Transcriptome analysis

2.17.1 RNA-seq by Novogene

Novogene performed quality control checks on the RNA samples. Once QC was passed, an mRNA library was prepared with poly-A enrichment. Sequencing was carried out with the Illumina NovaSeq PE150 platform and 150 bp paired-end reads were generated.

2.17.2 Read alignment, quantification and DEG analysis

Fastp v0.20.1 was used for quality filtering of the raw reads, adaptor trimming, base correction and over-representation analysis (Chen et al., 2018). Filtered reads

were mapped to the reference transcriptome (Phypa_V3), obtained from Ensembl Plants (Lang et al., 2018). Mapping and quantification was done using Salmon v1.5.1 (Patro et al., 2017). Differential expression analysis was done in R v4.0.2 (R Core Team, 2021). The packages biomaRt v2.46.3 and GenomicFeatures were used to create the TxDb (Durinck et al., 2009; Lawrence et al., 2013). A Tx2gene data frame was created using the TxDb. The Tx2gene dataframe was converted to a matrix using tximport (Soneson et al., 2015). Differentially Expressed Gene (DEGs) analysis was done using DeSeq2 v1.30.1 (Love et al., 2014). LFC was shrunk (normal shrinkage type) to adjust for lowly expressed genes. Statistical comparisons were made using the Wald test with an FDR <0.05. Gene set enrichment analysis was done with clusterProfiler 4.10.0 (Wu et al., 2021). Example code can be found in the Appendix.

2.18 Protein extraction

Protein Extraction buffer (PEB; 100 mM Tris pH6.8, 10% (v/v) glycerol, 0.5% (w/v) SDS, 0.1% (v/v) Triton X-100, 5 mM EDTA, 10 mM DTT, 0.5% (v/v) protease inhibitors (Sigma; P9599) was prepared fresh. Tissue was harvested into a 2 ml eppendorf tube and frozen in liquid nitrogen. On ice, and with a cooled pestle, the tissue was ground to a powder. 200-400 μ l cold PEB was added and the tissue was homogenised further with the pestle. Samples were then centrifuged at full speed at 4°C. The liquid supernatant was decanted into a fresh tube and the protein extract was stored at -80°C until use. Protein concentration was measured by Bradford assay, with BSA protein standard (Sigma; P0834) used to make the standard curve. Protein extracts were diluted 20x in PBS and 100 μ l of protein dilution was added to 3 ml Bradford reagent (Sigma; B6196). Absorbance of the protein/Bradford mixture was measured on a Jenway Genova spectrophotometer at 595 nm, using semi-micro cuvettes.

2.19 SDS-PAGE

5X protein loading buffer (12.5% (v/v) Tris-HCl pH6.8, 2% (w/v) SDS, 10% (v/v) glycerol, 5% (v/v) β -mercaptoethanol, 0.001% (w/v) Bromophenol blue) was added at the right proportion to the protein samples and they were boiled at 100°C for 10 min. The denatured protein was loaded alongside PageRuler Plus pre-stained protein marker (Thermo Scientific; 26619) into a 10% or 12% mini-PROTEAN TGX precast Tris-glycine gel (Bio-Rad; 4561034) in TGS running buffer (25 mM Tris base, 192 mM glycine, 0.1% SDS). Two identical gels, containing the same samples, were run together. The samples were run at 150 - 250 V until the dye front had reached the bottom of the gel. After the SDS-PAGE, one of the identical gels was stained with InstantBlue[®] Coomassie Protein Stain (Abcam; ab119211) as a loading control to show the total amount of protein loaded in each well.

2.20 Western blotting

The remaining gel was used for protein transfer to a PVDF membrane, using the Trans-Blot Turbo RTA Transfer Kit (Bio-Rad; 170-4272) and the Trans-Blot Turbo Transfer System (Bio-Rad), according to manufacturer instructions. Once the proteins were transferred to the membrane, it was incubated for 30-60 min in TBST + 5% skimmed milk blocking buffer (TBST; 10 mM Tris-HCl pH8, 150 mM NaCl, 0.1% Tween-20). The membrane was then incubated with TBST + 5% skimmed milk containing anti-GFP, rabbit IgG fraction, horseradish peroxidase conjugate (Invitrogen; A10260) at a 1:1000 dilution for at least 1 h at room temperature. If the incubation occurred overnight it was done in the cold room. The antibody solution was then removed and kept at -20 °C for further use. The membrane was washed 3 times for 10 min in TBST and once for 10 min in TBS (10 mM Tris-HCl pH8, 150 mM NaCl). Pierce ECL Western Blotting Substrate (Thermo Scientific; 32109) was applied to the membrane and chemiluminescence was visualised on an ImageQuant LAS 4000. The membrane was stained with Ponceau solution (5% v/v

glacial acetic acid and 0.1% w/v Ponceau S) to show efficient transfer of proteins onto the membrane.

2.21 Heat shock and proteasome inhibitor treatments

Tissues from wild type and *Ppnog1R pHSP::PpNOG1-GFP* were grown on BCD for one week. Then one plate of protonema for each line and each treatment was transferred to 5 ml liquid BCD containing either mock (5 μ l ml⁻¹ DMSO), 5 μ M bortezomib (Sigma; 5.04314) or 50 μ M bortezomib in a 6-well plate. Treatments for each plant line were duplicated in another 6-well plate. After acclimatisation for 1 h, both plates were wrapped in foil. One plate was heat shocked at 37 °C for 1 h, while the other remained in standard conditions. After 1 h the heat shocked plate was returned to standard conditions and the foil was removed from both plates. After a further 4 h, tissue was harvested from each of the wells, blotted dry and frozen in liquid nitrogen. In another experiment, tissues from wild type and WT *pUBI::Pp3c8_19720-GFP* were grown on BCD for a week before being transferred to liquid BCD containing either mock (5 μ lml⁻¹ DMSO) or 50 μ M bortezomib in a 6-well plate. They were returned to standard growth conditions for 5 h, before the tissue was harvested, blotted dry and frozen in liquid nitrogen.

Chapter 3: Genetic manipulation of the *Ppnog1R* mutant allows further characterisation of the function of *PpNOG1*.

3.1 Introduction

PpNOG1 is a positive regulator of the transition to 3D growth and is required for the establishment and maintenance of the tetrahedral apical cell which generates the 3D gametophore. *Ppnog1R* mutants have suppressed gametophore initiation as fewer gametophore buds form in the mutant compared to wild type. Those gametophore buds that do form do not follow the prescribed program of oriented cell divisions that is required to establish the tetrahedral apical cell. Consequently, all buds arrest at an early stage of development. Thus *PpNOG1* promotes the initiation of buds and is required for correctly orientating the vital asymmetric cell divisions within the bud (Moody et al., 2018).

PpNOG1 is an early acting component of the 3D regulatory pathway and has been shown to positively regulate the expression of the *PpAPB* genes, which are considered to be the master regulators of bud initiation (Aoyama et al., 2012; Moody et al., 2018). Several other regulators have been shown to be transcriptionally downstream of *PpNOG1*, including *PpNOG2* and a few *CLAVATA* genes (Whitewoods et al., 2018; Moody et al., 2018, 2021). However, the mechanism by which the PpNOG1 protein influences the expression of downstream genes is unknown. The actual function of the protein remains obscure. Due to the presence of a C-terminal ubiquitin-associated (UBA) domain, it is predicted to interact with

a ubiquitinated target protein (or proteins) and shuttle them to the proteasome, but this is yet to be verified (Moody et al., 2018).

Previous attempts have been made to isolate the protein for further characterisation but it has proved very challenging. In Moody et al. (2018), the *Ppnog1R* mutant was complemented with a pAct::PpNOG1-GFP construct, which used the rice actin promoter to drive high expression of the *PpNOG1* transcript. This resulted in a rescue of the *Ppnog1R* phenotype, but despite the strong promoter, PpNOG1-GFP could not be detected on a western blot (Evans and Moody, unpublished). This evidence indicates that protein levels of PpNOG1 are kept very low, possibly due to a high rate of PpNOG1 protein degradation. This is supported by the fact that AtNOG1 was identified in a high-throughput proteomics screen as a target of ubiquitination (as well as being a putative interactor of other ubiquitinated targets, via its UBA domain). Furthermore, AtNOG1 was shown to be stabilised by treatment with a proteasome inhibitor, indicating that it is indeed degraded by the proteasome (Kim et al., 2013). If the same is true for *PpNOG1*, that may explain the difficulties in visualising the protein.

One of my project objectives was to further characterise the function of the *PpNOG1* gene. I tried alternative overexpression strategies with the aim of isolating the PpNOG1 protein for further analysis. Additionally, I tested candidate protein-protein interactions using yeast two-hybrid. Also, various versions of *PpNOG1* were generated and transformed into *Ppnog1R*, with a range of outcomes that were informative. *PpDEK1* is another important early regulator of the transition to 3D growth. *Ppdek1* mutants initiate more buds than in wild type, but all of them undergo developmental arrest at a very early stage, before establishment of the tetrahedral apical cell. Therefore, *PpDEK1* negatively regulates gametophore initiation but is required for correct gametophore development (Perroud et al., 2014). I was interested in the potential genetic interactions between *PpNOG1* and *PpDEK1* as they antagonise each other at the gametophore initiation stage but both positively regulate the establishment of the tetrahedral apical cell. To investigate this I attempted to generate *Ppdek1* deletions in the *Ppnog1R* background.

3.2 Results

3.2.1 *NOG1*-like sequences are found in the green algae

It was previously reported that *NOG1* was a land plant specific gene (Moody et al., 2018). However, there have been several new algal genomes released in recent years. So, I performed another search using Orthofinder for *PpNOG1* homologues across plant and green algal genomes as well as a brown algae, yeast, *Drosophila* and the human genome. There were no *NOG1*-like sequences detected in the animal, yeast or brown algal genomes. There were *PpNOG1* homologues detected in all of the land plant species included in the analysis (27 species representing all of the major land plant clades). The green algae species I used were seven chlorophytes (*Botryococcus braunii*, *Chlamydomonas reinhardtii*, *Coccomyxa subsellipsoidea*, *Micromonas pusilla*, *Volvox carteri* and *Dunaliella salina*) and six charophytes (*Chara braunii*, *Chlorokybus atmophyticus*, *Klebsormidium nitens*, *Spirogloea muscicola*, *Mesotaenium endlicherianum* and *Mesostigma viride*). There were *NOG1*-like sequences detected in *C. subsellipsoidea*, *M. pusilla*, *C. atmophyticus* and *M. endlicherianum* (Figure 3.1). This means that *NOG1* is not a land plant specific gene as it is found in both chlorophytes and charophytes, although it appears to have been lost in several of the lineages from each of those clades (for an alignment of the algal *NOG1*-like sequences see Figure S21).

3.2.2 PpNOG1-GFP but not PpNOG1-GUS can rescue the *Ppnog1R* phenotype

The *Ppnog1R* mutant can be rescued by a PpNOG1-GFP fusion protein with expression driven either by the rice actin promoter or the endogenous *PpNOG1* promoter (Moody, 2018; Moody, unpublished). This confirms that the PpNOG-GFP fusion protein is functional and is capable of restoring 3D growth. Previous experiments have indicated that the PpNOG1 protein is usually maintained at

very low levels in *P. patens* tissue, regardless of the level of expression of the transcript. When PpNOG1-GFP expression was driven by the overexpressing rice actin promoter, PpNOG1-GFP was not detectable on a western blot with anti-GFP antibodies (Evans and Moody, unpublished). This has limited our ability to study and characterise the function of this protein.

I decided to try a strategy whereby *PpNOG1* expression was driven to high levels using an inducible heat shock promoter from soybean. This heat shock promoter is capable of driving high expression in short periods of time, upon induction with a heat shock at 37°C. This induction can be enhanced by repeated cycles of heat shock treatment, with a period of recovery in between, which can produce extremely high accumulations of protein product (Saidi et al., 2005). Therefore, I hoped to use this inducible system to quickly induce high expression of PpNOG1-GFP, hopefully overwhelming whatever mechanism maintains the low protein levels of PpNOG1, so that I could detect the protein.

Two constructs were generated; *pHSP::PpNOG1-GFP* and *pHSP::PpNOG1-GUS*. Both were transformed into *Ppnog1R*. This successfully produced a *Ppnog1R pHSP:PpNOG1-GUS* and a *Ppnog1R pHSP::PpNOG1-GFP* line. As expected in the *Ppnog1R pHSP::PpNOG1-GFP* line, PpNOG1-GFP rescued the *Ppnog1R* phenotype, confirming that it was a functional PpNOG1 protein (Figure 3.2A). Rescue of the *Ppnog1R* phenotype occurred even when heat shock was not applied, which indicated that the heat shock promoter was leaky. In *Ppnog1R pHSP::PpNOG1-GUS*, despite repeated heat shocks, gametophores never developed (Figure 3.2A). However, GUS staining of this line after heat shock revealed that PpNOG1-GUS was being expressed at high levels (Figure 3.2B). Therefore, the *Ppnog1R* mutation was not complemented by the expression of PpNOG1-GUS, indicating that this fusion protein is non-functional. I hypothesise that this difference is due to the size difference between GFP (27 kDa) and GUS (68 kDa). As GFP or GUS were fused to the C-terminus of PpNOG1, where the UBA domain resides, perhaps the greater size of GUS resulted in inhibition of UBA domain functioning, whereas the smaller GFP did not.

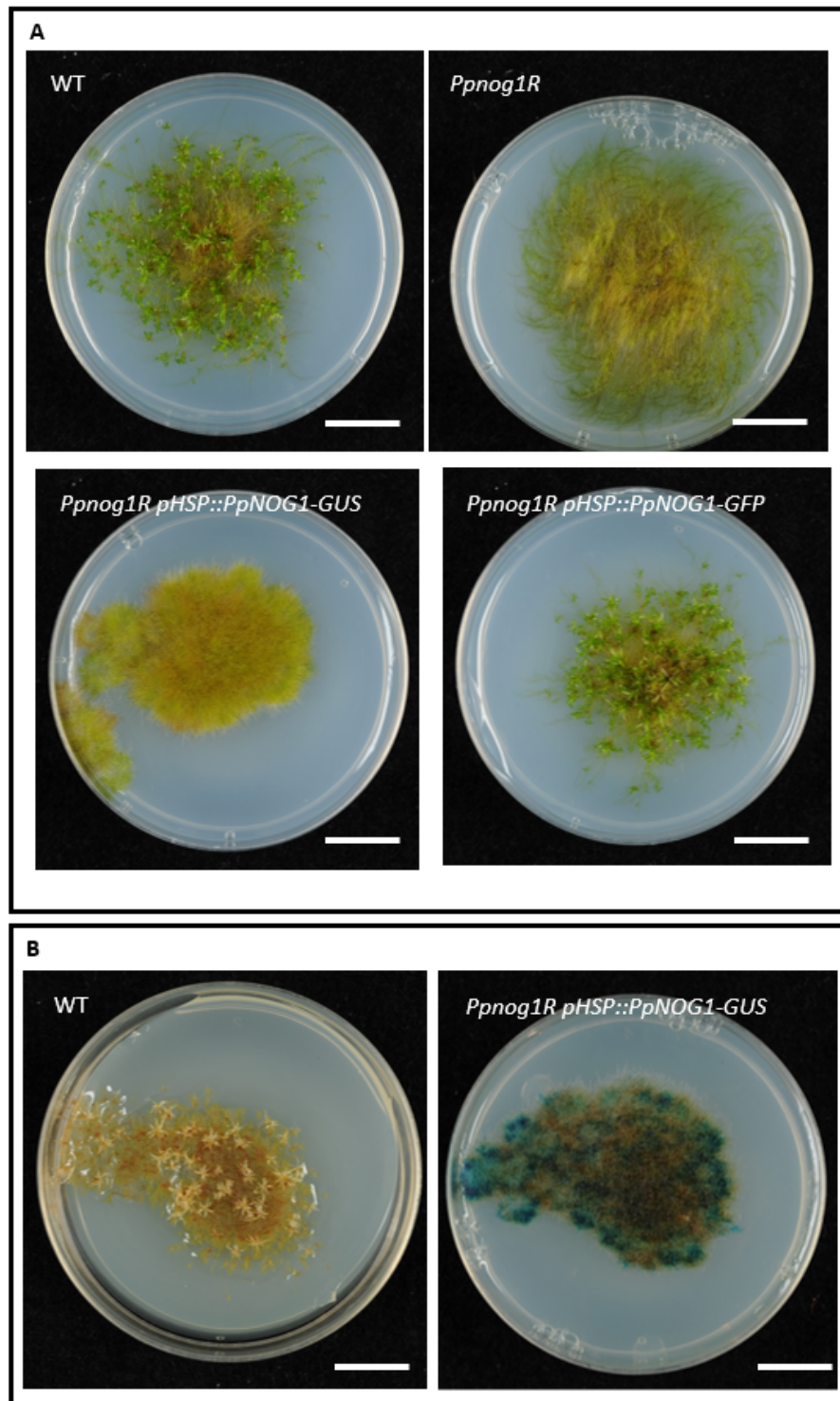


Figure 3.2: A) Representative images of *P. patens*. Scale = 1 cm. These plants did not experience heat shock while growing. B) GUS-stained tissues that have been cleared with ethanol. Both plants were treated with 1 hr heat shock at 37°C prior to GUS staining. Blue colouration indicates the presence of PpNOG1-GUS. Scale = 1 cm

3.2.3 PpNOG1 has nuclear and cytosolic localisation

Despite numerous attempts to get reliable overexpression of PpNOG1-GFP, with either high constitutive or inducible promoters, the protein still remains rather elusive. The majority of the time when these lines (either *pACT::PpNOG1-GFP* or *pHSP::PpNOG1-GFP*) were examined by fluorescence microscopy, there was no convincing signal. On very rare occasions a signal was detected in transiently transformed protoplasts. For example, in Figure 3.3, the *pACT::PpNOG1-GFP* construct was transiently transformed into wild type protoplasts. PpNOG1-GFP signal can be seen clearly in the cytosol and the nucleus. However, repeated efforts to visualise PpNOG1-GFP in stably transformed plant tissues have failed, despite a persistent ‘rescue’ phenotype when *pACT::PpNOG1-GFP* or *pHSP::PpNOG1-GFP* is transformed into the *Ppnog1R* background.

3.2.4 PpNOG1 is degraded by the proteasome

In a preliminary experiment, I extracted protein from *Ppnog1R pHSP::PpNOG1-GFP* tissues that had either not been treated, or had been heat shocked at 37°C. When I did a western blot on these protein extracts, with an anti-GFP antibody, the amount of detectable protein was extremely low in both cases and only visible after exposing the membrane for 45 minutes (data not

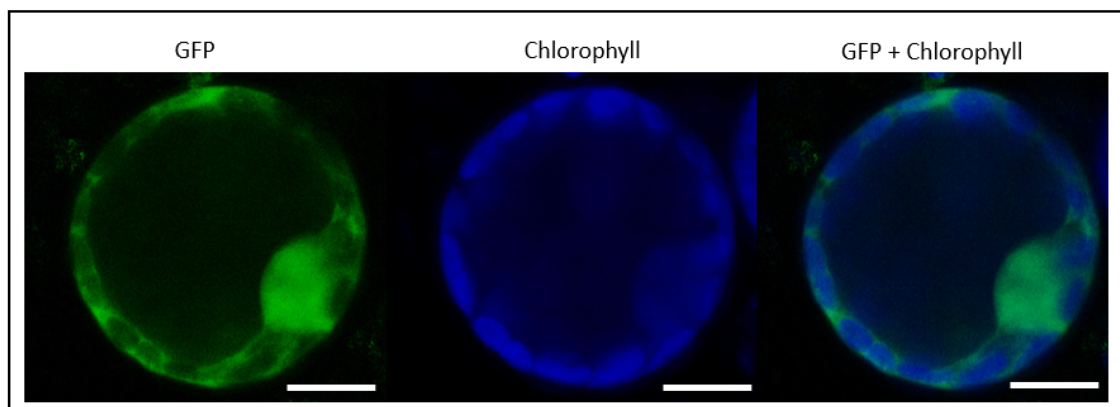


Figure 3.3: Confocal images of wild type protoplasts transiently transformed with *pACT::PpNOG1-GFP*. Green is PpNOG1-GFP. Blue is chlorophyll autofluorescence. Scale = 10 μ m. Transformation and imaging performed by Laura Moody.

shown). Therefore, to test whether PpNOG1 is degraded by the proteasome, I designed another experiment in which *Ppnog1R pHSP::PpNOG1-GFP* (and wild type) were treated with either heat shock or no heat shock and either mock, 5 μ M bortezomib or 50 μ M bortezomib. Bortezomib is a potent proteasome inhibitor.

The results showed that the heat shock treatment didn't noticeably affect the amount of protein detected. This indicates that protein levels for PpNOG1 were maintained at the same level regardless of the level of transcript expression. There was a slight increase in the amount of detectable protein in the 5 μ M bortezomib treatment and then another slight increase in the 50 μ M bortezomib treatment. Therefore PpNOG1 is degraded by the proteasome. However, the size of the protein being detected was not what I expected. Free GFP is 27 kDa, as can be seen in the protein extracts from GdGFP (a fluorescent marker line that expresses high levels of cytosolic free GFP). The molecular weight of PpNOG1 is 23 kDa. Therefore the expected molecular weight of the whole fusion protein is 50 kDa. However there was no protein detected that was that size. Instead there was a band of approximately 28-34 kDa and one of approximately 18-23 kDa. This 18-23 kDa band is smaller than free GFP but must contain the GFP epitope that the anti-GFP antibody recognises. The larger protein band must contain at least a portion of GFP to be detected and also a portion of PpNOG1, but is not the whole fusion protein. It is unlikely that this is due to alternative splicing producing a smaller polypeptide as the current gene model predictions for *PpNOG1* show no alternatively spliced coding sequence exons. Therefore, the evidence suggests that there is another proteolytic process that is acting on PpNOG1-GFP, which is independent of the proteasome. That there was no detectable full-length PpNOG1-GFP suggests that it is cleaved as soon as it is produced. The molecular weight of the C-terminal end of the PpNOG1 protein from the start of the UBA domain is around 7 kDa. Therefore the predicted molecular weight of the UBA domain fused to GFP is around 34 kDa. Assuming that the entire GFP protein is present, the molecular weight of the 28-34 kDa fragment is consistent with cleavage of the protein around or after the start of the UBA domain, resulting in a peptide that consists of some or all of the UBA domain attached to

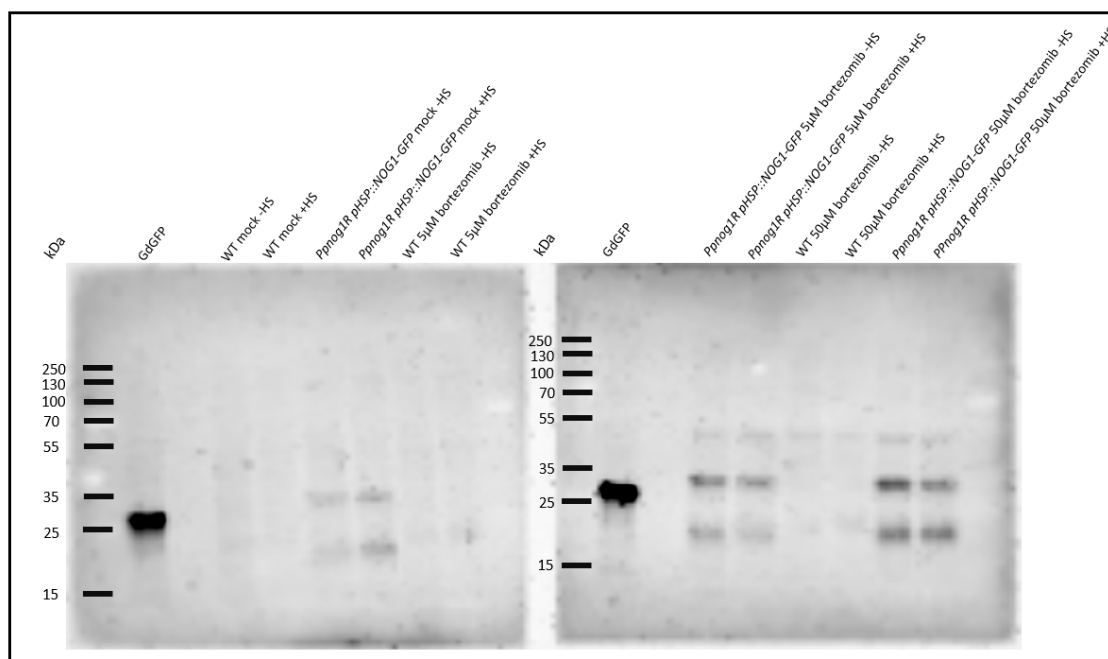


Figure 3.4: Western blot to detect PpNOG1-GFP from the wild type *pHSP::PpNOG1-GFP* line +/- heat shock treatment and with either mock, 5 μ M bortezomib or 50 μ M bortezomib.

GFP.

The evident instability of PpNOG1, both due to proteasomal degradation and apparent cleavage has made working with this protein rather challenging. As it stood I was not ready to attempt methods such as a pull-down to identify direct interaction partners. Therefore, I used alternative approaches to probe the function of the protein; genetics and a yeast two-hybrid assay.

3.2.5 Complementation of *Ppnog1R* with NOG1 variants results in some partial rescues

In the *Ppnog1R* mutant, I ectopically expressed several different variants of NOG1, driven by the moderately strong, constitutive, rice actin promoter. The first *NOG1* variant was simply the wild type *PpNOG1* coding sequence. As expected, *Ppnog1R PpNOG1* displayed a fully restored phenotype and was visually indistinguishable from wild type (Figure 3.5A and C).

Secondly, to evaluate functional conservation of *NOG1* across the land plants, the *AtNOG1* coding sequence was cloned for expression in *Ppnog1R*. The protein

sequences of *PpNOG1* and *AtNOG1* are fairly well conserved (Figure S22). When *AtNOG1* was expressed, the phenotype showed only a mild restoration of 3D growth (Figure 3.5D). Very small gametophores developed, which suggested an ability to establish the tetrahedral apical cell, but failure of those gametophores to develop beyond a very small size may indicate that the apical cell is not maintained.

Several variants of the *PpNOG1* coding sequence in which conserved amino acid residues in the UBA domain had been mutated were generated. Transformation into *Ppnog1R* was successful for two of these variants: *PpNOG1^{M181A}* and *PpNOG1^{L192A}*. In *PpNOG1^{M181A}*, the highly conserved methionine in the ‘MGF’ loop of the UBA domain was converted to an alanine. In *PpNOG1^{L192A}* a conserved leucine residue in the second α -helix of the three α -helix bundle that makes up the structure of the UBA was mutated to an alanine. Interestingly, when *PpNOG1^{M181A}* was expressed in *Ppnog1R*, there was no visible change in the phenotype and *Ppnog1R PpNOG1^{M181A}* was indistinguishable from *Ppnog1R* (Figure 3.5B and E). *Ppnog1R PpNOG1* and *Ppnog1R PpNOG1^{M181A}* are identical except for that mutated methionine residue, yet the former displays wild type-like 3D growth and the latter is still 3D-defective. This means that that single methionine residue in the UBA is essential for the role of *PpNOG1* in promoting 3D growth. Contrastingly, *Ppnog1R PpNOG1^{L192A}* does exhibit a restoration of 3D growth (Figure 3.5F). The gametophores produced in *Ppnog1R PpNOG1^{L192A}* are of a similar scale to wild type gametophores but do have some apparent developmental abnormalities. Therefore, this leucine residue is not absolutely required for establishing or maintaining 3D growth but mutating it does slightly impair the functioning of *PpNOG1* during gametophore development.

Finally, the unmutated UBA domain alone (*PpNOG1^{UBA}*) was also expressed in *Ppnog1R*. When the UBA domain alone was expressed in *Ppnog1R*, it produced a mild restoration of 3D growth, similar to that seen in the *Ppnog1R AtNOG1* line (Figure 3.5G). Therefore the UBA domain seems to be sufficient to establish 3D growth, but not to maintain it.

The *Ppnog1R PpNOG1* line was verified by PCR amplification of the *PpNOG1* coding sequence from *Ppnog1R PpNOG1* genomic DNA. This confirmed the

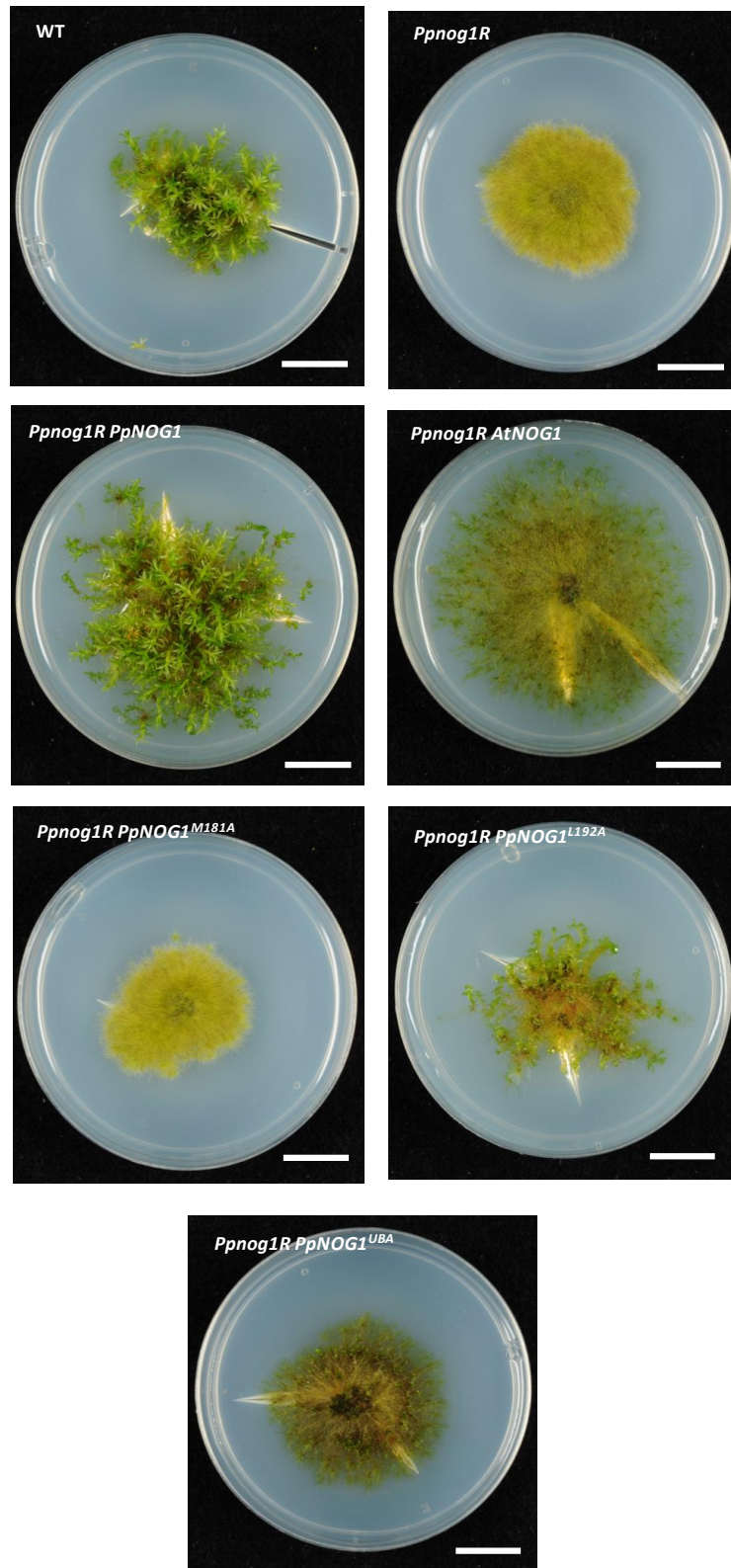


Figure 3.5: Representative images of *P. patens* lines. Plants were all 5 weeks old. Scale = 1 cm. A) Wild type. B) *Ppnog1R*. C) *Ppnog1R PpNOG1*. D) *Ppnog1R AtNOG1*. E) *Ppnog1R PpNOG1^{M181A}*. F) *Ppnog1R PpNOG1^{L192A}*. G) *Ppnog1R PpNOG1^{UBA}*.

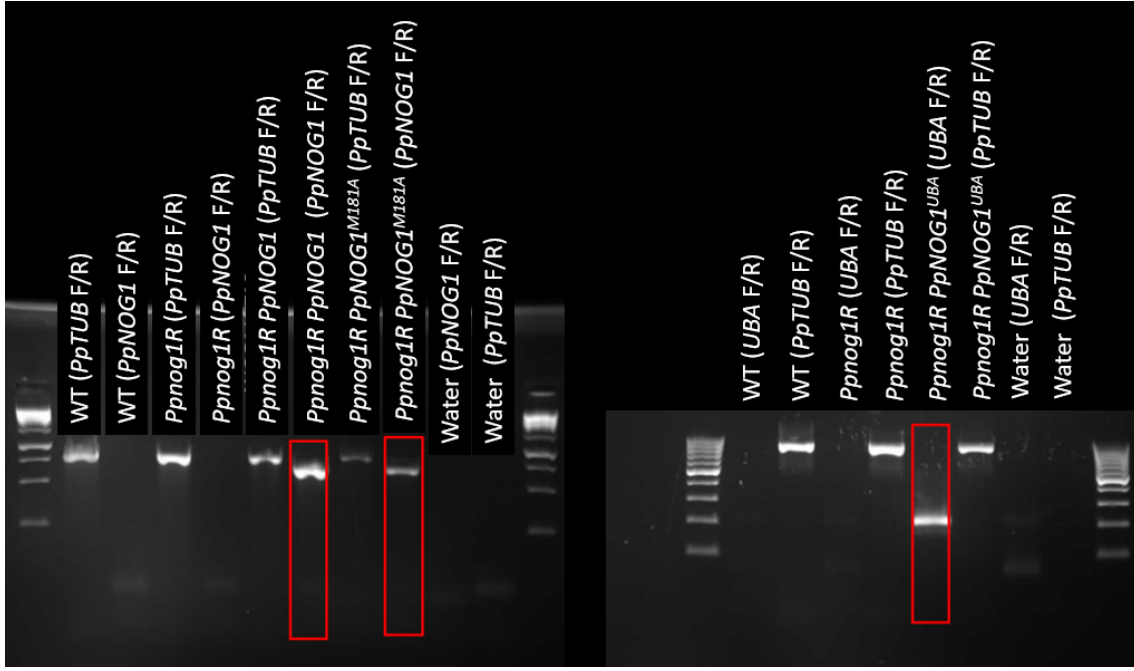


Figure 3.6: Genotyping result Left: *Ppnog1R PpNOG1* and *Ppnog1R PpNOG1^{M181A}* are confirmed by PCR amplification of the *PpNOG1* coding sequence from genomic DNA. Right: *Ppnog1R PpNOG1^{UBA}* was confirmed by PCR amplification of the UBA domain coding sequence from genomic DNA. This PCR genotyping was carried out by Abir Patwary

integration of the overexpression construct. Similarly for *Ppnog1R PpNOG1^{M181A}*, the *PpNOG1^{M181A}* sequence was amplified from genomic DNA using the same primer pair as for the *PpNOG1* coding sequence. For *Ppnog1R PpNOG1^{UBA}*, the coding sequence of the UBA domain was successfully amplified from genomic DNA to verify the line (Figure 3.6). Genotyping has not yet been carried out for *Ppnog1R PpAtNOG1* or *Ppnog1R PpNOG1^{L192A}*, but they both displayed a phenotype that was distinct from the *Ppnog1R* phenotype, so it seems likely that the lines are genuine.

I assayed the number of gametophores produced by each line when the same amount of plant material is allowed to grow for the same amount of time. For this, I only counted gametophores which I could extract whole from the media with forceps and count. Therefore, those structures produced in *Ppnog1R AtNOG1* and *Ppnog1R PpNOG1^{UBA}* were not counted as they were too small to reliably quantify in this way, and therefore were recorded as zero.

There was no significant difference in the number of gametophores produced per

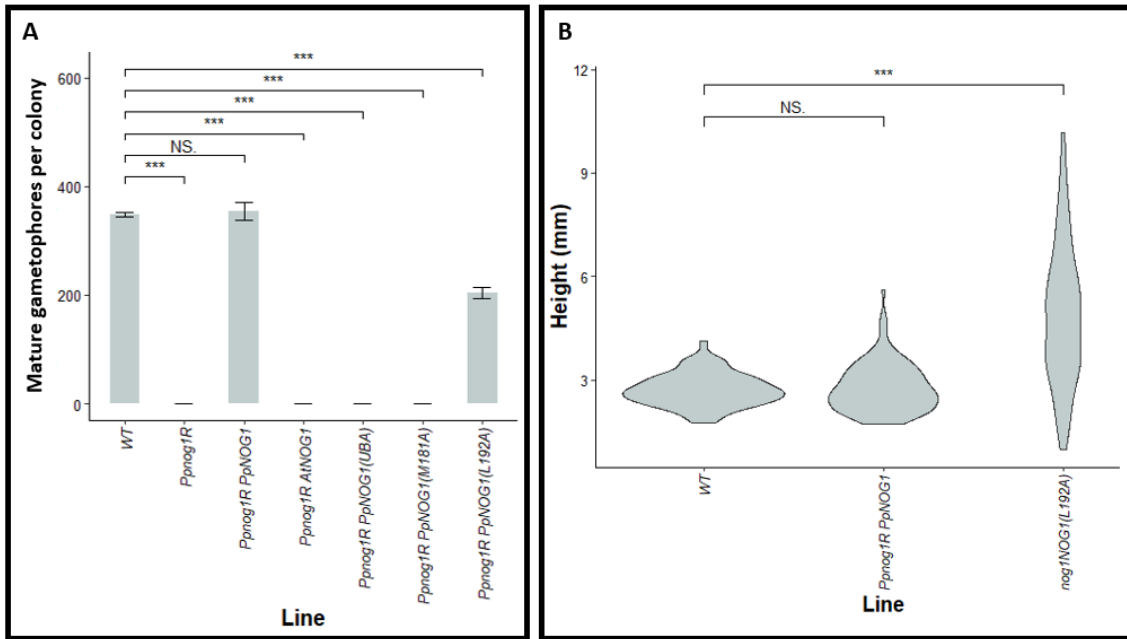


Figure 3.7: A) Number of gametophores per colony ($n = 5$). Standard error is shown for each bar. Horizontal lines indicate result of t-test for each generated line compared to wild type. *** means p -value < 0.001 . B) Violin plot showing the measured heights of gametophores from each line. For wild type $n = 121$. For *Ppnog1R PpNOG1* $n = 125$. For *Ppnog1R PpNOG1^{L192A}* $n = 140$.

colony between wild type and *Ppnog1R PpNOG1* (Figure 3.7A). When I measured the heights of the gametophores from wild type and from *Ppnog1R PpNOG1* there was again no significant difference (Figure 3.7B). Additionally, there was no visual difference in the patterning of gametophores from each line (Figure 3.8A). This once again indicated that *Ppnog1R PpNOG1* was a full rescue of the *Ppnog1R* phenotype.

Ppnog1R PpNOG1^{L192A} produced significantly fewer gametophores per colony than wild type (Figure 3.7A). Therefore, it is possible that the L192A mutation caused a suppression of gametophore initiation, as well as impacting gametophore development. *Ppnog1R PpNOG1^{L192A}* gametophores were, on average, taller than wild type gametophores, although there was a large variance in the heights of *Ppnog1R PpNOG1^{L192A}* gametophores (Figure 3.7B). When I examined the individual gametophores, I noticed they displayed some unusual patterning (Figure 3.8A). In many cases the base of the gametophore had fairly normal patterning but the apex was very elongated, with smaller phyllids than were produced at the base. The transition between the normal sized phyllids at the base to the elongated stem with small phyllids occurred at a distinct point on the stem. Some of the

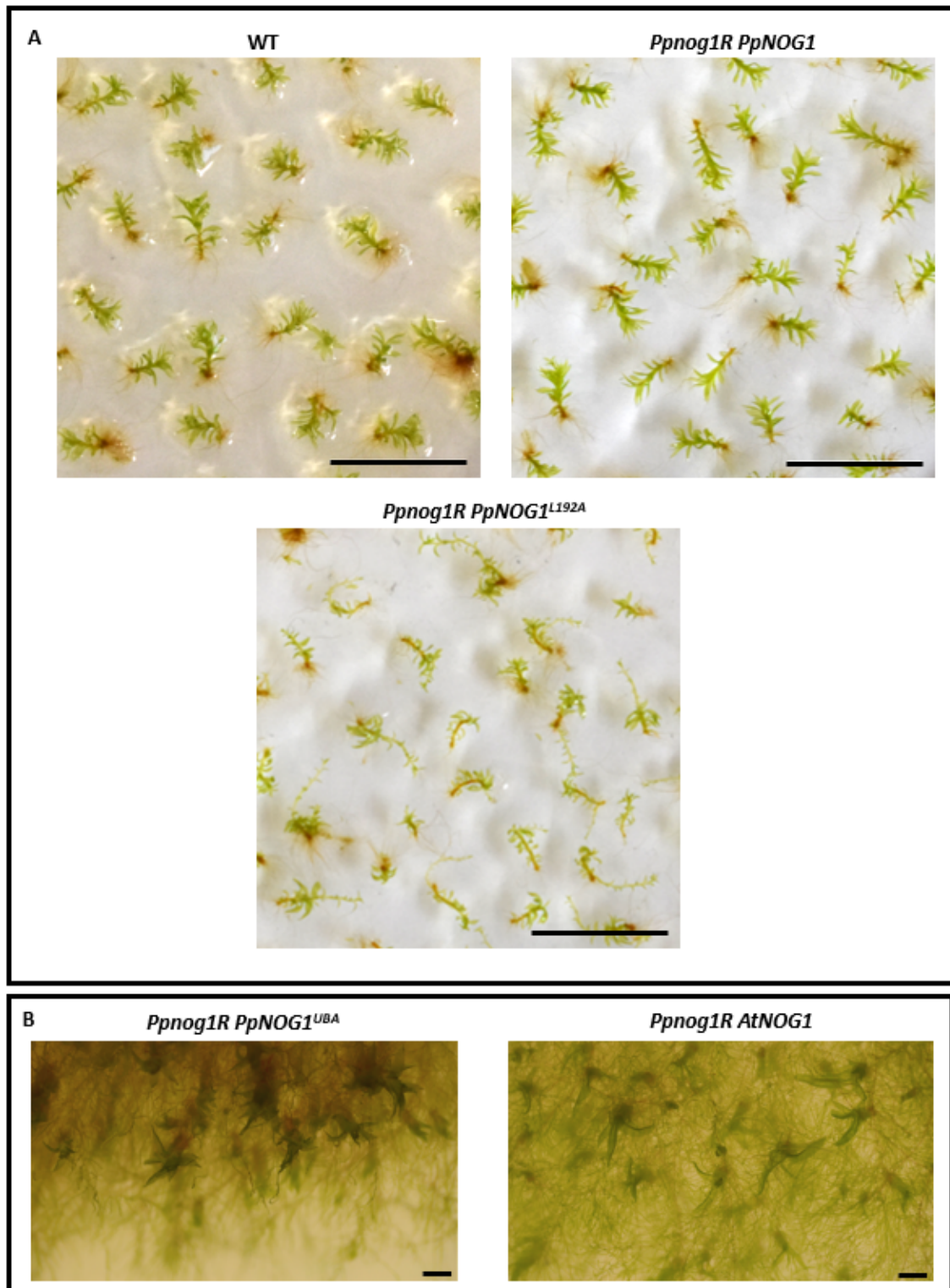


Figure 3.8: A) Photographs showing representative selection of gametophores from wild type, *Ppnog1R PpNOG1* and *Ppnog1R PpNOG1^{L192A}*. Scale = 0.5 cm. B) Light microscope images of *Ppnog1R PpNOG1^{UBA}* and *Ppnog1R AtNOG1*. Scale = 1 mm

gametophores had only the elongated stem with small phyllids for the entire length of the gametophore. This phenotype may indicate a problem with regulating growth

rates in the gametophores of *Ppnog1R PpNOG1^{L192A}*, possibly due to misregulation of the balance between the growth of the stem versus the phyllids.

When viewed under a light microscope, *Ppnog1R AtNOG1* seems to produce gametophores with up to three phyllids before development aborts. In *Ppnog1R PpNOG1^{UBA}* they develop a little further, producing a few more phyllids before also aborting (Figure 3.8B).

In wild type *P. patens*, the first division of the gametophore initial cell was oblique and yielded an apical and a basal cell (Figure 3.9A). Next, the apical and basal cells both divided obliquely at an angle that was roughly perpendicular to the initial oblique division (Figure 3.9B and C). The apical cell then underwent a series of rotating cell divisions (Figure 3.9D shows the first of these divisions), to establish the tetrahedral apical cell (Figure 3.9E). The tetrahedral apical cell divided to maintain itself and generate the phyllid initials (Figure 3.9F). The phyllids were arranged around the gametophore central axis in a spiral phyllotaxy (Figure 3.9G).

In *Ppnog1R*, the divisions in the gametophore bud are consistently defective, which causes very early developmental arrest (Moody et al., 2018). Therefore, I examined the early bud development for each of those lines in which 3D growth was partially restored, to see if this defect had been corrected.

In the *Ppnog1R PpNOG1^{UBA}* line it was evident that there was some aberrant development of the protonema. In many filaments, cells were much shorter than in wild type, with disorganised cell division patterns that were hard to characterise as either straight or oblique. It was consequently difficult to distinguish between cell types (chloronema are characterised by straight division planes between cells whereas caulonema have oblique divisions between cells; Figure 3.9H). Buds in *Ppnog1R PpNOG1^{UBA}* appeared to follow the normal prescribed pattern of cell division (Figure 3.9I and J) and seemed to correctly establish a tetrahedral apical cell that could produce phyllid initials (Figure 3.9K). That led to the development of gametophores that bore a few phyllids, but without fail, development halted at this stage (Figure 3.9L).

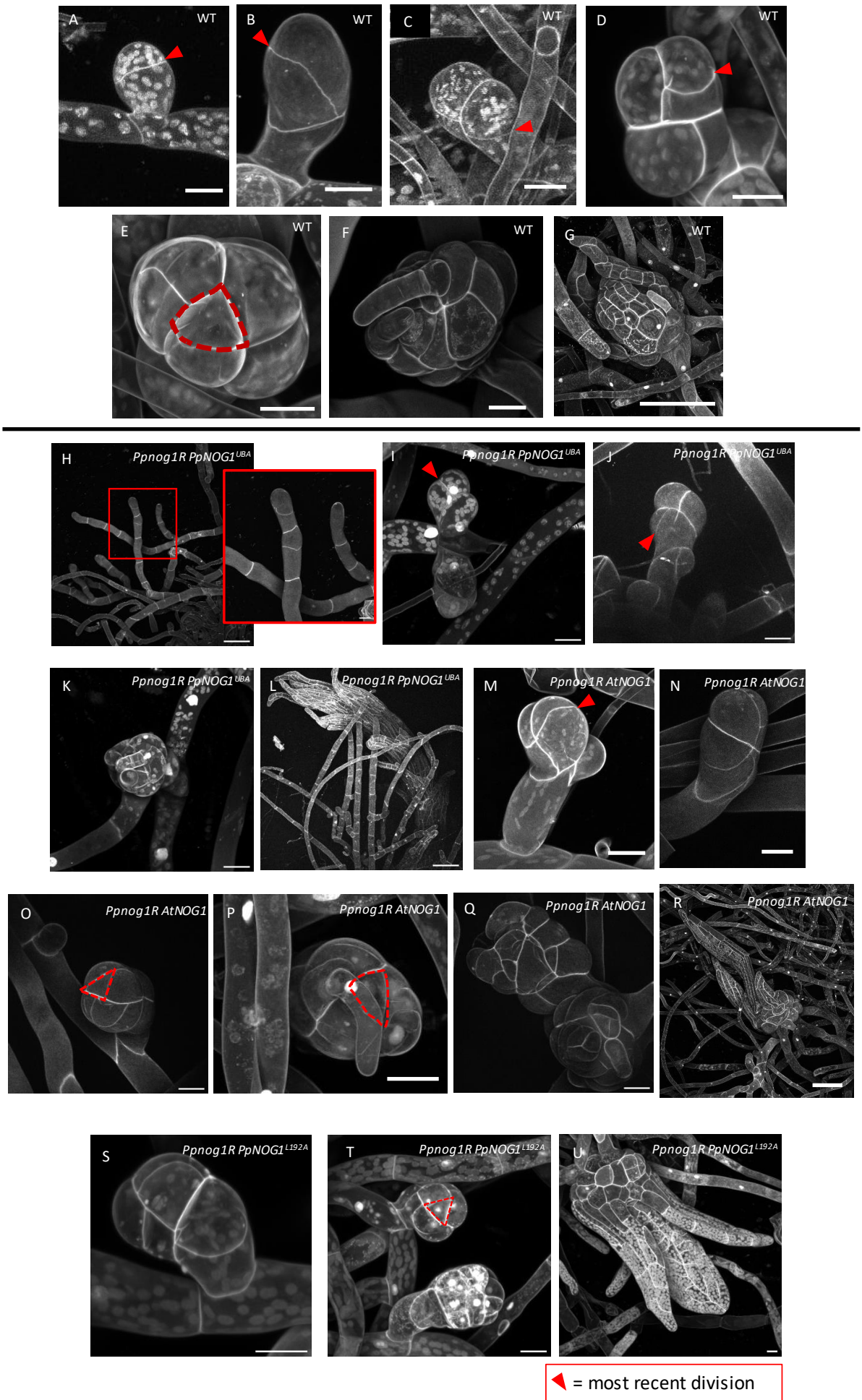


Figure 3.9: Confocal microscopy images of propidium iodide stained tissues. Scale = 20 μm unless otherwise stated. Red arrows indicate site of most recent cell division. Red dotted lines indicate the location of a tetrahedral apical cell. A-G) wild type. A) 2-cell bud. B) 3-cell bud. C) 4-cell bud. D) 5-cell bud. E) Bud with a newly-formed tetrahedral apical cell. F) Bud that has produced two phyllid initials. G) Small gametophore with phyllids forming in spiral phyllotaxy. Scale = 100 μm . H-L) *Ppnog1R PpNOG1^{UBA}*. H) Protonema with a region enhanced to show the aberrant cell division planes. Scale = 100 μm . I) 3-cell bud. J) 6-cell bud. K) Bud that has produced two or three phyllid initials. K) Small gametophore. Scale = 100 μm . M-R) *Ppnog1R AtNOG1*. M) 6-cell bud. N) 7-cell bud. O) Bud with a newly-formed tetrahedral apical cell. P) Bud with a newly-formed tetrahedral apical cell and one phyllid initial. Q) Bifurcated bud. R) Small gametophore. Scale = 100 μm . S-U) *Ppnog1R PpNOG1^{L192A}*. S) 6-cell bud. T) Bud with newly-formed tetrahedral apical cell. U) Small gametophore with phyllids in spiral phyllotaxy.

In *Ppnog1R AtNOG1* I observed some early buds that seemed defective in the patterning of their early divisions. For example in Figure 3.9M and N, after the initial oblique division, there appeared to be a delay in the divisions of the basal cell, whereas the apical cell seemed to be following the normal pattern of divisions. Nevertheless, *Ppnog1R AtNOG1* buds were able to establish a tetrahedral apical cell (Figure 3.9O), that could produce phyllid initials (Figure 3.9P). Occasionally, I observed instances where buds had bifurcated, which was caused by the formation of ectopic apical cells (Figure 3.9Q). *Ppnog1R AtNOG1* gametophores could produce two or three phyllids but at that point, development halted (Figure 3.9R).

In *Ppnog1R PpNOG1^{L192A}*, the first five divisions appeared to follow the normal order and orientation (Figure 3.9S). Thereafter, there was successful establishment of a tetrahedral apical cell (Figure 3.9T) and the formation of phyllids in spiral phyllotaxy (Figure 3.9U). Unlike in *Ppnog1R PpNOG1^{UBA}* and *Ppnog1R AtNOG1*, development of the gametophores in *Ppnog1R PpNOG1^{L192A}* does not halt after the formation of a few phyllids.

In wild type, treatment with high concentrations of the synthetic cytokinin analogue 6-benzylaminopurine (BAP), elicits an overproduction of gametophore initials that become callus-like buds due to disorganised overproliferation of cells. *Ppnog1R* does not form gametophore initials in response to BAP treatment. Therefore *Ppnog1R* is insensitive to BAP (Moody et al., 2018). I found the same to be true when I treated wild type and *Ppnog1R* with 1 μM BAP for three days.

As in wild type, *Ppnog1R PpNOG1* demonstrated strong induction of gametophore bud formation upon BAP treatment. Both *Ppnog1R PpNOG1^{M181A}* and *Ppnog1R PpNOG1^{L192A}* also showed induction of buds with BAP treatment, although the callus-like buds appeared perhaps a little smaller than in wild type, but this was not quantified. Herein was revealed a key difference between *Ppnog1R*, which was insensitive to BAP and *Ppnog1R PpNOG1^{M181A}*, which was not insensitive to BAP. Therefore, while methionine¹⁸¹ is required for the development of gametophore initials into gametophores (as *Ppnog1R PpNOG1^{M181A}* remains 3D-defective), it is not required for the BAP-mediated induction of gametophore initial formation. Interestingly, *Ppnog1R PpNOG1^{UBA}* was insensitive to BAP, like *Ppnog1R*. This indicates that the UBA domain is not involved in the response to BAP (Figure 3.10). Therefore, responsiveness to BAP is contingent on some feature in the uncharacterised region of the PpNOG1 protein, not on the functioning of the UBA domain. The BAP response of *Ppnog1R AtNOG1* has not yet been examined.

3.2.6 PpNOG1 may interact with polyubiquitin.

To test the predicted interaction between PpNOG1 and polyubiquitin, a yeast two-hybrid assay was performed. For this I cloned *PpNOG1* to pair with three *PpUBQ3* genes that I refer to as *PpUBQ3A*, *PpUBQ3B* and *PpUBQ3C*. These three genes represent a subset of the polyubiquitin genes in the *P. patens* genome. I also cloned the *PpNOG1^{M181A}* variant. The inclusion of *PpNOG1^{M181A}* in the assay was to test if any potential interactions with polyubiquitin were abolished with this mutated version of *PpNOG1*. Ectopic expression of *PpNOG1^{M181A}* does not complement the *Ppnog1R* phenotype (Figure 3.5). I wanted to test whether this non-functioning of *PpNOG1^{M181A}* is due to an inability to interact with polyubiquitin. I also cloned *AtNOG1* and *AtUBQ3* for inclusion in this experiment, to assess the evolutionary conservation of any potential interactions. Due to reports of UBA domain dimerisation in other UBA domain containing proteins, I also tested whether PpNOG1 is able to interact with itself or even AtNOG1 (Long et al., 2010).

The yeast two-hybrid assay did suggest a possible weak interaction between

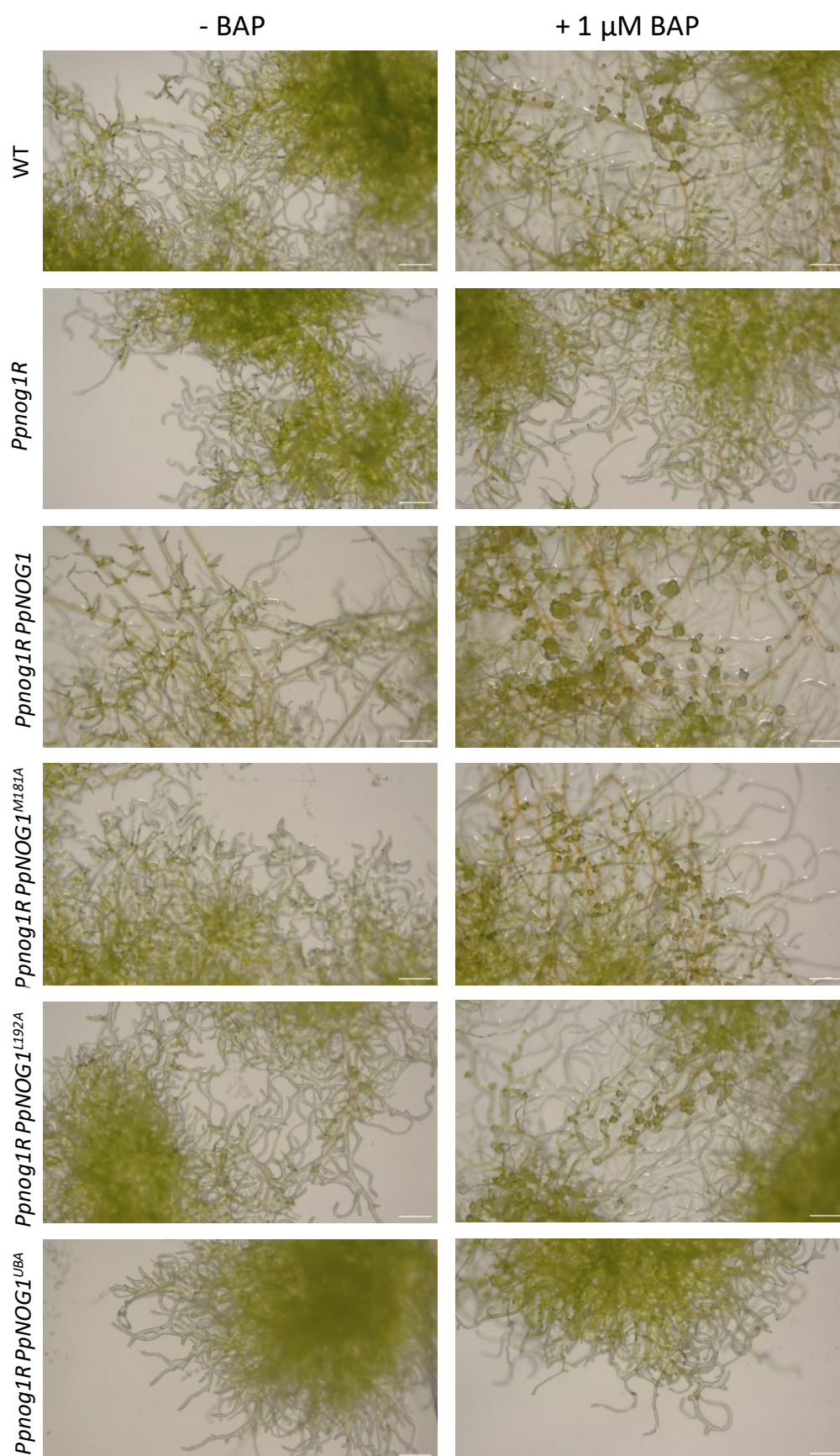
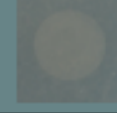
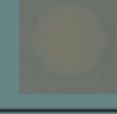
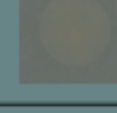
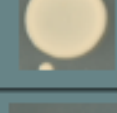
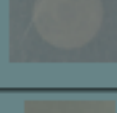
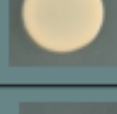
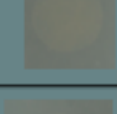
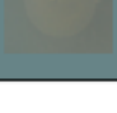


Figure 3.10: Representative images of wild type, *Ppnog1R*, *Ppnog1R PpNOG1*, *Ppnog1R PpNOG1^{M181A}*, *Ppnog1R PpNOG1^{L192A}*, *Ppnog1R PpNOG1^{UBA}* in the presence or absence of 1 μ M BAP. Tissue was grown for one week on BCD medium, before being transferred either to fresh BCD medium (control) or BCD medium containing 1 μ M BAP. After a further three days the tissues were photographed. Scale = 500 μ m.

PpNOG1 and PpUBQ3A and PpUBQ3C, but not PpUBQ3B. PpNOG1^{M181A} showed a possible interaction with PpUBQ3A but not PpUBQ3C. This implies that any possible affinity of PpNOG1 for PpUBQ3A may be unaffected by mutation of the ‘MGF’ loop in the UBA domain. The possible affinity for PpUBQ3C did seem to be decreased by the mutated ‘MGF’ loop. Overall, the evidence for a direct association between PpNOG1 and PpUBQ3 is not entirely convincing due to the apparent weakness of the potential interactions, indicated by poor yeast growth on the quadruple dropout medium. An interaction between PpNOG1 and PpUBQ3 cannot be ruled out but more evidence is required. AtNOG1 and AtUBQ3 did not display an interaction. There were also no interactions between any of the NOG1/NOG1 pairs, hence they do not appear to dimerise (Figure 3.11).

3.2.7 Knockdown of *PpDEK1* expression in *Ppnog1R* produces an overbudding phenotype

I attempted to generate *Ppnog1R/ dek1 del* double disruptant lines by transforming *Ppnog1R* with a *PpDEK1* deletion construct. 16 lines survived selection indicating stable integration of the construct. However, PCR genotyping to confirm the correct targeting of the construct seemed to indicate that none of the 16 lines had a genuine deletion of the *PpDEK1* locus and the construct seemed to have been mistargeted and not integrated correctly at the desired locus (data not shown). However, a few of the lines displaying phenotypes that were unlike *Ppnog1R*. Therefore, for two of these lines, I did semi-quantitative RT-PCR on extracted RNA and amplified different regions of the *PpDEK1* transcript (Figure 3.12A). The results indicated

pGADT7 AD (prey)	pGBKT7 (bait)	SD- Leu/Trp	SD- Ade/His/ Leu/Trp
empty	empty		
<i>AtNOG1</i>	empty		
<i>PpNOG1</i>	empty		
<i>PpNOG1</i> ^{M181A}	empty		
empty	<i>AtUBQ3</i>		
empty	<i>PpUBQ3A</i>		
empty	<i>PpUBQ3B</i>		
empty	<i>PpUBQ3C</i>		
empty	<i>PpNOG1</i>		
<i>AtNOG1</i>	<i>AtUBQ3</i>		
<i>AtNOG1</i>	<i>PpUBQ3A</i>		
<i>AtNOG1</i>	<i>PpUBQ3B</i>		

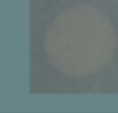
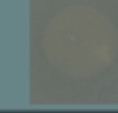
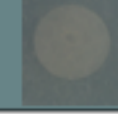
<i>AtNOG1</i>	<i>PpUBQ3C</i>		
<i>AtNOG1</i>	<i>PpNOG1</i>		
<i>PpNOG1</i>	<i>AtUBQ3</i>		
<i>PpNOG1</i>	<i>PpUBQ3A</i>		
<i>PpNOG1</i>	<i>PpUBQ3B</i>		
<i>PpNOG1</i>	<i>PpUBQ3C</i>		
<i>PpNOG1</i>	<i>PpNOG1</i>		
<i>PpNOG1^{M181A}</i>	<i>AtUBQ3</i>		
<i>PpNOG1^{M181A}</i>	<i>PpUBQ3A</i>		
<i>PpNOG1^{M181A}</i>	<i>PpUBQ3B</i>		
<i>PpNOG1^{M181A}</i>	<i>PpUBQ3C</i>		
<i>PpNOG1^{M181A}</i>	<i>PpNOG1</i>		

Figure 3.11: Yeast two-hybrid assay results. Protein interactions are indicated by growth of the cotransformed yeast on SD-Ade/His/Leu/Trp.

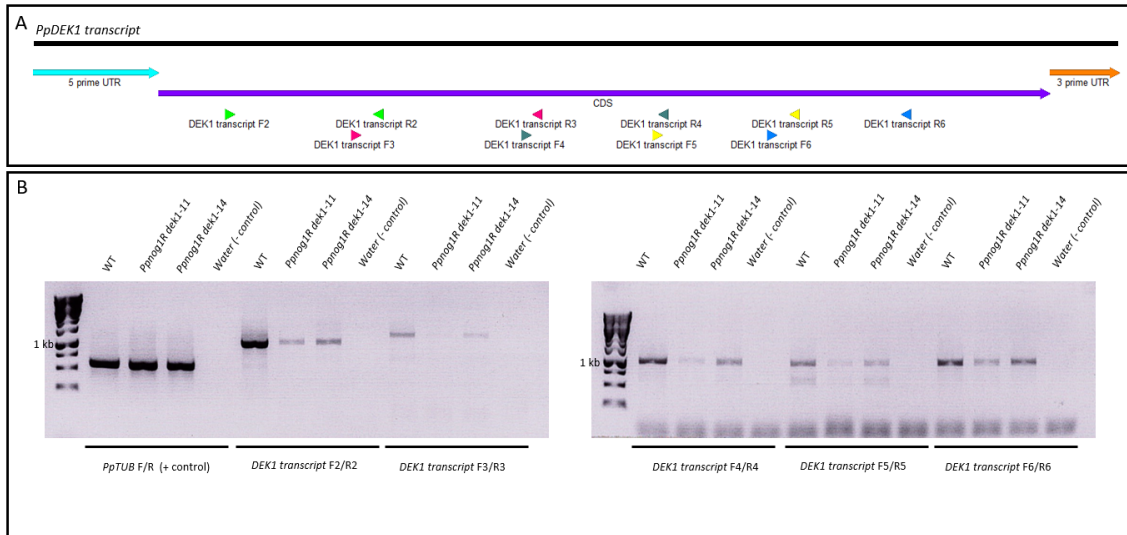


Figure 3.12: A) Schematic showing amplified regions of the *PpDEK1* transcript. B) semi-quantitative RT-PCR result for wild type and two *Ppnog1R dek1* lines.

that *PpDEK1* expression was suppressed in these lines (Figure 3.12B). It seems that I had generated *PpDEK1* knockdown mutants rather than *PpDEK1* deletion mutants. Therefore, these two lines are henceforth referred to as *Ppnog1R dek1 kd-11* and *Ppnog1R dek1 kd-14*.

At the macroscopic level, *Ppnog1R dek1 kd-11* and *Ppnog1R dek1 kd-14* resembled *Ppnog1R* (Figure 3.13). However, a preliminary examination of *Ppnog1R dek1 kd-11* and *Ppnog1R dek1 kd-14* under a light microscope revealed deviations from the *Ppnog1R* phenotype. *Ppnog1R* initiates fewer buds than wild type, and all buds abort after a few disorganised cell divisions (Moody et al., 2018). However, in *Ppnog1R dek1 kd-11* and *Ppnog1R dek1 kd-14* I could see more buds forming than in *Ppnog1R*. In fact *Ppnog1R dek1 kd-11* and *Ppnog1R dek1 kd-14* seemed to be exhibiting an excessive budding phenotype. In addition they formed unusual stack-like structures that had little resemblance to a gametophore but produced some very small phyllid-like outgrowths at the edges (Figure 3.14). However the phenotype of these lines requires much more detailed analysis to fully characterise the effect of *PpDEK1* suppression in *Ppnog1R*.

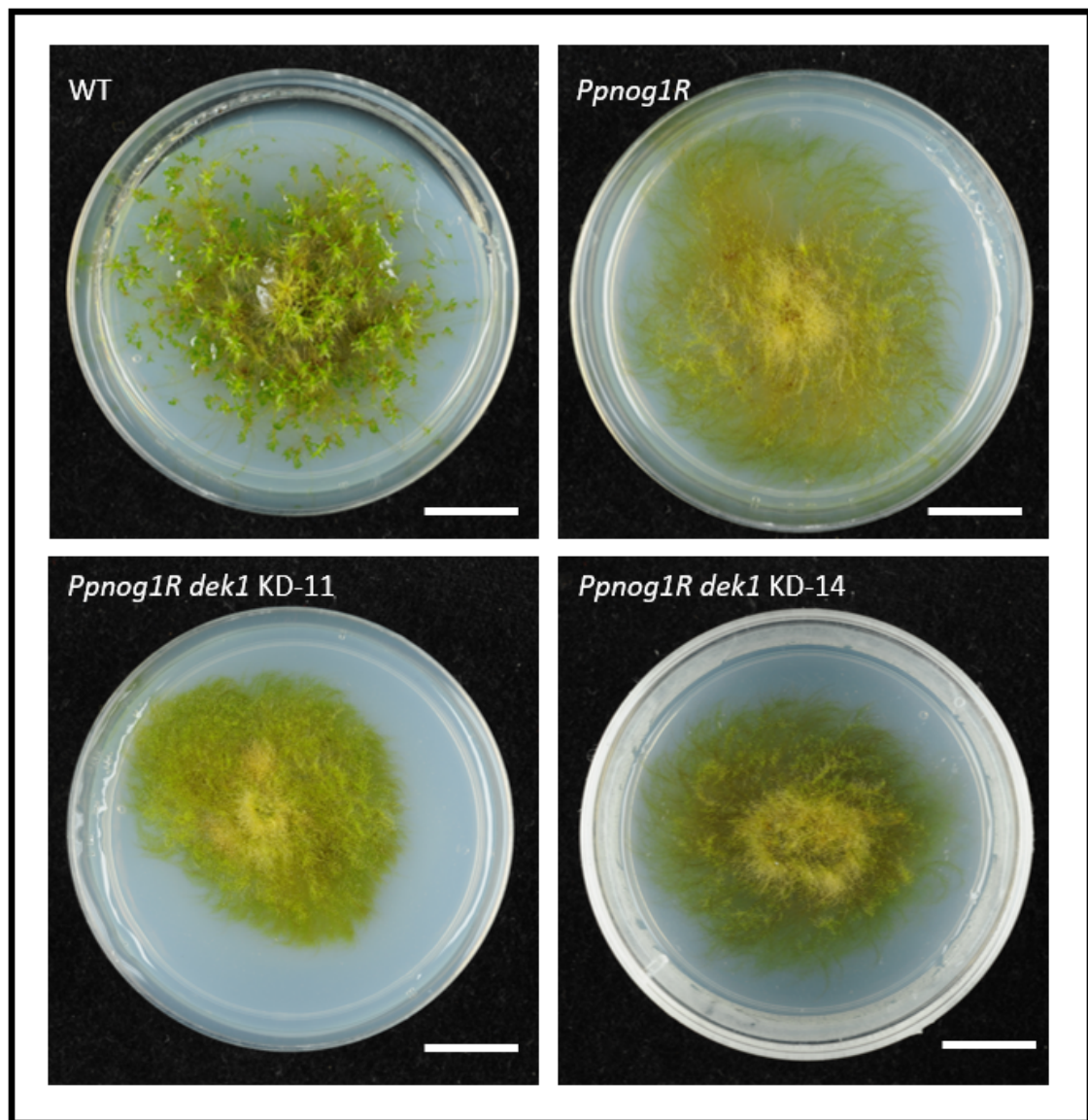


Figure 3.13: Representative images of *P. patens* lines. Scale = 1 cm.

3.3 Discussion

With the release of more and more whole genome sequences over the last few years, I was able to broaden the search for NOG1-like proteins into more algal species and did find convincing homologues in both charophytes and chlorophytes. Interestingly, three out of four algae with NOG1 homologues are single celled, and all four of them lack any form of apical growth. Therefore *PpNOG1* precedes the evolution of 3D growth and was possibly co-opted from another function. It would be very interesting to see if any of the algal *NOG1* sequences can complement the *Ppnog1R* mutant.

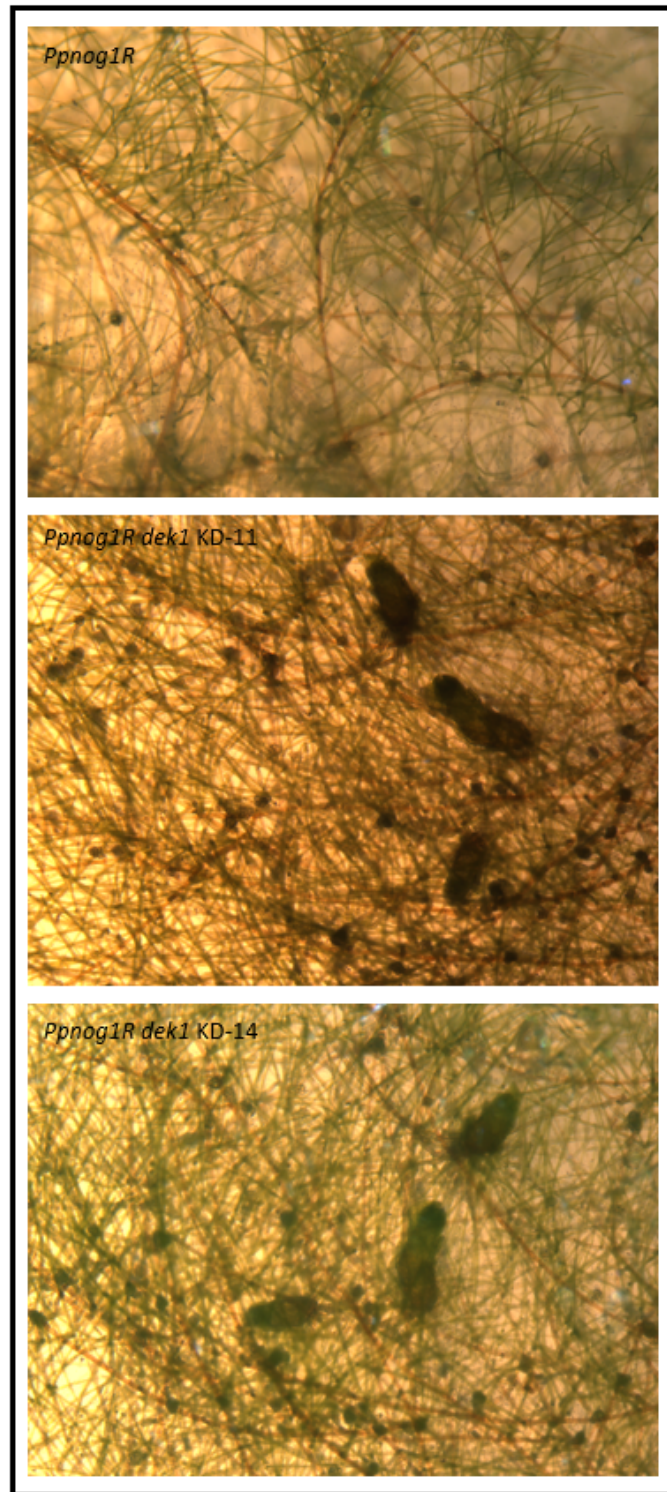


Figure 3.14: Representative light microscope images of *Ppnog1R*, *Ppnog1R dek1 kd-11* and *Ppnog1R dek1 kd-14* protonemal tissue

Expressing variants of NOG1 in *Ppnog1R*, has given some further clues as to the mechanisms of PpNOG1 function. Firstly, the ‘MGF’ loop in the UBA domain is essential for the function of PpNOG1 in prescribing the proper development of the gametophore bud, but is not required for the initiation of gametophore initials in response to BAP. Secondly, leucine¹⁹² in the UBA domain is not required for the early development of the gametophore bud and the establishment of the tetrahedral apical cell, nor is it required to initiate buds in response to BAP. However, the absence of leucine¹⁹² does affect the later development of the gametophore, producing a defective morphology with elongated stems and small phyllids. Next, the UBA domain alone is sufficient to promote establishment of the tetrahedral apical cell, but not sufficient to allow development of gametophores to maturity. Early developmental arrest may be due to an inability to maintain the tetrahedral apical cell. Additionally, expression of the UBA domain does not allow gametophore initiation in response to BAP. Other than the UBA domain, the PpNOG1 protein is as yet poorly characterised. Some region of the protein, other than the UBA domain is required for response to cytokinin. Most likely that region resides in the half of the protein that is not translated in *Ppnog1R* due to the premature STOP codon, or the region is disrupted by the truncation. To investigate this theory, I had started cloning alternative fragments of *PpNOG1* for expression in *Ppnog1R*. For example, I amplified a fragment from the location of the *Ppnog1R* mutation to the end of the transcript. Unfortunately I did not have sufficient time to pursue this avenue of investigation further. Expressing AtNOG1 in *Ppnog1R* results in a very mild rescue of the *Ppnog1R* phenotype, yielding gametophores that do establish a tetrahedral apical cell and produce two or three phyllids, but then abort development. The partial complementation of PpNOG1 with AtNOG1, points to a partially conserved function of the NOG1 protein across land plants. Performing this cross-species complementation in reverse would be informative. I did design constructs for expression of PpNOG1 in Arabidopsis *Atnog1* CRISPR mutants, but due to complications with crossing and genotyping those lines I was unable to perform this experiment. Ideally, for these experiments where NOG1 variants were

ectopically expressed in *Ppnog1R*, I would have analysed two independent lines for each construct. I did make some attempts at generating additional lines, but these were unsuccessful. Additionally it would be beneficial to transform constructs into an independent *Ppnog1* line, in which none of the *PpNOG1* sequence is expressed. This would rule out that the expressed truncated protein was a confounding factor in any of the manifested phenotypes. Additionally, other complementation experiments where *NOG1* sequences from other plant lineages such as ferns and lycophytes could be informative to assess functional conservation more broadly in land plants.

My yeast two-hybrid assay indicated that PpNOG1 may be able to interact with POLYUBIQUITIN3. However, the evidence for an interaction seemed to be rather weak, evident by poor growth of cotransformed yeast on quadruple dropout media. I hypothesise that this is because of the specificity that many UBA domains have for different polyubiquitin topologies (Raasi et al., 2005). For this experiment I cloned polyubiquitin genes from the genomes of *P. patens* and *Arabidopsis*. POLYUBIQUITIN genes produce a polypeptide consisting of tandem ubiquitin repeats, which are quickly proteolitically processed into monoubiquitin units (Callis et al., 1995). Therefore it is likely that the PpNOG1 gene does not have as strong affinity for the unprocessed POLYUBIQUITIN polypeptide as it would to a synthesised polyubiquitin chain with certain linkages. Also, it is possible that the PpUBQ3 and AtUBQ3 genes were being cleaved into monoubiquitin in the yeast cells as ubiquitin is highly conserved across eukaryotes (Callis et al., 1995).

I have determined that PpNOG1 is degraded by the proteasome as detection of the PpNOG1-GFP fusion protein on a western blot increased after treatment with the proteasome inhibitor bortezomib. It was not possible to increase protein level further by induction of transcription by the heat shock promoter. Regardless of heat shock treatment, the level of protein remained the same indicating there is very little correlation between transcript level and protein level for PpNOG1. What was also interesting was that PpNOG1 is apparently cleaved by another factor, possibly a cytosolic protease, as only fragments of the NOG1-GFP protein were detected. This cleavage occurs very quickly after translation as no full-length protein was

detectable. The cleavage of PpNOG1 likely has a regulatory function rather than a degradatory one. It should be noted that PpNOG1 is not a predicted target of the PpDEK1 calpain protease domain (Demko et al., 2024). One of the peptides I detected was at a molecular weight consistent with a peptide containing GFP and some or all of the UBA domain. I also detected a peptide that was consistent with cleavage within the GFP sequence, as the peptide was smaller than free GFP. This implies that the GFP itself is being degraded, which is very unusual as GFP is usually very resistant to proteolytic cleavage (Chiang et al., 2001). This degradation of the fluorescent protein may explain why it is so difficult to visualise PpNOG1-GFP in plant tissues. Generating an N-terminal fusion protein would be interesting as the N-terminal fragment of PpNOG1 may display different properties to the C-terminal fragment. For example it may be less subject to degradation by the proteasome and therefore may be easier to visualise.

It is interesting that PpNOG1 is targeted to the proteasome as it has a putative role in the ubiquitin-proteasome system, and such proteins are often protected from proteasomal degradation. The presence of C-terminal UBA domains has been shown to protect the protein from proteasomal degradation (Heinen et al., 2011). In fact, the addition of a UBA domain has been used to stabilise Cas9 to improve the efficiency of CRISPR-Cas9 genome editing in rice (Zheng et al., 2020).

PpNOG1 and PpDEK1 are both essential regulators of gametophore initiation and development. The lack of either one of these proteins results in plants that are 3D defective. However, it is not known whether PpNOG1 and PpDEK1 operate in independent pathways or if they are in the same pathway. The *PpAPB* transcription factors are the master regulators of gametophore initiation (Aoyama et al., 2012). In *Ppnog1R*, in which there are fewer gametophore initials formed, the expression of *PpAPBs* are suppressed (Moody et al., 2018). In *Ppdek1*, which form more gametophore initials than usual, the expression of *PpAPBs* is elevated (Demko et al., 2014). It has been hypothesised therefore, that PpDEK1 and PpNOG1 somehow regulate the expression of the *PpAPBs*, with PpNOG1 promoting *PpAPB* expression and PpDEK1 suppressing it. If this is the case, the question remains

as to whether they each independently regulate the *PpAPBs* or if one of them regulates the other which regulates *PpAPB* expression, in a more linear manner. My very preliminary examination of *Ppnog1R dek1 kd-11* and *Ppnog1R dek1 kd-14* revealed that they seemed to produce a high number of bud initials. This means they resemble the *Ppdek1* phenotype more than they do the *Ppnog1R* phenotype. This could indicate an epistatic interaction whereby PpDEK1 is downstream of, and inhibited by, PpNOG1. However, my analysis was very preliminary and much more work is required to ascertain if this is the case. The unusual structures seen in *Ppnog1R dek1 kd-11* and *Ppnog1R dek1 kd-14* also do not resemble the phenotype of either *Ppnog1R* or *Ppdek1*. However, they do resemble the phenotype of *Ppdek1Δloop* mutants, which also produce unusual stack-like structures (Demko et al., 2014). The PpDEK1 protein comprises 23 transmembrane domains, a C-terminal cytosolic calpain protease domain and a cytosolic loop that sits between two of the transmembrane domains (Perroud et al., 2014). In *Ppdek1Δloop* mutants, just the cytosolic loop has been removed, leaving the rest of the protein intact (Demko et al., 2014). The significance of this is yet to be explored. The initial evidence does seem to suggest that there is some kind of genetic interaction between *PpNOG1* and *PpDEK1*, although I did not have sufficient time to explore this in much detail. It is definitely worth pursuing for further investigation. It would be of value to generate genuine *dek1* deletion mutants rather than knock down mutants for future studies as this would be more robust.

Chapter 4: A suppressor screen reveals a new regulator of the 2D-to-3D growth transition.

4.1 Introduction

Previously, a forward genetics screen was carried out that aimed to identify genes required for successful transition to 3D growth. The haploidy of the *P. patens* gametophyte and the fact that the presence or absence of gametophores is a very easily screenable phenotype makes this forward genetics approach a very powerful tool for investigating 3D growth in a completely unbiased way. This allows us to identify new genetic regulators of the transition from 2D to 3D growth, even if those genes are not well annotated or characterised or are not obvious candidates for investigation. Plants were mutagenised with UV radiation and screened for a 3D-defective phenotype. The first such mutant identified was called *no gametophores 1* (*Ppnog1R*). *Ppnog1R* is completely unable to form gametophores. Few gametophore initials formed, and those that did displayed disorganised cell division planes that resulted in aborted development due to a failure to establish the critical tetrahedral apical cell. Additionally, *Ppnog1R* was completely insensitive to treatment with the cytokinin analogue 6-benzylaminopurine (BAP), which usually induces the formation of excessive gametophore initials in wild type. By mapping the causative mutation in the *Ppnog1R*, the novel positive regulator of the transition to 3D growth, *PpNOG1*, was identified. The PpNOG1 protein bears a C-terminal ubiquitin-associated (UBA) domain, which suggests a role in protein degradation processes. The expression of the *PpAPB* genes, which are master regulators

of gametophore initial formation, is decreased in *Ppnog1R*, which suggests that *PpNOG1* may be operating upstream of the *PpAPBs* (Aoyama et al., 2012; Moody et al., 2018).

To identify potential genetic interactors of *PpNOG1*, a suppressor screen was designed. The goal was to mutagenise an independent *nog1* mutant to generate mutations that alleviated the *nog1* phenotype, *i.e.* restored 3D growth. This approach was successful and two suppressor mutants were generated. For one of these mutants, *suppressor of nog1 a* (*snog1a*), the causative mutation was mapped and the responsible gene identified.

4.2 Results

4.2.1 The *nog1dis* mutant resembles *Ppnog1R*

An initial forward genetics screen was performed by UV-mutagenising protoplasts isolated from the *Ppnog1R* mutant. This yielded eight ‘*suppressor of nog1*’ mutants that had recovered the ability to form gametophores. However, further examination revealed that in all eight mutants the original causative point mutation in *Ppnog1R* (R⁷⁶Ter) had simply been corrected. Thus, it was necessary to generate an independent *nog1* mutant in which the *PpNOG1* locus was disrupted in such a way that could not be easily rectified by a UV-induced SNP. Additionally, an independent *nog1* mutant would be clean of all the background mutations that were introduced into *Ppnog1R* from the initial mutagenesis. Therefore, attempts were made to generate a *PpNOG1* deletion line, by replacing the *PpNOG1* open reading frame with a hygromycin resistance cassette (Figure 4.1A).

Many attempts were made to generate a *PpNOG1* full deletion mutant. However, as has been the case previously (see Moody et al., 2018), it was only possible to create a mutant where the locus was disrupted but not deleted. From this and previous attempts, it is apparent that there may be some inherent difficulty with deleting the *PpNOG1* locus. The homology-directed strategy that was used works well for deleting other loci, but is not effective for *PpNOG1*. Nevertheless, a mutant

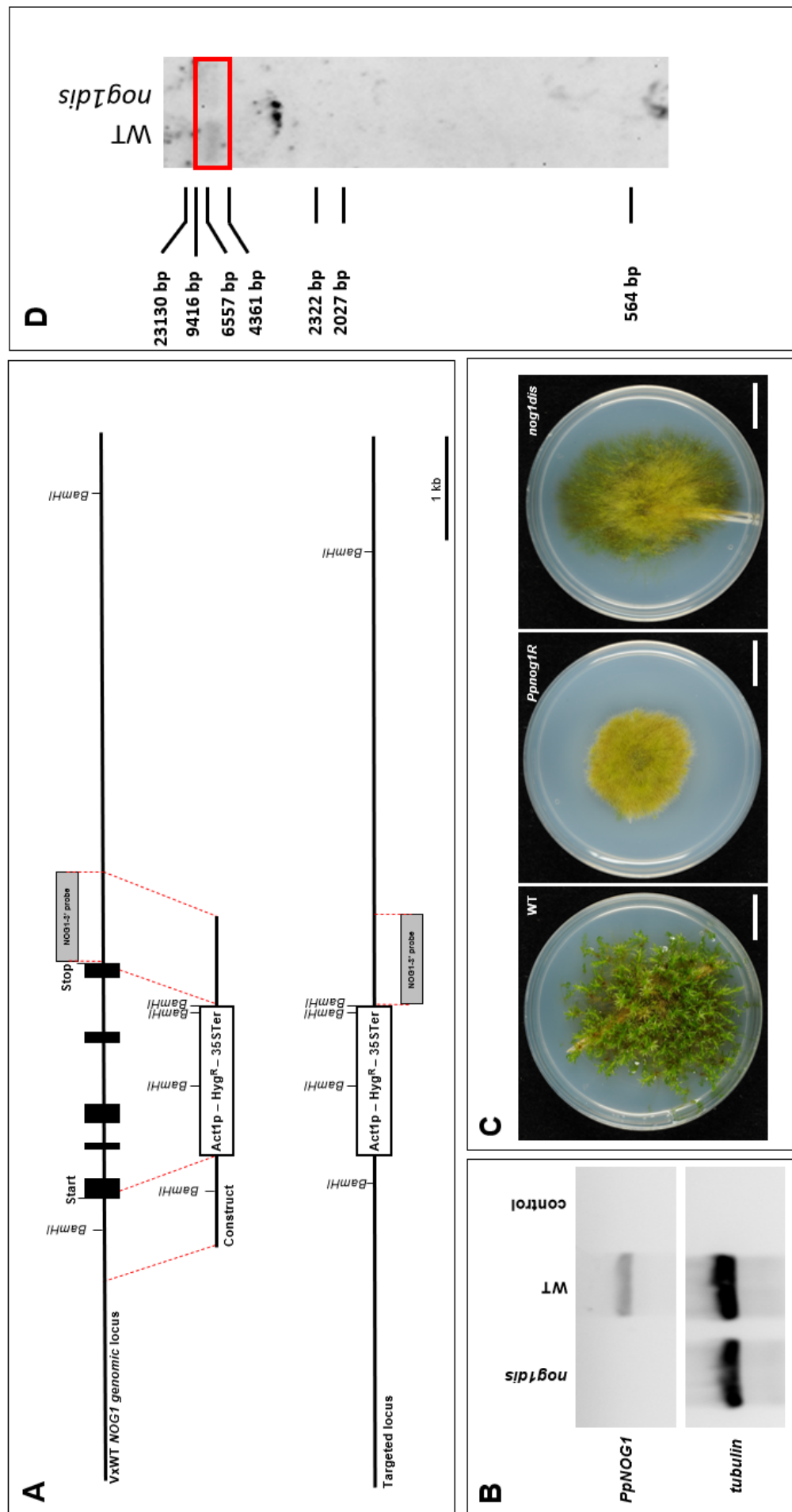
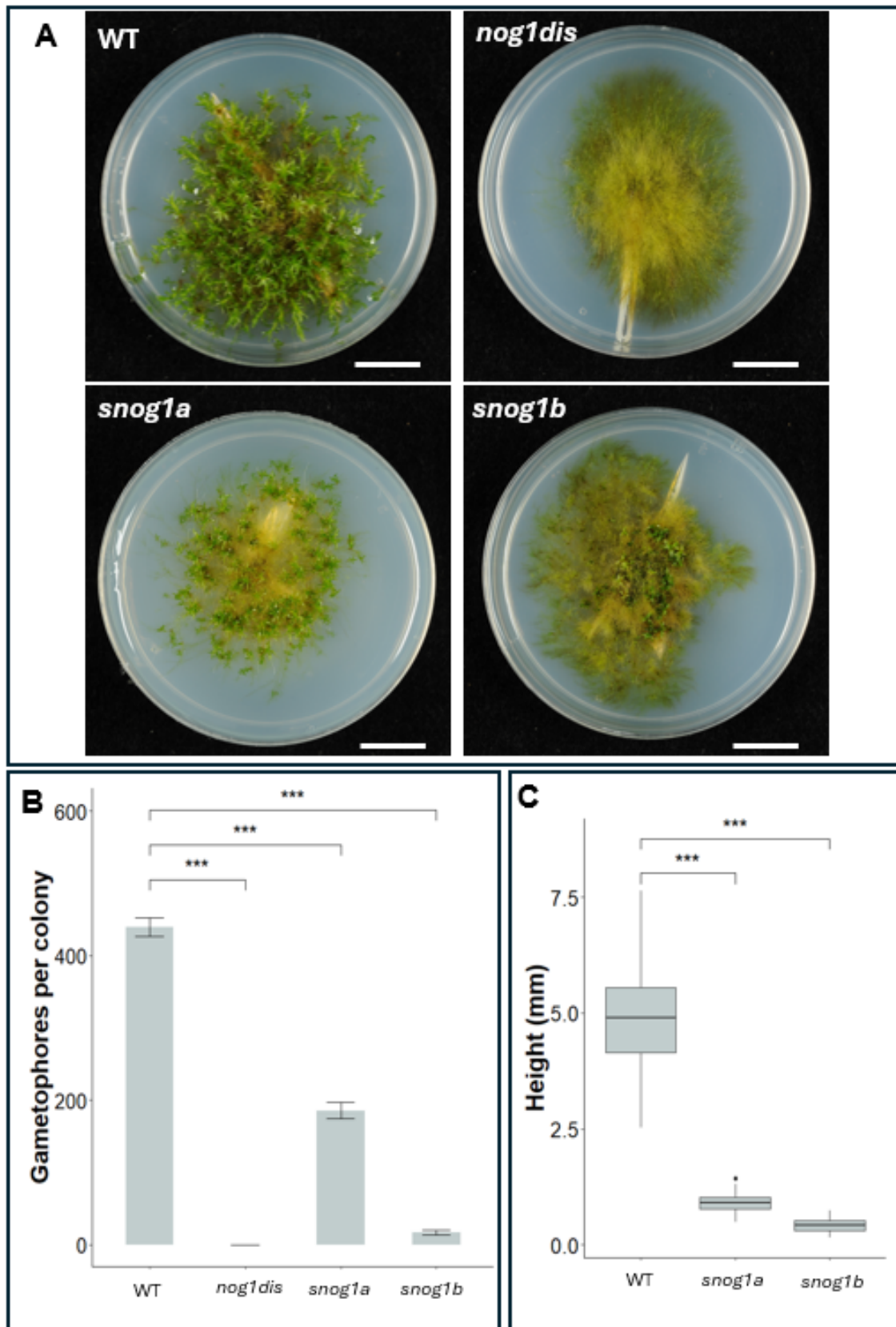


Figure 4.1: A) *PpNOG1* deletion strategy. BamHI restriction sites are highlighted. The region targeted by the NOG1-3' Southern blotting probe is indicated. B) RT-PCR on cDNA from wild type and *nog1dis* showing the absence of the *PpNOG1* transcript in *nog1dis*. RT-PCR carried out by Gargi Chaturvedi. C) Representative images of wild type, *Ppnog1R* and *nog1dis*. Scale = 1 cm. D) Southern blot result. genomic DNA from wild type and *nog1dis* was digested with BamHI and probed with a DIG-labelled NOG1-3' probe.

(*nog1dis*) was obtained in which *PpNOG1* cannot be expressed. RT-PCR could not detect the *PpNOG1* transcript in cDNA from *nog1dis*, but it was evident in wild type (Figure 4.1B). Furthermore, the *nog1dis* mutant exhibited a robust ‘*no gametophores 1*’ phenotype (Figure 4.1C). However, later treatments with hygromycin B revealed that *nog1dis* was no longer resistant and had apparently lost the hygromycin resistance cassette. A Southern blot confirmed this was the case (Figure 4.1D). If the hygromycin resistance cassette had stably integrated the expected band size would be 4828 bp for *nog1dis* vs 7498 bp for the wild type locus. The blot shows two bands of similar size, consistent with the wild type locus. Nevertheless, given the difficulty in generating *nog1* mutants and the consistent absence of gametophores and *PpNOG1* transcripts in *nog1dis*, it was decided to proceed with the suppressor screen using *nog1dis*. I hypothesise that the integration and subsequent loss of the hygromycin resistance cassette caused sufficient sequence change at some point in the *PpNOG1* locus as to prevent all transcription of the gene. I attempted to diagnose the precise disruption with some targeted sequencing. A preliminary examination revealed a lesion in the promoter region of *PpNOG1* (Figure S24). The lesion is a 6 bp deletion (ACCAGC) and this sequence has previously been described as an important binding motif for zinc finger transcription factors Ace2p and Swi5p in yeast, so loss of this sequence may explain why *PpNOG1* expression is completely abolished in *nog1dis* (Dohrmann et al., 1996; Kato et al., 2004; Wang and Zhang, 2006).

4.2.2 UV mutagenesis produced two *suppressor of nog1* mutants in which 3D growth is restored

Forward mutagenesis was performed by exposing *nog1dis* protoplasts to a high dose of UV radiation, sufficient to kill approximately 90% of the cells. The surviving protoplasts were allowed to regenerate and the resulting plants were then screened for the presence or absence of gametophores. Two of the recovered individuals exhibited a reversion of the *nog1dis* phenotype and were therefore named *suppressor of nog1 a* (*snog1a*) and *suppressor of nog1 b* (*snog1b*; Figure 4.2A).



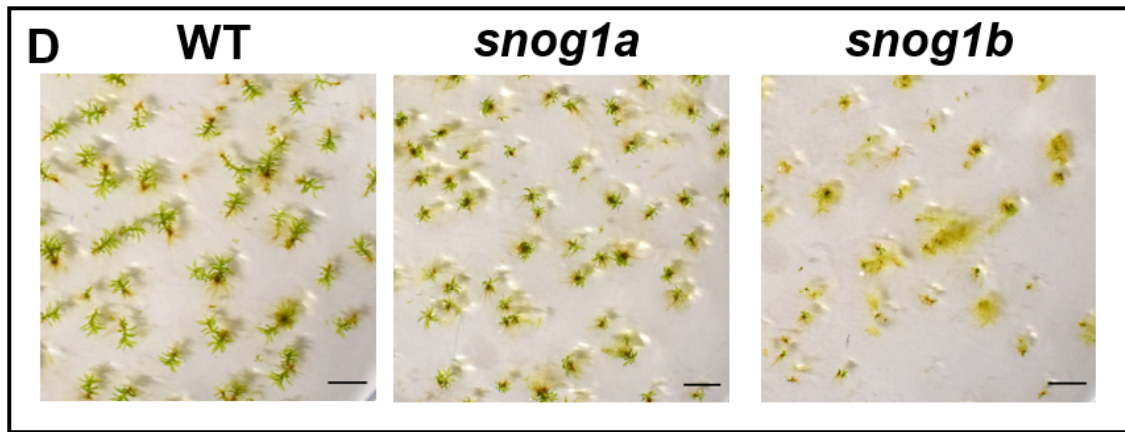


Figure 4.2: A) Representative images of wild type, *nog1dis*, *snog1a* and *snog1b* after 6 weeks of growth. Scale = 0.5 cm. B) Gametophore number per colony for wild type (n=5), *nog1dis* (n=5), *snog1a* (n=5) and *snog1b* (n=5) after 2 months of growth. Standard error is shown for each bar. Horizontal lines indicate result of t-test for each generated line compared to wildtype. *** means p-value < 0.001. C) Gametophore heights of wild type (n=100), *snog1a* (n=94) and *snog1b* (n=51) after 2 months of growth. Horizontal lines indicate result of t-test for each generated line compared to wildtype. *** means p-value < 0.001. D) Representative images of 2-month-old gametophores from wild type, *snog1a* and *snog1b*. Scale = 0.5 cm.

The *snog1a* and *snog1b* mutants yielded approximately 42% and 4% the number of gametophores as wild type, respectively (Figure 4.2B). The gametophores of *snog1a* were rather stunted, reaching only about 1/5 the height of wild type gametophores (Figure 4.2C and D) and emerge later than in wild type. The gametophores of *snog1b* were very severely stunted, only producing a few phyllids and only growing to approximately 0.5 mm (Figure 4.2C and D).

Treatment with the cytokinin analogue 6-benzylaminopurine (BAP) elicits a strong bud initiation response in wild type. It was previously reported that the *Ppnog1R* mutant was completely unable to respond to BAP (Moody et al., 2018). As expected, I found that the same was true of *nog1dis*, which did not produce buds in response to treatment with a high concentration of BAP (1 μ M). However, *snog1a* and *snog1b* both displayed high rates of bud initiation in response to BAP (Figure 4.3). Thus, the mutant phenotypes observed in *nog1dis*, *snog1a* and *snog1b* mutants can possibly be attributed to changes in cytokinin perception.

Next I examined the early stages of bud development in wild type, *nog1dis*, *snog1a* and *snog1b*. The first few cell divisions within the bud should follow

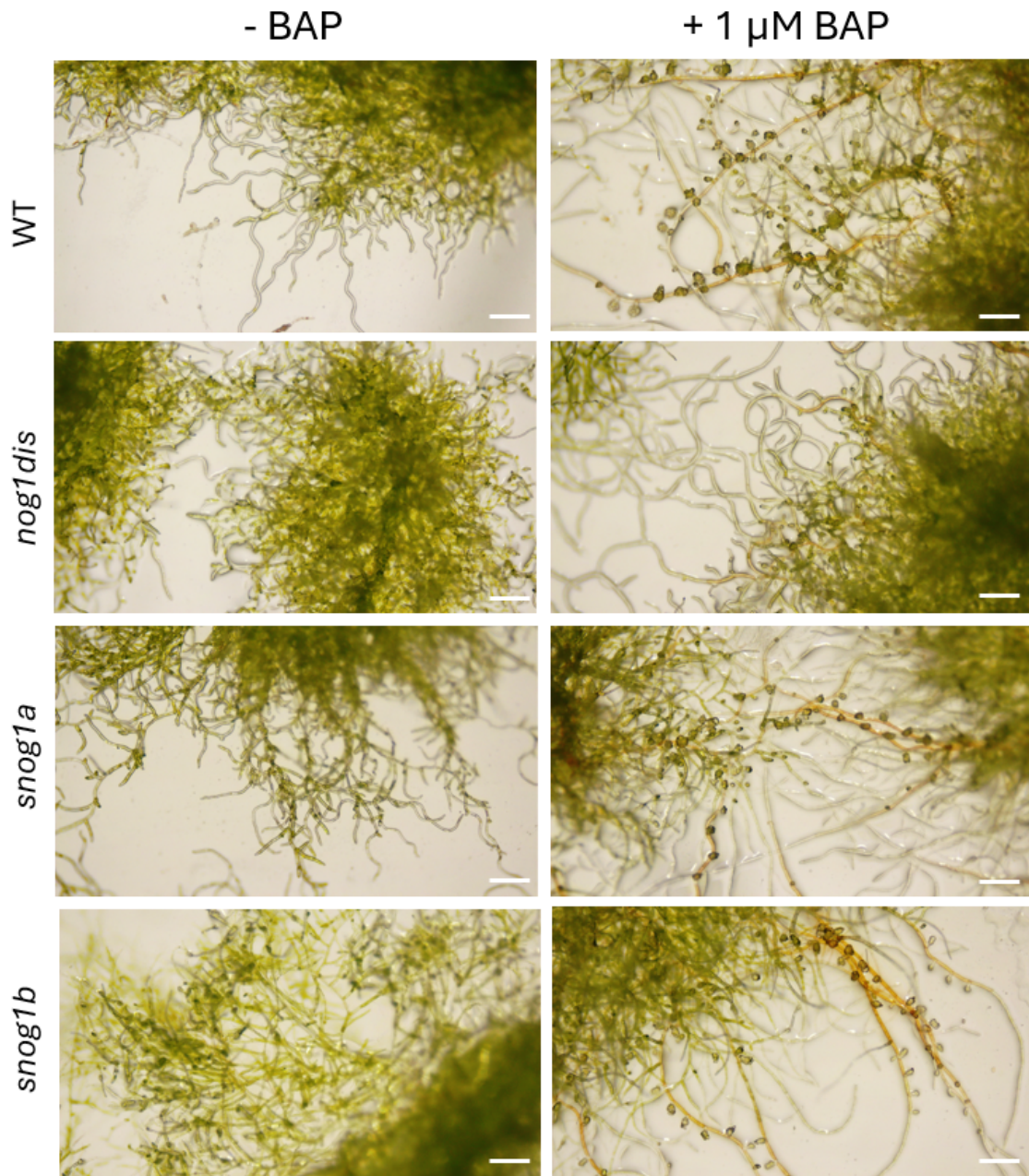


Figure 4.3: Representative images of wild type, *nog1dis*, *snog1a* and *snog1b*, in the presence of absence of 1 μ M BAP. Tissue was grown for one week on BCD medium, before being transferred either to fresh BCD medium (control) or BCD medium containing 1 μ M BAP. After a further three days the tissues were photographed. Scale = 500 μ m.

a well-documented stereotypical pattern, which is crucial for establishing the tetrahedral apical cell (Harrison et al., 2009). In *Ppnog1R*, all divisions within the bud are misoriented, resulting in very early developmental arrest due to failure to establish the tetrahedral apical cell (Moody et al., 2018). I set out to determine what defects in bud development are exhibited in *nog1dis* and whether these are corrected in *snog1a* and *snog1b*. This was done by staining tissues with propidium iodide, which allowed visualisation of cell boundaries using fluorescence microscopy.

In wild type *P. patens*, the first division of the gametophore initial cell was invariably oblique and yielded an apical and a basal cell (Figure 4.4A). Two additional oblique divisions of the apical and basal cells then occurred at an angle roughly perpendicular to that of the first division. The order in which the apical and basal cells divided was inconsistent, but the timing of the second and third divisions have previously been observed to be rather synchronous, so this is not unexpected (Harrison et al., 2009; Kosetsu et al., 2017). It seemed to me that the apical cell divided before the basal cell as often as the basal cell divided before the apical cell (Figure 4.4B and C). Successive rotating divisions at the apex specify the critical tetrahedral apical cell, which divides to self-renew and produce phyllid initials. The phyllids that then develop are arranged in spiral phyllotaxy around the central axis of the gametophore (Figure 4.4D; Harrison et al., 2009).

The early bud development of *nog1dis* shows similarities but also differences to that seen in *Ppnog1R*. In both, fewer gametophore initial cells are formed. Often the first division plane of the *nog1dis* gametophore initial cell was not characteristically oblique and instead was roughly parallel to the axis of the parental caulonemal cell as was also observed in *Ppnog1R* (Figure 4.4E). This was followed by random, misplaced division planes that resulted in early developmental abortion due to a failure to establish the tetrahedral apical cell. This series of events was the same as was observed in the *Ppnog1R* mutant. In some cases the *nog1dis* gametophore initial cells swelled and elongated excessively but could then divide in a somewhat oblique manner at the first and second divisions (Figure 4.4F). However, subsequent divisions were usually misplaced (Figure 4.4G). In the *nog1dis* mutant I occasionally

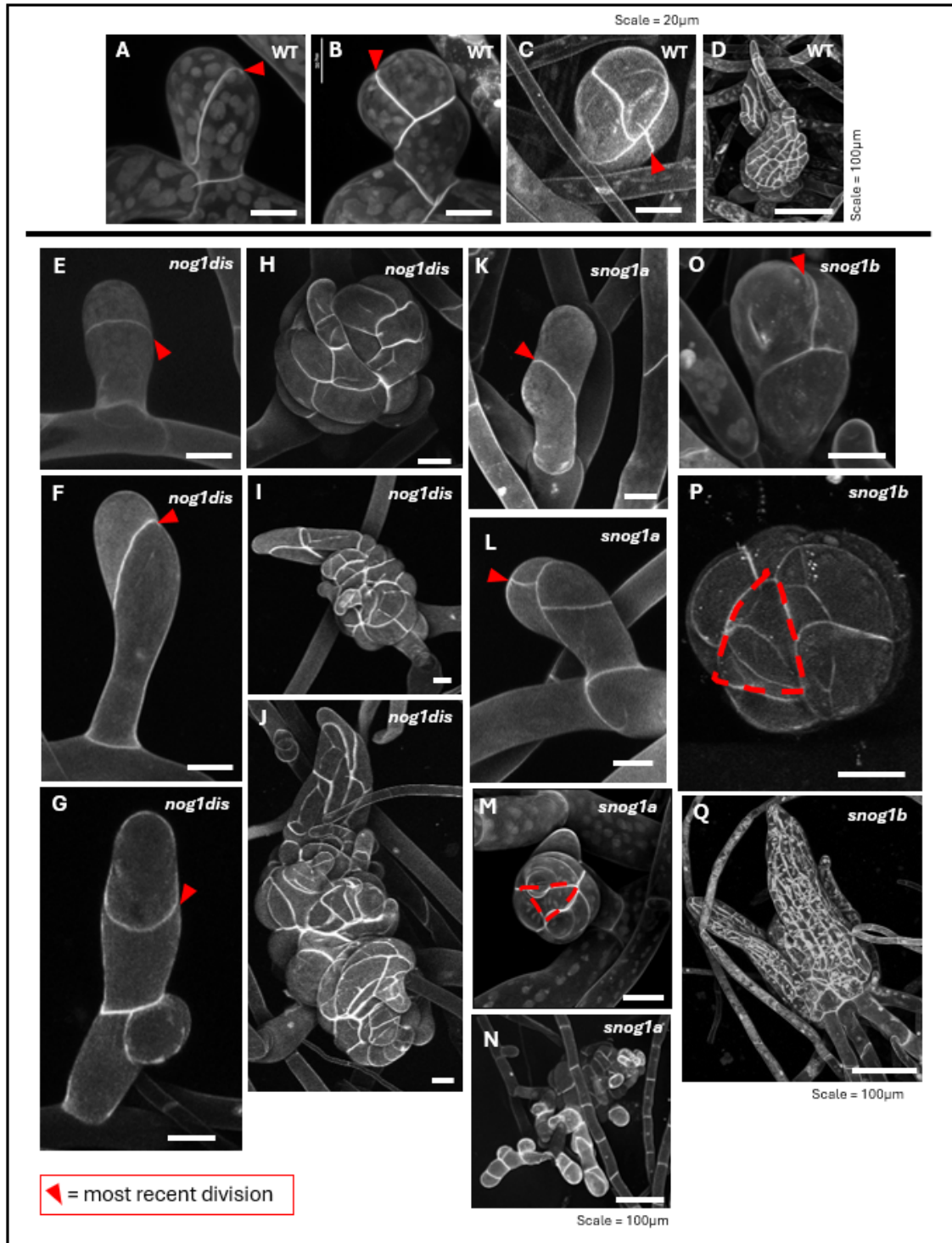


Figure 4.4: Z-stack images of early buds stained with 10 μgml^{-1} propidium iodide. Red arrows indicate the most recent cell division. Scale = 20 μm unless otherwise stated. A-D) wild type. A) 2-cell bud. B) 3-cell bud. C) 4-cell bud. D) Early gametophore with two phyllids. Scale = 100 μm . E-J) *nog1dis*. E-F) 2-cell buds. G) 4-cell bud. H) Callus-like bud. I) Disorganised bud with one phyllid. J) Bifurcated bud. K-N) *snog1a*. K) 2-cell bud. L) 4-cell bud. M) Bud with tetrahedral apical cell highlighted with red outline. N) Disorganised bud. Scale = 100 μm . O-Q) *snog1b*. O) 3-cell bud. P) Bud with tetrahedral apical cell highlighted with red outline. Q) Early gametophore with about four phyllids. Scale = 100 μm .

observed the developmental progression of buds beyond the 5-10 cell stage. For example, I observed a few callus-like buds that appeared to have formed due to uncontrolled proliferation (Figure 4.4H). I also occasionally observed buds that had attempted to form a single phyllid although those buds had noticeably disorganised cell divisions (Figure 4.4I). The presence of a single phyllid suggests that in some rare cases a tetrahedral apical cell was established briefly but that it was not maintained. Furthermore, I also observed a few bifurcation events (Figure 4.4J). The bifurcations suggest confused cell fate that may have resulted in the specification of ectopic apical cells. Nevertheless, if apical cells were produced, they were not maintained long enough to produce a gametophore. Indeed I did not observe any buds with more than one very small phyllid. These observations were quite rare and *nog1dis* buds were usually impaired at the first division stage. Therefore, *nog1dis* shares the main characteristics of *Ppnog1R*; reduced bud initiation frequency and misplaced cell division planes within the bud resulting in a failure to form or maintain the tetrahedral apical cell and therefore to specify 3D growth.

In the *snog1a* mutant, the initiation of gametophore initial cells was partially restored, and these largely followed a wild type pattern of development, with some exceptions. Mostly, I observed that the first division of the bud was characteristically oblique (Figure 4.4K) and the following divisions of the apical and basal cells also followed the wild type pattern (Figure 4.4L). Subsequently buds were able to form and maintain the tetrahedral apical cell to produce viable gametophores (Figure 4.4M). However I did occasionally observe buds in which there appeared to be some aberrant disorganised cell divisions (Figure 4.4N). It is likely that those buds aborted development. For the most part, it appeared that early bud development had been restored in *snog1a*. For *snog1b* the picture was less clear. At least the first two divisions seemed to mostly occur as seen in wild type (Figure 4.4O) and buds did seem to be able to establish the tetrahedral apical cell (Figure 4.4P). The resulting gametophores were very stunted and did not reach maturity, only forming at most around 2-6 phyllids (Figure 4.4Q).

4.2.3 The mutated *snog1a* locus was mapped to a region on chromosome 8

The simplest way to identify the causative mutations in the suppressor mutants was to outcross the mutant and perform bulk segregant analysis. However, this required the mutants to be reproductively fertile. The *nog1dis* mutant, like all other 3D defective mutants, was completely infertile due to the lack of gametophores, from which the gametangia develop. Due to the incredibly stunted growth of the *snog1b* gametophore, *snog1b* was also sterile. Therefore I focused on *snog1a*, which had more developed-looking gametophores and therefore might have regained fertility. To assess this I examined mature gametophores of *snog1a* that had been exposed to sporophyte-inducing conditions and identified normal-looking archegonia and antheridia (the female and male gamete bearing organs, respectively). Therefore I proceeded with a cross. The *nog1dis* mutant was generated in the Villersexel wild type background. For bulk segregant analysis I needed to outcross *snog1a* with a male parent from a different genetic background. Therefore, I performed a cross between the *snog1a* mutant and the highly fertile Reute strain. I used a line that expresses mCherry (Reute::mCherry; ReCh; Perroud et al., 2019). Attempts at fertilisation were successful, confirming that *snog1a* was indeed fertile and could produce sporophytes, although fertility was low and the number of sporophytes obtained was very few. I verified that these were outcrossed sporophytes by checking that they expressed mCherry. Fortunately, these few sporophytes yielded a sufficient number of germinating spores to proceed with the bulk segregant analysis.

If the *snog1a* phenotype was caused by a mutation in one gene, rather than multiple genes, there would be two segregating mutated loci (the mutation within the *PpNOG1* gene, and the unknown *snog1a* causative mutation). If this were the case, it was expected that 25% of the progeny would not be able to produce gametophores as the genotype would be the same as *nog1dis* (mutated *nog1* allele and wild type *SNOG1A* allele). The remaining 75% of the progeny were expected to be capable of 3D growth, exhibiting wild type-like or *snog1a*-like phenotypes (Figure

Observed			Expected			Chi-square	Critical chi-square value
3D growth	no 3D growth	Total	3D growth	no 3D growth	Total		
574	80	654	490.5	163.5	654	56.9	3.841 (df = 1, p<0.001)

Table 4.1: Observed and expected frequencies of the progeny phenotypes

4.5A).

The frequencies that I observed were generally consistent with the mutation of a single genetic locus in the *snog1a* mutant, although this was not statistically significant as I observed fewer *nog1dis*-like progeny than expected (12.2% rather than 25%; Table 4.1). One possible explanation for the reduced *nog1dis* mutant phenotype frequency was that disruption of the *PpNOG1* gene had affected haploid spore germination and/or the viability of young sporelings. It is unlikely that two genes were mutated as this would have resulted in an expected 1/16 of the progeny exhibiting an inability to produce gametophores.

I prepared genomic DNA from 80 individuals that exhibited the *nog1dis* phenotype and pooled these in equimolar amounts, and this became the ‘wild type pool’ (all individuals resembled the non-mutagenized parental line, *nog1dis*). Additionally, I prepared genomic DNA from 98 individuals with the capacity for 3D growth, showing a preference for those that most closely resembled the *snog1a* mutant rather than the true wild type, and this became the ‘mutant pool’. These two pools were sequenced at 44X coverage alongside both of the parental lines *nog1dis* and ReCh. The *nog1dis* mutant was used as the reference female parent rather than *snog1a* as it allowed me to distinguish between single nucleotide polymorphisms (SNPs) that were present in the different genetic backgrounds (Villersexel vs Reute) and mutations that were induced by the UV mutagenesis. To identify the genomic region containing the *snog1a* mutation, bulk segregant analysis was performed. SNPs that differed between the two parental lines (*nog1dis* and ReCh) were identified as markers and the frequency of each SNP variant allele was mapped across the chromosomes to show the parental origin of each region. For regions not associated with the phenotypes of the two pools, the expected SNP allele frequency was around 0.5, showing equal contribution from both parents, due to random recombination and

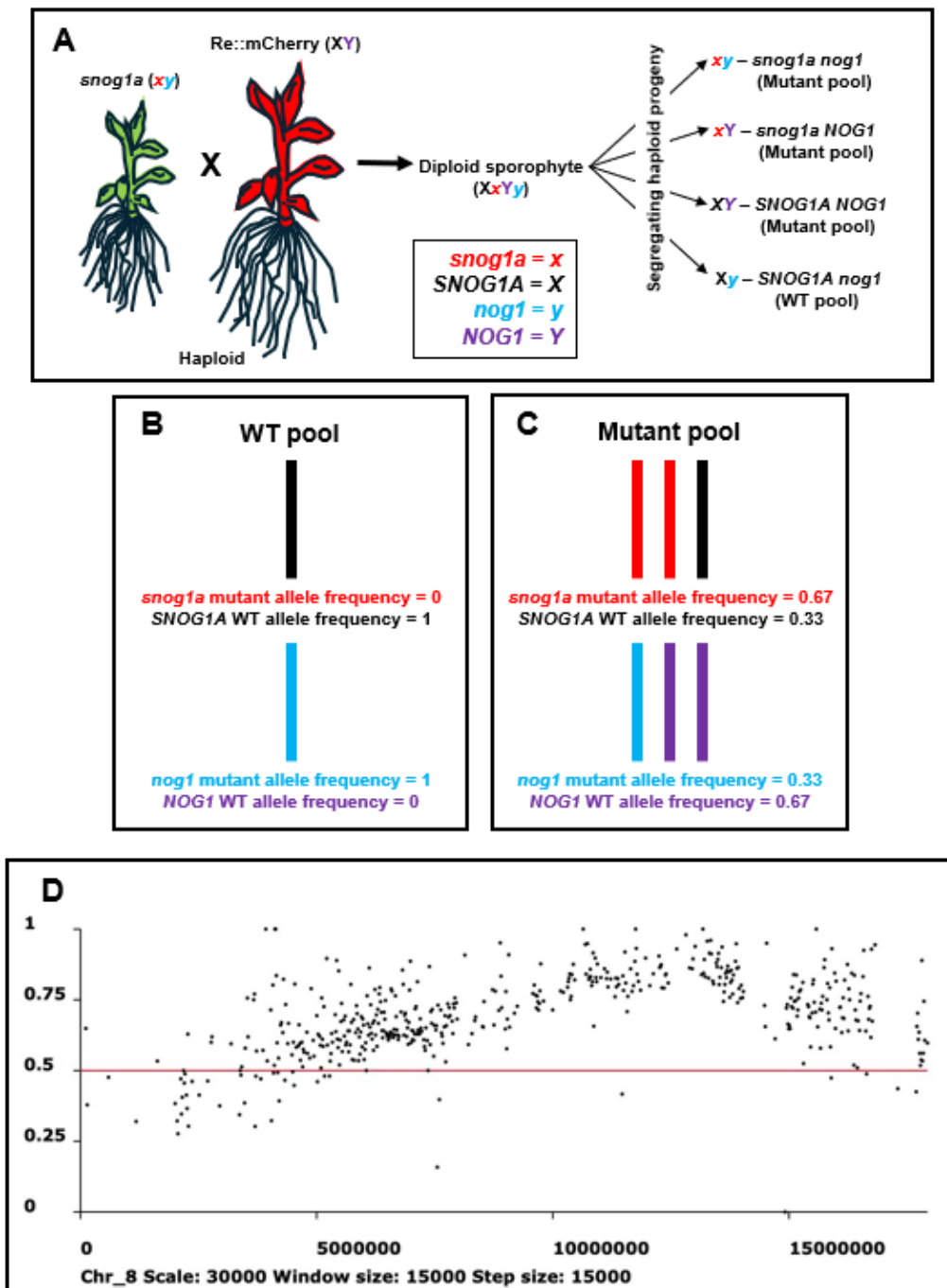


Figure 4.5: A) Schematic of the cross to generate the mapping population for bulk segregant analysis to identify the locus of the causative mutation in *snog1a*. B) Predicted allele frequencies in the wild type pool. C) Predicted allele frequencies in the mutant pool. D) The allele frequency of SNP markers in the mutant pool mapped across chromosome 8.

segregation. In the ‘wild type pool’, at the *SNOG1A* locus, the expected frequency of the SNP variants originating from *nog1dis* was 0 (Only the wild type SNOG1A allele from ReCh present; Figure 4.5B). In the ‘mutant pool’, at the *SNOG1A* locus, the expected frequency of the *nog1dis* SNP variants was 0.67 (2/3 of individuals in this pool were expected to carry the mutated *snog1a* allele and 1/3 the wild type SNOG1A allele from ReCh; Figure 4.5C). When the SNP allele frequencies for the two pools were plotted across all 27 chromosomes in the *P. patens* genome assembly, a peak of the expected allele frequency was revealed on chromosome 8 (Figure 4.5D), which provided a region of the genome to interrogate for causative mutations.

4.2.4 Point mutations in the coding sequence of *Pp3c8_19720* resulted in the introduction of two premature STOP codons

When looking for the causative mutation in *snog1a*, mutations within coding sequences (CDS) were the most likely candidates. Any SNPs that originated from either of the parent lines (ReCh and *nog1dis*) could be filtered out, leaving only the mutations induced by UV treatment. Of these, there were only two mutations within a CDS and both were located in the same gene: *Pp3c8_19720* (XP_024382948.1, https://phytozome-next.jgi.doe.gov/report/gene/Ppatens.v3.3/Pp3c8_19720).

These point mutations were both C → T transitions that converted two separate glutamine residues to STOP codons (Gln³³⁵Ter and Gln³⁷⁴Ter; Table 4.2; Figure 4.6). These premature STOP codons signify a dramatic truncation of the Pp3c8_19720 protein with a loss of 289 amino acid residues from the C-terminus, almost half the length of the entire protein (Figure 4.6). The severity of this mutation meant that *Pp3c8_19720* was definitely a viable candidate as the mutated gene responsible for the *snog1a* phenotype.

Targeted sequencing of a approximately 500bp region surrounding the two SNPs confirmed that these mutations are present in *snog1a*, but not in *nog1dis* or ReCh (Figure S25). In both the *Ppnog1R* and *Ppnog2R* mutants, it was observed that

Chromosome	Position	Reference	Variant	Gene	Location	Amino Acid Change
Chr08	13077134	C	T	<i>Pp3c8_19720</i>	exon 3	CAG (Gln) to UAG (stop)
Chr08	13077251	C	T	<i>Pp3c8_19720</i>	exon 3	CAA (Gln) to UAA (stop)

Table 4.2: In *snog1a*, two mutations in *Pp3c8_19720* result in two premature STOP codons

the UV-induced premature STOP codons introduced into *PpNOG1* and *PpNOG2* triggered alternative splicing patterns of the *PpNOG1* and *PpNOG2* transcripts, likely to prevent the expression of a potentially deleterious truncated protein (Moody et al., 2018, 2021). Sequencing of *Pp3c8_19720* transcripts from *snog1a* and wild type did not reveal any splicing defects in the transcripts from *snog1a* (Figure S26).

To determine whether *snog1b* is caused by mutation in the same gene as *snog1a*, I sequenced the whole *Pp3c8_19720* genomic sequence from *snog1b*. This revealed one SNP in an intronic region which was determined to simply be a difference between the Villersexel ecotype (the *snog1b* genetic background) and the Gransden ecotype (the reference genome; Figure S27). There were no exonic mutations, therefore the phenotype of *snog1b* is not caused by a mutation in *Pp3c8_19720*. This was expected as *snog1a* and *snog1b* do not exhibit the same phenotype.

4.2.5 Overexpression of *Pp3c8_19720* in *snog1a* complements the *snog1a* mutant phenotype

To conclusively confirm that the *snog1a* mutant phenotype was caused by mutations in *Pp3c8_19720*, I transformed *snog1a* with a construct designed to overexpress the wild type version of *Pp3c8_19720* using the rice actin promoter. Successful transformation was confirmed by the presence of the *Pp3c8_19720* CDS sequence in *snog1a pACT::Pp3c8_19720-GFP* genomic DNA (Figure 4.7A and B). The *snog1a pACT::Pp3c8_19720-GFP* line displayed a reversion to the ‘no gametophores’ phenotype seen in *nog1dis* (Figure 4.7C-E). Therefore, introducing a functional version of *Pp3c8_19720* complements the UV-induced mutation by restoring the inhibition of 3D growth. This confirms that the phenotype is caused by the mutations within *Pp3c8_19720*.

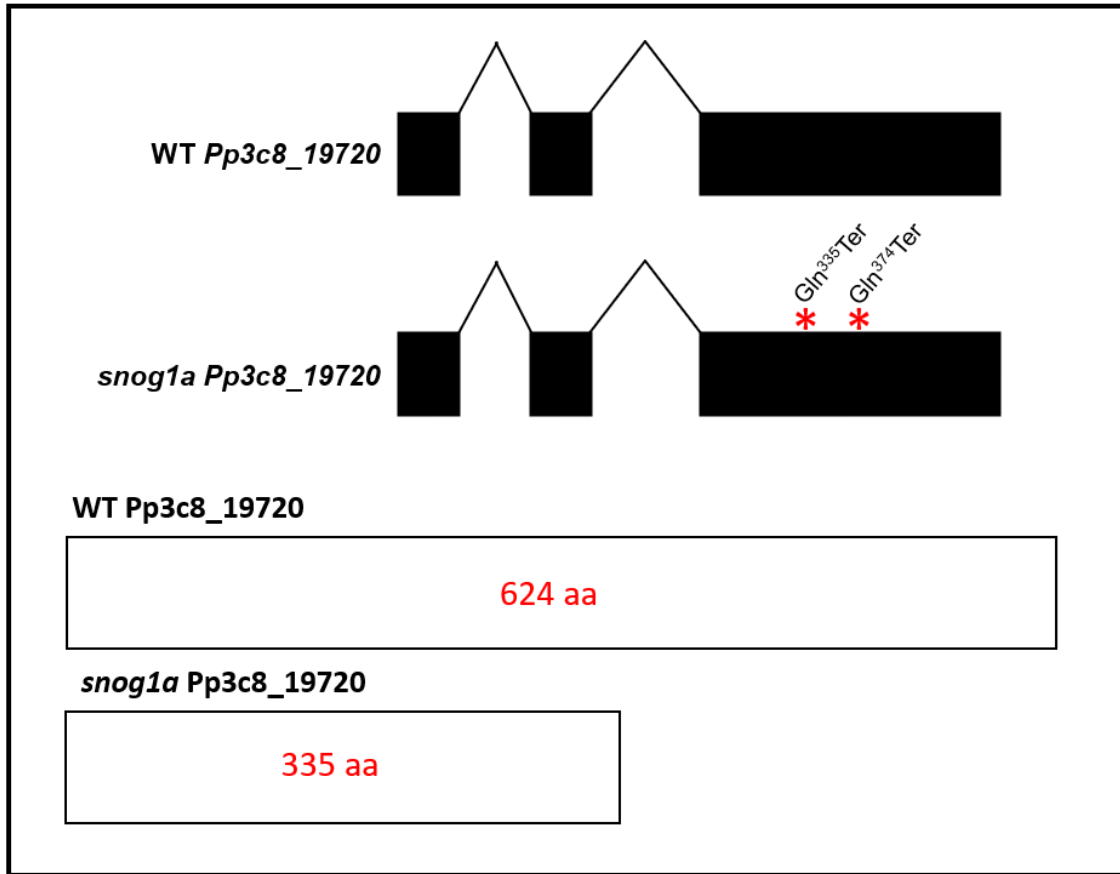


Figure 4.6: Gene and protein structure of *Pp3c8_19720* from wild type and *snog1a*. The gene sequence consists of 3 exons and two introns. The location of the two UV-induced mutations found in the *snog1a Pp3c8_19720* sequence are highlighted with asterisks. The wild type protein is 624 amino acids in length. The *snog1a* protein is severely truncated, only 335 amino acids in length.

I also overexpressed the *Pp3c8_19720*-GFP fusion protein in wild type, driven by the maize ubiquitin promoter. Again, the presence of the *Pp3c8_19720* CDS in the genomic DNA of wild type *pUBI::Pp3c8_19720-GFP* confirmed a successful transformation (Figure 4.7A and B). Overexpression of *Pp3c8_19720*-GFP in wild type did not produce an obvious change in the phenotype (Figure 4.7F and G), but a closer examination and quantitative analysis is required to ascertain whether this is indeed the case. In this line, *Pp3c8_19720* was driven by the maize ubiquitin promoter which generates very high overexpression, whereas in the *snog1a* complementation line, *Pp3c8_19720* was driven by the rice actin promoter, which is less highly expressing. The reason for this was that I was initially using the construct with the maize ubiquitin promoter to transform both wild type and *snog1a*, but after many rounds of transformation I had only obtained a successful transformant

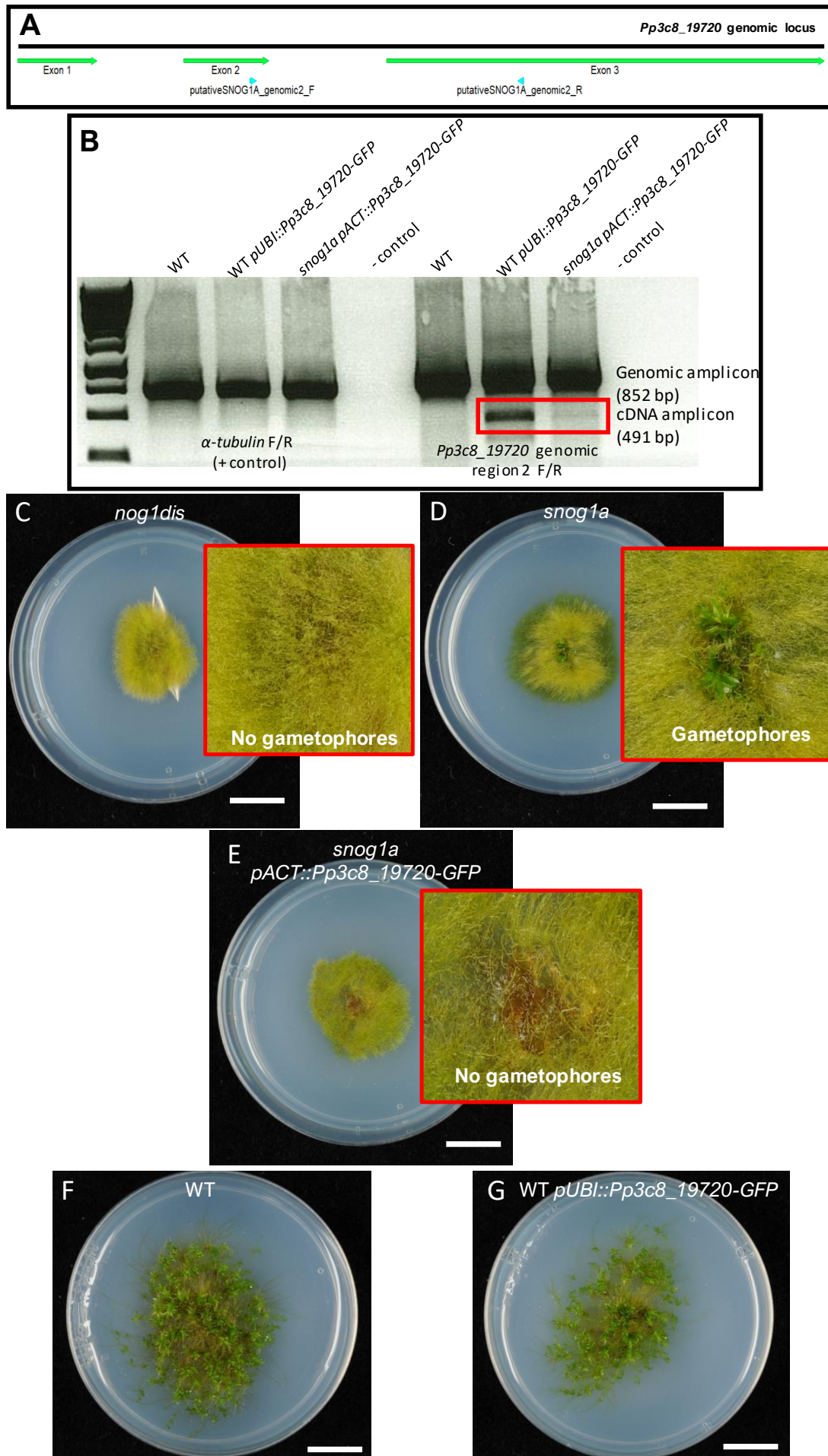


Figure 4.7: A) Genotyping strategy for lines overexpressing *Pp3c8_19720*. B) Genotyping result for wild type *pUBI::Pp3c8_19720-GFP* and *snog1a pACT::Pp3c8_19720-GFP*, confirming the presence of the *Pp3c8_19720* CDS in genomic DNA extracted from those lines. C-G) Representative images of *P. patens* lines. Scale = 1cm. C) *nog1dis*. D) *snog1a*. E) *snog1a pACT::Pp3c8_19720-GFP*. C-E were 1 month old when photographed. This is sooner than I would usually photograph *snog1a* as it is slower to develop gametophores than wild type. However, it was necessary due to time constraints. F) wild type. G) wild type *pUBI::Pp3c8_19720-GFP*. F and G were 5 weeks old when photographed.

in the wild type background. In case this low transformation efficiency was due to the effects of the very strong ubiquitin promoter, I redesigned the construct with the less strong actin promoter. I then used the actin promoter construct for further transformations, yielding *snog1a pACT::Pp3c8_19720-GFP*. I examined both *snog1a pACT::Pp3c8_19720-GFP* and wild type *pUBI::Pp3c8_19720-GFP* with a confocal microscope, but was unable to detect Pp3c8_19720-GFP in any tissues.

4.2.6 Deletion of *Pp3c8_19720* in *nog1dis* recapitulates the *snog1a* phenotype

To provide more evidence that the correct gene had been identified, I generated two lines in which *Pp3c8_19720* was deleted in *nog1dis*. To confirm successful deletion of the *Pp3c8_19720* open reading frame, I PCR-amplified a region of genomic DNA to check for successful integration at the targeted locus (Figure 4.8A). This produced a band of the correct size (approximately 2kb) for both of the lines (Figure 4.8B). However, as there was also some non-specific amplification I sequenced the 2kb amplicon, which confirmed that it was the targeted region (Figure S28). Thus, I confirmed that the deletion construct had integrated at the correct locus and deleted the *Pp3c8_19720* open reading frame. As expected, *nog1dis snog1a del-4* and *nog1dis snog1a del-6* did produce gametophores and therefore recapitulated the phenotype of the *snog1a* mutant (Figure 4.8C-F). Unfortunately, I did not have sufficient time to verify that the *PpSNOG1A* transcript is not expressed in these lines, but given the convincing phenotype, they are likely to be genuine.

I have demonstrated by complementation and by recapitulation that the

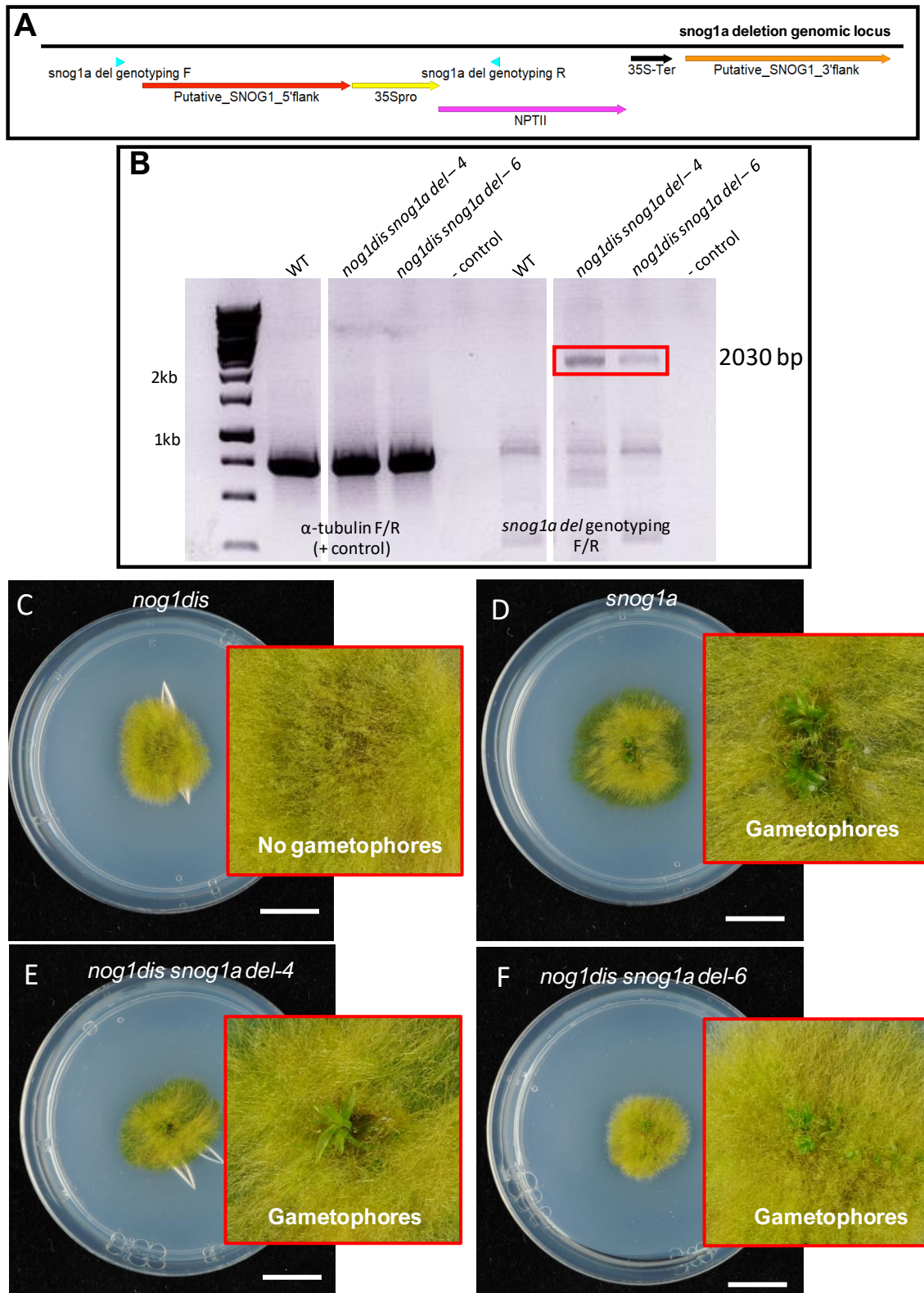


Figure 4.8: A) Genotyping strategy for *snog1a* deletion lines. B) Genotyping result for *nog1dis snog1a del-4* and *nog1dis snog1a del-6* showing successful integration of the deletion construct. Representative images of *P. patens* lines. Scale = 1cm. C) *nog1dis*. D) *snog1a*. E) *nog1dis snog1a del-4*. F) *nog1dis snog1a del-6*. C-F were 1 month old when photographed. This is sooner than I would usually photograph *snog1a* and the recapitulation mutants as they are slower to develop gametophores than wild type. However, it was necessary due to time constraints.

causative mutations in *snog1a* were those in *Pp3c8_19720*. Therefore, I could confidently ascribe the name *SUPPRESSOR OF NOG1 A (PpSNOG1A)* to *Pp3c8_19720*.

4.2.7 SUPPRESSOR OF NOG1 A is an uncharacterised protein that shares some structural similarities with PpNOG1

SNOG1A is an uncharacterised protein that contains 624 amino acid residues. Automatic annotation of the protein sequence in publicly available databases (<https://www.uniprot.org/uniprotkb/A0A2K1K818/entry>) indicates that SNOG1A has a C-terminal ubiquitin-associated (UBA) domain at position 561-605. This is interesting because NOG1 is also an uncharacterised protein with a UBA domain at the C-terminus (Moody et al., 2018). SNOG1A also has several automatically annotated disordered domains, the structure of which cannot be predicted with confidence. PpSNOG1A is about three times the length of PpNOG1 and has a long coiled-coil in the middle of the sequence whereas PpNOG1 does not (Figure 4.9).

As PpSNOG1A has a C-terminal UBA domain, I set out to determine whether PpSNOG1A can physically interact with polyubiquitin, using yeast two-hybrid. I paired PpSNOG1A with PpUBQ3A, PpUBQ3B and PpUBQ3C. Some UBA domains dimerise so I also tested whether PpSNOG1A could interact with itself or PpNOG1 (Long et al., 2010). The results of this showed only a potential weak interaction between PpSNOG1A and PpUBQ3A (Figure 4.10).

To detect the PpSNOG1A-GFP fusion protein, I performed a western blot using protein extracts isolated from the wild type *pUBI::Pp3c8_19720-GFP* (wild type *pUBI::PpSNOG1A-GFP*) line. To investigate whether PpSNOG1A was degraded by the proteasome, I exposed the plant tissues to either a mock treatment or 50 μ M of the proteasome inhibitor bortezomib. Two bands were detected using an anti-GFP antibody, but neither band corresponded to the expected size of a PpSNOG1A-GFP fusion protein (GFP is 27 kDa and PpSNOG1A is 68 kDa, total

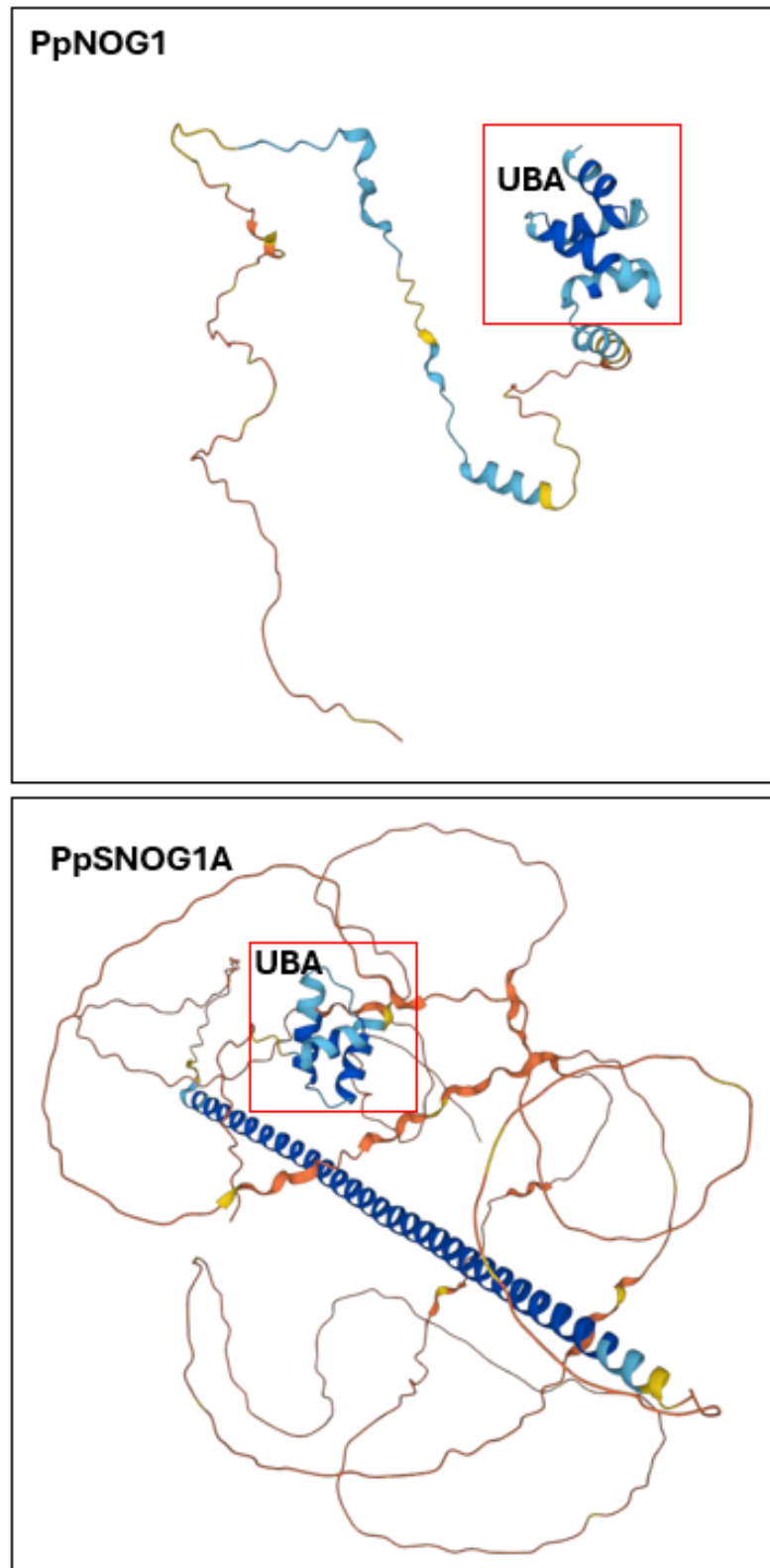


Figure 4.9: AlphaFold prediction of PpNOG1 (Uniprot ID A9SZ24) and PpSNOG1A (Uniprot ID A0A2K1K818) protein structure. UBA domain is highlighted. Dark blue indicates a very high model confidence (pLDDT > 90). Light blue indicates high model confidence (90 > pLDDT > 70). Yellow indicates low model confidence Low (70 > pLDDT > 50) and orange indicates very low confidence (pLDDT < 50).

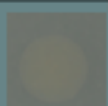
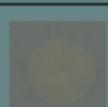
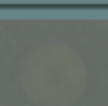
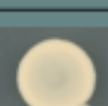
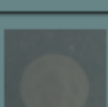
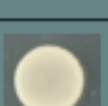
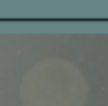
pGADT7 AD (prey)	pGBKT7 (bait)	SD- Leu/Trp	SD- Ade/His/ Leu/Trp
empty	empty		
<i>PpSNOG1A</i>	empty		
empty	<i>PpUBQ3A</i>		
empty	<i>PpUBQ3B</i>		
empty	<i>PpUBQ3C</i>		
empty	<i>PpSNOG1A</i>		
empty	<i>PpNOG1</i>		
<i>PpSNOG1A</i>	<i>PpUBQ3A</i>		
<i>PpSNOG1A</i>	<i>PpUBQ3B</i>		
<i>PpSNOG1A</i>	<i>PpUBQ3C</i>		
<i>PpSNOG1A</i>	<i>PpSNOG1A</i>		
<i>PpSNOG1A</i>	<i>PpNOG1</i>		

Figure 4.10: Yeast 2-hybrid assay results. Protein interactions are indicated by growth of the cotransformed yeast on SD-Ade/His/Leu/Trp.

= 95 kDa). Instead, a faint band of around 32-36 kDa and a much stronger band of around 18-23 kDa were present. GFP is 27 kDa and the smallest protein detected on the PpSNOG1A-GFP western blot was 18-23 kDa. This protein must at least contain the GFP epitope that the anti-GFP antibody recognises, suggesting that this detected protein is a fragment of GFP. The 32-36 kDa protein was larger than GFP, but not the size of the full fusion protein. There is no indication from current gene models that this size difference may be due to alternative splicing. As GFP is 27 kDa and the UBA domain alone is around 7 kDa, the 32-36 kDa peptide is consistent with a peptide consisting of GFP plus the UBA domain, give or take. There was some change in the intensity of the bands between the mock and bortezomib treated tissues, indicating that these fragments of PpSNOG1A-GFP are likely degraded by the proteasome (Figure 4.11).

4.2.8 PpSNOG1A protein homologues are found across land plants and green algae

Using Orthofinder, I identified protein homologues of PpSNOG1A (Emms and Kelly, 2019). I found them in plants and green algae (both charophytes and chlorophytes), but not in the yeast or animal genomes I interrogated (*Saccharomyces cerevisiae*, *Drosophila melanogaster* and *Homo sapiens*). There were also no PpSNOG1A homologues in the brown alga *Porphyra umbilicalis*. The green algae species I used were 7 chlorophytes (*Botryococcus braunii*, *Chlamydomonas reinhardtii*, *Coccomyxa subellipsoidea*, *Micromonas pusilla*, *Volvox carteri* and *Dunaliella salina*) and 6 charophytes (*Chara braunii*, *Chlorokybus atmophyticus*, *Klebsormidium nitens*, *Spirogloea muscicola*, *Mesotaenium endlicherianum* and *Mesostigma viride*). I found PpSNOG1A homologues in all of these green algal genomes except *C. subellipsoidea* and *M. viride*.

There are four homologues of PpSNOG1A in *P.patens*, but only one detected in the liverwort *Marchantia polymorpha*, which seems to indicate an expansion of the gene family in the lineage leading to *P. patens*. Interestingly, I could find no PpSNOG1A homologues in the hornwort *Anthoceros agrestis*. Apart from *A.*

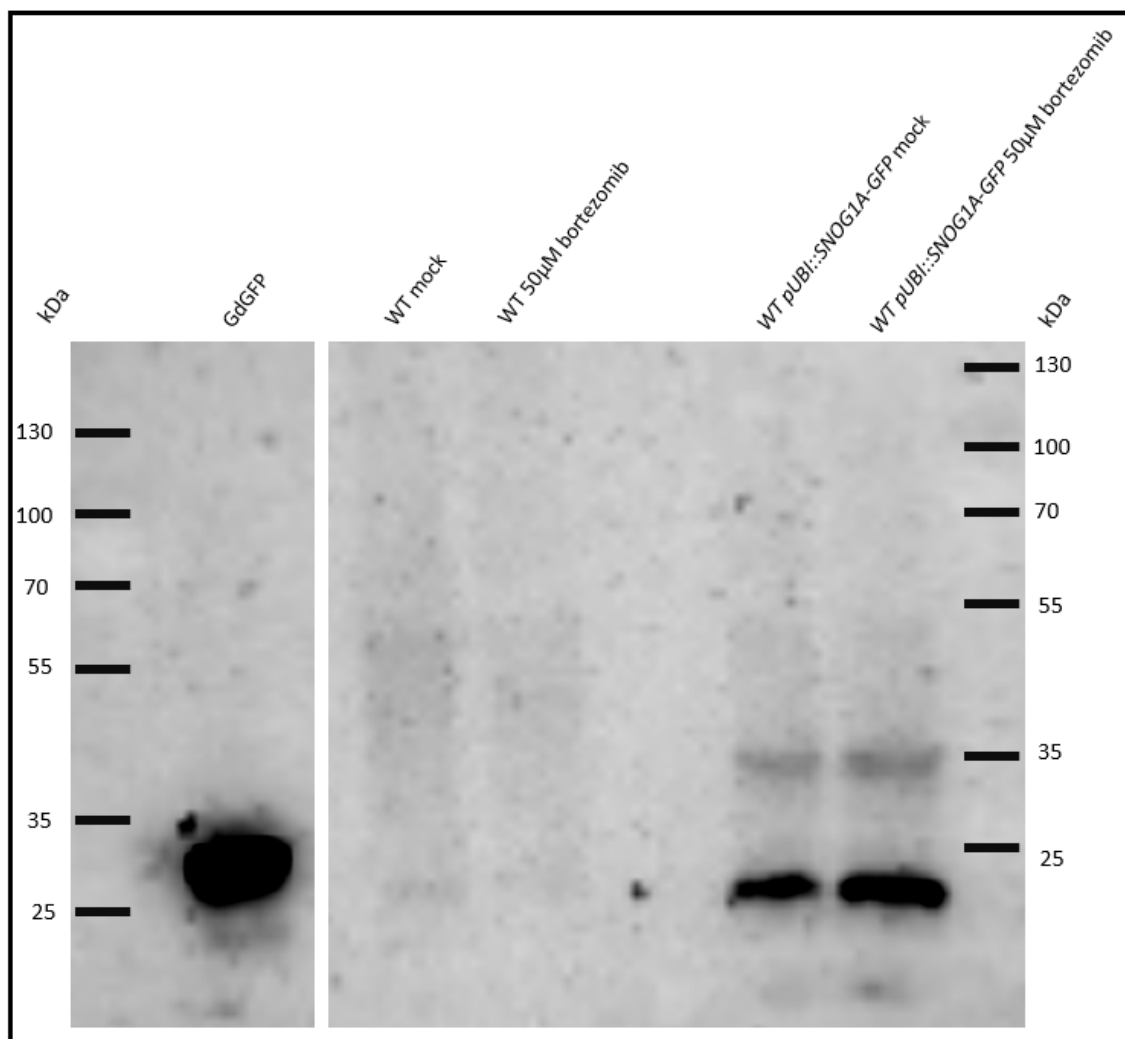


Figure 4.11: western blot to detect PpSNOG1A-GFP +/- treatment with 50 μ M bortezomib. Note that the right side of the gel ran a little slower than the left and the ladders are different.

agrestis, I found PpSNOG1A homologues in every plant species that I included in the analysis (Figure 4.12).

I identified three Arabidopsis proteins that are homologous to PpSNOG1A: AT4G28300, AT5G14540 and AT3G01560. A recent study by Dorone et al. (2021) characterised AT4G28300, which they named AtFLOE1. AT5G14540 and AT3G01560 were designated AtFLOE2 and AtFLOE3, respectively. AtFLOE1 is prion-like protein that can undergo hydration-dependent phase separation to regulate seed germination. AtFLOE2 and AtFLOE3 were also shown to undergo phase separation when heterologously expressed in tobacco, but their functions remain unknown (Dorone et al., 2021). Their phylogenetic analysis agreed with my own in that the FLOE/SNOG1A protein family consists of two large clades.

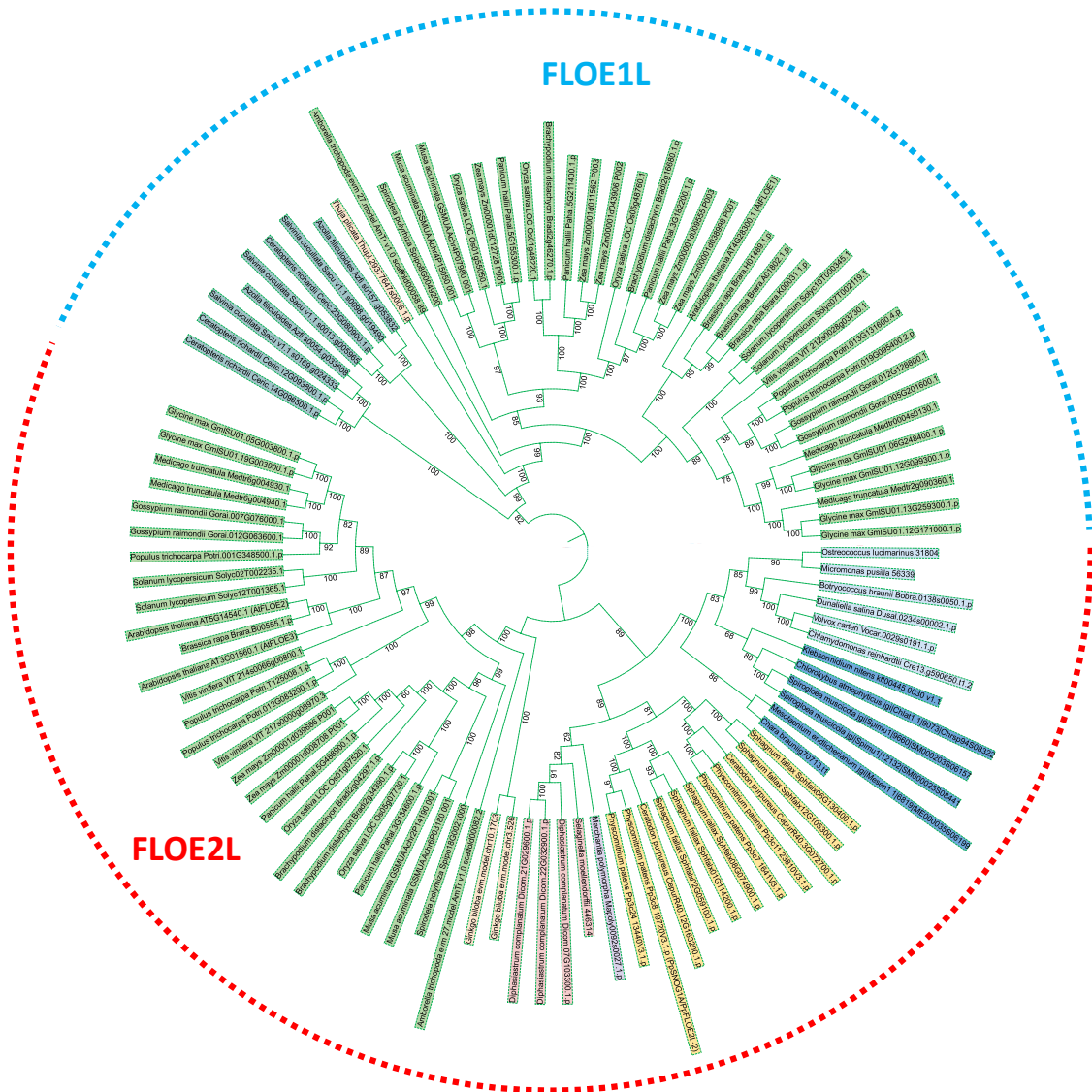


Figure 4.12: Phylogenetic tree based on a multiple sequence alignment of SNOG1A-like protein homologues. Numbers on branches are bootstrap values. Light blue = chlorophytes. Dark blue = charophytes. Yellow = mosses. Purple = liverworts. Teal = ferns. Pink = lycophytes. Orange = gymnosperms. Green = angiosperms. The gene family forms two clades which Dorone et al. (2021) classified as FLOE1-like (FLOE1L; cyan) and FLOE2-like (FLOE2L; red).

Dorone et al. (2021) found that the clade containing AtFLOE1, which they called the FLOE1-like clade, was restricted to seed plants. However, their analysis did not contain any fern species, whereas in my analysis, the fern homologues were found in the FLOE1-like (FLOE1L) clade. Both our analyses found that the clade containing AtFLOE2 and AtFLOE3 (the FLOE2L clade) contained homologues from seed plants, non-seed plants (including PpSNOG1A) and the green algae (Figure 4.12).

The AtFLOE1, AtFLOE2 and AtFLOE3 proteins contain a predicted folded domain (known as the nucleation domain), a coiled-coil domain, a UBA domain at the C-terminus and two disordered regions; one enriched in aspartic acid and serine (DS-rich domain) and one enriched in glutamine, proline, and serine (QPS-rich domain; Dorone et al., 2021). Alignment of PpSNOG1A with AtFLOE1, 2 and 3 confirmed that these protein domains are conserved in PpSNOG1A (Figure 4.13). Therefore, the truncated *SNOG1A* protein that is expressed in *snog1a* is missing not only its UBA domain, but also a large portion of the QPS-rich disorder domain (Figure 4.14). Dorone et al. (2021) reported that in AtFLOE1, the QPS-rich domain is required for the formation of phase-separated protein condensates.

4.3 Discussion

The use of forward genetics mutagenesis screens has once again proven useful in identifying novel regulators of the transition to 3D growth. After generating an independent *nog1* mutant, called *nog1dis*, a suppressor screen produced two ‘*suppressor of nog1*’ mutants, *snog1a* and *snog1b*. Both mutants had regained the ability to produce gametophores. In *snog1a*, the gametophores were stunted but proved to be reproductively fertile, whereas the gametophores of *snog1b*, were severely stunted and sterile. Therefore I focused on *snog1a* and proceeded to map the causative mutation using bulk segregant analysis. The results of this were very exciting as a candidate gene (*Pp3c8_19720*) was revealed with two introduced premature STOP codons in the coding sequence, which resulted in a truncated protein product. Complementation and recapitulation lines were generated which

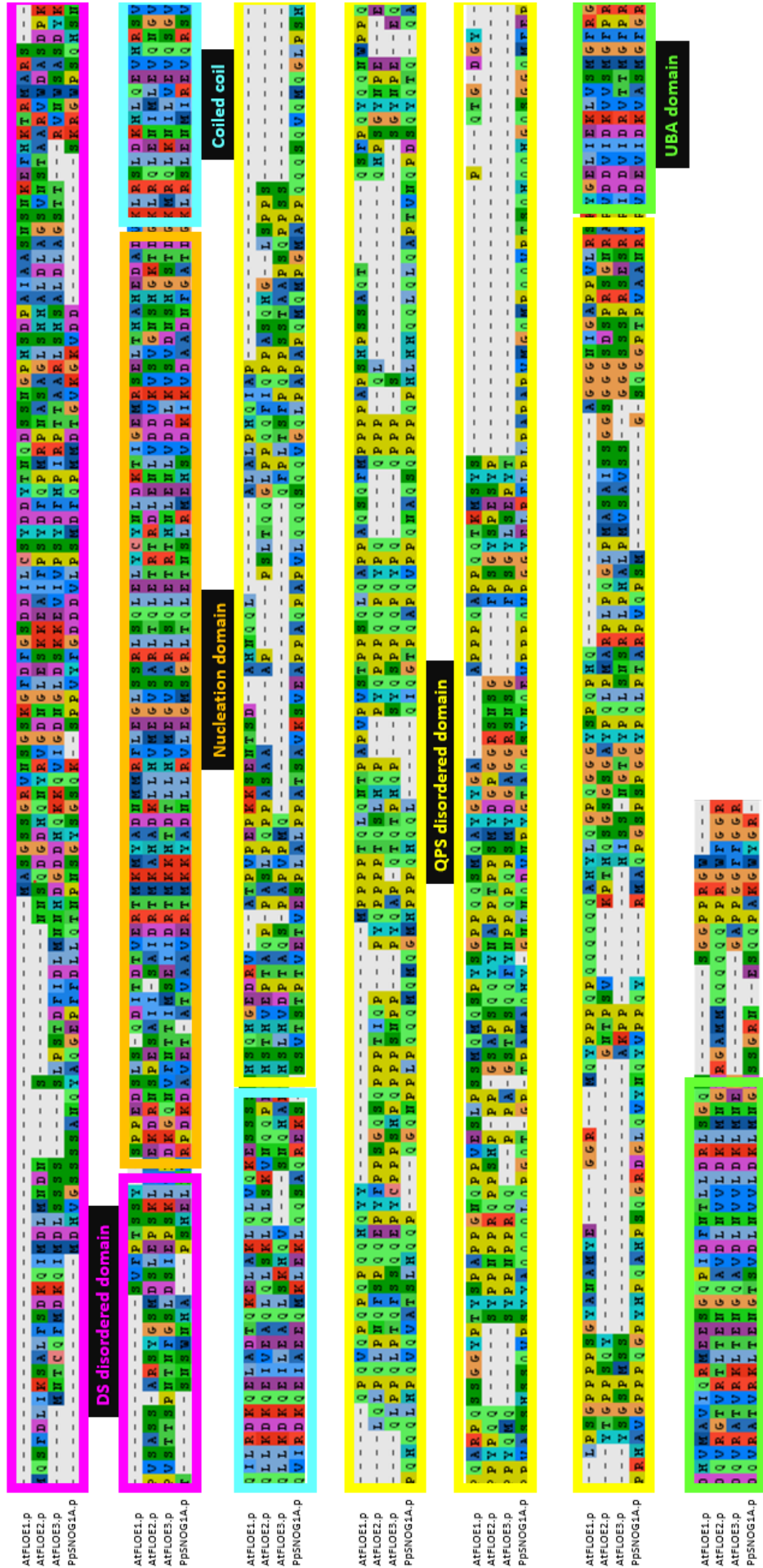


Figure 4.13: Protein alignment of AtFLOE1, AtFLOE2, AtFLOE3 and PpSNOG1A/Pp3c8_19720, showing conservation of domain structure (Dorone et al., 2021). The domains highlighted are: DS-rich disordered domain (magenta), nucleation domain (orange), coiled coil (cyan), QPS-rich disordered domain (yellow) and the UBA domain (green).

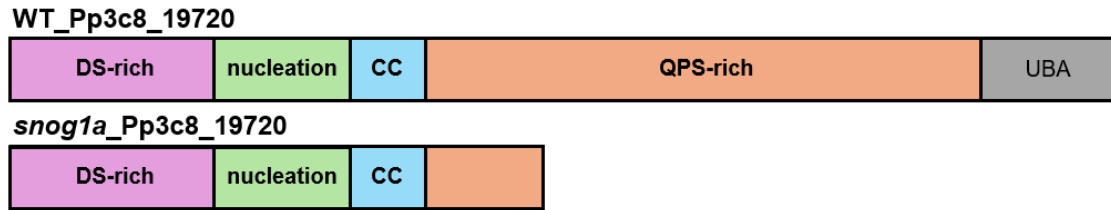


Figure 4.14: Diagram showing the domain architecture of PpSNOG1A (Pp3c8_19720): DS-rich domain, nucleation domain, coiled coil, QPS-rich domain and ubiquitin associated domain (UBA). In *snog1a*, the UBA and a large portion of the QPS-rich domain are lost from the C-terminus, due to the presence of two UV-induced premature STOP codons.

confirmed that *Pp3c8_19720*, now dubbed *PpSNOG1A*, was the gene responsible for the *snog1a* phenotype. Therefore *PpSNOG1A* is an inhibitor of the transition to 3D growth and antagonises the action of *PpNOG1*, which positively regulates 3D growth.

One of the things that is particularly interesting about PpSNOG1A, is that it is remarkably similar to PpNOG1, although they are not related. Firstly, they share some common structural features, most notably a C-terminal UBA domain. The presence of the UBA domain makes it likely that both PpNOG1 and PpSNOG1A function in protein degradation processes, as this is a typical role for a UBA domain-containing protein (Rao and Sastry, 2002; Raasi et al., 2005; Su and Lau, 2009). As such, I used yeast two-hybrid to test whether *PpSNOG1A* can interact with PpUBQ3 A, B or C (polyubiquitin genes). The result was a potential weak interaction with PpUBQ3A only. This shows that PpSNOG1A can possibly interact with ubiquitin *in vivo*. However, as discussed in the previous chapter, there were some flaws to this approach as UBA domains tend to have stronger affinities for specific ubiquitin chain topologies (Raasi et al., 2005). The *PpUBQ3* genes used for the yeast two-hybrid code for polypeptides containing tandem ubiquitin repeats, which does not have the same structure as a polyubiquitin chain assembled out of covalently linked ubiquitin monomers (Callis et al., 1995). I had hypothesised that PpNOG1 and PpSNOG1A may interact via their UBA domain, as UBA homodimerisation and heterodimerisation has been reported (Long et al., 2010).

This kind of relationship may have explained their seemingly antagonistic roles, for example if they mutually inhibited each other by dimerisation of their UBA domains, resulting in occlusion of the ubiquitin-binding site. However, the yeast two-hybrid results indicated that PpNOG1 does not interact directly with PpSNOG1A and neither does PpSNOG1A homodimerise with itself.

When I attempted to detect the PpSNOG1A-GFP fusion protein on a western blot, using an anti-GFP antibody, I detected two peptides. Neither was the expected molecular weight of the full PpSNOG1A-GFP fusion protein (95 kDa). One peptide was approximately 32-36 kDa, likely GFP attached to a portion of PpSNOG1A containing the UBA domain, and the other peptide was approximately 18-23 kDa, likely a cleaved fragment of GFP. What is very interesting about this is that I got a very similar result when I attempted to detect PpNOG1-GFP in the same way. In that case I also only detected a roughly 18-23 kDa protein and a roughly 28-34 kDa protein, but there was no protein detected that was consistent with the molecular weight of PpNOG1-GFP (50 kDa). The 28-34 kDa portion of PpNOG1-GFP that was detected seemed slightly smaller than the 32-36 kDa fragment of PpSNOG1A-GFP that was detected on a separate western blot. It would be useful to run both samples on the same gel to directly compare the size of this fragment. If they are indeed different sizes, that suggests that they are being cleaved at different points in the protein, relative to the start of the UBA domain, which is approximately equidistant from the C-terminus of each protein.

When trying to detect PpSNOG1A-GFP, the majority of detected protein was the GFP fragment, suggesting that the C-terminal GFP is proteolytically cleaved soon after the synthesis of the protein. This may explain why I was unable to visualise PpSNOG1A-GFP, using fluorescence microscopy. One difference between PpNOG1-GFP and PpSNOG1A-GFP was that for PpSNOG1A-GFP the amount of detectable protein seemed to be only slightly increased by treatment with 50 μ M bortezomib, suggesting a possible minor role of the proteasome in degrading PpSNOG1A. As with PpNOG1, these results suggest the action of a proteasome-independent proteolytic process that is acting on PpSNOG1A, to

produce these 18-23 kDa and 32-36 kDa fragments of PpSNOG1A-GFP.

I identified homologues of the PpSNOG1A protein across land plants and green algae, but not in other eukaryotes. The only land plant that did not have a PpSNOG1A homologue was *Anthoceros agrestis*. Homologues were also present in all green algal species examined except *C. subellipsoidea* and *M. viride*. This could indicate loss of *PpSNOG1A* genes in this particular lineages or it may be due to incomplete assembly or annotation of those genomes. The presence of PpNOG1 and PpSNOG1A in many algal species, indicates that genes involved in the regulation of 3D growth were already present prior to the colonisation of land and were likely co-opted from other functions.

I also identified the three Arabidopsis protein homologues (named AtFLOE1, 2 and 3) which Dorone et al. (2021) have recently investigated. In the Arabidopsis embryo, AtFLOE1 undergoes reversible liquid-liquid phase separation (LLPS) in response to water availability. When the seed is hydrated, AtFLOE1 forms phase separated protein condensates, but when the seed is dessicated, the protein is dispersed. An *Atfloe1* knockout line is more likely to germinate when the seed is experiencing osmotic stress. Therefore, AtFLOE1 acts to inhibit germination in unfavourable water conditions. This role was shown to be linked to its ability to form condensates. The functional relevance of each of the AtFLOE1 domains was tested by expressing domain deletion variants of the protein. When expressing these protein variants heterologously in tobacco, it was observed that loss of the UBA domain did not affect the ability of AtFLOE1 to form condensates. However, loss of the QPS-rich disorder domain resulted in an inability of the protein to form condensates. Conversely loss of the N-terminal DS-rich disorder domain resulted in constitutive formation of condensates under any hydration conditions. Therefore, these disorder domains are important for the liquid-liquid phase separation (LLPS) properties of the protein, whereas the UBA domain is not. When expressing the protein domain variants in the *Atfloe1* mutant, the variant missing the UBA domain completely complemented the mutant, indicating that the UBA domain is not important for the role of AtFLOE1 in regulating germination. In contrast the variant lacking

the QPS-rich disorder domain was unable to complement the mutant, whereas the variant lacking the DS-rich disorder domain exhibited much more severe mutant phenotype, with a much higher rate of germination under osmotic stress than the *Atfloe1* mutant. Therefore the ability to reversibly undergo LLPS is vital for the function of AtFLOE1 (Dorone et al., 2021).

I must consider what this work done in Arabidopsis might be able to tell me about the function of PpSNOG1A in *P. patens*. There is great evolutionary distance between the two and AtFLOE1 is functioning in the sporophyte embryo/seed, whereas I am studying the function of PpSNOG1A in the gametophyte. Additionally AtFLOE1 is in the FLOE1-like clade whereas PpSNOG1A is in the FLOE2-like clade. The functions of AtFLOE2 and AtFLOE3 have not been investigated although they have been shown to undergo LLPS. Indeed many representatives of both clades across the land plants and green algae have been shown to undergo LLPS when heterologously expressed in tobacco leaves. This includes a *P. patens* paralogue of PpSNOG1A; Pp3c11_23810 (Dorone et al., 2021).

The question remains as to whether the *snog1a* mutant phenotype is caused by the loss of the PpSNOG1A UBA domain or a large portion of the QPS-rich domain (or both). The evidence from AtFLOE1 suggested that the QPS-rich domain was more important for function than the UBA domain, in that particular context. My data indicated that the PpSNOG1A UBA domain can interact with polyubiquitin, but it is not yet known whether it does in fact interact with polyubiquitinated substrates *in vivo*, and what significance that may have for its function. I was also unable to verify whether PpSNOG1A forms condensates in *P. patens* cells. My PpSNOG1A-GFP expressing lines did not display any convincing signal for either dispersed or condensed protein. This could be related to the result of the western blot, which seemed to indicate cleavage of the GFP. It is possible that proteolytic processing of the fusion protein is preventing its visualisation. However, whatever is happening to the PpSNOG1A-GFP does not appear to be affecting its function as PpSNOG1A-GFP was able to successfully complement the *snog1a* phenotype and cause a reversion to the ‘no gametophores’ phenotype. It would be informative

to express PpSNOG1A-GFP heterologously in tobacco leaves to see if the fusion protein can be visualised in that context, and whether or not it forms condensates. If this doesn't work perhaps an N-terminal GFP fusion would be more successful. Additionally it will be important to assess the roles of each of the domains. To that end I did design and assemble domain deletion constructs for expression in *P. patens*, but unfortunately did not have the time to successfully generate the lines in the *snog1a* background. I also assembled a construct for the heterologous expression of PpSNOG1A-GFP in tobacco leaves, although I did not have sufficient time to carry out the experiment.

One way to assess whether the functions of AtFLOE1 and PpSNOG1A are conserved would be to do cross-species complementation. It would also be interesting to try complementing *snog1a* with proteins from algal species to see if the function of this 3D regulator is the same in non-3D species. The presence of 4 closely related PpSNOG1A paralogues in *P. patens* might indicate that there is some functional redundancy between these proteins. The generation of single, double, triple and quadruple knockout mutants in the *snog1a* background would be informative in diagnosing whether this is the case.

Another question that should be addressed following on from this work is; by what mechanism do PpNOG1 and PpSNOG1A antagonise each other in the regulation of the transition to 3D growth? A direct interaction between PpNOG1 and PpSNOG1A was not supported by the yeast 2-hybrid assay. Therefore, this regulatory relationship seems to be an indirect one, with a myriad of potential mechanisms. The consistent similarities between PpNOG1 and PpSNOG1A are a very interesting result. Does the similarities in their architecture indicate a similar function? Is that function in the same pathway? Are they being acted on by the same proteolytic agent resulting in the similar patterns of protein fragments observed on the western blots. Both PpNOG1 and PpSNOG1A are challenging proteins to work with, due to their apparent instability. But their central role in this important developmental transition makes it a very interesting prospect for future work.

Chapter 5: RNA-seq experiments yield insights into gene expression in 3D-defective mutants.

5.1 Introduction

Previously, Frank and Scanlon (2015) conducted a cell type-specific transcriptomic analysis of the 2D-to-3D transition in *P. patens*. In this study, they compared the transcriptomes of protonemal tip cells and 4-cell buds. They reported that there were distinct gene expression profiles in the 2D protonemal tip cells versus the gametophore buds that were undergoing the transition to 3D growth. The protonema tip cells were enriched for transcripts of genes associated with photosynthesis, whereas the bud cells were enriched for transcripts associated with developmental patterning. In particular they found that bud cells had increased expression of genes related to the development of shoot apical meristems as well as the cytoskeleton, asymmetric cell division and cell wall remodelling (Frank and Scanlon, 2015). This trend is not surprising as the protonema carry out nutrient acquisition and photosynthesis, whereas the gametophore buds are undergoing active morphogenesis to develop the 3D shoot-like structures of the plant.

Targeted quantitative RT-qPCR experiments have demonstrated that transcript levels of known 3D regulatory genes, such as the *PpAPBs*, *CLV* genes and various auxin responsive genes are perturbed in *Ppnog1R* and/or *Ppnog2R* mutants, compared to wild type (Moody et al., 2018, 2021). Therefore, to gain insights into the regulatory mechanisms that underpin 3D growth, I conducted transcriptome analyses using available 3D-defective mutants *Ppnog1R* and *Ppnog2R*, as well as the

nog1 suppressor mutant, *snog1a*.

5.2 Results

5.2.1 The transcriptional response to cytokinin was severely perturbed in *Ppnog1R*

The first RNA-seq experiment was designed to compare wild type with *Ppnog1R*, with and without treatment with 1 μ M of the synthetic cytokinin 6-benzylaminopurine (BAP). In wild type, BAP treatment induces high levels of bud initiation, but *Ppnog1R* is insensitive to BAP treatment. In this experiment I wanted to examine the baseline transcriptional differences between wild type and the *Ppnog1R* mutant, but also investigate their differential responses to BAP to try and understand why *Ppnog1R* is unable to respond to BAP treatment.

Triplicate biological samples of two-week-old wild type and *Ppnog1R*, either untreated or treated with 1 μ M BAP for one week were harvested and RNA was extracted from the whole plant. After two weeks of growth on BCD medium, the tissue primarily consisted of protonema, but bud induction was underway. Therefore, this timepoint is ideal for studying the transcriptional regulation of the transition to 3D growth i.e., the initiation and early patterning of the gametophore bud, before the gametophores become established and mature and other developmental pathways become more dominant.

With the RNA-seq data, I performed differential gene expression (DGE) analysis for wild type versus *Ppnog1R* (Figure 5.1A), wild type cultured in the presence or absence of BAP (Figure 5.1B) and *Ppnog1R* cultured in the presence or absence of BAP (Figure 5.1C). In each comparison there was quite a substantial perturbation in the transcriptome. In the wild type versus *Ppnog1R* dataset, 19.24% of all genes were significantly differentially expressed ($\text{padj} < 0.05$). For wild type versus wild type + BAP, 17.74% of all genes were differentially expressed. The greatest perturbation was observed in the *Ppnog1R* versus *Ppnog1R* + BAP comparison, as 27.23% of

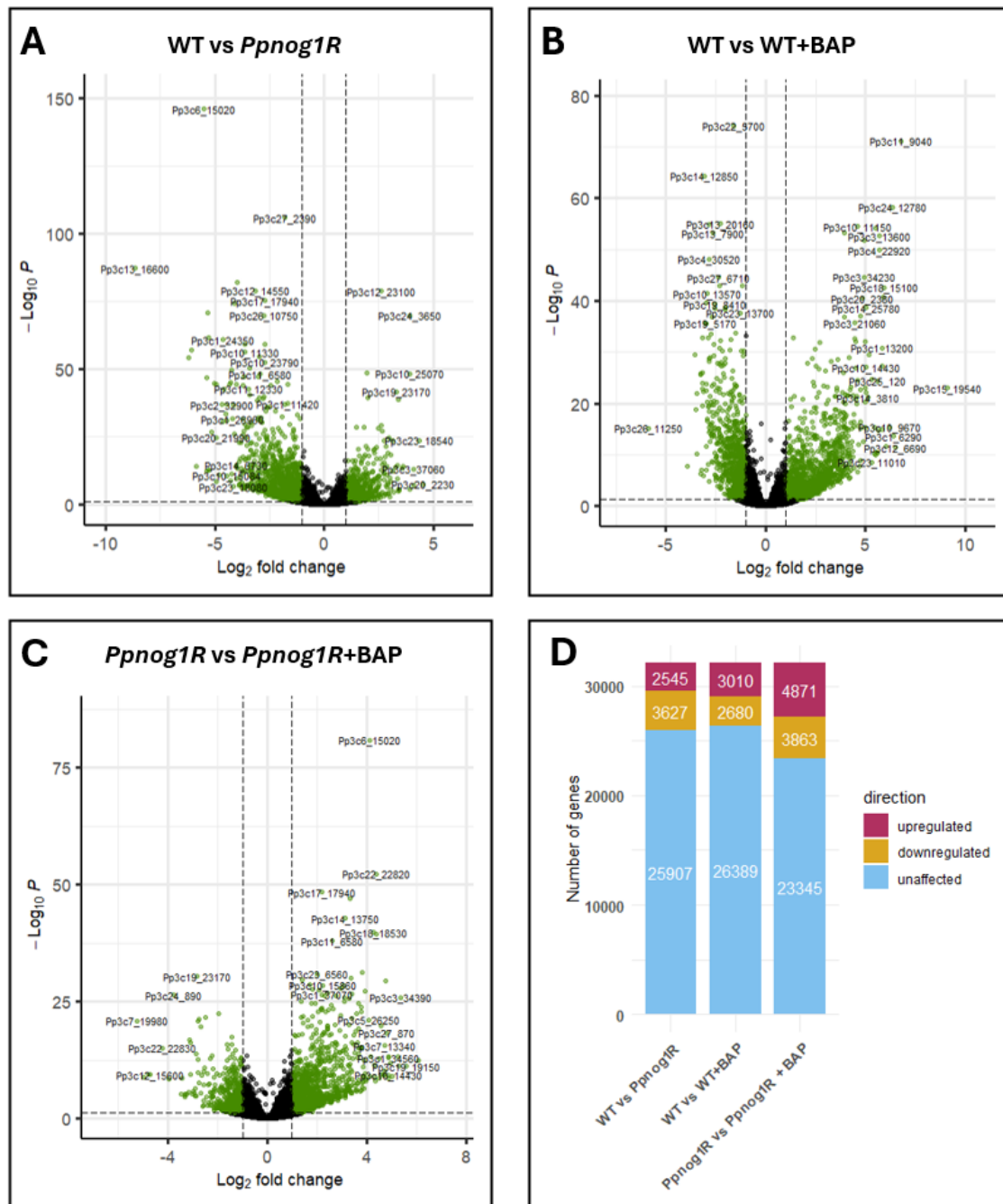


Figure 5.1: A-C) Volcano plot showing differential gene expression. Green points indicate that the gene was significantly differentially expressed (adjusted p-value < 0.05) with a Log₂ fold change of at least 1. Some strongly affected genes are labelled with their gene accession number. For each line comparison (eg. X vs Y), X is the baseline, and Y is the line for which gene expression changes are being shown. A) wild type versus *Ppnog1R*. B) wild type versus wild type + BAP. C) *Ppnog1R* versus *Ppnog1R* + BAP. D) Graph showing the number of unaffected, upregulated and downregulated genes for each comparison.

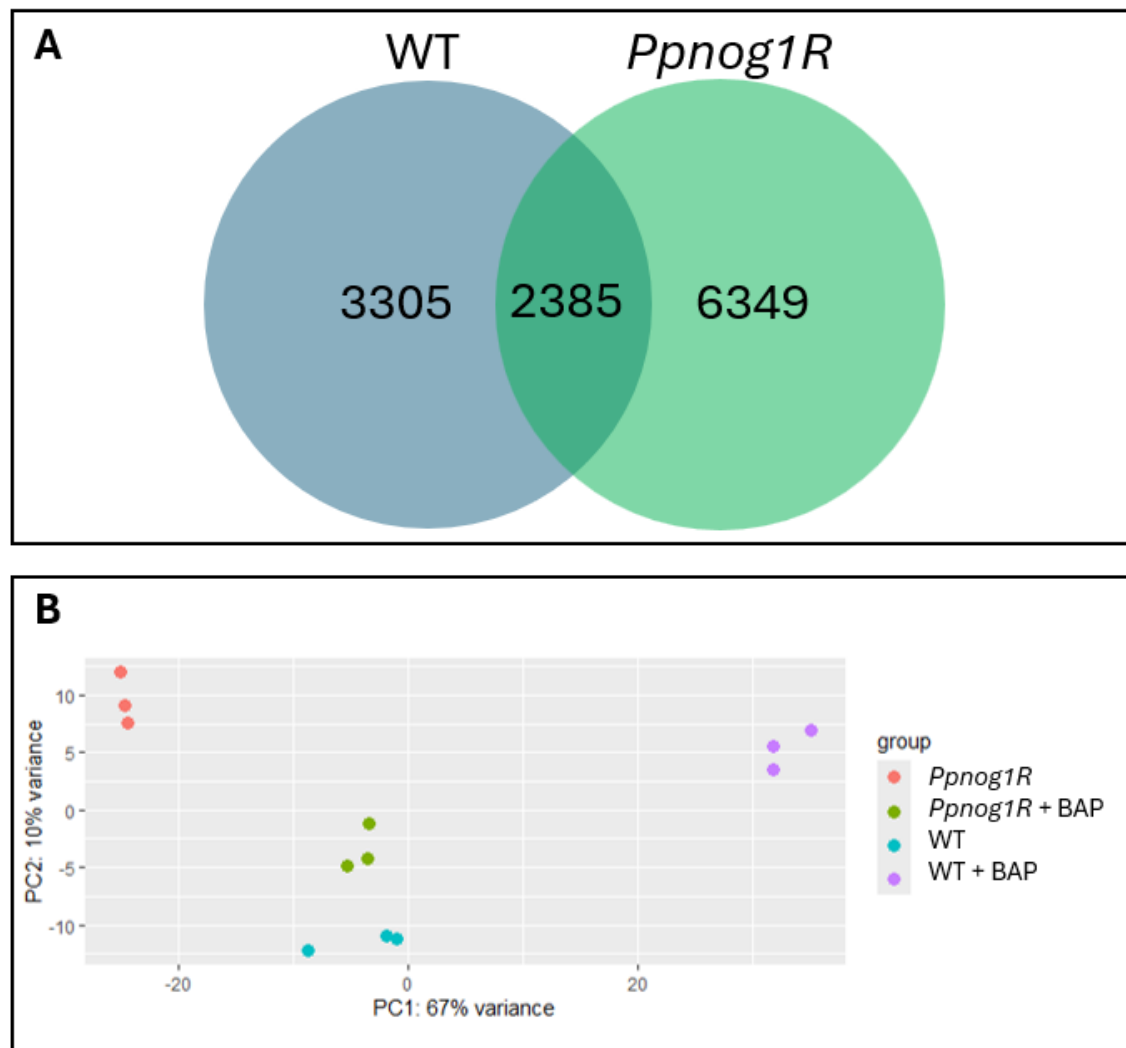


Figure 5.2: A) Venn diagram showing the number of shared and unique differentially expressed genes in response to BAP treatment for wild type and *Ppnog1R*. B) Principal Components Analysis (PCA) plot showing the relative similarity of the individual samples.

all genes were differentially expressed. In each case there were similar numbers of upregulated and downregulated genes indicating no overall trend of activation or suppression in any of the comparisons (Figure 5.1D).

In wild type, BAP treatment elicits a strong phenotypic response; the induction of excessive numbers of buds, which fail to develop normally but instead form a callus-like structure. With such a strong treatment (1 μ M BAP), almost every potential side branch acquires bud fate. Contrastingly, *Ppnog1R* exhibits no change in phenotype upon treatment, even when cultured in the presence of highly concentrated BAP. It was therefore interesting to note that at the transcriptional level, BAP induces much more differential gene expression in *Ppnog1R* than in wild

type. Additionally, when I compared the differentially expressed genes (DEGs) in response to BAP treatment in wild type and *Ppnog1R*, it revealed that the number of shared DEGs was 2385, whereas the number of DEGs responding to BAP only in wild type was 3305 and the number responding only in *Ppnog1R* was 6349. Therefore, the majority of genes responding to BAP in wild type and *Ppnog1R* were unique to each line (Figure 5.2A). This means that, contrary to what the phenotype suggests, *Ppnog1R* is not insensitive to BAP, but rather hypersensitive at the transcriptional level, as the number of BAP responsive genes was approximately 1/3 higher in *Ppnog1R* compared with wild type. Additionally, the identity of those BAP responsive genes was very distinct between wild type and *Ppnog1R*. This indicates that BAP was eliciting a very different response in *Ppnog1R* compared to wild type so the usual 3D-transition pathways were not being activated in *Ppnog1R*. When a principal components analysis was performed, the biological replicates clustered together, as expected. While wild type + BAP and *Ppnog1R* clustered at opposite ends of the principal component 1 axis, wild type and *Ppnog1R* + BAP clustered quite close together, indicating that they are more transcriptionally similar to each other than either was to their treated or untreated partner respectively (Figure 5.2B).

To identify biological processes that were enriched in the differentially expressed genes, Gene Set Enrichment Analysis (GSEA) was performed using Gene Ontology (GO) biological processes. To do this, BLAST searches were carried out for every *P. patens* gene against Arabidopsis and the best hit retrieved. This meant the well-annotated Arabidopsis gene set database could be used for GSEA. In genes that were upregulated in *Ppnog1R* compared to wild type, the biological processes that were enriched were those associated with developmental transitions, as well as genes associated with transcription and responses to osmotic stress. Processes associated with the synthesis of secondary metabolites, such as terpenes, carotene and xanthophyll, were all suppressed in *Ppnog1R* relative to wild type. Also suppressed were genes associated with regulation of ubiquitin protein ligase activity (Figure 5.3A).

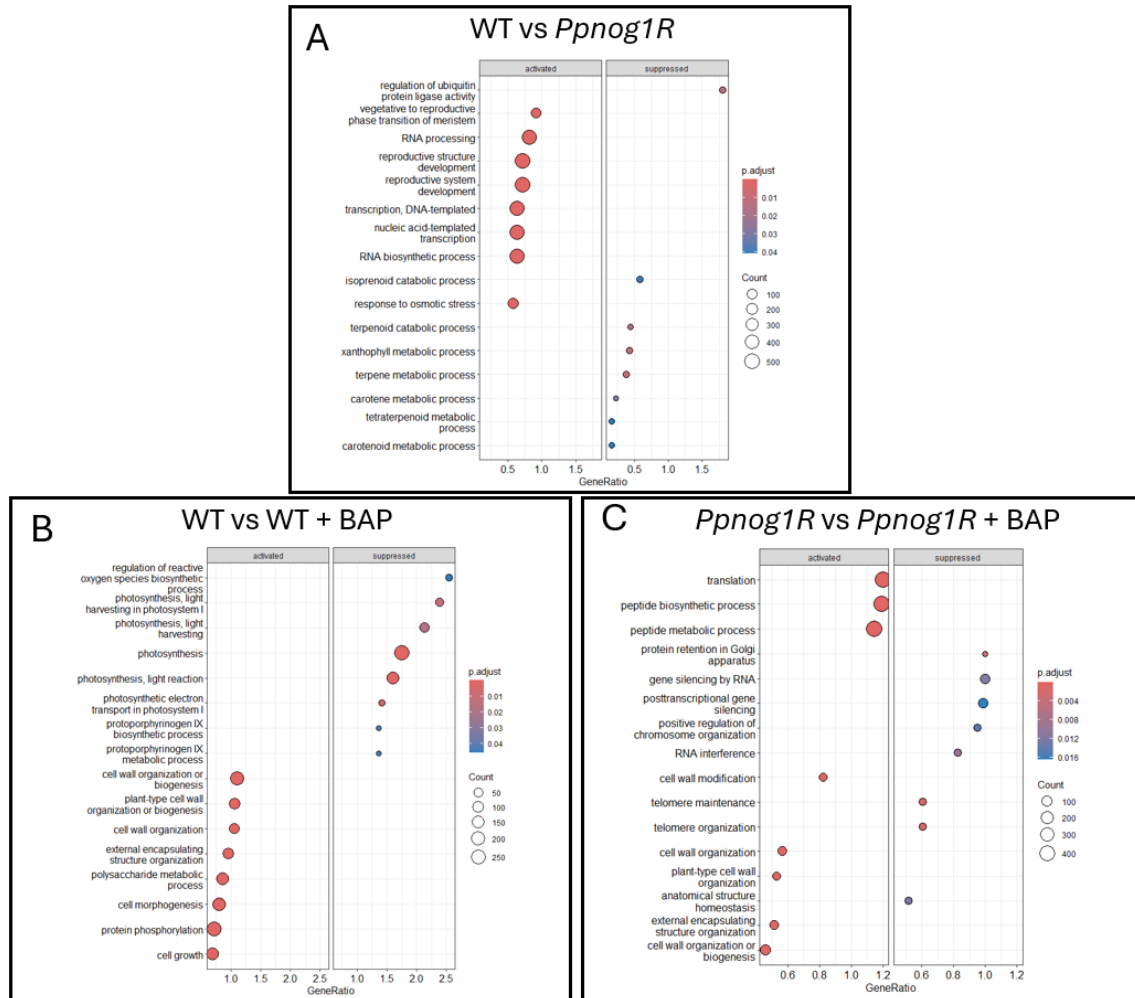


Figure 5.3: Gene Set Enrichment Analysis (GSEA) with Gene Ontology (GO) biological process terms. For each comparison there are 8 of the most significant activated and suppressed biological processes. The colour indicates the adjusted p-value for the test for enrichment. Count indicates the number of input genes. Gene ratio is the ratio of input genes to the total number of genes in the gene set. Note that because *P. patens* genes were BLASTed against *A. thaliana* for this analysis, there were cases where *P. patens* had multiple genes mapping to the same Arabidopsis gene, resulting in some gene ratios exceeding 1. For each line comparison (eg. X vs Y), X is the baseline, and Y is the line for which gene expression changes are being shown. A) wild type versus *Ppnog1R*. B) wild type versus wild type + BAP. C) *Ppnog1R* versus *Ppnog1R* + BAP.

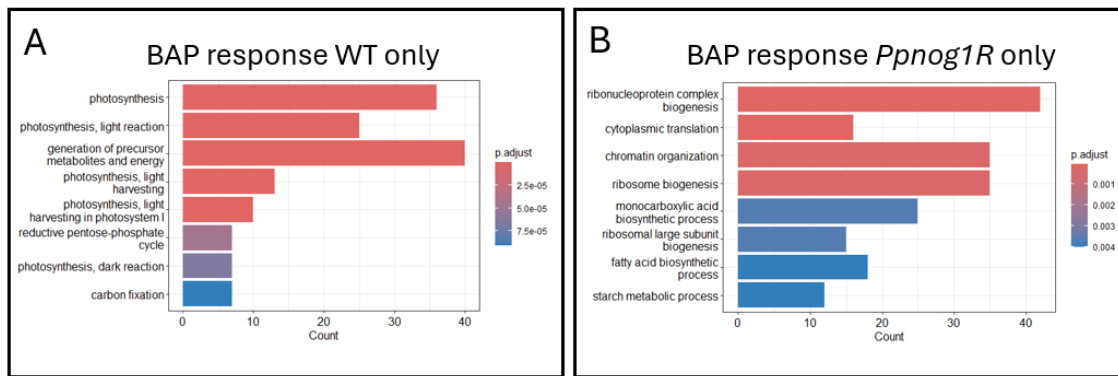


Figure 5.4: Over-representation Analysis (ORA) with GO biological process terms. Colour indicates the adjusted P value for statistical over-representation. Count is the number of input genes for each biological process. A) Genes that were differentially expressed in wild type but not *Ppnog1R* in response to BAP. B) Genes that were differentially expressed in *Ppnog1R* but not wild type in response to BAP.

When wild type was treated with BAP there was suppression of genes associated with photosynthesis. Conversely, genes associated with cell growth, cell wall synthesis, cell wall organisation and protein phosphorylation (often involved in signaling transduction) were activated (Figure 5.3B). This aligns with expected responses to cytokinin. There is a suppression of the photosynthetic activities of the protonema and a switch towards growth and morphogenesis ie. the induction of buds. Those buds do not grow normally as the excessive amount of cytokinin stimulates uncontrolled cell growth that produces callus-like buds.

In *Ppnog1R* treated with BAP, there was activation of processes associated with cell wall synthesis and cell wall remodelling. This was similar to wild type treated with BAP. Other than the cell wall, the processes affected in *Ppnog1R* treated with BAP differ to those in wild type treated with BAP. In BAP-treated *Ppnog1R*, genes associated with protein synthesis were activated. The expression of genes associated with chromosome maintenance were suppressed, as were genes associated with gene silencing and protein retention in the Golgi apparatus (Figure 5.3C).

The differences between processes affected in wild type and *Ppnog1R* reflects the observation that only a subset of genes were differentially expressed in response to cytokinin in both lines. That subset of genes was enriched in those genes associated with the cell wall synthesis processes that were upregulated in response to BAP in both *Ppnog1R* and wild type. I examined those genes that were differentially

expressed in only one line (either wild type or *Ppnog1R*, but not both) in response to BAP and performed Over-Representation Analysis (ORA) on these subsets. ORA was performed rather than GSEA as the lists had already been filtered for significant differential expression and GSEA must be used on ranked unfiltered DEG lists, whereas ORA is used with filtered lists. The subset of genes that only responded to BAP in wild type (upregulated or downregulated) was enriched in genes associated with photosynthesis (Figure 5.4A). The subset of genes that only responded to BAP treatment in *Ppnog1R* (upregulated or downregulated) was enriched in genes associated with ribonucleoprotein complex biogenesis, translation, chromatin organisation, and the metabolism of monocarboxylic acid, fatty acids and starch (Figure 5.4B). Therefore, the regulation of photosynthetic processes by BAP did not occur in *Ppnog1R* as it did in wild type.

The GSEA for wild type versus *Ppnog1R* highlighted that genes associated with the ubiquitin-proteasome system (UPS) were suppressed in *Ppnog1R*. This interested me because *PpNOG1* has a putative role in the UPS. When I investigated which genes were associated with that GO term, I found several that were associated with the Anaphase-Promoting Complex/Cyclosome (APC/C), an E3 RING-cullin ubiquitin ligase that catalyses the ubiquitination of cell-cycle regulatory proteins and targets them to the proteasome to bring about anaphase followed by exit from mitosis (Peters, 2006). I identified a homologue of the gene *CELL CYCLE SWITCH 52 A* (*CCS52A*; also known as *Cdh1*, *FIZZY-RELATED*, *TE* and *TAD1*), a component of the APC/C that functions in substrate recognition and binding (Xu et al., 2012; Lin et al., 2012). This gene was suppressed in *Ppnog1R* compared to wild type. In addition, the expression of this gene was induced by BAP in wild type but not in *Ppnog1R* (Figure 5.5). In rice, *CCS52A* binds to and ubiquitinates the gene *MONOCULM1* (*MOC1*), a positive regulator of tillering (Xu et al., 2012; Lin et al., 2012). The expression of two *MOC1*-related genes (also known as *LATERAL SUPPRESSOR* in Arabidopsis) was also suppressed in *Ppnog1R* (Figure 5.5). In Arabidopsis, *LATERAL SUPPRESSOR* positively regulates the formation of axillary meristems and therefore branching (Greb et al., 2003). Therefore, these

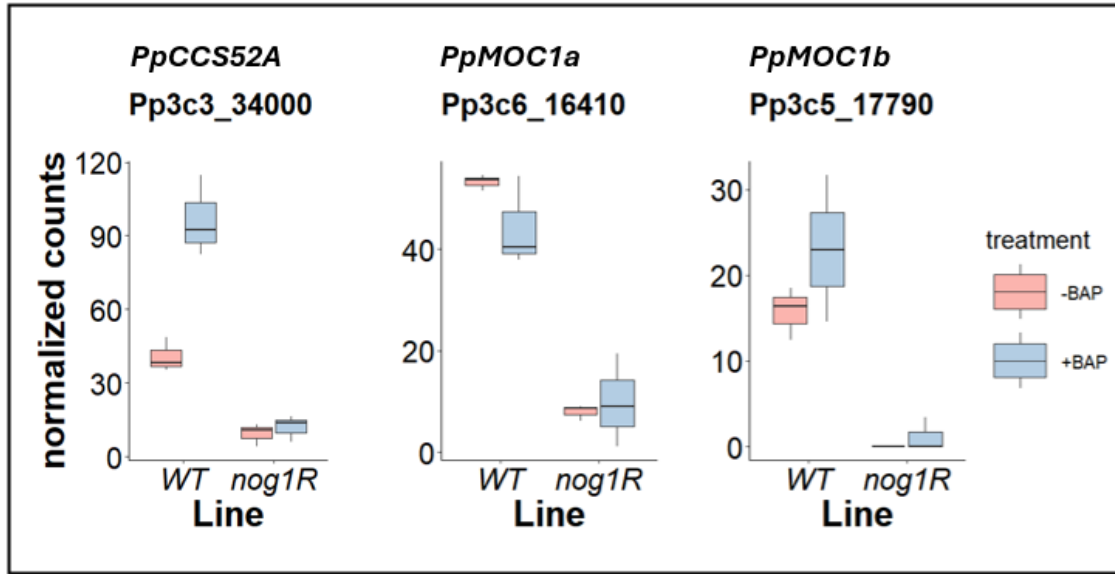


Figure 5.5: The normalised read count data showing expression levels of *PpCCS52A*, *PpMOC1a* and *PpMOC1b* in wild type, wild type + BAP, *Ppnog1R* and *Ppnog1R* + BAP.

three genes in *P. patens*, *PpCCS52A*, and two *MOC1*-related genes that I will refer to as *PpMOC1a* and *PpMOC1b*, seemed like interesting candidates for further investigation. As they regulate axillary meristem formation and the expression of each is suppressed in *Ppnog1R*, I was curious whether restoring their expression in *Ppnog1R* might in any way alleviate the 3D-defective phenotype of *Ppnog1R*.

I assembled individual constructs in which the expression of the wild type transcripts of *PpCCS52A*, *PpMOC1a* or *PpMOC1b* were driven by the strong constitutive maize ubiquitin promoter. I attempted to transform all of these constructs into both *Ppnog1R* and wild type. Despite many attempts I could not recover successful overexpression lines for *PpCCS52A*, *PpMOC1a* or *PpMOC1b*. This could suggest that overexpressing these genes was lethal to the plant. Given their role in regulating the cell cycle, this is very possible. It is possible that the maize ubiquitin promoter used was too strong given that the wild type expression levels were relatively low. Perhaps a less strong promoter would be more successful.

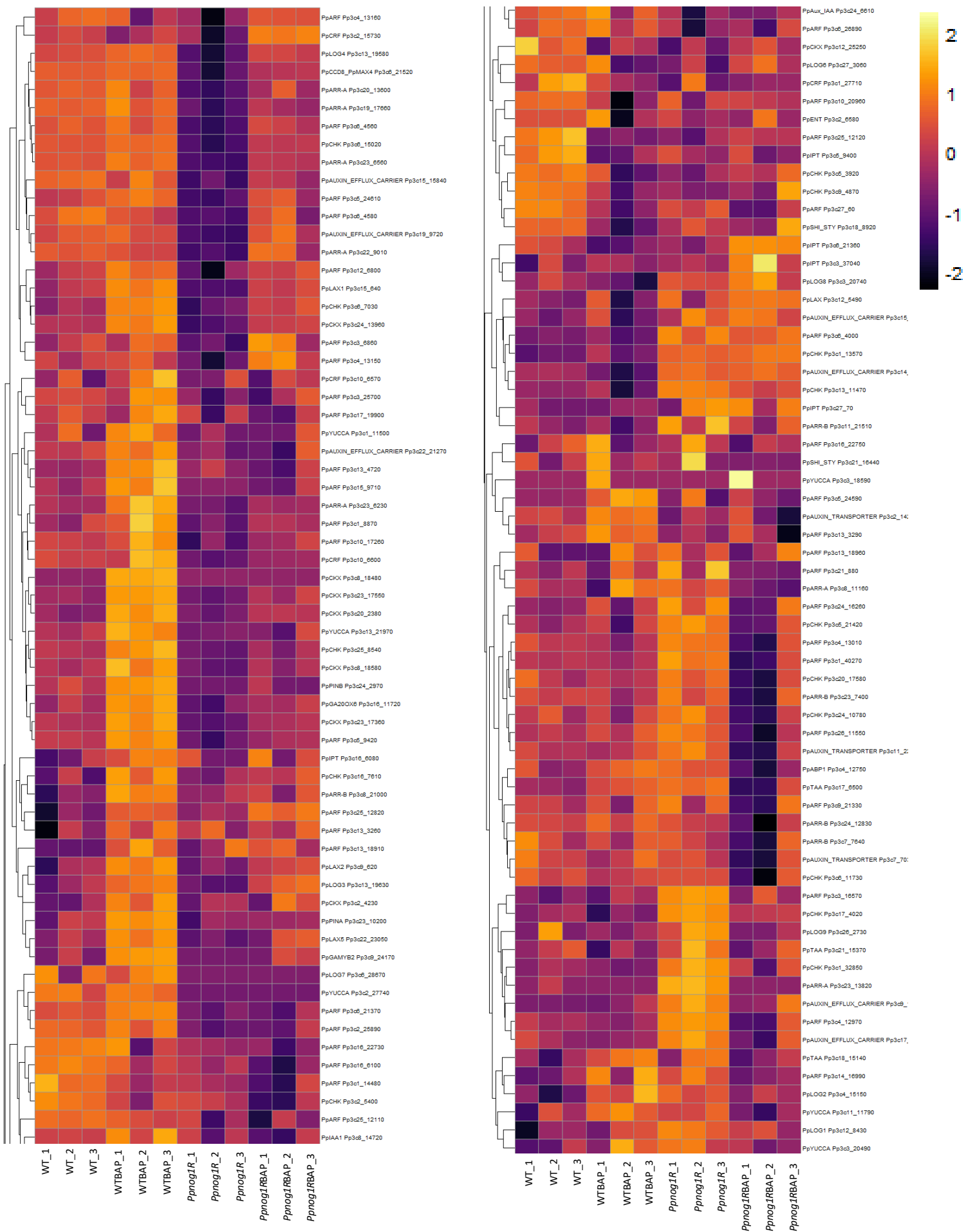
5.2.2 Many known 3D regulators were differentially expressed in *Ppnog1R*

Based on a thorough search of the literature, I compiled a list of any genes that had a confirmed or putative role in the transition to 3D growth. This included some genes that had been implicated from previous RNA-seq experiments (Frank and Scanlon, 2015; Demko et al., 2024). I also included genes associated with auxin and cytokinin synthesis, perception and signal transduction, as well as genes involved in critical downstream processes like cell wall synthesis, cell wall remodelling and cuticle formation. I examined the expression patterns of these gene of interest in this RNA-seq data set (Figure 5.6).

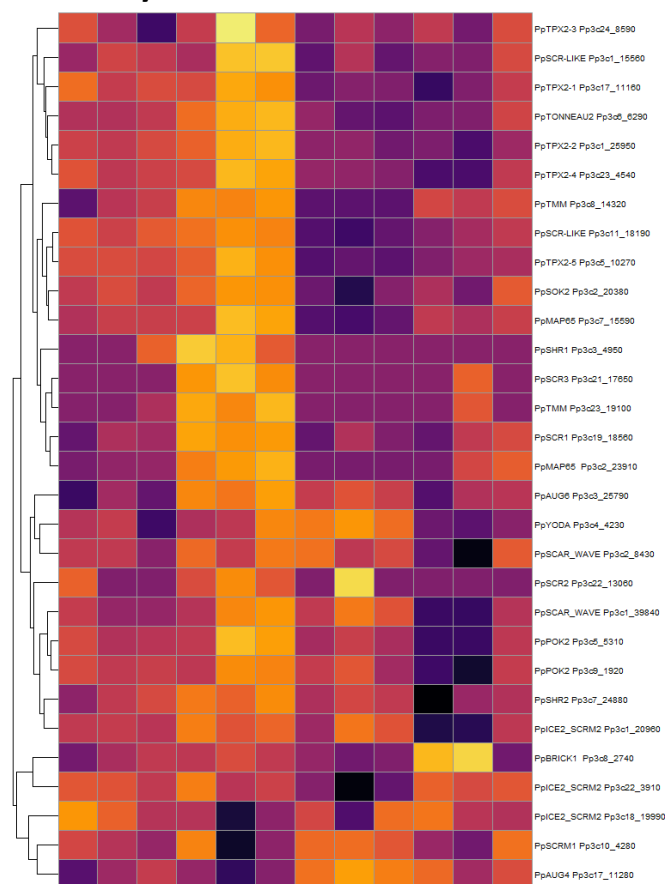
The hormone signalling genes examined included auxin and cytokinin biosynthesis genes, receptors, response factors and transporters. A large subset of these genes had lower expression in *Ppnog1R* than wild type. There was a smaller subset of genes that had an elevated expression level in *Ppnog1R* compared to wild type, some of which were downregulated in response to BAP in *Ppnog1R*, but unaffected in wild type. Overall, there was perturbed expression of many genes related to both cytokinin and auxin signalling, indicating that *Ppnog1R* was generally defective in its response to both hormones.

Precisely orientated asymmetric cell divisions are crucial for the early patterning of the gametophore bud and for the establishment of the tetrahedral apical cell. The cytoskeleton has been shown to be crucial in controlling the orientation of division planes and for the generation of asymmetry (Kosetsu et al., 2017; Kozgunova et al., 2022). Other factors are required for the establishment of polarity in cells that need to undergo asymmetric cell division. In the gametophore bud, it is not yet known what these polarity signals may be but several genes that are known to play this role in *Arabidopsis*, such as *TOO MANY MOUTHS*, *SOSEKI* or *YODA*, have been highlighted as potential candidates for this role in *P. patens* (Frank and Scanlon, 2015; Moody, 2019; van Dop et al., 2020). Around half of the selected

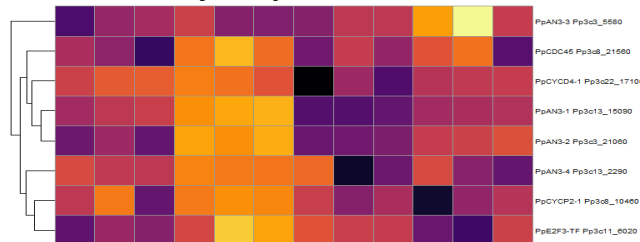
Hormone signalling



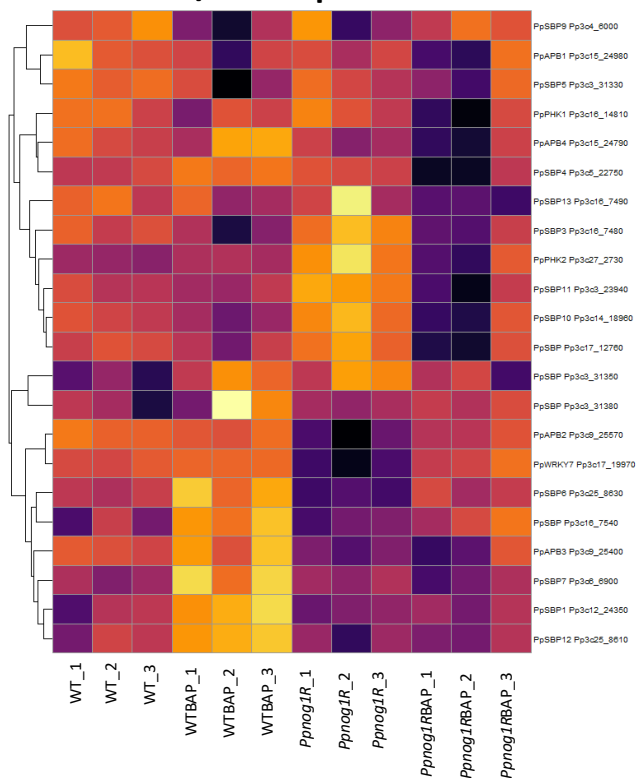
Cytoskeleton and/or asymmetric cell division



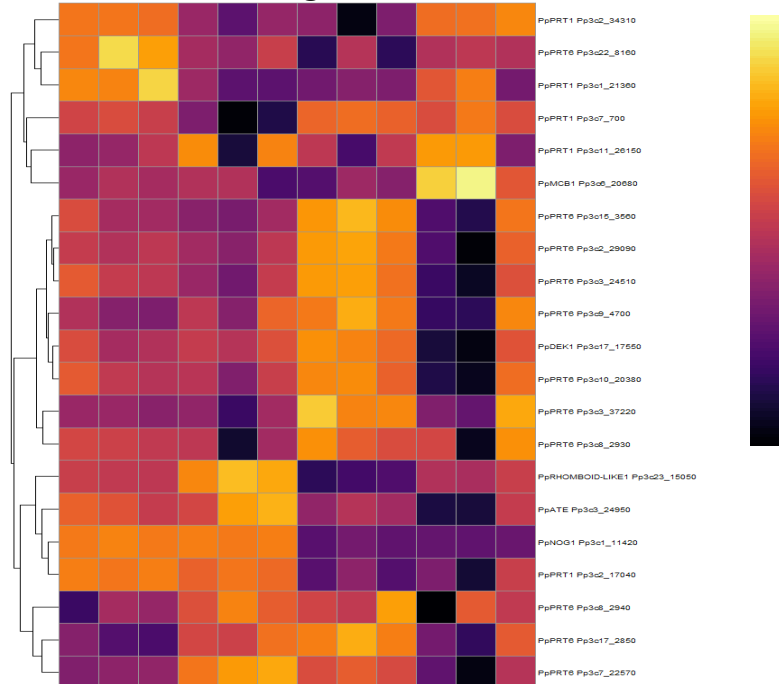
Cell cycle/proliferation



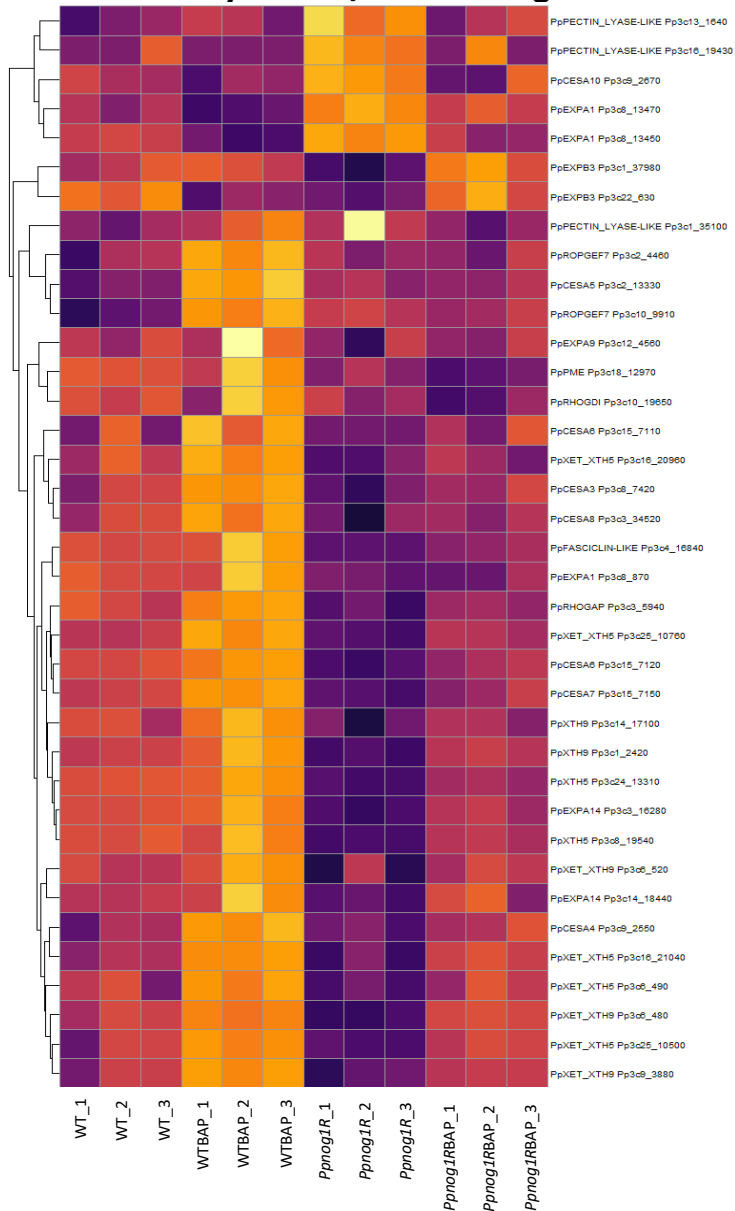
Maturation/developmental transition



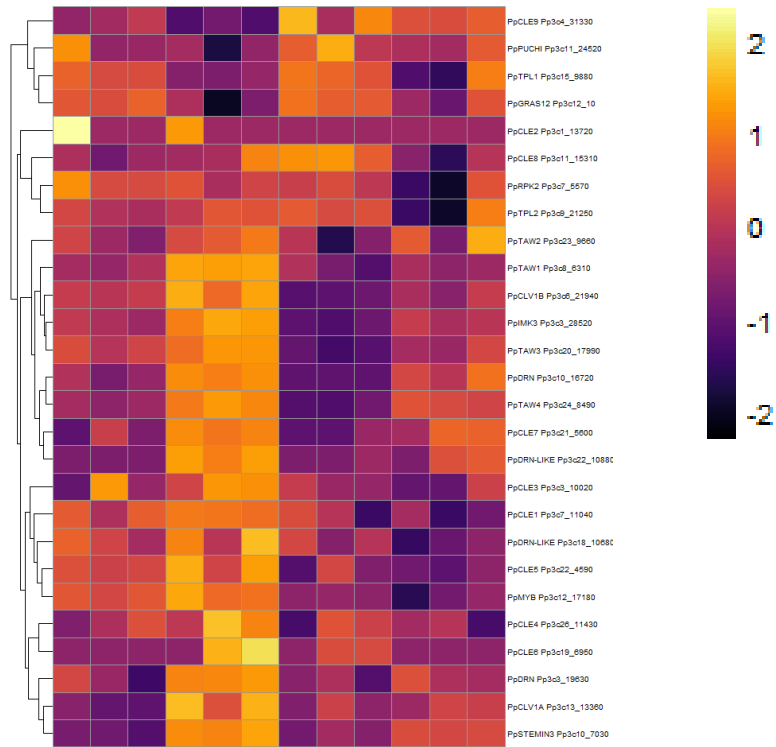
Protein degradation



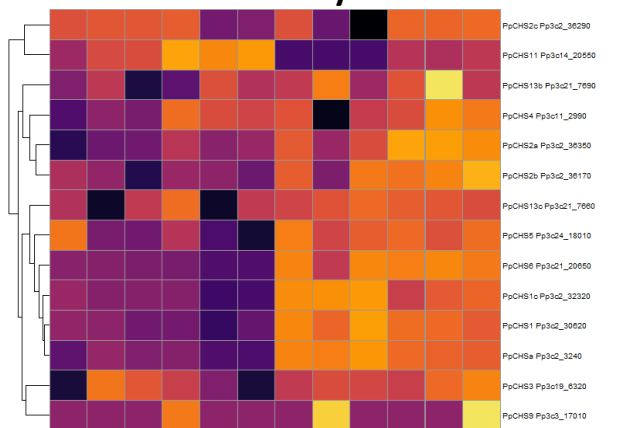
Cell wall synthesis/remodelling



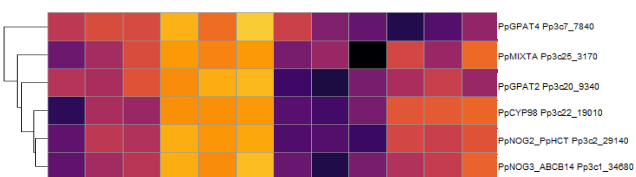
Apical meristem development



Flavonoid biosynthesis



Cuticle formation



Cell trafficking



Epigenetic modification



Post-transcriptional modification

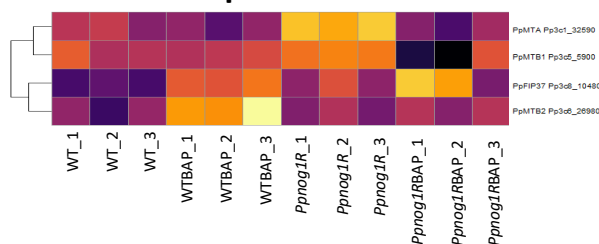


Figure 5.6: Heatmaps showing the normalised counts for each sample, $\log_2$ transformed and scaled by row. This visualises relative expression levels of selected genes in different functional categories for three biological replicates for wild type, wild type + BAP, *Ppnog1R* and *Ppnog1R* + BAP.

genes associated with cytoskeleton and/or asymmetric cell division were downregulated in *Ppnog1R* compared to wild type. This includes all of the *PpTPX2* genes, which have been shown to be important for controlling the orientation of divisions within the gametophore bud (Kozgunova et al., 2022). The suppression of *PpTPX2* expression could possibly be contributing to the misoriented cell divisions exhibited in *Ppnog1R* buds (Moody et al., 2018).

Regulation of the cell cycle and therefore the rate of cell proliferation is vital for development. The spatial and temporal regulation of cell proliferation is essential for the maintenance of the apical cell. Cytokinin promotes cell proliferation and inhibits differentiation (Ashton et al., 1979; Hata and Kyoizuka, 2021). Therefore it was unsurprising to see that, for the most part, cell cycle-related genes were upregulated by BAP treatment in wild type. In *Ppnog1R*, some of the genes exhibit a weak upregulation in response to BAP, but not as strong as was seen in wild type. The baseline levels of all these genes appear broadly similar in wild type and *Ppnog1R*, with a few exceptions such as *PpCYCD4-1* and *PpAN3-1* having suppressed expression in *Ppnog1R* compared to wild type.

The initiation of the transition to 3D growth (i.e., the developmental decision to form a gametophore bud initial rather than a protonema side branch initial) is determined by the parental caulonemal cell. It has previously been demonstrated that the *PpAPB* genes are master regulators of the formation of gametophore initials (Aoyama et al., 2012). The *PpSBP* genes are also known regulators of maturation and *PpSBP3* has been demonstrated to be involved in the decision to initiate 3D growth (Cho et al., 2012). Previous RT-qPCR experiments have shown that all of the *PpAPBs* are downregulated in *Ppnog1R* (Moody et al., 2018). This RNA-seq data corroborated that finding for *APB2,3* and *4*, but not for *PpAPB1* (Figure 5.6). However, out of the four *PpAPB* genes, *PpAPB4* is known to exert the most control over bud initiation. *PpAPB4* alone is sufficient to drive gametophore induction

and loss of *PpAPB4* expression causes a significant decrease in gametophore initial formation (Aoyama et al., 2012). The downregulation of *PpAPBs* corresponds to the lower bud initiation frequency observed in *Ppnog1R* (Moody et al., 2018). *SBP3* is known to inhibit bud initiation (Cho et al., 2012). *PpSBP3* was slightly upregulated in *Ppnog1R* compared to wild type, but suppressed by BAP treatment in both *Ppnog1R* and wild type, which aligns with the role of cytokinin in promoting bud induction.

Protein degradation is vital for the spatiotemporal control of protein populations in a given cell/tissue. Genes associated with protein degradation processes that are known to play role in the 3D transition are *PpNOG1*, *PpDEK1*, *PpATE* and *PpMCB1* (Girod et al., 1999; Perroud et al., 2014; Hoernstein et al., 2016; Moody et al., 2018). I also included homologues of *PROTEOLYSIS 1* (*PpPRT1*) and *PROTEOLYSIS 6* (*PpPRT6*) which recognise N-degrons and target proteins to the proteasome via the N-degron pathway (Stary et al., 2003; Holman et al., 2009). In this experiment, all but one of the *PpPRT6* genes appears to be downregulated in *Ppnog1R*, whereas all but one of the *PpPRT1* genes were upregulated in *Ppnog1R*, compared to wild type. *PpATE* and *PpMCB1* were downregulated in *Ppnog1R* compared to wild type, whereas *PpDEK1* was upregulated.

Cell wall synthesis and remodelling is strongly linked to growth and development. Cell proliferation requires the production of new cell walls and anisotropic cell expansion requires the remodelling of cell walls. A large subset of cell wall-related genes showed upregulation upon BAP treatment in wild type, whereas in *Ppnog1R*, expression was suppressed compared to wild type and showed little or no upregulation in response to BAP. A small subset of these cell wall genes were expressed more highly in *Ppnog1R*, than wild type. Those same genes were downregulated in response to BAP in *Ppnog1R* so that their expression was similar to that in wild type not treated with BAP.

The formation and maintenance of the tetrahedral apical cell (meristem), which organises 3D growth in the gametophore is known to be regulated by *CLAVATA* signalling, as well as the *TOPLESS* transcriptional corepressor and the *PpGRAS12*

transcription factor (Whitewoods et al., 2018; Beheshti et al., 2021; Causier et al., 2023). Members of the Arabidopsis light-dependent short hypocotyl (LSH) and Oryza G1 (ALOG) gene family in *P. patens*, known as *PpTAW1,2,3* and *4*, have also been implicated in apical cell maintenance in a recent preprint (Hata et al., 2023). Other potential regulators have been highlighted in previous transcriptomics experiments (Frank and Scanlon, 2015). All four *PpTAW* genes were downregulated in *Ppnog1R* compared to wild type, and showed upregulation in response to BAP in both wild type and *Ppnog1R*. Both *PpTPL* genes and *PpGRAS12* displayed similar levels of expression in *Ppnog1R* as in wild type, although *TPL2* was downregulated in response to BAP in *Ppnog1R*, but not in wild type. Of the *CLV* genes, *PpCLV1b* and *CLE9* were clearly downregulated in *Ppnog1R* compared to wild type.

Generally, the chalcone synthase genes, which operate in the flavonoid biosynthesis pathway, don't seem to respond strongly to BAP treatment, with the exception of *PpCHS11* and *4*. Also, for most of the chalcone synthase genes there is higher expression in *Ppnog1R* than in wild type. One possible reason for this could be related to the fact that *PpNOG2* was suppressed in *Ppnog1R*. The substrate for *PpNOG2*, p-coumaroyl CoA feeds both the cuticle biosynthesis pathway, in which *PpNOG2* operates, and the flavonoid biosynthesis pathway, where chalcone synthases catalyse the conversion of p-coumaroyl CoA to chalcone. The downregulation of *PpNOG2* would result in more p-coumaroyl CoA entering the flavonoid biosynthesis pathway. This could explain the upregulation of chalcone synthases as a response to increased flux through that pathway (Moody et al., 2021; Kriegshauser et al., 2021).

Cuticle formation is emerging as a very important process for the transition to 3D growth. The cuticle is absent in the protonema but is present in the gametophore. Several mutants that are defective in cuticle deposition are also defective in the 3D transition (Renault et al., 2017; Lee et al., 2020; Moody et al., 2021; Kriegshauser et al., 2021; Zhang et al., 2023). Of the cuticle-related genes that I included in this analysis, there seemed to be a similar trend across all of them. In wild type, these genes were upregulated in response to BAP. In *Ppnog1R*, expression of all the genes

was lower than in wild type and most of them were slightly upregulated in response to BAP in *Ppnog1R*, but did not reach the same level of expression as in wild type + BAP.

Cell trafficking is integral to many important processes including cell division and deposition of the new cell wall and the cuticle. The trend of expression among these selected trafficking genes generally seemed to indicate that they are upregulated in wild type when treated with BAP, but downregulated in *Ppnog1R*, in the presence of BAP.

Two subunits of the Polycomb Repressive Complex 2 (PRC2), which modifies histones to silence genes, are important for inhibiting the formation of sporophyte-like apical cells during the transition to 3D growth; *PpNOG4/PpCLF* and *PpFIE* (Mosquna et al., 2009; Okano et al., 2009; Raquid et al., 2023). *PpNOG4/PpCLF* was upregulated in *Ppnog1R* compared to wild type, but suppressed by BAP treatment in *Ppnog1R*, while remaining unaffected in wild type. *PpFIE* was less clear but it appears to be upregulated by BAP in *Ppnog1R* but otherwise remained unaffected across the other conditions.

Finally, several genes involved in post-transcriptional modification have been implicated in the 3D transition. The *PpMTA*, *PpMTB1*, *PpMTB2* and *PpFIP37* have been shown to stabilise the transcripts of *PpAPBs* by addition of N⁶-methyladenosine, to allow sufficient transcript accumulation to initiate the transition to 3D growth (Garcias-Morales et al., 2023). *PpMTA* and *PpMTB1* were upregulated in *Ppnog1R* compared to wild type, and downregulated in response to BAP in *Ppnog1R*, but unaffected by BAP in wild type. The other two genes seemed less affected in *Ppnog1R*, although *PpMTB2* was not upregulated in *Ppnog1R* treated with BAP as it was in wild type.

Overall, the general trend for these genes that are known or hypothesised to be involved in regulating the transition to 3D growth, was that most of them do show perturbed expression in *Ppnog1R* compared to wild type. Many also fail to respond to BAP treatment in *Ppnog1R*, in the same way as they do in wild type. Expression was often unaffected in *Ppnog1R* + BAP compared to *Ppnog1R*, or shows differential

Gene	Description	WTvsWTBAP LFC	<i>Ppnog1RvsPpnog1RBAP</i> LFC
Pp3c13.16300	related to AP2 9	4.178990612	0.204149859
Pp3c5.20560	basic helix-loop-helix (bHLH) DNA-binding superfamily protein	3.677394289	0
Pp3c25.8610	squamosa promoter binding protein-like 8	3.422739531	0.200675652
Pp3c26.6050	Homeodomain-like superfamily protein	3.248095765	0.163454178
Pp3c1.4970	myb domain protein 106	3.22101613	0.522786738
Pp3c13.23440	Homeodomain-like superfamily protein	3.14680282	-0.231075856
Pp3c22.9600	homeobox protein 31	2.962903409	0.396728187
Pp3c26.1390	myb domain protein 55	2.952403913	-0.147611583
Pp3c4.10850	HSI2-like 1	2.826287817	-0.022461829
Pp3c13.7150	BEL1-like homeodomain 6	2.67183358	0.278078428
Pp3c6.22390	BES1/BZR1 homolog 1	2.572682534	-0.160632293
Pp3c12.24350	squamosa promoter binding protein-like 8	2.426125247	0.417328325
Pp3c14.13110	GRAS family transcription factor	2.368732965	-0.156568536
Pp3c11.14240	WRKY family transcription factor	2.176760047	-0.148401677
Pp3c10.19020	basic helix-loop-helix (bHLH) DNA-binding superfamily protein	2.157235909	-0.294391963
Pp3c8.17290	basic helix-loop-helix (bHLH) DNA-binding superfamily protein	2.118182241	-0.025707038
Pp3c5.20890	basic helix-loop-helix (bHLH) DNA-binding superfamily protein	2.075490712	0.594578202
Pp3c14.21750	ovate family protein 4	-2.183379327	0.554681962
Pp3c7.1850	basic helix-loop-helix (bHLH) DNA-binding superfamily protein	-2.220664324	0.149202005

Table 5.1: A list of putative transcription factors that were differentially expressed in response to BAP in wild type but unaffected in *Ppnog1R*. LFC = Log₂ Fold Change

expression in the opposite direction as was seen in wild type + BAP. Some genes were affected by BAP in *Ppnog1R* but not in wild type.

5.2.3 The overexpression of the transcription factor HECATE1 in *Ppnog1R* resulted in partial rescue of the *Ppnog1R* phenotype

To identify new possible regulators that could control the cytokinin-dependent induction of bud formation, I filtered the DEG lists to generate a list of putative transcription factors that responded to BAP in wild type but not in *Ppnog1R* (Table 5.1). Exploring these genes could provide clues as to why the *Ppnog1R* mutant does not produce buds in response to BAP.

From this list one of the basic helix-loop-helix (bHLH) DNA-binding genes (*Pp3c10-19020*) seemed a potentially interesting gene to investigate. *Pp3c10-19020* is a homologue of *HECATE1* (*AtHEC1*) from Arabidopsis. In the Arabidopsis SAM, *AtHEC1* expression is directly suppressed by WUSCHEL and *AtHEC1* regulates both cytokinin and auxin signalling to control meristem size. Expression of *AtHEC1* is not influenced by cytokinin (Gaillochet et al., 2017). Unlike *AtHEC1*, *Pp3c10-19020* (which I will hence refer to as *PpHEC1a*), was upregulated by cytokinin in wild type. Using Orthofinder, I identified the *P. patens* homologues of *AtHEC1*. There were six putative homologues (Figure S30). For one of

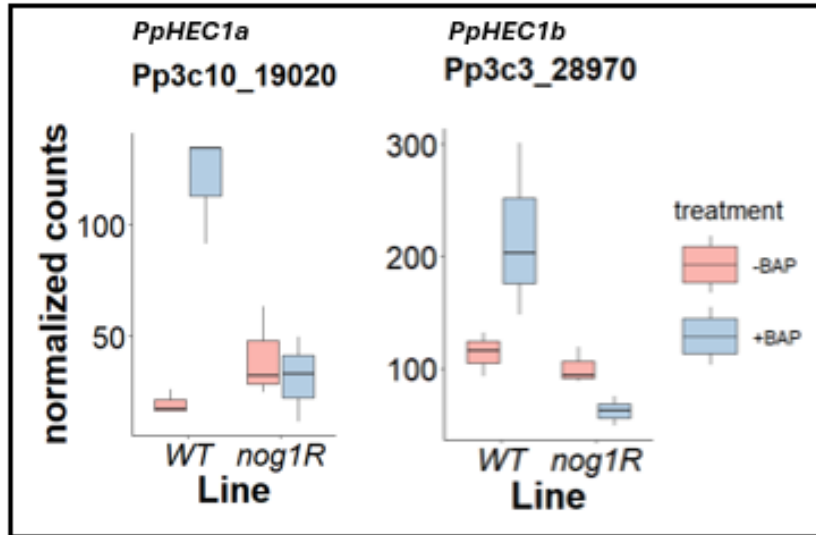


Figure 5.7: The normalised read count data showing expression levels of *PpHEC1a* and *PpHEC1b* in wild type, wild type + BAP, *Ppnog1R* and *Ppnog1R* + BAP.

these homologues, there was no expression detected in this RNA-seq dataset. Interrogation of the BAR eFP browser for *P. patens* indicated that this gene is only expressed in the sporophyte and the spores (Winter et al., 2007). Of the remaining homologues, the most closely related sequence to *PpHEC1a* (*Pp3c3_28970*; henceforth referred to as *PpHEC1b*), had the same expression pattern as *PpHEC1a*, wherein it was upregulated in response to BAP in wild type, but insensitive to BAP in *Ppnog1R* (Figure 5.7). The other three homologues (*Pp3c17_15380*, *Pp3c14_16620* and *Pp3c1_40040*) are closely related to each other and more distantly related to *PpHEC1a/b*. These three genes all had the same expression pattern, which was distinct from the expression pattern in *PpHEC1a* and *PpHEC1b*; they had high expression in wild type, which was suppressed in wild type + BAP and in *Ppnog1R* expression was elevated compared to wild type and also downregulated in response to BAP (Figure S31).

I decided to focus on *PpHEC1a* and *PpHEC1b* for further investigation. I assembled constructs in which the CDS of either *PpHEC1a* or *PpHEC1b* was driven by the highly active maize ubiquitin promoter. I attempted to transform each into the *Ppnog1R* and wild type backgrounds. I recovered a *PpHEC1a* line in the wild type background but not in *Ppnog1R* (data not shown). However, I successfully

generated two independent lines for both *Ppnog1R pUBI::HEC1b-GFP* and wild type *pUBI::HEC1b-GFP*.

To validate the lines, I performed a PCR with genomic DNA extracted from each of the lines, using primers that amplified the region from the START codon to just before the STOP codon. For the genomic gene sequence this amplified the 3 exons and 2 introns. For the coding sequence (CDS) used in the overexpression construct, this amplified just the 3 exons (Figure 5.8A). Therefore, successful integration of the construct was evident by the presence of a band corresponding to the size of the CDS. This confirmed the presence of the *PpHEC1b* overexpression construct in the generated lines (Figure 5.8B).

Interestingly, both *Ppnog1R pUBI::HEC1b-GFP* lines displayed a very severe phenotype. *Ppnog1R* develops normally prior to the stage at which bud initiation occurs, producing tip-growing protonema, consisting of chloronemal and caulonemal cells (Figure 5.9B). In contrast, both *Ppnog1R pUBI::HEC1b-GFP-1* and *-2* exhibited strong defects in protonemal growth (Figure 5.9C and D). Instead of forming filaments, the cells were rounded, displaying isotropic growth instead of directional tip growth (Figure 5.9E). The cells divided slowly, resulting in slow growth and limited plant spread. For example, in Figure 5.9C and D) *Ppnog1R pUBI::HEC1b-GFP-2* had grown only approximately 2-3 mm over the course of three months.

What was even more intriguing was that the two wild type *pUBI::HEC1b-GFP* had no obvious defects. Both grew normally, produced gametophores and appeared broadly similar to wild type (Figure 5.9F). Therefore the severity of the defect in the *Ppnog1R pUBI::HEC1b-GFP* phenotype requires both the absence of functional *PpNOG1* and the overexpression of *PpHEC1b*.

The initial phenotype of both *Ppnog1R pUBI::HEC1b-GFP* lines was characterised by a complete inability to produce protonema, let alone gametophores. The tissues appeared like disorganised clumps of spherical cells. However, over a period of several months there was recovery of the *Ppnog1R pUBI::HEC1b-GFP* phenotype. This recovery occurred sooner in *Ppnog1R pUBI::HEC1b-GFP-1* than

Ppnog1R pUBI::HEC1b-GFP-2, but the latter was following the same trajectory as the former. First, protonema were able to grow from the spherical cell clusters. Gradually, protonema became more prevalent and the tissue began to resemble *Ppnog1R*. When looking at the protonema under the microscope, I still found some spherical cells in amongst the filaments, but the protonema dominated. Eventually *Ppnog1R pUBI::HEC1b-GFP-1* even started to produce gametophores. As of last observation *Ppnog1R pUBI::HEC1b-GFP-2* had not developed any gametophores, but as it lagged behind *Ppnog1R pUBI::HEC1b-GFP-1* in the changes to the

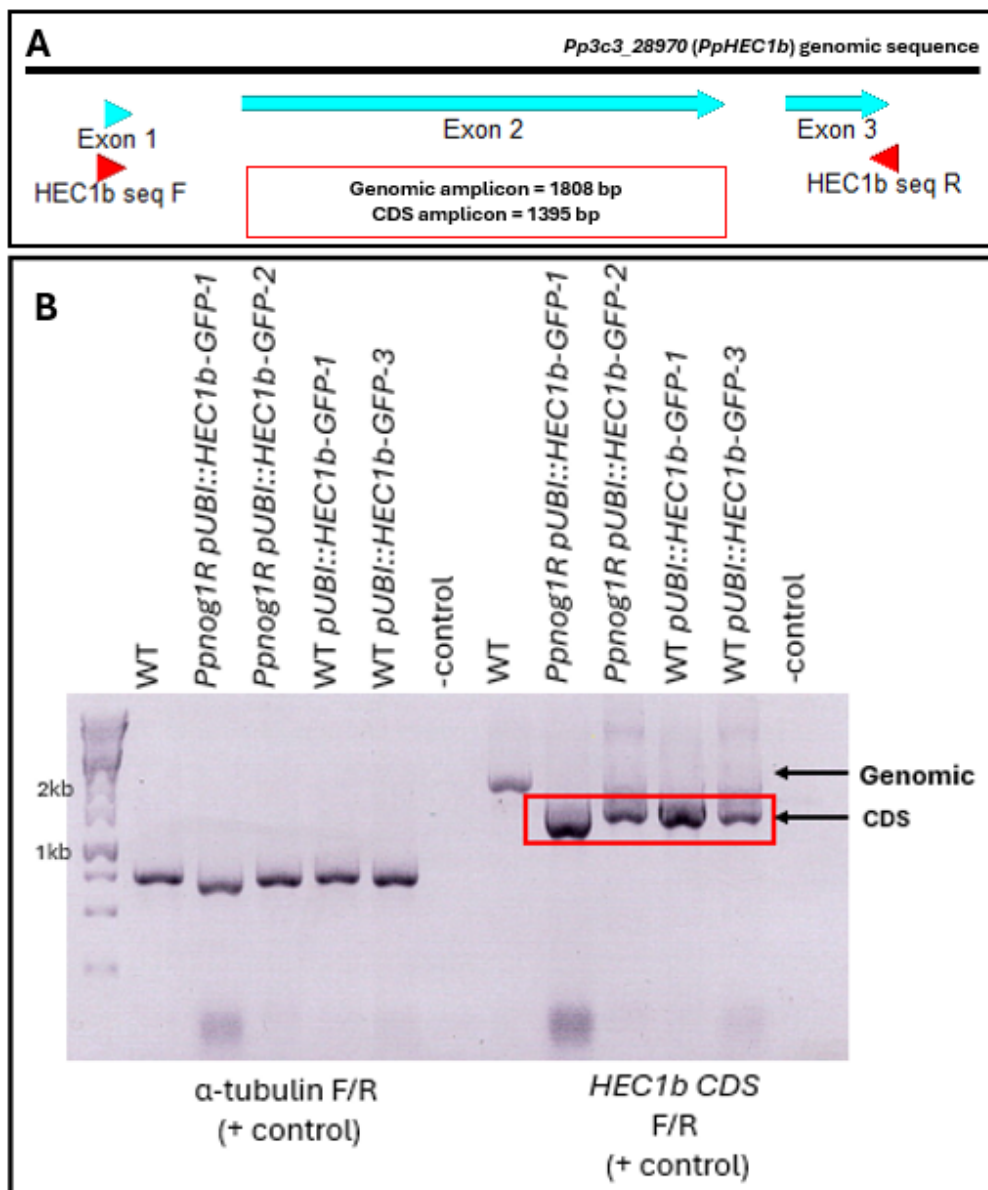


Figure 5.8: A) Genotyping strategy for *PpHEC1b* overexpressing lines. B) Genotyping result for *PpHEC1b* overexpressing lines confirming the presence of the *PpHEC1b* CDS in the genomic DNA.

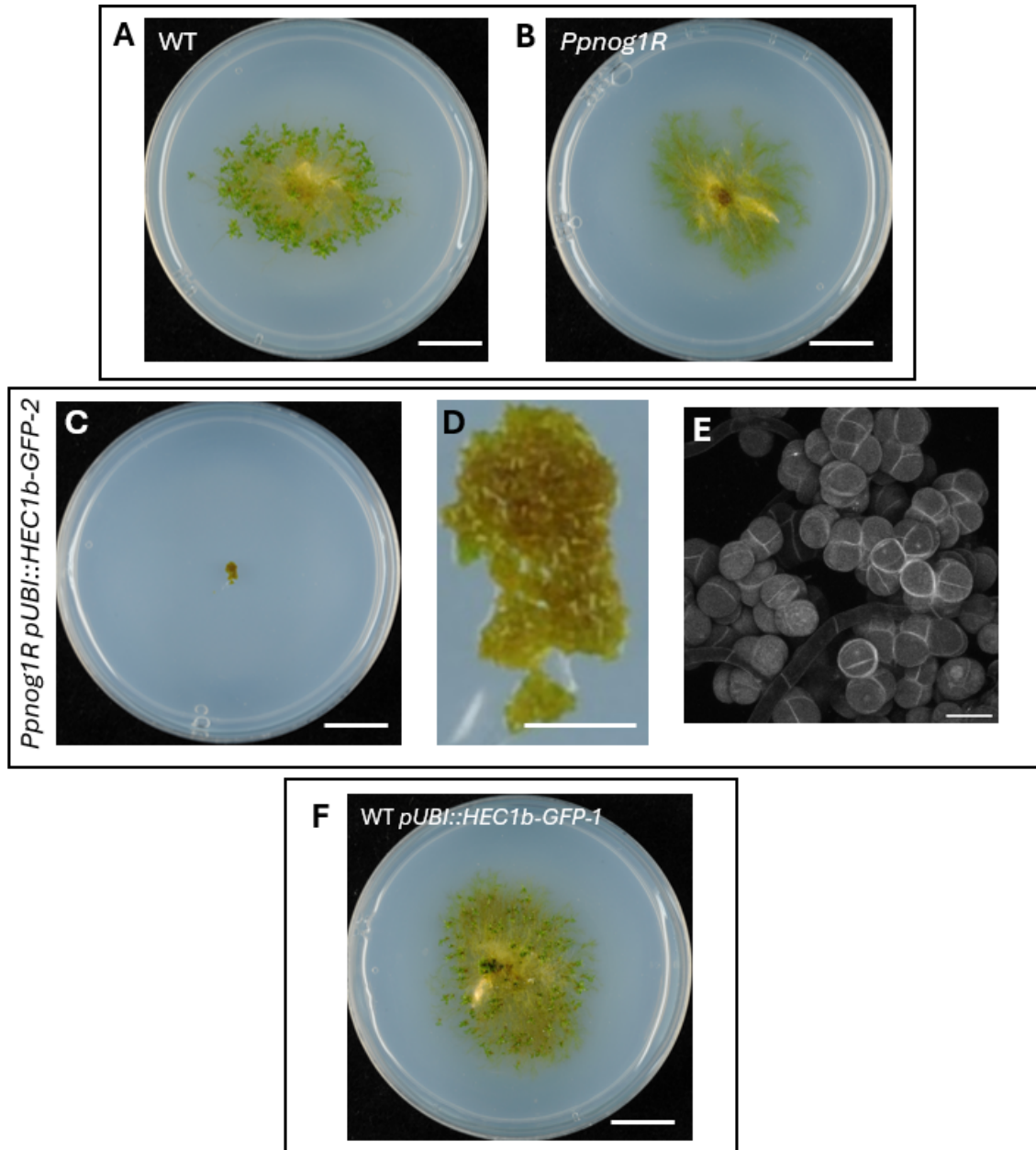


Figure 5.9: A-C and F) Representative images of plants grown on BCD. A,B and F were grown for 2 months. C was grown for 3 months. A) wild type. Scale = 1 cm. B) *Ppnog1R*. Scale = 1 cm. C-D) *Ppnog1R pUBI::HEC1b-GFP-2*. C) Scale = 1 cm. D) Zoomed in version of C. Scale = 1 mm. E) Confocal image of PI-stained *Ppnog1R pUBI::HEC1b-GFP-2*, showing its spherical cells and few protonema. Scale = 50 μm. F) wild type *pUBI::HEC1b-GFP-1*. Scale = 1 cm.

phenotype, it may still yet (Figure 5.10).

When examining the gametophores of *Ppnog1R pUBI::HEC1b-GFP-1* under the microscope I could see that the gametophores developed abnormally, appearing stunted in height, with a dense arrangement of phyllids. I also observed some examples of gametophore bifurcation, suggesting the formation of ectopic apical

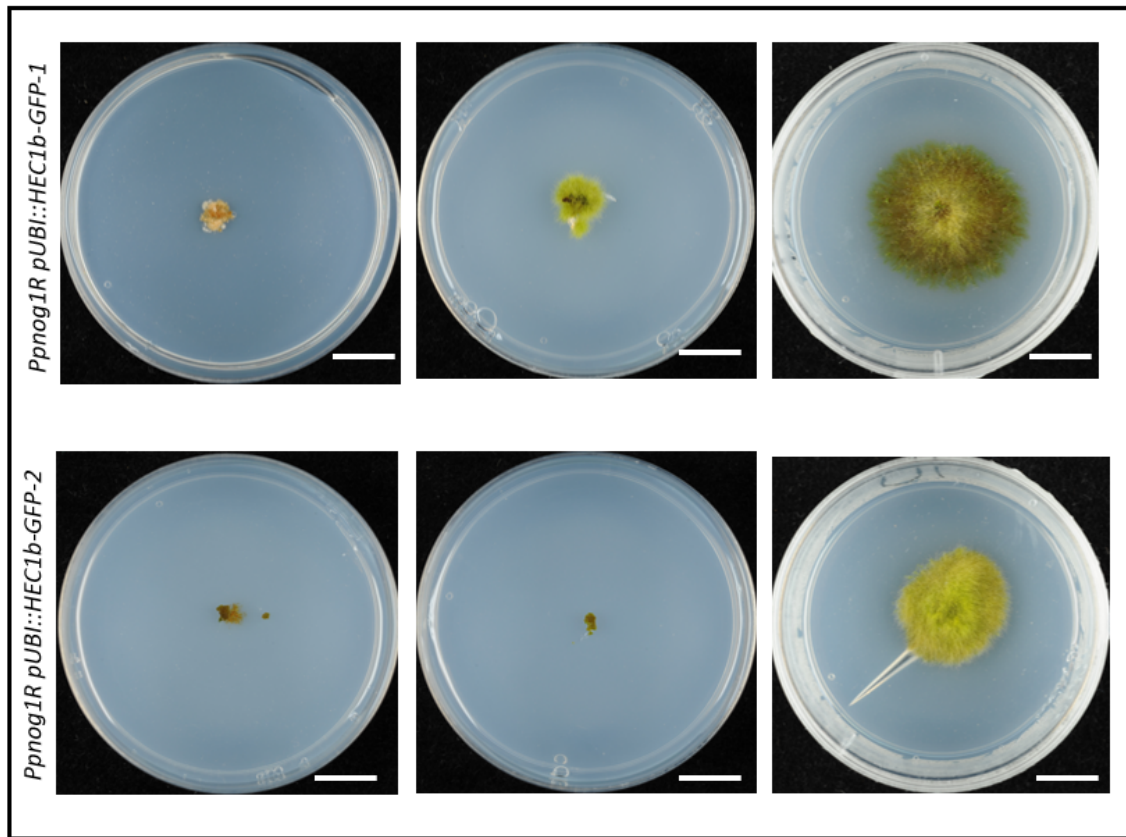


Figure 5.10: Representative images of *Ppnog1R pUBI::HEC1b-GFP-1* and *Ppnog1R pUBI::HEC1b-GFP-2* propagated at different times. Left is earliest and right is latest. For each line, each pair of plates was propagated at the same time and photographed at the same time so the representative images show that the phenotype of *Ppnog1R pUBI::HEC1b-GFP-1* changed earlier than that of *Ppnog1R pUBI::HEC1b-GFP-2*. Scale = 1 cm.

cells (Figure 5.11A). The tissues of *Ppnog1R pUBI::HEC1b-GFP-2* consisted primarily of protonema, but there were still clusters of spherical cells persisting amongst the protonemal filaments. Making up the protonemal filaments were both chloronemal (perpendicular cell divisions) and caulonemal cells (oblique cell divisions; Figure 5.11B). Caulonemal cells give rise to gametophore buds so *Ppnog1R pUBI::HEC1b-GFP-2* does develop the prerequisite cell type to produce gametophores.

Overall, this showed that expressing *PpHEC1b* in *Ppnog1R*, ultimately resulted in a partial rescue of the *Ppnog1R* phenotype, in the case of *Ppnog1R pUBI::HEC1b-GFP-1*. Whether it also does so in *Ppnog1R pUBI::HEC1b-GFP-2* is yet to be determined. The phenotype may still be changing in both lines and it will be necessary to wait for stabilisation of the phenotype before fully characterising

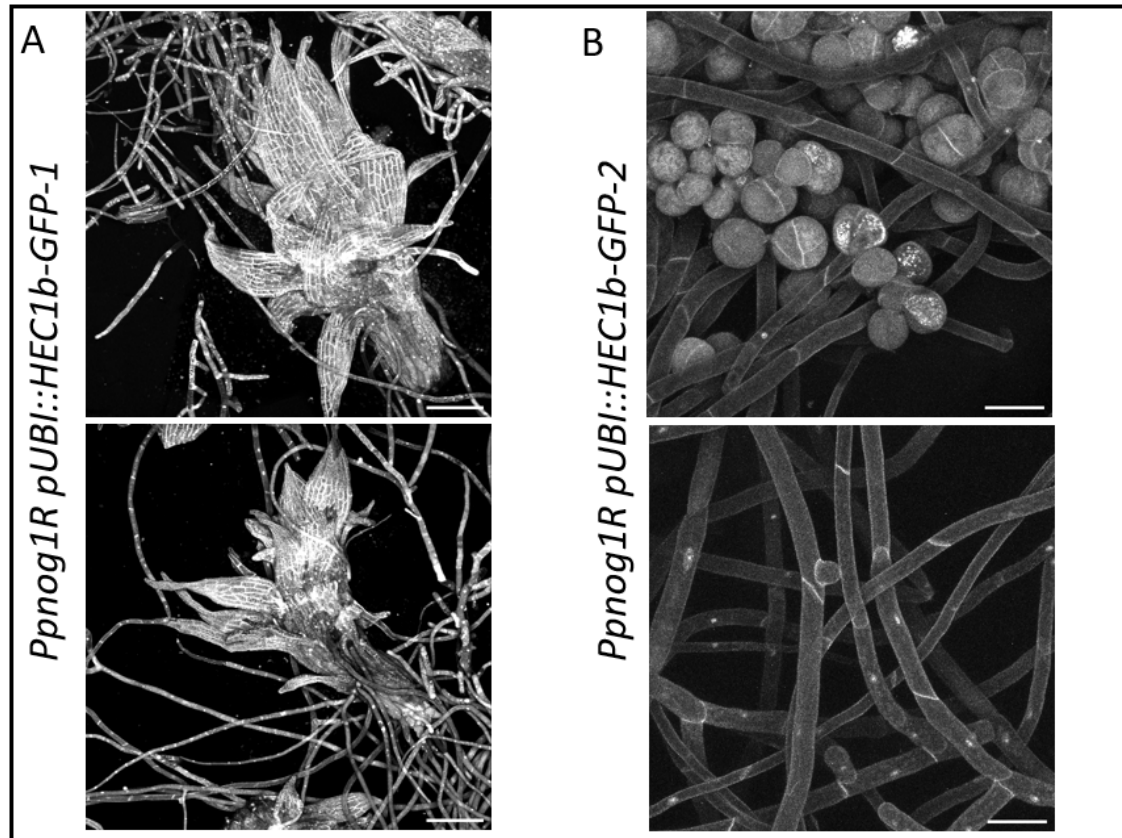


Figure 5.11: Confocal images of PI-stained tissues A) *Ppnog1R pUBI::HEC1b-GFP-1*. Scale = 200 μm . B) *Ppnog1R pUBI::HEC1b-GFP-2*. Scale = 50 μm .

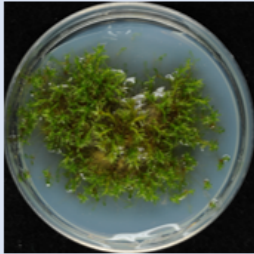
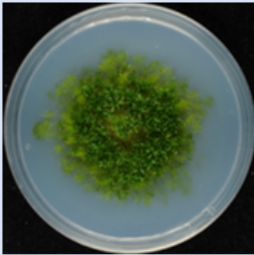
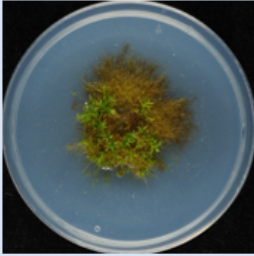
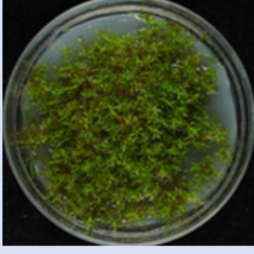
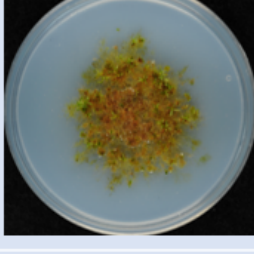
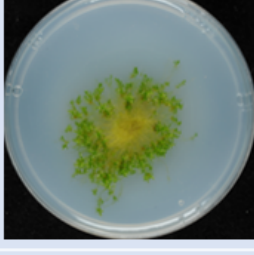
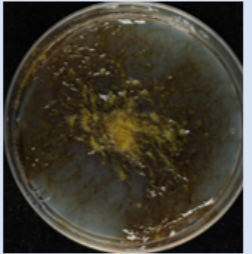
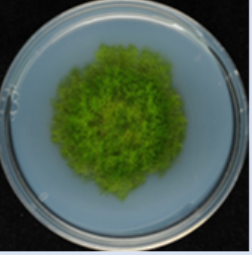
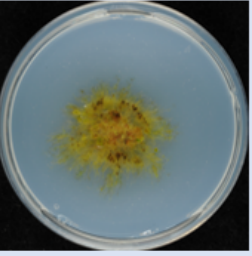
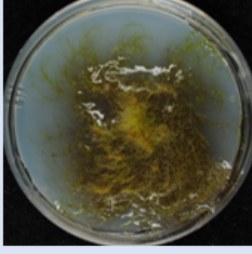
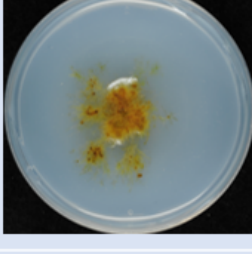
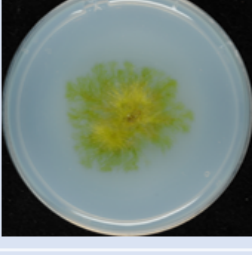
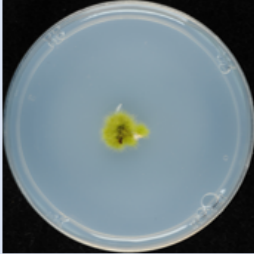
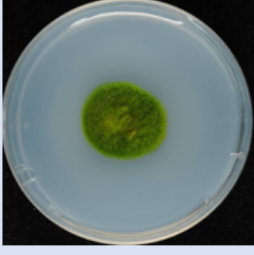
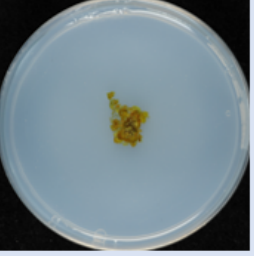
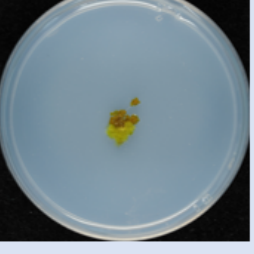
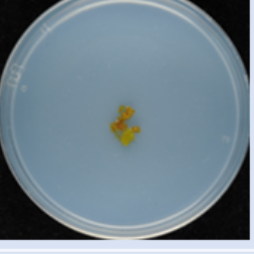
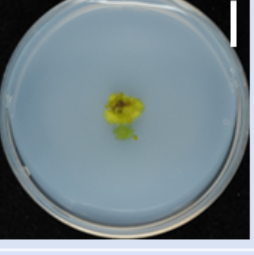
the effect that *PpHEC1b* has on *Ppnog1R*.

I conducted a preliminary hormone treatment experiment on the *PpHEC1b* overexpressing lines, to observe how their responses to various treatments compared to wild type responses. Note that when I did this experiment, the *Ppnog1R pUBI::HEC1b-GFP* lines were still in the early stages of their phenotype progression. Each overexpression line along with wild type and *Ppnog1R* was grown on BCD, BCDAT, BCD + 100 nM BAP, BCD + 100 nM 1-Naphthaleneacetic acid (NAA: a synthetic auxin), BCD + 100 nM BAP and 100 nM NAA and BCD + 5 μM N-1-naphthylphthalamic acid (NPA: a polar auxin transport inhibitor). Note that the tissues on BCDAT and BCD + 5 μM NPA were four weeks younger than all the other treatments, which were 3 months old. Ideally I would have repeated the experiment, but there was not sufficient time. Growth on BCD medium promotes normal development. BCDAT is the same as BCD, with the addition of ammonium

tartrate, which accelerates growth and promotes the formation of chloronema, but inhibits the formation of gametophores. In wild type, treatment with BAP induces bud initiation but persistent BAP treatment perturbs gametophore development. Treatment with lower levels of cytokinin results in small gametophores with broad, short phyllids and the formation of ectopic apical cells (Cammarata et al., 2023). Higher levels of cytokinin results in over-proliferation of the cells in the gametophore bud, resulting in a callus-like structure (Ashton et al., 1979). In this experiment I treated the plants with an intermediate concentration of BAP (100 nM). I also applied a treatment of 100 nM NAA. 100nM NAA treatment results in mild defects. Usually, gametophores appear normal or slightly elongated with additional phyllids. Treatment with 5 μ M NPA has been shown to cause similar effects to treatment with NAA, with some mild phyllid morphological defects (Bennett et al., 2014).

Figure 5.12, shows the affect of these treatments on all lines. On BCD, wild type tissue that was three months old, gametophores dominated. On BCDAT, the colony was noticeably greener due to the proliferation of chloronemal filaments and the initiation and development of gameophores was partially suppressed (although to what degree it is not possible to say as these were younger than the plants on BCD). With the BAP treatment, bud initiation was high, but few were able to develop into gametophores. With the NAA treatment growth appeared normal, and possibly enhanced by the treatment. The combined BAP and NAA treatment resulted in inhibition of gametophore development. With the NPA treatment, development appeared to be progressing normally (it appears smaller than the tissue growing on BCD with no treatment because it was younger). The *Ppnog1R* mutant, as expected, did not produce gametophores under any conditions. It appeared similar grown on BCD and BCD + NAA, spreading outward through the media. On BCDAT, *Ppnog1R* formed a smaller, denser and greener colony. The BAP treatment induced the formation of a few callus-like structures on *Ppnog1R*, as did the combined 100 nM BAP and 100 nM NAA treatment. The NPA treatment resulted in no obvious deviations from normal *Ppnog1R* growth.

Both *Ppnog1R pUBI::HEC1b-GFP* lines exhibited similar responses to the

	BCD	BCDAT	BCD + 100 nM BAP	BCD + 100 nM NAA	BCD + 100 nM BAP + 100 nM NAA	BCD + 5 μ M NPA
WT						
<i>Ppnog1R</i>						
<i>Ppnog1R</i> <i>pUBI::</i> <i>HEC1b-</i> <i>GFP-1</i>						

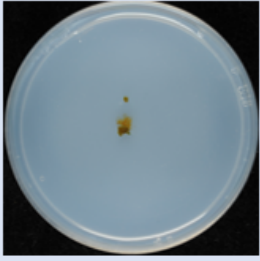
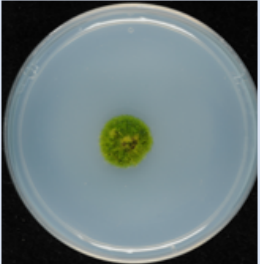
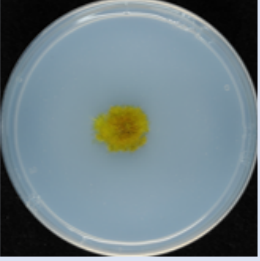
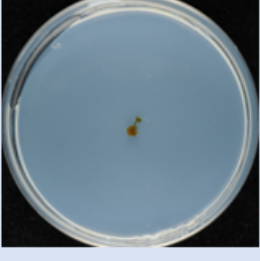
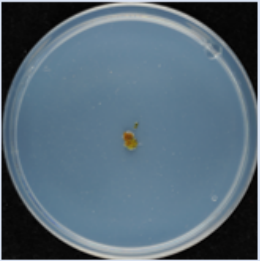
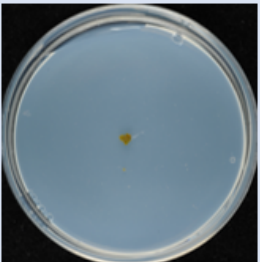
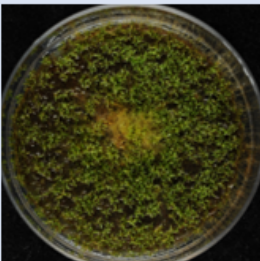
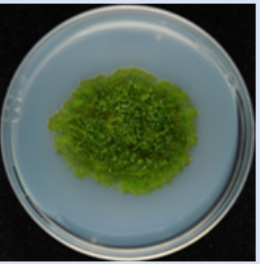
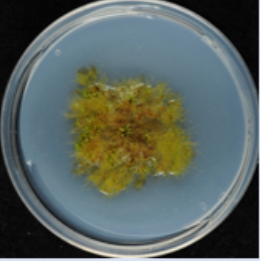
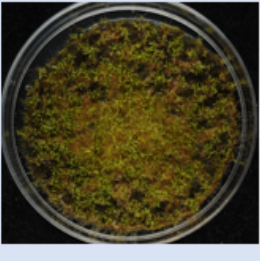
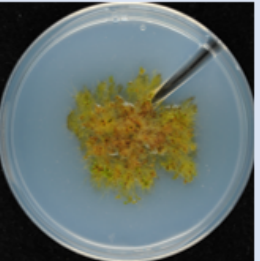
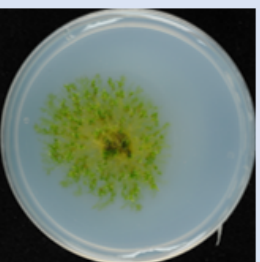
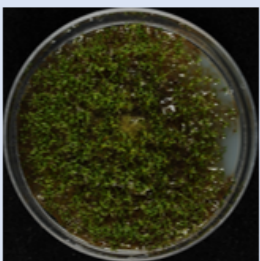
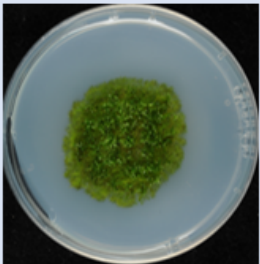
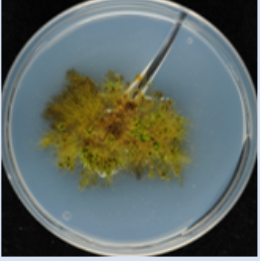
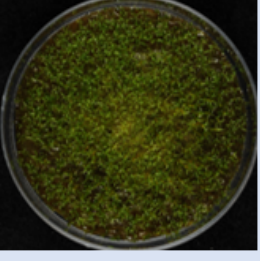
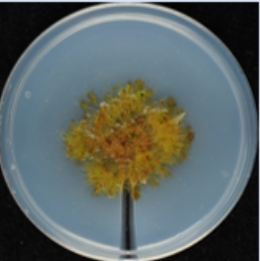
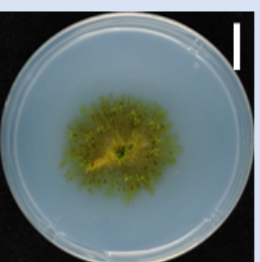
	BCD	BCDAT	BCD + 100 nM BAP	BCD + 100 nM NAA	BCD + 100 nM BAP + 100 nM NAA	BCD + 5 μ M NPA
<i>Ppnog1R</i> pUBI:: HEC1b- GFP-2						
WT pUBI:: HEC1b- GFP-1						
WT pUBI:: HEC1b- GFP-2						

Figure 5.12: Images of wild type, *Ppnog1R*, *Ppnog1R pUBI::HEC1b-1* and *-2* and wild type *pUBI::HEC1b-1* and *-3*. These lines were grown on BCD, BCDAT, BCD + 100 nM BAP, BCD + 100 nM NAA, BCD + 100 nM BAP and 100 nM NAA and BCD + 5 μ M NPA. Scale = 1 cm. Note: the tissues on BCDAT and BCD + 5 μ M NPA were four weeks younger than all the other treatments, which were 3 months old.

hormone treatments. On BCD, *Ppnog1R pUBI::HEC1b-GFP-1* grew as a clumps of spherical cells with green protonema radiating outwards from it. *Ppnog1R pUBI::HEC1b-GFP-2* consisted of spherical cells with some brownish protonema emerging from one side of the tissue. Interestingly, growth on BCDAT i.e. the presence of ammonium tartrate, promoted protonemal growth in both *Ppnog1R pUBI::HEC1b-GFP* lines. Both appeared similar to *Ppnog1R* grown on BCDAT, although with a smaller radius of growth. The BAP treatment promoted protonemal outgrowth in *Ppnog1R pUBI::HEC1b-GFP-2*, but seemingly less so in *Ppnog1R pUBI::HEC1b-GFP-1*. Treatment with NAA, BAP + NAA and NPA resembled plants grown on BCD, with mostly spherical cells and limited protonemal outgrowth. The wild type *pUBI::HEC1b-GFP* lines both phenocopied wild type in response to all the treatments, except that they seem to have an enhanced rate of growth, producing more tissue over the same time period on the BCD and BCD + NAA plates.

As PpHEC1b is a transcription factor, I had expected that it would localise to the nucleus. When I examined all four lines on the confocal microscope I did observe PpHEC1b-GFP in the nucleus. However, PpHEC1b-GFP also appears to localise to the cell periphery, most likely to the plasma membrane (Figure 5.13).

5.2.4 Fewer genes were differentially expressed in *Ppnog2R*, than in *Ppnog1R*

The *Ppnog2R* mutant was the second mutant to be characterised from the forward genetics screen that also yielded *Ppnog1R*. The *Ppnog2R* mutant produces an excess of bud initials, but due to division orientation defects all fail to develop into gametophores. RT-qPCR experiments have shown that in *Ppnog2R*, the expression

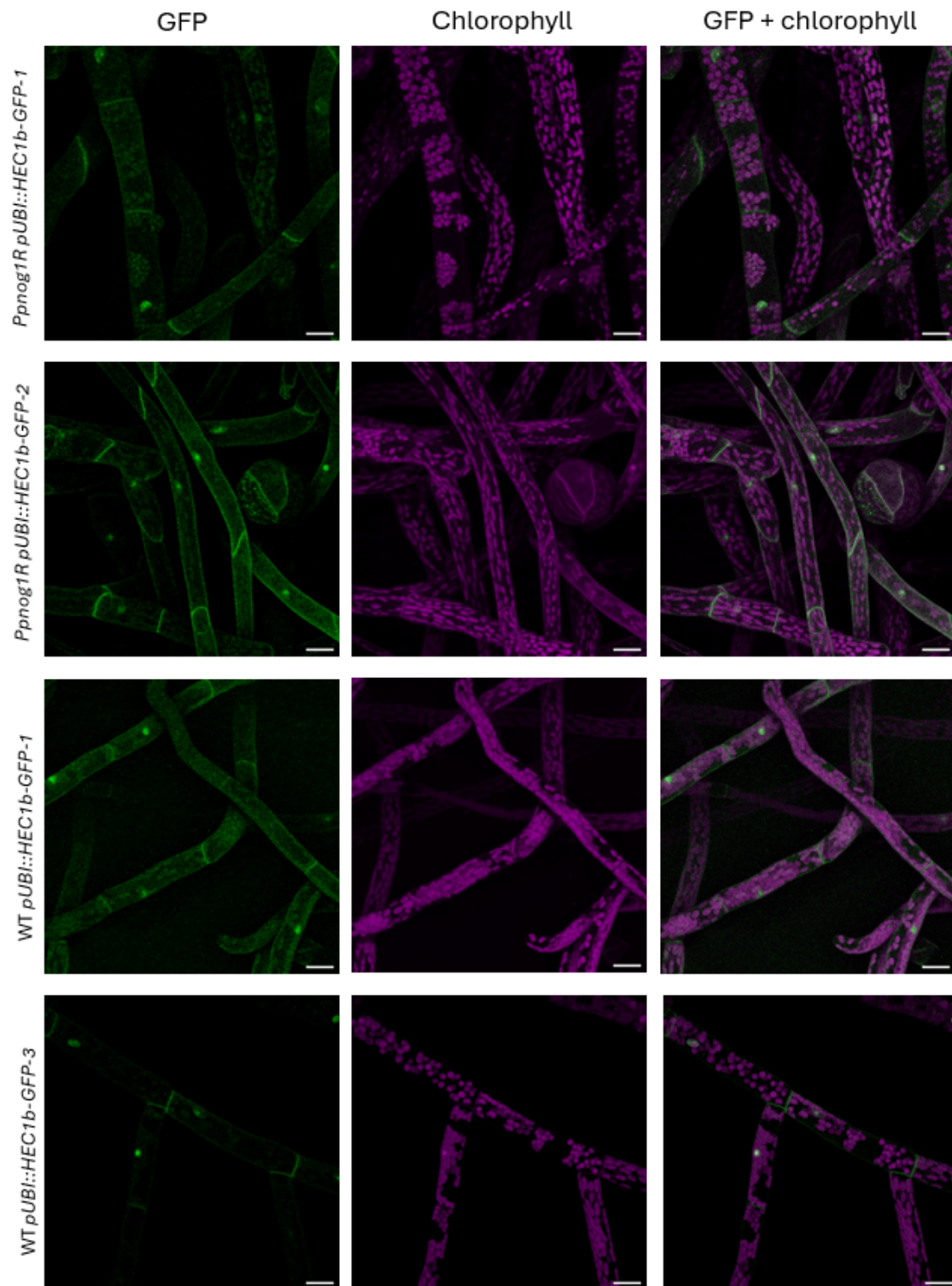


Figure 5.13: Confocal images of *Ppnog1R pUBI::HEC1b-GFP-1*, *Ppnog1R pUBI::HEC1b-GFP-2*, wild type *pUBI::HEC1b-GFP-1* and wild type *pUBI::HEC1b-GFP-3*. L-R: GFP only, chlorophyll only, GFP + chlorophyll overlay.

of some auxin-responsive genes were shown to be perturbed in a way that suggests lower levels of auxin than wild type. In addition, the expression of *PpNOG2* was shown to be suppressed in *Ppnog1R*, indicating that *PpNOG2* functions downstream of *PpNOG1* (Moody et al., 2021). In the second RNA-seq experiment, I compared the transcriptomes of *Ppnog2R* and a *PpNOG2* overexpression mutant (hereafter *PpNOG2* OE) with wild type. The aim of generating this dataset was to hopefully identify genes that are directly influenced by the manipulation of *PpNOG2*.

Before conducting this RNA-seq experiment, the *PpNOG2* OE line had to be generated. The construct containing the *PpNOG2* CDS driven by the rice actin promoter was transformed into wild type. The resulting *PpNOG2* OE line did not have an obvious phenotype and resembled wild type (Figure 5.14A). Overexpression of *PpNOG2* was confirmed by RT-qPCR (Figure 5.14B). Although the difference in expression level was not statistically significant, due to high variability between biological samples, I judged it to be a convincing overexpressor and proceeded with this line for the RNA-seq experiment.

I then extracted RNA from triplicate samples of wild type, *Ppnog2R* and *PpNOG2* OE tissues, that had been grown for 2 weeks. After sequencing of the mRNA, I performed DGE analysis. In *Ppnog2R*, 8.68% of all genes were significantly differentially expressed compared to wild type ($\text{padj} < 0.05$; Figure 5.15A and C). In *PpNOG2* OE 5.21% of all genes were differentially expressed compared with wild type (Figure 5.15B and C). Compared to the previous experiment, lower numbers of genes were differentially expressed in response to changes in the level of functional *PpNOG2*. Therefore, *PpNOG2* does not exert as much control over the transcriptome as *PpNOG1*. This was not surprising as it is already known that *PpNOG2* is not as early acting as *PpNOG1* in the regulatory network that controls the transition to 3D growth.

I performed GSEA on the differentially expressed genes for each comparison, to identify biological processes that were significantly perturbed in *Ppnog2R* and *PpNOG2* OE, compared to wild type. For both comparisons, there were no significantly suppressed biological processes. The top ten significant activated

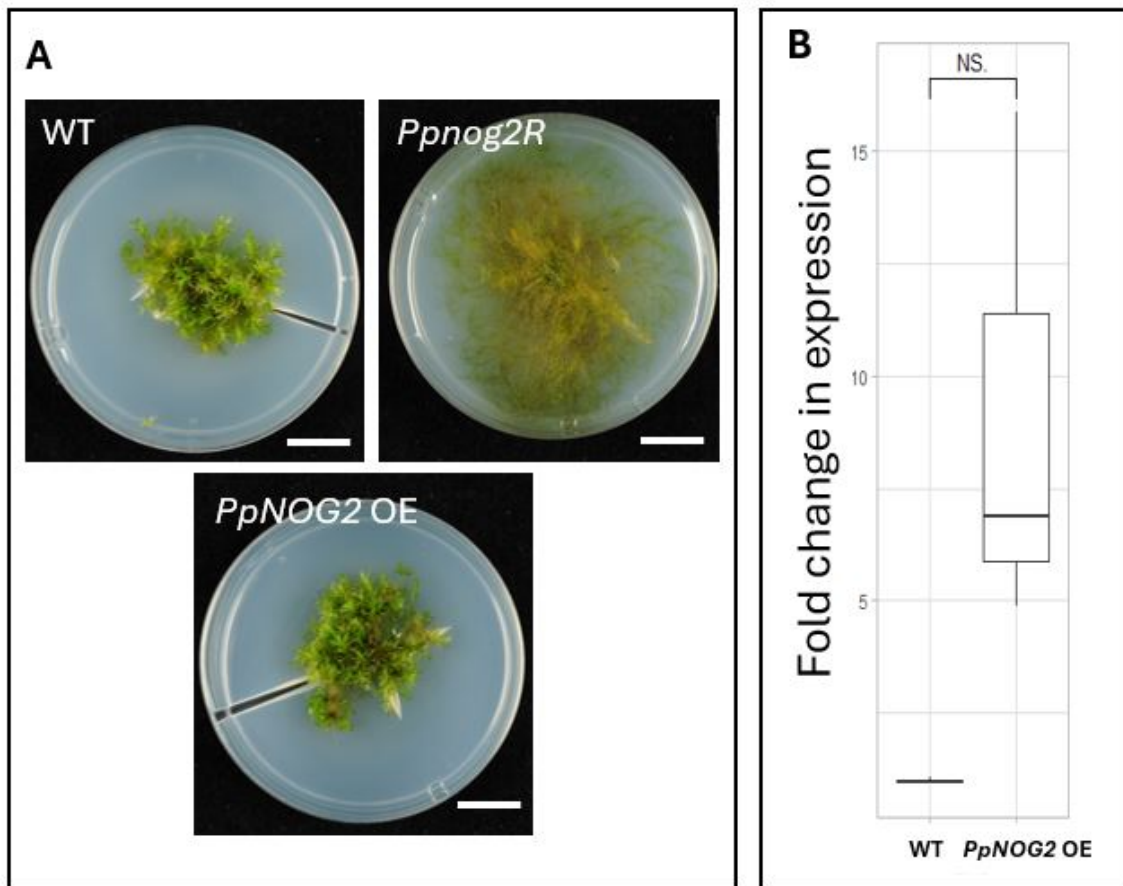


Figure 5.14: A) Representative images of wild type, *Ppnog2R* and *PpNOG2* OE at 7 weeks of growth. Scale = 1 cm. B) RT-qPCR data showing the fold change of expression of *PpNOG2* in wild type versus *PpNOG2* OE. A t-test was used to compare the fold change.

processes for wild type versus *Ppnog2R* included several related to protein synthesis, photosynthesis, amide biosynthesis, ribosome biogenesis and primary metabolism (Figure 5.16A). For wild type versus *PpNOG2* OE, the processes that were significantly activated were all related to responses to biotic and abiotic stressors (Figure 5.16B).

5.2.5 *PpNOG2* controls the expression of genes associated with cellular transport

To identify genes that were directly affected by the expression level of *PpNOG2*, I filtered the DEGs for genes that were oppositely differentially expressed in *Ppnog2R* and *PpNOG2* OE (either downregulated in *Ppnog2R* and upregulated in *PpNOG2* OE or *vice versa*). This search yielded 122 genes (Figure 5.17).

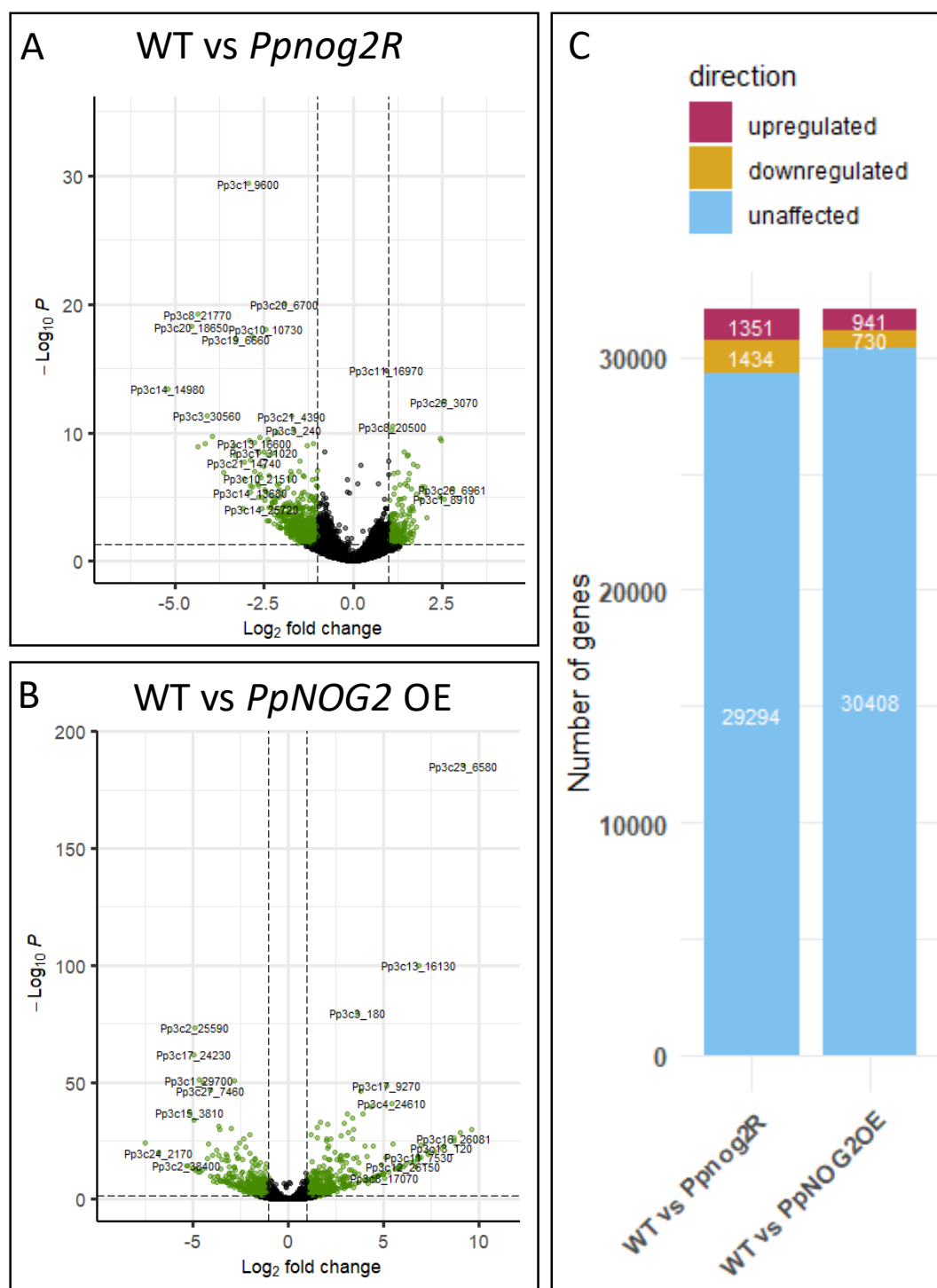


Figure 5.15: A-B) Volcano plot showing differential gene expression. Green points indicate that the gene was significantly differentially expressed (adjusted p-value < 0.05) with a $\log_2$ fold change of at least 1. Some strongly affected genes are labelled with their gene accession number. For each line comparison (eg. X vs Y), X is the baseline, and Y is the line for which gene expression changes are being shown. A) wild type versus *Ppnog2R*. B) wild type versus *PpNOG2* OE. C) Graph showing the number of unaffected, upregulated and downregulated genes for each comparison.

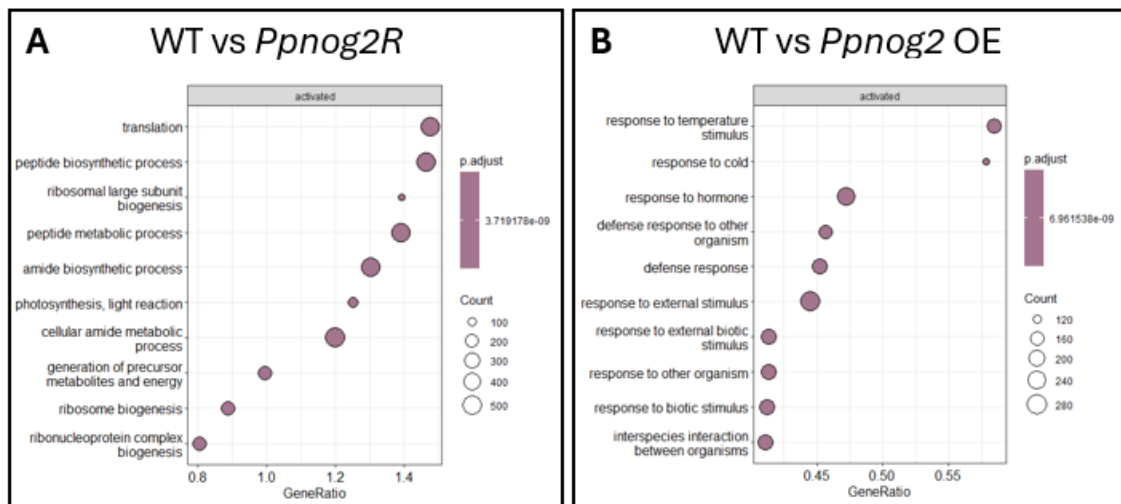


Figure 5.16: A-B) Gene Set Enrichment Analysis (GSEA) with Gene Ontology (GO) biological process terms. For each comparison there are ten of the most significant activated biological processes. In both cases there were no significantly suppressed biological processes. The colour indicates the adjusted p-value for the test for enrichment. Count indicates the number of input genes. Gene ratio is the ratio of input genes to the total number of genes in the gene set. Note that because *P. patens* genes were BLASTed against *A. thaliana* for this analysis, there were cases where *P. patens* had multiple genes mapping to the same Arabidopsis gene, resulting in some gene ratios exceeding 1. For each line comparison (eg. X vs Y), X is the baseline, and Y is the line for which gene expression changes are being shown. A) wild type versus *Ppnog2R*. B) wild type versus *PpNOG2* OE.

Those genes included two basic helix-loop-helix (bHLH) DNA-binding proteins that were putative transcription factors. One of these bHLH transcription factors was actually *PpHEC1b*. There was also one auxin response factor gene and several genes associated with responses to biotic and abiotic stress, particularly heavy metal tolerance and disease resistance. There were two cellulose synthase genes, which are involved in cell wall synthesis, and one pectin lyase, which is important for cell wall remodelling. There were also two genes that are closely related to *PpNOG2* (*Pp3c22_3160* and *Pp3c22_4140*; both hydroxycinnamoyl-CoA shikimate/quinate hydroxycinnamoyl transferases), which were both downregulated in *Ppnog2R* and upregulated in *PpNOG2* OE (Moody et al., 2021).

I performed over-representation analysis on the list of oppositely differentially expressed genes in wild type versus *Ppnog2R* and wild type versus *PpNOG2* OE. I used GO terms and instead of limiting the analysis to biological processes, I extended the analysis to all three GO categories: biological processes, molecular functions

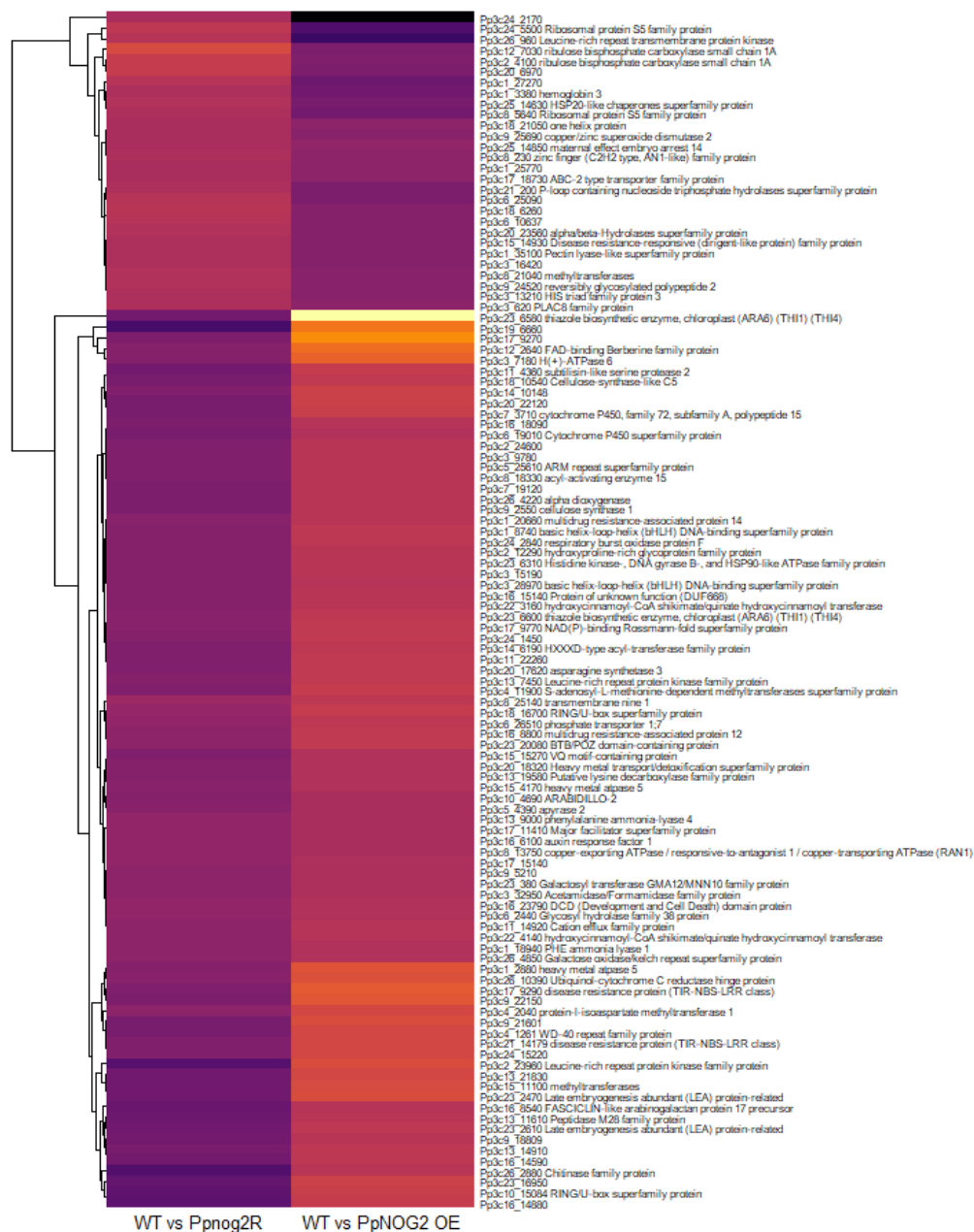


Figure 5.17: Log₂ fold change for oppositely differentially expressed genes in wild type versus *Ppnog2R* and wild type versus *PpNOG2* OE. In both cases WT is the baseline and the plot is showing changes in gene expression in *Ppnog2R* and *PpNOG2* OE respectively, compared to the WT levels

and cellular components as there were no significantly over-represented biological processes (adjusted p-value < 0.05). This showed that these oppositely differentially expressed genes were enriched in genes associated with cellular trafficking (Figure 5.18).

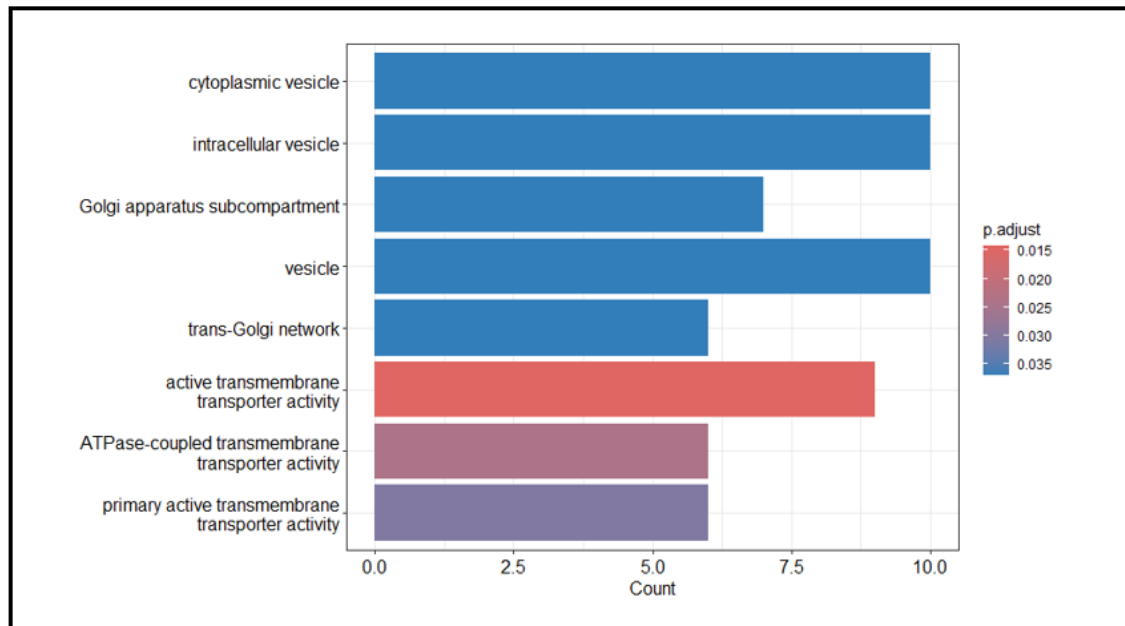


Figure 5.18: Over-representation analysis on genes that were oppositely differentially expressed in wild type versus *Ppnog2R* and wild type versus *PpNOG2* OE. Colour represent p-value from the over-representation analysis.

5.2.6 Many known 3D regulators were differentially expressed in *Ppnog2R* and *PpNOG2*OE

As with the previous RNA-seq experiment, I examined the expression of known or putative 3D regulators in wild type, *Ppnog2R* and *PpNOG2* OE (Figure 5.19)

The expression of hormone signalling genes was somewhat perturbed in *Ppnog2R* and *PpNOG2* OE. For many of the genes there wasn't a clear pattern due to variability between biological replicates. There was a subset of genes with expression that was positively regulated by the presence of functional *PpNOG2* as they were downregulated in *Ppnog2R* and upregulated in *PpNOG2* OE relative to wild type. This subset included auxin response factors, auxin transporters, including *PpPINA*, auxin receptors and cytokinin oxidases. There was also a small subset of genes that

appeared to be slightly upregulated in *Ppnog2R* and downregulated in *PpNOG2* OE relative to wild type. This subset included three *PpLOG* genes, which are involved in the final step of cytokinin biosynthesis.

For genes involved in the cytoskeleton and asymmetric cell division and also those involved in the cell cycle, no strong patterns emerged, again due to variability among biological replicates. For those genes associated with maturation and developmental transition, *APB1,2,3* and *4* were downregulated in *Ppnog2R*, although not much affected in *PpNOG2* OE, relative to wild type. This is surprising as *Ppnog2R* has an overbudding phenotype, which often correlates with enhanced expression of the *PpAPBs*. In previous RT-qPCR experiments *PpAPB1* was upregulated in *Ppnog2R*, whereas the other three *PpAPB* genes were not significantly affected (Moody et al., 2021).

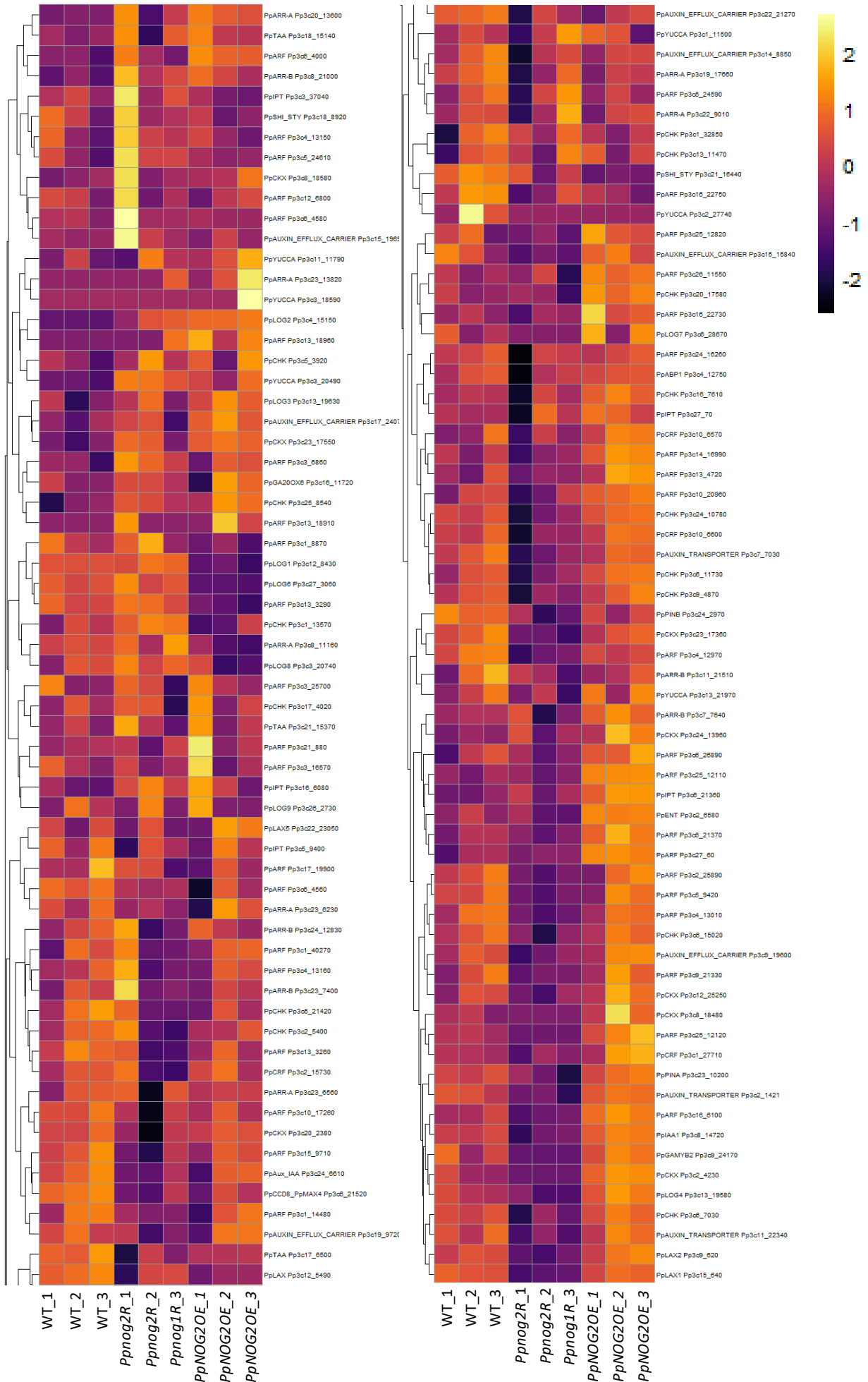
For the selected protein degradation-related genes, several components of the N-degron pathway, including a *PpPRT1* gene, *PpATE*, and several *PpPRT6* genes were downregulated in *Ppnog2R* compared to wild type. Notably, *PpDEK1* was downregulated. *Ppdek1* mutants also produce an increased frequency of buds like *Ppnog2R* (Perroud et al., 2014). Perhaps the overproduction of buds in *Ppnog2R* is related to the suppression of *PpDEK1* expression.

Around half of the cell wall-related genes had suppressed expression in *Ppnog2R* compared to wild type. A smaller subset were upregulated in *Ppnog2R* and downregulated in *PpNOG2* OE and this included all three selected *PECTIN LYASE-LIKE* genes, which are important for cell wall remodelling.

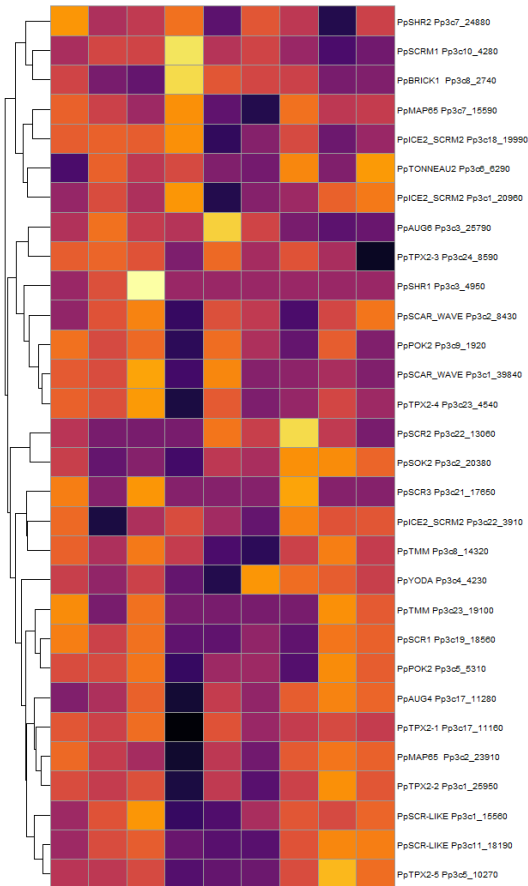
The only notable pattern in the apical meristem maintenance genes was that *PpCLE3,4,5,7,8, 9* of the *CLV* signalling pathway as well as *PpIMK3*, *PpTAW3* and two *PpDORNRÖSCHEN* genes were upregulated in *PpNOG2* OE compared to wild type and *Ppnog2R*.

Surprisingly, most of the chalcone synthases were downregulated in *Ppnog2R*, which refutes my earlier hypothesis that lower expression of *PpNOG2* results in upregulation of components of the flavonoid biosynthesis pathway due to increased flux through that pathway.

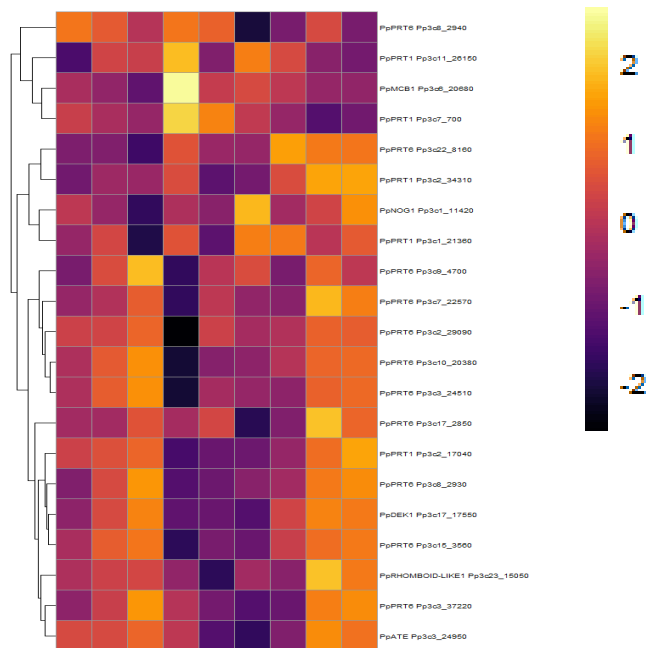
Hormone signalling



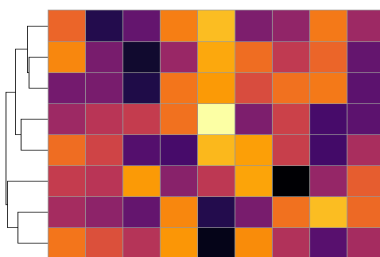
Cytoskeleton and/or asymmetric cell division



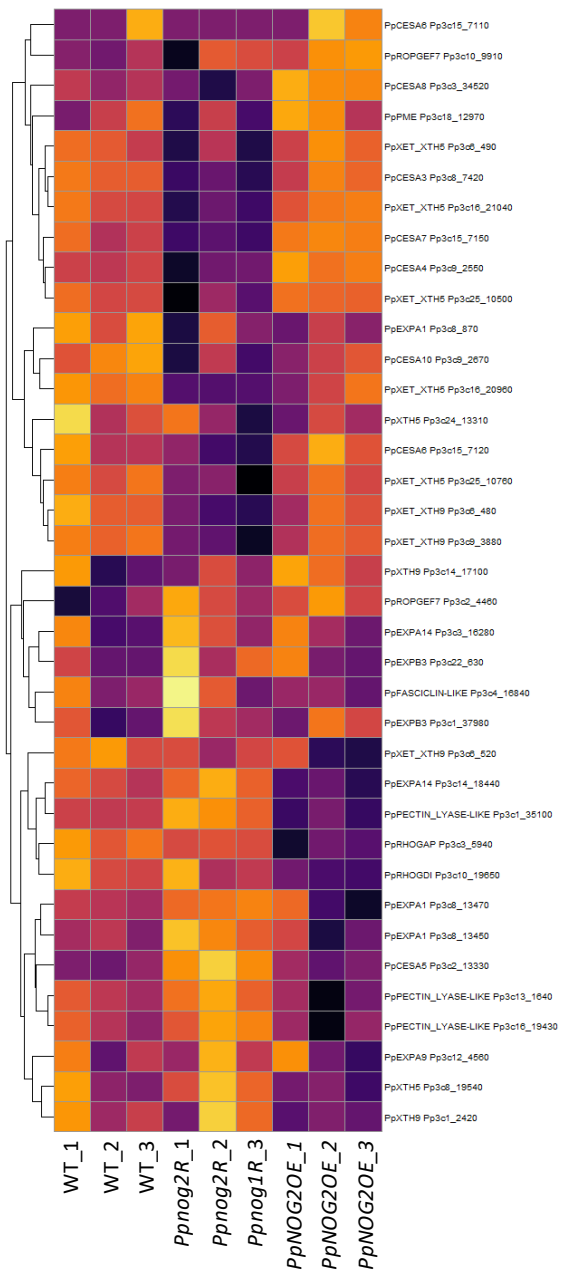
Protein degradation



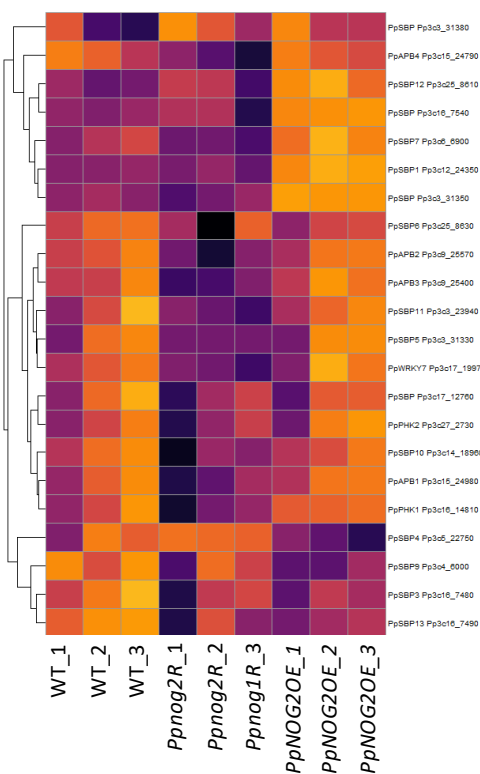
Cell cycle/proliferation



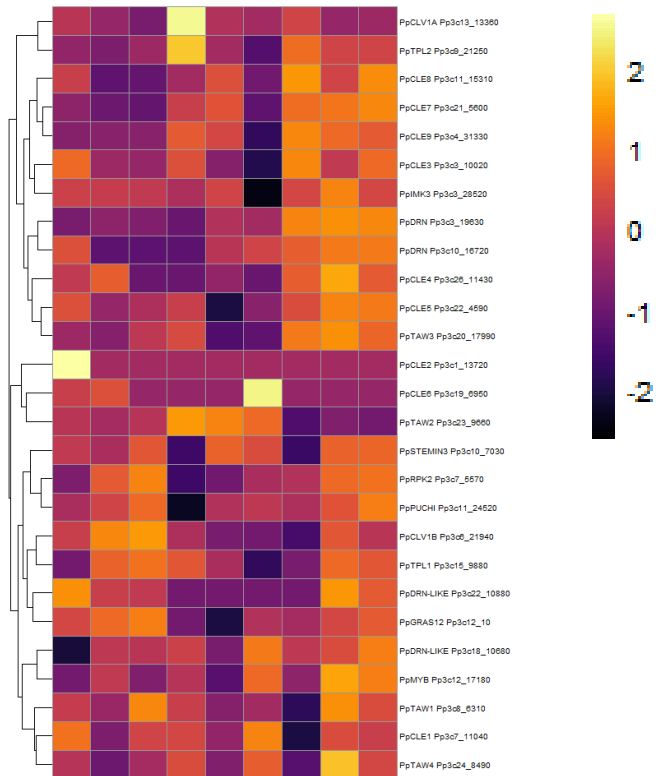
Cell wall synthesis/remodelling



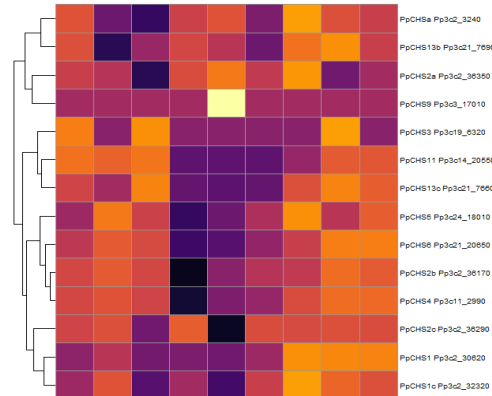
Maturation/developmental transition



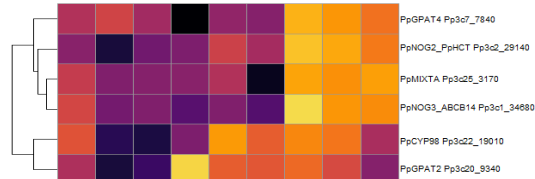
Apical meristem development



Flavonoid biosynthesis



Cuticle formation



Cell trafficking



Epigenetic modification



Post-transcriptional modification

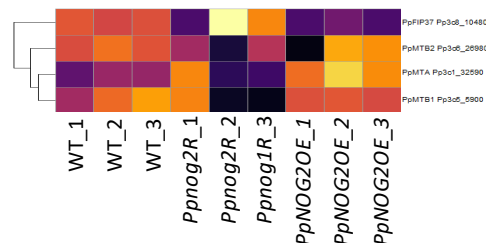


Figure 5.19: Heatmaps showing the normalised counts for each sample, $\log_2$ transformed and scaled by row. This visualises relative expression levels of selected genes in different functional categories for three biological replicates for wild type, *Ppnog2R* and *PpNOG2* OE.

PpNOG2 functions in the cuticle biosynthesis pathway (Moody et al., 2021; Kriegshauser et al., 2021). Therefore, it was not surprising that all the other selected cuticle biosynthesis genes were upregulated in *PpNOG2* OE. Again due to sample variability there were no clear expression trends for the cell trafficking, epigenetic modification and post-transcriptional modification genes.

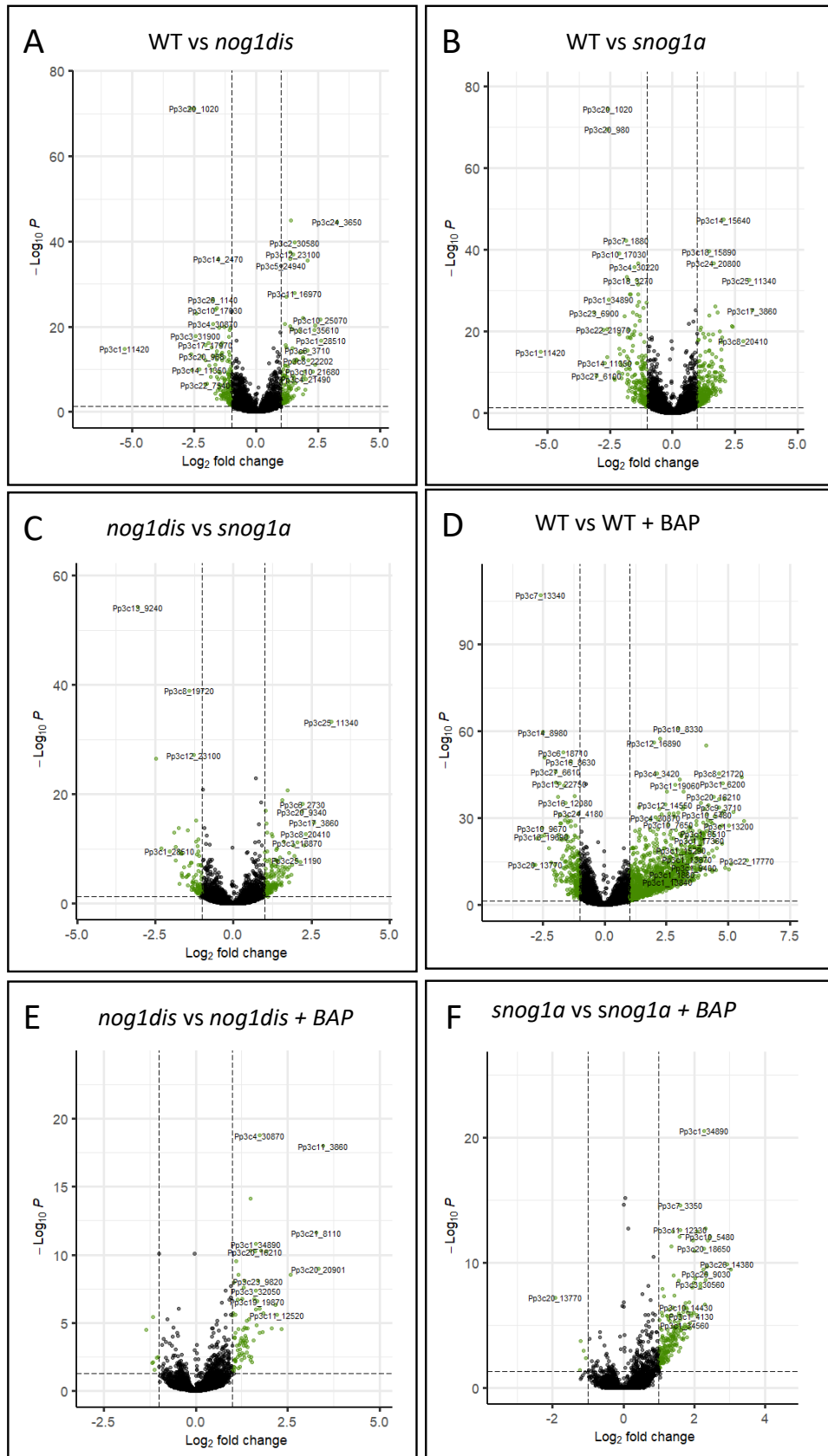
5.2.7 Fewer genes were differentially expressed in *nog1dis* compared to *Ppnog1R*

Finally, after the success of identifying the causative mutation in *snog1a*, I decided to do one final RNA-seq experiment in which I compared wild type, *nog1dis* and *snog1a*, with and without 1 μ M BAP treatment. Like *Ppnog1R*, *nog1dis* is insensitive to BAP treatment. However, *snog1a* has regained the ability to induce bud formation in response to BAP treatment, hence I wanted to explore the gene expression patterns underlying this to understand more about how the mutation in *PpSNOG1A* antagonises the effect of a mutation in *PpNOG1*.

In this final RNA-seq experiment I generated triplicate biological samples of two-week-old wild type, *nog1dis* and *snog1a*, grown in the presence or absence of 1 μ M BAP for one week. I then performed DGE analysis with six comparisons; wild type versus *nog1dis* (Figure 5.20A), wild type versus *snog1a* (Figure 5.20B), *nog1dis* versus *snog1a* (Figure 5.20C), wild type versus wild type + BAP (Figure 5.20D), *nog1dis* versus *nog1dis* + BAP (Figure 5.20E) and *snog1a* versus *snog1a* + BAP (Figure 5.20F). In wild type versus *nog1dis*, 5.80% of all genes were differentially expressed (adjusted p-value < 0.05). This was considerably fewer genes than were affected in wild type versus *Ppnog1R* (19.24%) in the first RNA-seq experiment. In wild type versus *snog1a*, 8.46% of genes were affected. In *nog1dis* versus *snog1a*, it was 3.56% of genes. For wild type versus wild type + BAP, 19.14% of genes

were impacted, which was only 1.4% difference from the 17.74% affected in the wild type versus wild type + BAP comparison in the first RNA-seq experiment. For *nog1dis* versus *nog1dis* + BAP, only 1.11% of genes were affected, with a clear skew towards the upregulation of genes. This was surprising because in the first RNA-seq experiment, *Ppnog1R* versus *Ppnog1R* + BAP showed the greatest perturbation with 27.23% of genes affected. Finally, for *snog1a* versus *snog1a* + BAP, 1.52% of genes were affected, again with a skew towards upregulation (Figure 5.20G).

I performed GSEA for each of the comparisons, using GO biological process terms. For wild type versus *nog1dis*, there was activation of processes associated with the localisation and transport of proteins. Only one process was significantly suppressed (p-value < 0.05), which was ubiquitin-dependent protein catabolism via the multivesicular body sorting pathway (Figure 5.21A). In wild type versus *snog1a*, ubiquitin-dependent protein catabolism via the multivesicular body sorting pathway was also suppressed. However, this was simply because that GO term is associated with *AtNOG1*, to which *PpNOG1* was mapped when performing this analysis. Most of the processes that were upregulated in *snog1a* were related to responses to abiotic and biotic stress or carbohydrate metabolism (Figure 5.21B). When comparing *nog1dis* and *snog1a*, there were no significantly suppressed biological processes, but some hormone response pathways were activated (Figure 5.21C). The comparison of wild type versus wild type + BAP produced similar results to the same comparison in the first experiment: suppression of photosynthesis and activation of growth and morphogenesis processes (Figure 5.21D). Even though the number of differentially expressed genes in *nog1dis* versus *nog1dis* + BAP was much less than was previously observed in *Ppnog1R* versus *Ppnog1R* + BAP, the same processes were enriched across both datasets. Genes involved in protein synthesis and cell wall processes were upregulated. Post-transcriptional regulation, chromosome organisation and protein retention in the Golgi apparatus were all suppressed (Figure 5.21E). This indicates that although the total number of genes differentially expressed differs, the processes that were being most affected remained the same. Notably, similarly to *Ppnog1R*, the *nog1dis* mutant did not exhibit a similar response to BAP as wild



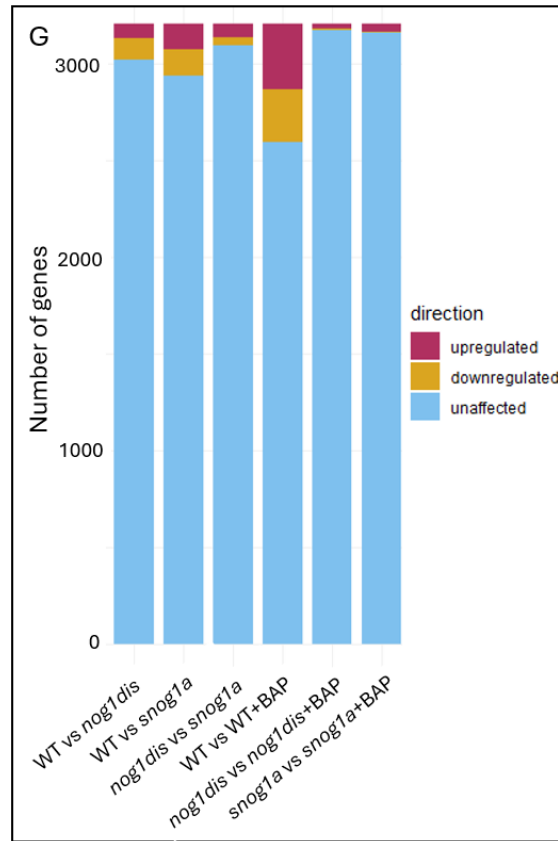


Figure 5.20: A-F) Volcano plot showing the differential gene expression. Green points indicate that the gene is significantly differentially expressed (adjusted p-value < 0.05) with a Log₂ fold change of at least 1. Some strongly affected genes are labelled with their gene accession number. For each line comparison (eg. X vs Y), X is the baseline, and Y is the line for which gene expression changes are being shown. A) wild type versus *nog1dis*. B) wild type versus *snog1a*. C) *nog1dis* versus *snog1a*. D) wild type versus wild type + BAP. E) *nog1dis* versus *nog1dis* + BAP. F) *snog1a* versus *snog1a* + BAP G) Graph showing the number of unaffected, upregulated and downregulated genes for each comparison.

type. For *snog1a* versus *snog1a* + BAP, no processes were significantly suppressed. Processes associated with the cell wall and growth dominate the activated processes (Figure 5.21F). This was similar to the processes that were activated in the wild type response to BAP.

The reason why *snog1a* is able to produce gametophores and *nog1dis* is not may be because the expression of certain genes is perturbed in *nog1dis*, and this expression perturbation is corrected in *snog1a* by the lack of functional *PpSNOG1A*. I identified genes which had altered expression in wild type versus *nog1dis*, but that were not significantly affected in wild type versus *snog1a*. This produced a list of 719 genes. I performed over-representation analysis with GO biological process terms.

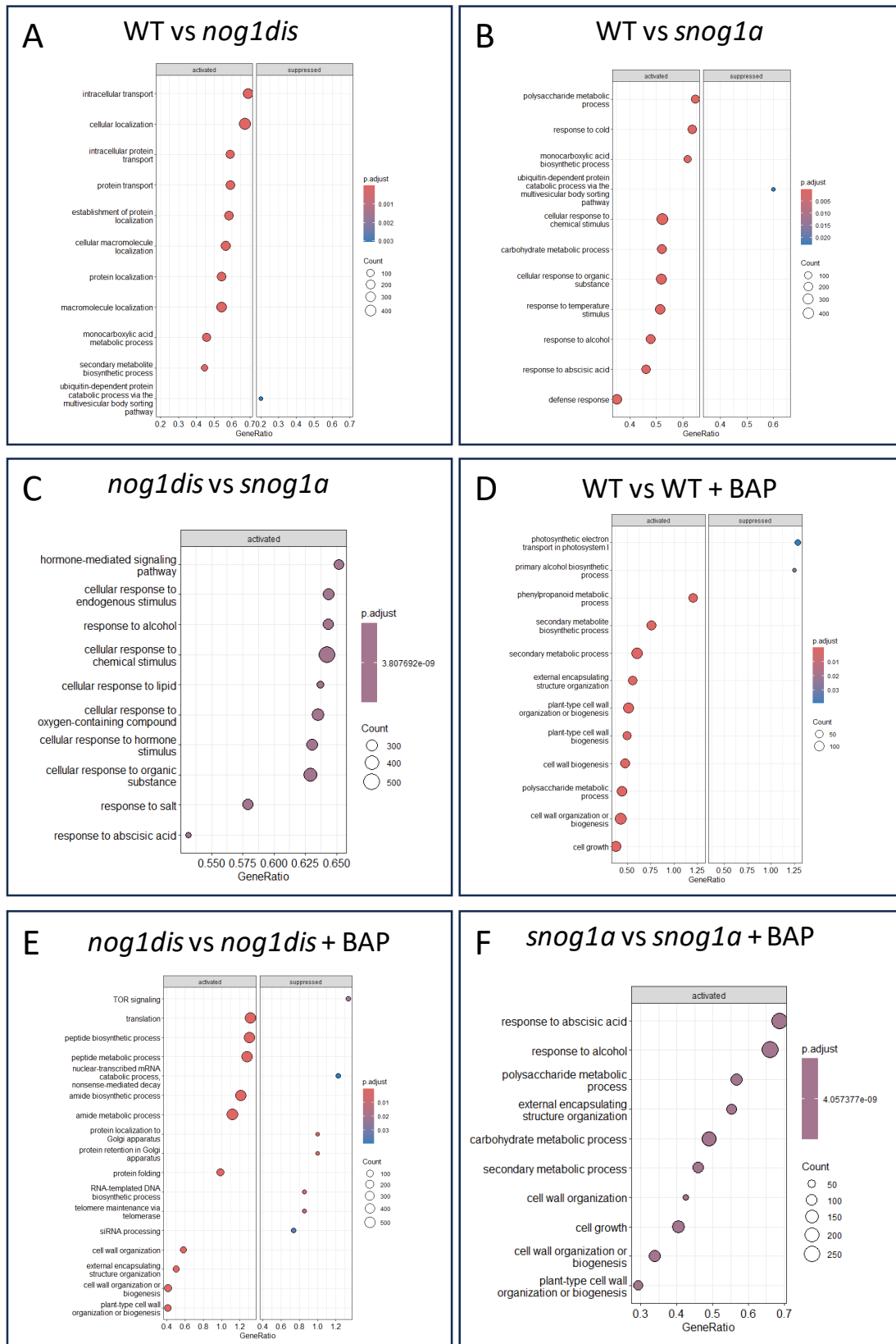


Figure 5.21: A-F) Gene Set Enrichment Analysis (GSEA) with Gene Ontology (GO) biological process terms. For each comparison there are up to ten of the most significant activated and suppressed biological processes. The colour indicates the adjusted p-value for the test for enrichment. Count indicates the number of input genes. Gene ratio is the ratio of input genes to the total number of genes in the gene set. Note that because *P. patens* genes were BLASTed against *A. thaliana* for this analysis, there were cases where *P. patens* had multiple genes mapping to the same Arabidopsis gene, resulting in some gene ratios exceeding 1. For each line comparison (eg. X vs Y), X is the baseline, and Y is the line for which gene expression changes are being shown. A) wild type versus *nog1dis*. B) wild type versus *snog1a*. C) *nog1dis* versus *snog1a*. D) wild type versus wild type + BAP. E) *nog1dis* versus *nog1dis* + BAP. F) *snog1a* versus *snog1a* + BAP

This highlighted various stress-response processes as being perturbed in *nog1dis*, but not in *snog1a* (Figure 5.22A).

Due to the differences between the transcriptomes of *Ppnog1R* and *nog1dis*, I tried to identify what I will refer to as ‘core *nog1* genes’. I compared the DEGs for wild type versus *Ppnog1R* and wild type versus *nog1dis* to identify those that were affected in both. Overall, 922 genes were differentially expressed in both wild type versus *Ppnog1R* (out of 6172) and wild type versus *nog1dis* (out of 1862). Next, I filtered out those that were differentially expressed in opposite directions and was left with 741 genes. I performed over-representation analysis on this list of genes to identify enriched biological processes that were affected in both *Ppnog1R* and *nog1dis*. This showed enrichment of processes associated with stress responses (Figure 5.22B). Interestingly, the processes that were enriched in the ‘core *nog1*’ perturbed genes had extensive overlap with the processes that were perturbed in *nog1dis* and not in *snog1a*. Furthermore, when looking at the whole dataset of differentially expressed genes in wild type versus *nog1dis*, these processes did not appear as the most significant, which means they were enriched in those subsets, rather than in the whole dataset (Figure 5.21A).

I also compared the BAP responses of *Ppnog1R*, which exhibits a very strong transcriptional BAP response and *nog1dis*, which exhibits a very weak one. In total only 73 genes responded the same way in *Ppnog1R* versus *Ppnog1R* + BAP and *nog1dis* versus *nog1dis* + BAP. Because the list is small, when I did over-representation analysis, there were no significantly affected biological processes

(p-value < 0.05). This was still the case when I extended the analysis to the other GO categories. When I raised the p-value threshold to < 0.1, the biological processes that seem to be enriched in this dataset were associated with growth and development (Figure 5.22C).

Next I compared the list of genes that have perturbed expression in *nog1dis*, but not in *snog1a* relative to wild type with the list of ‘core *nog1*’ genes and got a new list of 160 genes (Table S2). When I did ORA on this list and it showed that genes associated with phenylpropanoid metabolic processes, drought response, response to jasmonic acid, response to fatty acids and response to alcohol were over-represented (Figure 5.23). When I examined this list of genes, I noticed that it contained at least 13 putative transcription factors, a paralogue of *PpNOG2*, two cellulose synthases and two potential cytokinin receptors, among others (Table S2).

I had previously identified two *HECATE1* candidates for targeted investigation by probing the first experimental dataset. I examined the expression of those genes from the current experiment to see if they follow the same trends in expression in *nog1dis* as they did in *Ppnog1R*. Both *PpHEC1a* and *PpHEC1b* responded the same way in both experiments. *PpHEC1a* was upregulated by BAP in the wild type, but unaffected in *nog1dis*. *PpHEC1b* was upregulated by BAP in wild type, but downregulated by BAP in *nog1dis*. *PpHEC1a* was also elevated in *snog1a* and *snog1a* + BAP compared to the level in *nog1dis* and *nog1dis* + BAP (Figure S32).

5.2.8 A few known 3D regulators exhibit partial restoration of cytokinin responsiveness in *snog1a*.

As before I examined the expression of a curated list of genes known or putatively involved in the 3D transition or important downstream processes (Figure 5.24). The genes involved in hormone signalling were mostly affected by BAP treatment in wild type, and were both upregulated or downregulated. However, in both *nog1dis* and *snog1a*, there was very little BAP responsiveness among these genes. In *snog1a*, a few of the genes seemed to have regained some of the BAP responsiveness that was lacking in *nog1dis*. Several *ARR-A* genes were slightly upregulated in response

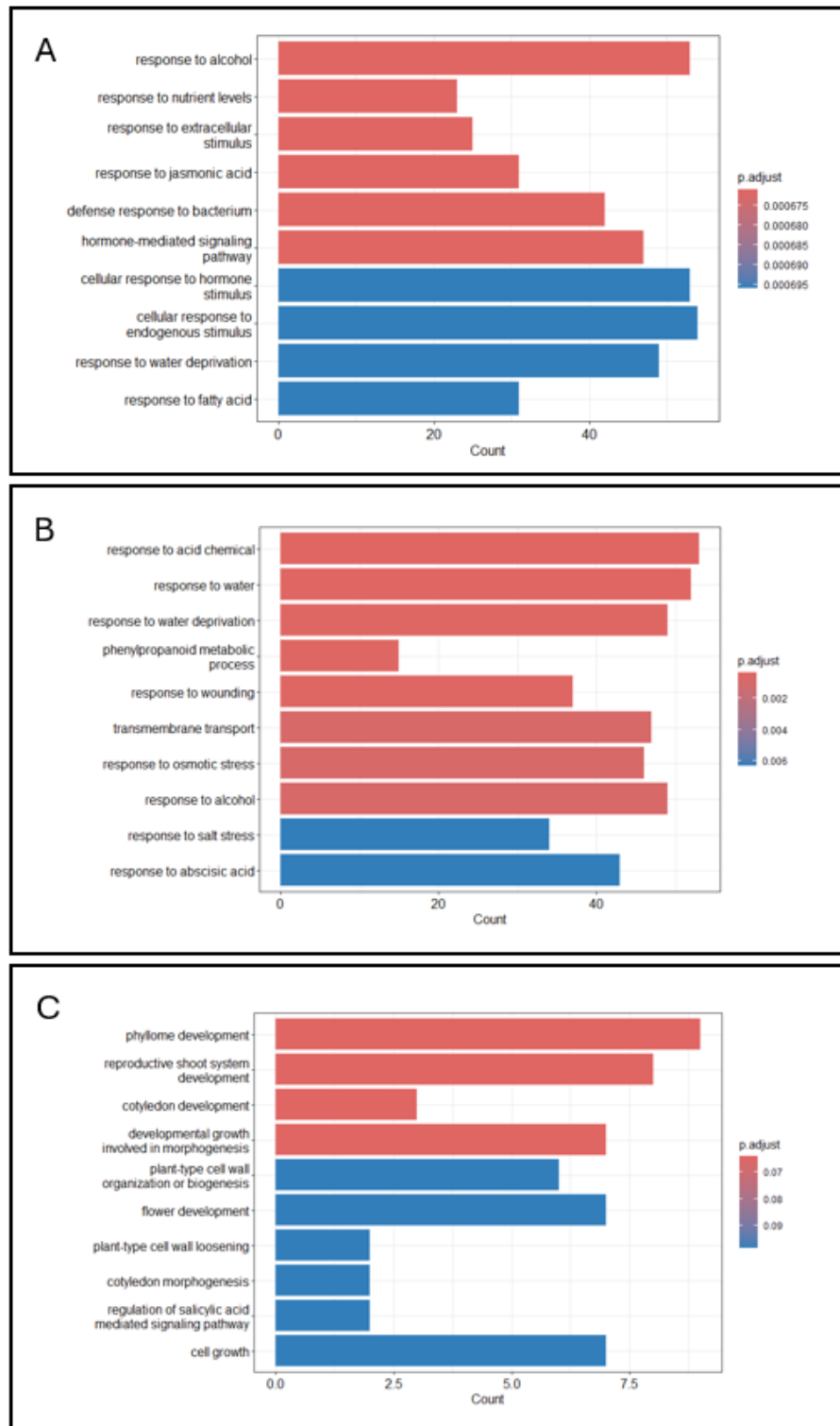


Figure 5.22: Over-representation analysis with GO biological process terms. Color represents p-value for the over-representation test and count refers to the number of genes in the input list associated with each biological process. A) genes significantly differentially expressed in wild type versus *nog1dis*, but not significantly differentially expressed in wild type versus *snog1a* (p-value < 0.05). B) ‘core *nog1*’ genes. (p-value < 0.05). C) ‘core *nog1*’ BAP-responsive genes. (p-value < 0.1)

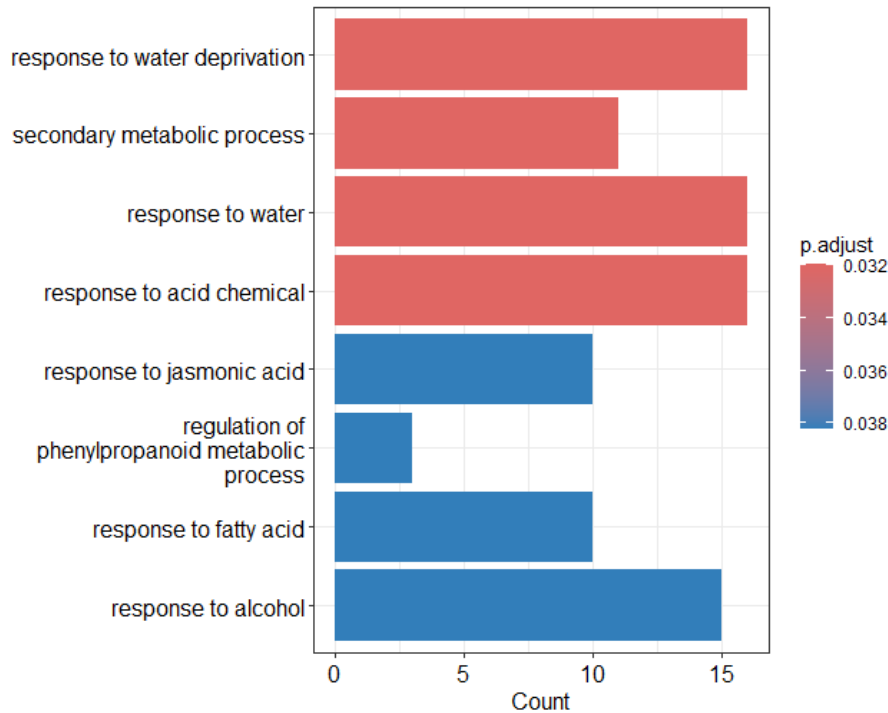


Figure 5.23: Over-representation analysis with GO biological process terms. Color represents p-value for the over-representation test and count refers to the number of genes in the input list associated with each biological process. Analysis performed on list of ‘core *nog1*’ genes that are also not differentially expressed in *snog1a* compared to wild type.

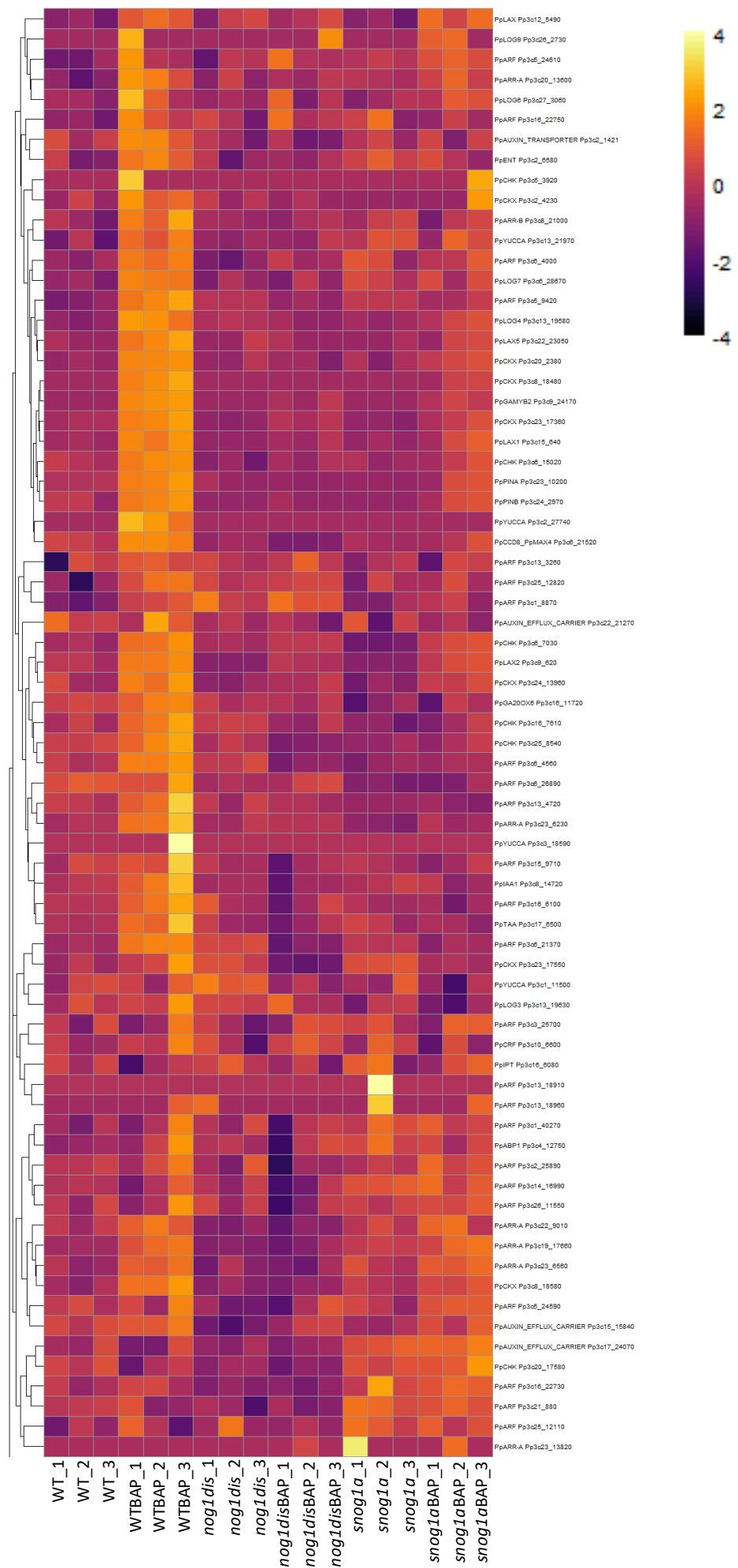
to BAP in *snog1a* as well as in wild type. Another subset of genes including two *IPT* genes and two *ARR-B* genes was downregulated in response to BAP in wild type and *snog1a*, but not *nog1dis*. Three *ARFs*, a *CKX* and an *AUXIN EFFLUX CARRIER* were not responsive to BAP in any of the lines but they were upregulated in *snog1a* compared to wild type, but not in *nog1dis*. A loss of BAP-responsiveness in *nog1dis* and *snog1a* was also seen in several of the cell cycle and proliferation genes and also the genes involved in the cytoskeleton, asymmetric cell division, apical meristem development, maturation, flavonoid biosynthesis, protein degradation, cuticle formation, cell trafficking and post-transcriptional modification. For the cell wall-associated genes, there was some weak recovery of BAP responsiveness in *snog1a*, that was lacking in *nog1dis*. A few of the cell wall genes were strongly induced by BAP in *nog1dis*, but not in wild type or *snog1a*. There was no clear patterns in the expression of epigenetic modification genes. Of the cuticle formation genes, *PpNOG3* and *PpGPAT2* were upregulated in *snog1a* compared to both wild

type and *nog1dis*. Overall, when examining known regulators, *snog1a* seems to exhibit broadly the same expression patterns as *nog1dis*, with only a few exceptions.

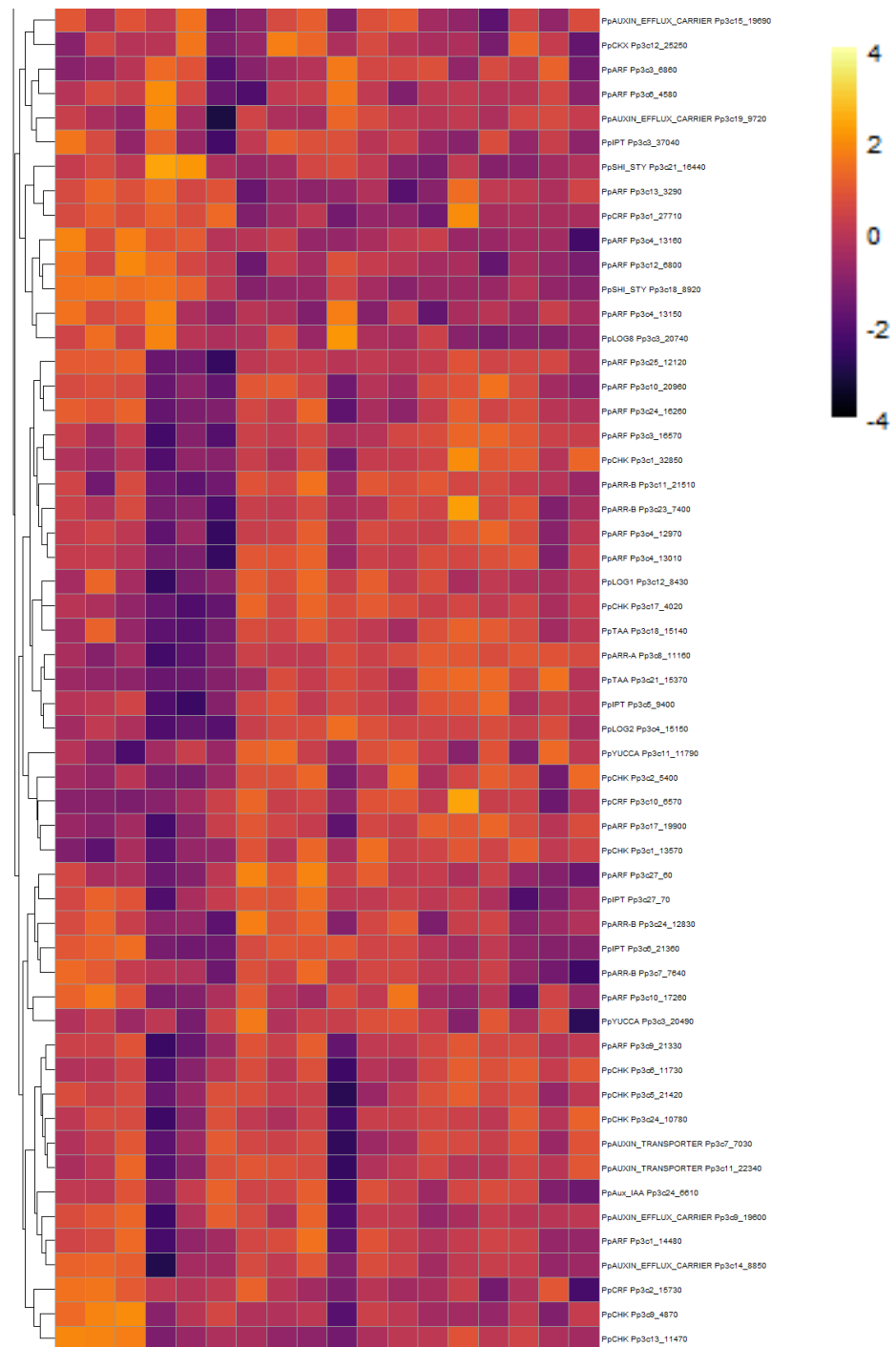
5.3 Discussion

The absence of functional *PpNOG1* in the *Ppnog1R* mutant caused widespread perturbation in the transcriptional landscape of *P. patens*. In *Ppnog1R*, nearly one fifth of all genes were differentially expressed in *Ppnog1R* versus wild type. Contrary to its phenotypic insensitivity to BAP, at the transcriptional level, *Ppnog1R* was hypersensitive, with many more genes being affected in *Ppnog1R* than were in wild type. Additionally, the majority of genes that respond to BAP in *Ppnog1R* do not do so in wild type. This means that many of the genes and therefore functions that were being activated or suppressed in *Ppnog1R* were not those that are usually induced by cytokinin. However, I have discovered that I must be careful in how I interpret this data as the strong impact that *Ppnog1R* has on gene expression is apparently not universal to all *nog1* mutants. *nog1dis* and *Ppnog1R* were both generated in the same genetic background; the Villersexel ecotype. Experimental conditions were consistent enough to produce broadly similar results for the wild type versus wild type + BAP comparisons in both experiments. *Ppnog1R* was generated by random UV-mutagenesis, which results in many single nucleotide polymorphisms (SNPs) throughout the genome. It is possible that these background SNPs have caused disruption to genes other than *PpNOG1* that affected the transcriptome, but that were not relevant to the *Ppnog1R* phenotype and therefore were not revealed as a mutated locus during the mapping of *Ppnog1R* (Moody et al., 2018). In contrast, *nog1dis* was generated in a clean, non-mutated background. Another possibility is that the presence of the truncated Ppnog1R protein product somehow caused the disruption. Truncated proteins can be deleterious to the cell. Transcripts containing premature STOP codons (nonsense mutations) are often actively degraded by the cell by the non-sense mediated decay pathway (Kalyna et al., 2012). From the

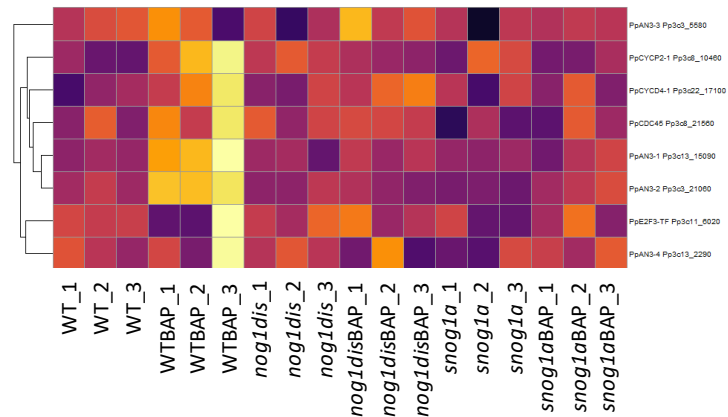
Hormone signalling



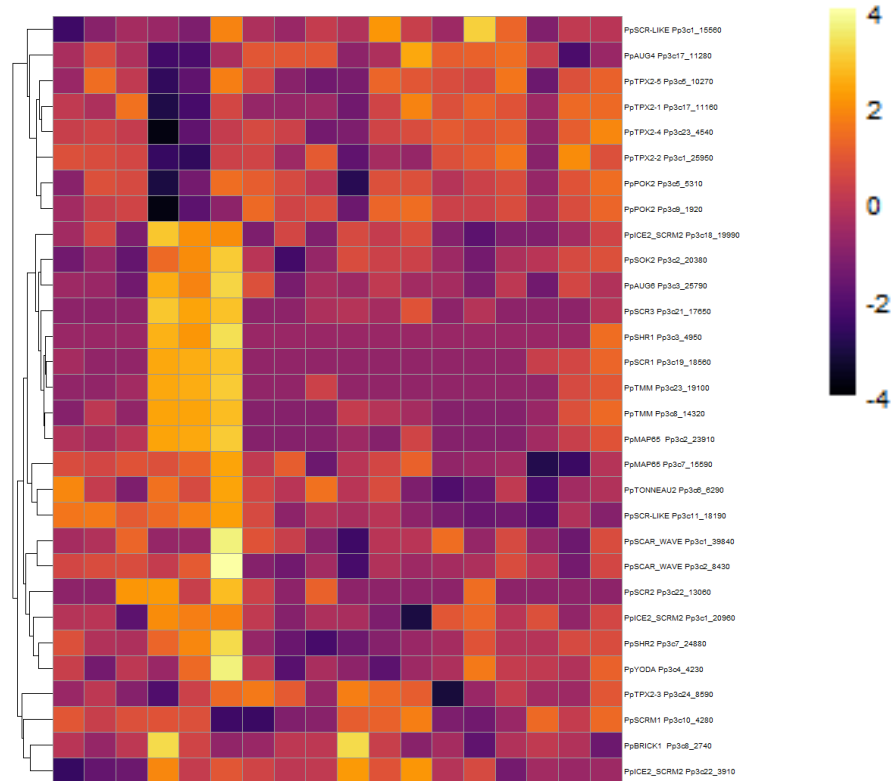
Hormone signalling



Cell cycle/proliferation



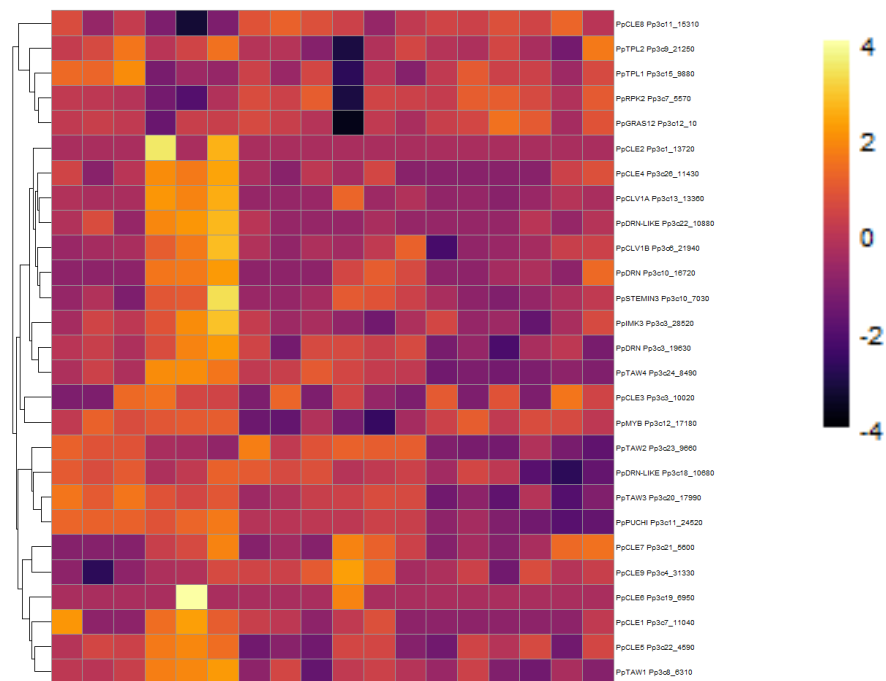
Cytoskeleton and/or asymmetric cell division



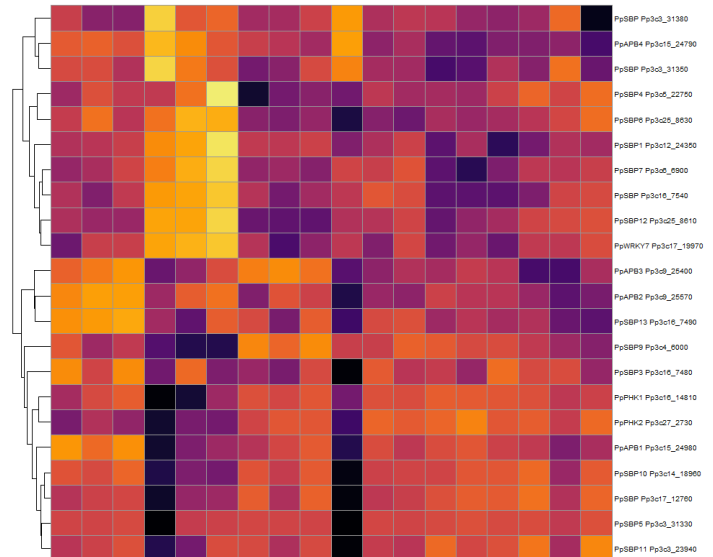
Cell wall synthesis/remodelling



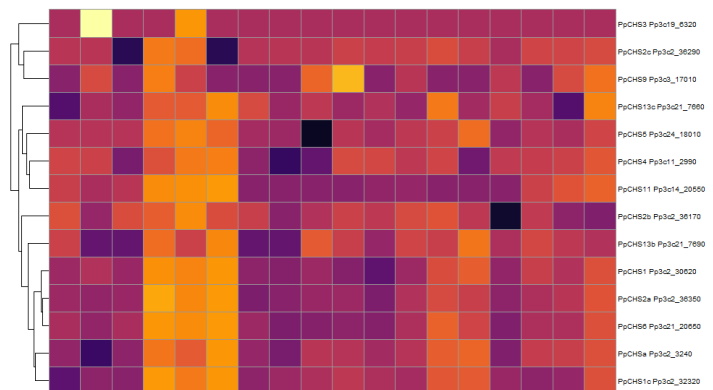
Apical meristem development



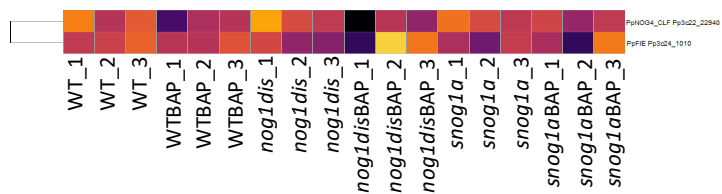
Maturation/developmental transition



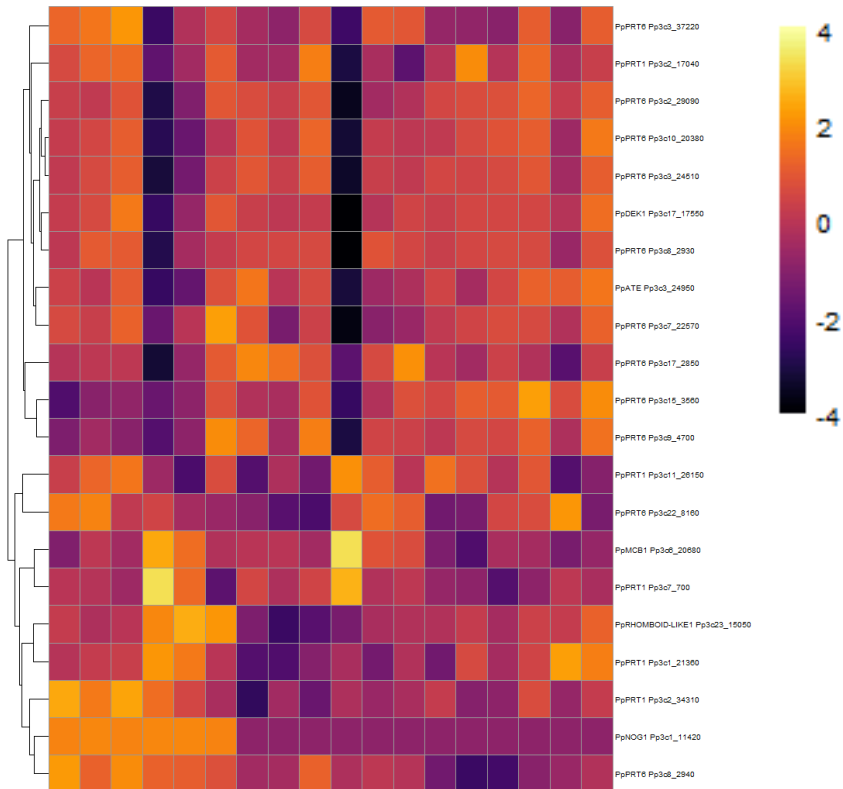
Flavonoid biosynthesis



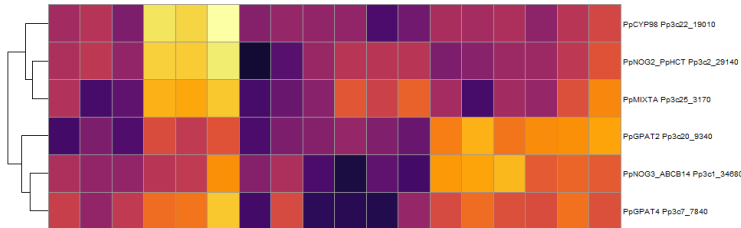
Epigenetic modification



Protein degradation



Cuticle formation



Cell trafficking



Post-transcriptional modification

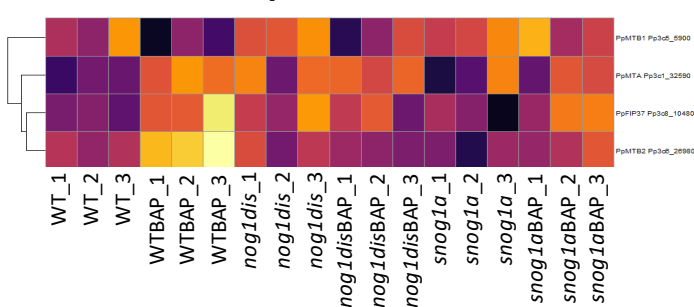


Figure 5.24: Heatmaps showing the normalised counts for each sample, $\log_2$ transformed and scaled by row. This visualises relative expression levels of selected genes in different functional categories for three biological replicates for wild type, wild type + BAP, *nog1dis*, *nog1dis* + BAP, *snog1a* and *snog1a* + BAP.

RNA-seq data, I could see that the *PpNOG1* transcript carrying the premature STOP codon is somewhat suppressed in *Ppnog1R*, but not entirely (data not shown). Therefore, the truncated protein could have been present in the cell and may have caused serious perturbations to the transcriptome. In *nog1dis* and *snog1a*, *PpNOG1* is not transcribed at all (no reads were mapped to the gene). The effect that BAP had on the transcriptome of *nog1dis* was much smaller than the effect on *Ppnog1R*. The small number of genes that were induced by BAP in *nog1dis* was more in line with the BAP insensitive phenotype, than the dramatic affects seen in *Ppnog1R*. Using both datasets I identified the gene expression changes that were common between *Ppnog1R* and *nog1dis*, as well as their common BAP-responsive genes. With this I hope to have narrowed down the list of potential genes that are key to unlocking the 3D regulatory network. The cytokinin response was still very dampened in *snog1a* compared to wild type as the majority of normally BAP responsive genes were unaffected in both *nog1dis* and *snog1a*. However, a few important genes showed some recovery of BAP responsiveness. What is still to be revealed is how the loss PpSNOG1A allows a restoration of BAP-induced gametophore induction. The answer to that may lie in the list of genes that are differentially expressed in *Ppnog1R* and *nog1dis* ('core *nog1*' genes) and also not differentially expressed in *snog1a*, compared to wild type. This list included many transcription factors and further investigation into these may help uncover why *snog1a* can produce gametophores and respond to cytokinin when *nog1dis* cannot.

My examination of the first dataset, comparing BAP responses in wild type and *Ppnog1R* led me to identify a transcription factor that is known to integrate hormone signals to regulate the shoot apical meristem in Arabidopsis; HECATE1. When I overexpressed *PpHEC1b* in *Ppnog1R* it resulted in a severely deleterious phenotype, which did not occur when I overexpressed the same gene in wild type. That deleterious phenotype showed recovery over time, resulting in first *nog1*-like

protonemal growth, but then the successful induction of a few gametophores. These gametophores were small and developmentally defective, but this still represents a mild rescue of the *Ppnog1R* phenotype. My hypothesis is that the expression of *PpHEC1b* driven by the maize ubiquitin promoter was too high and therefore produced a deleterious phenotype. The spherical cells of the early stages of the *Ppnog1R pUBI:PpHEC1b-GFP* lines resembled other described *P. patens* mutants that exhibit defective tip growth. These mutants are defective in the actin cytoskeleton which is essential for tip growth (Vidali et al., 2007, 2010). It is unclear how the overexpression of *PpHEC1b* resulted in the severe inhibition of tip growth in protonemal cells. I hypothesise that the gradual recovery of the phenotype was due either to gene silencing, or possibly due to some kind of sequestration of the excess transcripts or protein. However, this hypothesis needs to be verified. It would be informative to create a construct that drives the expression of *PpHEC1b* with a less strong promoter, to see if it produced as severe a phenotype when initially transformed into *Ppnog1R*. During the early stages, when the *Ppnog1R pUBI:PpHEC1b-GFP* lines were still exhibiting defective tip growth and existed primarily as spherical cells, tip growth was restored when the *Ppnog1R pUBI:PpHEC1b-GFP* lines were grown on BCDAT media, rather than BCD. It is unclear why this is the case as I could find no explanation for why the addition of ammonium tartrate would have this effect. BAP treatment also somewhat mitigated the severe initial *Ppnog1R pUBI:PpHEC1b-GFP* phenotype, as it promoted some protonemal growth from the spherical cells of the *Ppnog1R pUBI:PpHEC1b-GFP* mutants, suggesting that there may be a lack of, or inability to respond to, endogenous cytokinin in *Ppnog1R pUBI:PpHEC1b-GFP* at that stage.

In Arabidopsis, the HEC proteins have been shown to be important for modulating auxin and cytokinin responses in various developmental contexts. HEC1, 2 and 3 are important regulators of the SAM and are integrated into the WUS/CLAVATA signalling network which controls meristem size. The *AtHEC* genes operate semi redundantly, but in *Athec1,2,3* mutants, there is a decrease in SAM size (Schuster et al., 2014). With the TCSn cytokinin signalling reporter,

it was demonstrated that *Athe1/2/3* mutants have reduced cytokinin signalling (Zürcher et al., 2013; Gaillochet et al., 2017). The *Athe1,2,3* phenotype can also be complemented with BAP treatment. There is a corresponding expansion of the cytokinin signalling domain in the SAM of *AtHEC1* overexpression lines. Therefore HEC proteins promote cytokinin signalling in the SAM. HEC proteins were also shown to negatively regulate cell fate transitions of central zone cells to peripheral zone cells, thereby influencing meristem size by regulating recruitment from the central zone to the peripheral zone. HECs also interfere with auxin signalling feedbacks as they have a negative effect on the expression of auxin responsive genes. *AtHEC1*, 2 and 3 can all bind directly to the auxin response factor MONOPTEROS (MP). By regulating auxin signalling in the SAM, HECs influence the formation of local auxin maxima required for organ initiation and therefore regulate the transition from peripheral zone cell to differentiated lateral organ (Gaillochet et al., 2017). HECs are also known to regulate Arabidopsis gynoecium development and are required for setting up the required auxin gradients in the developing gynoecium as it binds directly to the promoter of *PIN3* to promote expression. Overexpression of *HEC1* in Arabidopsis can result in pin-like inflorescences (Gremski et al., 2007; Schuster et al., 2015). I did not observe that overexpression of *PpHEC1b* in *P. patens* wild type or *Ppnog1R* produced mutants that particularly resembled either *PpPIN* knockout or *PpPIN* overexpression mutants (Bennett et al., 2014; Viaene et al., 2014). In Arabidopsis HECs have been shown to act as versatile regulators that directly bind to a multitude of cofactors to produce different outcomes in different developmental contexts (Gaillochet et al., 2018). For example, HEC binds to another bHLH transcription factor called SPATULA as well as a RAV family transcription factor called NGATHA to regulate the SAM and gynoecium development (Gremski et al., 2007; Schuster et al., 2014, 2015; Ballester et al., 2021). HECs interact with DRINKME, a bZIP transcriptional repressor that negatively regulates growth by repressing the cell cycle (Lozano-Sotomayor et al., 2016). HECs also have a role in integrating environmental signals to modulate growth. They directly interact with PHYTOCHROME INTERACTING FACTORS (PIFs) to regulate

photomorphogenesis and germination as well as thermomorphogenesis (Zhu et al., 2016; Gaillochet et al., 2017, 2018; Lee et al., 2021). It may be interesting to investigate the function of some of these HEC interaction partners in *P. patens*, as they may have functions relevant to the transition to 3D growth. There was no severe phenotypic defect in the wild type *pUBI:PpHEC1b-GFP* lines, which suggests that the ability of *PpHEC1b* to cause a severely deleterious phenotype is dependent on the absence of functional *PpNOG1*. As cytokinin responsiveness is perturbed in *Ppnog1R*, perhaps that makes *Ppnog1R* more susceptible to overexpression of *PpHEC1b*, which is known to influence both cytokinin and auxin signalling.

The localisation of PpHEC1b-GFP to both the nucleus, where it is expected, but also to the plasma membrane, is interesting. There are many examples of transcription factors being sequestered at cellular membranes. However, the majority of those possess a transmembrane domain that allows them to integrate into the membrane and they have to be proteolytically cleaved to release the cytosolic segment of the protein, allowing it to then enter the nucleus (Kim et al., 2006; Seo et al., 2008; Kim et al., 2010; Ng et al., 2013; Iwata et al., 2017; Eysholdt-Derzsó et al., 2023). PpHEC1b does not have any predicted transmembrane domains. There are a few examples transcription factors that do not contains transmembrane domains localising to the plasma membrane. They are modified by the addition of a lipid anchor that holds them in the membrane until the transcription factor undergoes depalmitoylation to remove the lipid anchor and release the protein (Duan et al., 2017; Liu et al., 2018). However, none of the studies of HEC in Arabidopsis reported a plasma membrane localisation, so this finding needs to be verified.

It will be necessary to allow the phenotype of *Ppnog1R pUBI:PpHEC1b-GFP-1* and -2 to stabilise before it is possible to characterise the effect that PpHEC1b has on *Ppnog1R*. *HECATE* genes act redundantly in Arabidopsis, so it will be interesting to see if the same phenotype is achieved when overexpressing PpHEC1a in *Ppnog1* (Schuster et al., 2014). Deletions of *PpHEC1b* and/or *PpHEC1a* in the wild type background may also reveal more about the function of this gene and that approach will not be affected by potential gene silencing effects. It would also be important

to overexpress the *PpHEC1* genes in *nog1dis* as well as *Ppnog1R* to verify that the observed phenotype is common to both *nog1* mutants, not just *Ppnog1R*.

A general limitation of these experiments was that they only captured the transcriptional differences in whole-plant extracts. The transition from 2D to 3D growth often involves very localised patterns of gene expression in distinct cell types (Harrison et al., 2009; Aoyama et al., 2012; Frank and Scanlon, 2015; Thelander et al., 2019; Hyoungh et al., 2020; Hata et al., 2023). Using whole plant extracts allowed me to identify broad patterns of expression between lines and treatment types, but not the more subtle local differences occurring in the developing buds and the parent caulonemal cells. Therefore, gathering tissue-specific or single-cell RNA-seq datasets for some of these mutants would be of great value, but it was not possible for me to do this within my project.

Additionally, RNA-seq only allows examination of regulatory mechanisms at the transcriptional level. Post-transcriptional and post-translational regulation is known to be in effect during this developmental transition, and so quantification of mRNA only reveals some of the story. For example, the mRNA expression levels of *PpNOG1* is fairly high and not tissue-specific. In contrast the PpNOG1 protein seems to be maintained at low levels, even when strong promoters have been used to drive ectopic overexpression of the transcript. Therefore, future proteomics experiments would likely add some further insight. However, a limitation to proteomics is that it is often not sensitive enough to detect low abundance proteins. Therefore, there may be limits to what it can reveal about the regulation of 3D growth, as many of the proteins involved are thought to be present only at low levels.

Chapter 6: General discussion

Apical cells that can divide in 3 or more planes to generate complex plant shoot architectures, was an acquired trait that facilitated the evolutionary transition on to land. The moss *Physcomitrium patens* has become an effective model species in which to study the specification of 3D-capable apical cells. This has allowed researchers to investigate the genetics underpinning this fundamental and integral part of land plant development. In this work, I have been trying to further our knowledge of the genetic regulatory network that underpins the transition from 2D growth to 3D growth. Using 3D-defective mutants that were generated and characterised in our lab, I have used forward and reverse genetics approaches to further explore the functions of *PpNOG1* and uncover the novel regulator *PpSNOG1A*.

In the absence of PpNOG1, 3D growth does not occur. Therefore, PpNOG1 promotes the transition to 3D growth. The loss of PpSNOG1A in *nog1dis*, permitted 3D growth to occur again, although to a lesser extent than in wild type. Since the loss of PpSNOG1A reverses the 3D-defective phenotype of *nog1dis*, it is clear that PpSNOG1A normally inhibits 3D growth and possibly antagonises PpNOG1. I was not able to generate a PpSNOG1A deletion in the wild type background, however, the prevalence of gametophore-bearing progeny from the *snog1a* X Reute::mCherry cross suggests that individuals with wild type *PpNOG1* and mutated *Ppsnog1a* do develop gametophores, although whether they do so in a wild type-like manner or exhibit a phenotype, is yet to be determined. When I overexpressed *PpSNOG1A* in a wild type background, there was no obvious inhibition of 3D growth, but as I have discovered with *PpNOG1*, overexpression of the transcript does not always

translate as overexpression of the protein. I could not determine whether the levels of PpSNOG1-GFP that I detected in wild type pUBI::SNOG1A-GFP, were different to levels of the native protein in wild type.

PpNOG1 and PpSNOG1A have similar structures with a UBA domain at the C-terminus and a disordered N-terminus. I have generated lines that express PpNOG1-GFP or PpSNOG1A-GFP. In both cases, visualisation of the fusion protein has proven to not be straightforward. A reliable signal for PpNOG1-GFP was rare and I did not reliably detect PpSNOG1A-GFP at all. These difficulties are potentially explained by my findings when I detected each fusion protein on a western blot, with an anti-GFP antibody. For both fusion proteins, the results were very similar. There was no protein detected that corresponded with the expected molecular weight of the full-length fusion protein. Instead, in both cases, there was a protein that was approximately consistent with the size of the UBA domain (around 7 kDa) and the GFP (27 kDa). This indicates that both PpNOG1 and PpSNOG1A are likely cleaved in a region roughly near the start of the UBA domain, producing at least two fragments, which I will refer to as the UBA-fragment and the N-fragment. The N-fragment may be further cleaved but I could not know whether this was occurring. The fact that no full length fusion protein could be detected, indicates that cleavage is occurring co-translationally. The other protein detected was around 20 kDa. As GFP is 27 kDa, this band must correspond to a fragment of GFP containing the epitope recognised by the anti-GFP antibody. Therefore GFP is being degraded, which is very unusual as GFP is known to be very resistant to proteolysis (Chiang et al., 2001). GFP is sometimes excised from a fusion protein by cleavage in the linker peptide, but given the size of the protein detected, this doesn't appear to be the case for PpNOG1-GFP or PpSNOG1A-GFP. This cleavage of the GFP likely prevents fluorescence as it would disrupt the structure of the β -barrel (Stepanenko et al., 2013). GFP has a short maturation time, meaning the assembly of the fluorophore into a conformation that can fluoresce occurs in a matter of minutes (Balleza et al., 2018). This means that even proteins with a short half-life that are subject to high turnover, can normally be visualised when attached

to GFP. For PpSNOG1A-GFP, most of the protein detected was this GFP fragment, with very little of the UBA-fragment present. This indicates that cleavage of the GFP occurs fairly soon after protein synthesis, possibly before the maturation of the GFP. This could explain why I was unable to visualise the PpSNOG1A-GFP fusion protein. For PpNOG1-GFP, the amount of UBA-fragment and the GFP fragment was approximately the same, but both peptides were subject to strong proteasomal degradation unless tissues were treated with the proteasome inhibitor bortezomib. The degradation of the PpNOG1-UBA fragment by the proteasome is likely what prevented visualisation. Unfortunately, I did not have time to conduct an experiment whereby I visualised the fusion protein after bortezomib treatment, other than a very preliminary attempt that was inconclusive. The PpSNOG1A UBA-fragment was less subject to proteasomal degradation as bortezomib treatment did not much affect levels of detected protein.

The UBA domain is the only functionally annotated domain in both of the proteins. For both PpNOG1 and PpSNOG1A I demonstrated that they could potentially interact with polyubiquitin, most likely via the ubiquitin-associated domain. For PpNOG1, I demonstrated that the UBA domain is essential to prescribe 3D growth as mutating the crucial ‘MGF’ loop prevents complementation of *Ppnog1R*. However, the UBA domain is not sufficient for 3D growth as expressing the UBA domain alone only partially complements the *Ppnog1R* phenotype by promoting the formation of small gametophores that can develop a few phyllids before arresting development. Therefore, further work will be needed to examine the properties of the N-fragment, as it is clearly of functional importance. The same is likely true for PpSNOG1A as work on the Arabidopsis FLOE1 protein demonstrated the importance of the various disorder domains for modulating the physical properties of the protein, with a link to its role in regulating germination based on environmental cues (drought stress). Previous work on FLOE1 did not show the protein extracted from plant tissues and subsequently detected on a western blot, so there are no clues as to whether this cleavage occurs in homologues of PpSNOG1A, although the dynamics of the C-terminally fused fluorescent FLOE

fusion proteins does suggest that UBA cleavage is not prevalent for these proteins (Dorone et al., 2021). To visualise and characterise the N-fragments of both PpNOG1 and PpSNOG1A, N-terminally linked fluorescent protein fusions should be constructed. We know that the UBA-fragment of PpNOG1 is highly degraded by the proteasome as it is almost undetectable in tissues not treated with the proteasome inhibitor bortezomib. Meanwhile, the PpSNOG1A UBA-fragment is also targeted to the proteasome but not to the same degree and the UBA-fragment is still easily detectable in extracts from tissues that have not been treated with bortezomib. The question arises as to whether the N-fragment of PpNOG1 is also degraded by the proteasome, or whether cleavage from the C-terminal UBA allows it to escape the same fate. This would be an unusual paradigm as a C-terminal UBA usually protects a protein from proteasomal degradation (Heinen et al., 2011). The other question that arises, is whether the cleavage of the UBA domain from PpNOG1 and PpSNOG1A is constitutive, or whether it is conditional. If it is constitutive, this would indicate that the UBA-fragments and N-fragments operate completely independently and should be considered two proteins coded by one gene. For my western blots, I use whole-plant extracts from 1 week old tissues. These tissues consisted entirely of protonema and the transition to 3D growth would not have begun. I wonder if the same results would occur if extracts from older protonema or from gametophore tissues were used. Perhaps the cleavage (and/or proteasomal degradation) of PpNOG1 and PpSNOG1A occurs in immature tissues to prevent premature developmental progression, but then is repressed in older tissues or in certain cell types. These experiments should be included in future work.

I have determined that in both cases, the UBA domain is possibly capable of interacting with polyubiquitin. However, whether this occurs *in planta* is yet to be determined. To establish this, pull-down experiments with the UBA fragments might be informative. It would also be useful to do this with the N-fragments to discover more about their currently completely unknown functions. Another method that has proven to be useful for identifying the interaction partners of low abundance regulators is proximity labelling, using the tool TurboID (Mair et al.,

2019). This involves the fusion of a promiscuous biotin ligase to a protein of interest, which biotinylates all proteins within a close proximity to the protein of interest, upon addition of biotin. Biotinylated proteins are purified from extracts by affinity purification and identified by mass spectrometry (Guo et al., 2023). Interestingly, using this method in *Arabidopsis*, AtNOG1 was identified as a potential interactor (or close neighbor) of FAMA, a key regulator of stomatal development (Mair et al., 2019). In another proximity labelling study, AtNOG1 was found to be associated with the kinase BRASSINOSTEROID INSENSITIVE-2 (BIN2), which has broad developmental and physiological functions (Kim et al., 2023). To my knowledge, proximity labelling has yet to be applied in *P. patens*, but it has demonstrated to be useful across animals, plants and microbes so application in *P. patens* is likely possible (Guo et al., 2023).

My findings related to PpNOG1 and PpSNOG1A have highlighted the role of protein degradation in regulating the transition from 2D to 3D growth. PpNOG1 and PpSNOG1A are both degraded by the proteasome (PpNOG1 more so than PpSNOG1A) and they themselves likely play a role in the ubiquitin-proteasome system (UPS). Other UPS related genes such as *PpATE* and *PpMCB1* have also been shown to play a role in the transition to 3D growth (Girod et al., 1999; Schuessele et al., 2016). There is a growing body of research that is linking protein turnover to developmental processes. The ubiquitin-proteasome system can specifically target key proteins to quickly remodel the proteome in response to endogenous or exogenous cues (DeRenzo and Seydoux, 2004; Wang and Deng, 2011; Sadanandom et al., 2012; Xu et al., 2017; Aguilar-Hernández et al., 2017; Wang et al., 2022). There is much to still be uncovered about the role that protein degradation plays in coordinating the 3D transition.

The *snog1a* mutant was responsive to BAP, whereas the *nog1dis* mutant was persistently unresponsive to BAP. This could indicate that PpSNOG1A normally represses cytokinin responses. When I examined the transcriptomes of *nog1dis* and *snog1a*, indeed more genes responded to BAP treatment in *snog1a* than in *nog1dis*, but not nearly as many as respond to BAP in wild type. Similarly, when examining

genes with a known role in the 3D transition, those that responded to BAP in wild type, did not respond to BAP in *nog1dis*. In a few cases there was a weak restoration of a BAP response in *snog1a*, but it was not particularly striking. Therefore I can only attribute a weak suppressive effect of PpSNOG1A on cytokinin signalling.

When I searched for homologues of PpNOG1 and PpSNOG1A, I found them not only in land plants but also in the green algae which do not possess 3D-capable apical cells. Indeed some of those algae that had homologues of these genes were single-celled. *DEK1*-like genes have also been found in green algae (Liang et al., 2013). On the other hand, *CLAVATA* signalling genes are known to be land plant specific (Whitewoods et al., 2018). Also GRAS-family transcription factors, such as *PpGRAS12* also arose in the land plants, as did the *TOPLESS* genes (Plant et al., 2021; Beheshti et al., 2021). The SBP genes are present in both land plants and green algae but the regulation of their expression by miR156 only emerged in land plants (Guo et al., 2008). Therefore it seems that land plant ‘3D regulators’ are a combination of novel genes that emerged with the land plants, and genes already present in algal genomes that may have been co-opted from different functions and integrated into new regulatory networks to bring about this pivotal evolutionary transition.

In Figure 6.1, I have tried to integrate my findings into the current regulatory network. The network has been simplified for ease of interpretation. Firstly, I showed that PpNOG1 (or at least the UBA-fragment), is heavily degraded by the proteasome. Whether this degradation is conditional or constitutive, remains to be seen. My preliminary exploration of the *Ppnog1R dek1 kd* mutants in which *PpDEK1* expression is suppressed but not eliminated, indicated a possible epistatic relationship between *PpNOG1* and *PpDEK1*, that could indicate that they function in the same pathway, with *PpNOG1* preceding *PpDEK1*. This is a very tentative result as more analysis is required. Secondly, based on the results of my first RNA-seq experiment, I selected *PpHEC1a* and *PpHEC1b* as genes of interest that I wanted to investigate. In the *Ppnog1R* background, overexpression of PpHEC1b resulted in an unstable phenotype that initially presented as severe

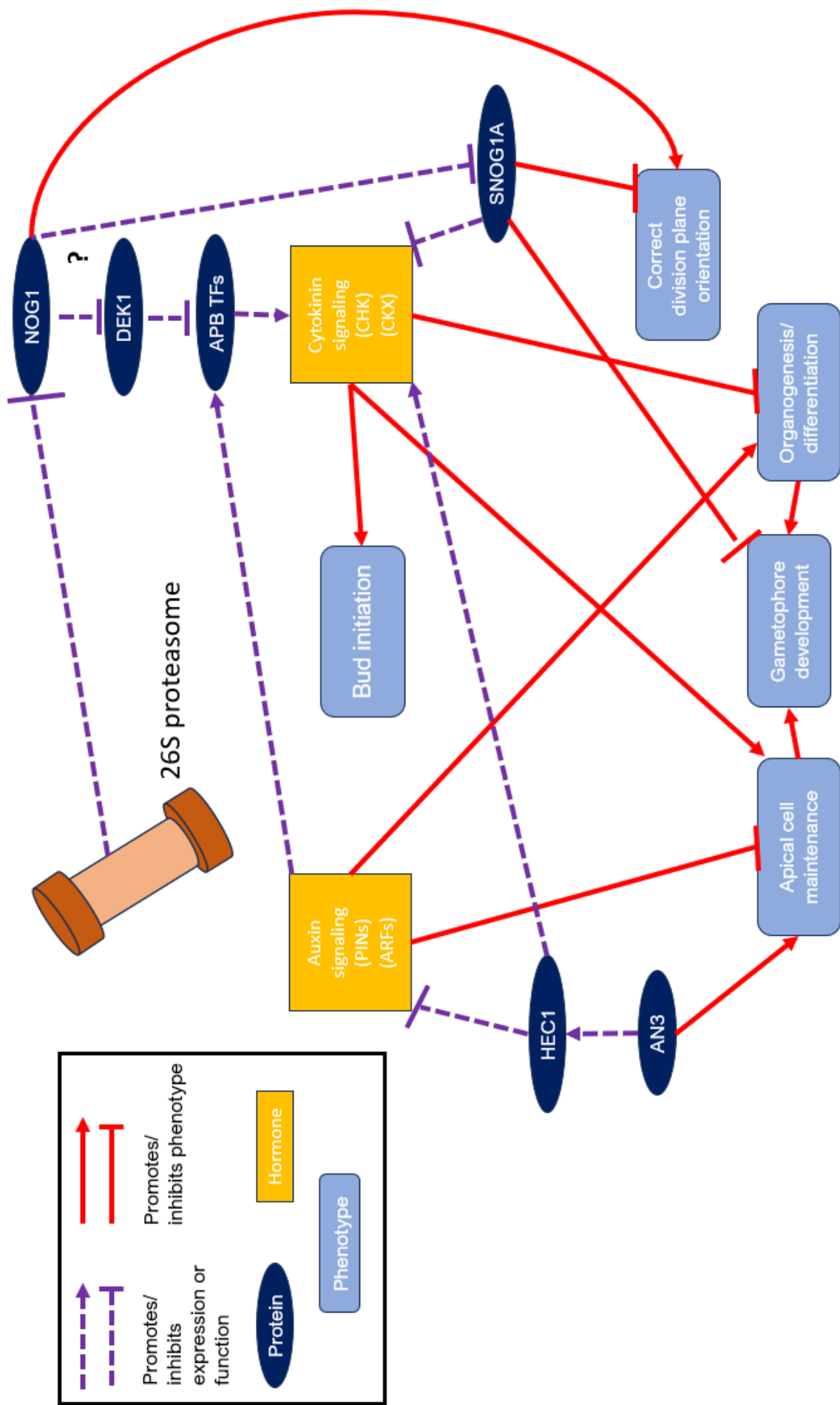


Figure 6.1: A schematic showing my findings and hypothesised interactions incorporated into a simplified version of the 2D-3D regulatory network.

inhibition of tip growth but that eventually produced a mild rescue of the *Ppnog1R* phenotype. In Arabidopsis, *HEC* genes are known to be important regulators of auxin and cytokinin signalling in the SAM, and are crucial for modulating the balance between the two to regulate SAM size versus lateral organ initiation. Arabidopsis HECs were shown to promote cytokinin signalling in the SAM, and interferes with auxin feedback mechanisms (Gaillochet et al., 2017). Interestingly, in Arabidopsis, *ANGUSTOLIA3* was shown to directly bind to the promoter region of *HEC1* to promote its expression (Vercruyssen et al., 2014). *PpAN3* is a known regulator of the 3D growth transition and it is possible that expression regulation of *PpHEC* by *PpAN3* is conserved, so I have included it in my regulatory network speculative model (Kawade et al., 2020).

My RNA-seq experiments revealed how cytokinin affects the transcription of many genes involved in the transition to 3D growth. In wild type, cytokinin induces the formation of gametophore initials, and the presence of the *PpAPB* transcription factors is required for this to occur (Aoyama et al., 2012). Application of high concentrations of exogenous cytokinin (BAP) results in the production of excessive numbers of gametophore initials (Brandes and Kende, 1968). In *nog1R* and *nog1dis*, this phenotype is not seen, upon treatment with BAP. Therefore, *PpNOG1* is required for a normal cytokinin response. When I rescued *Ppnog1R* with *PpNOG1* variants and treated those lines with BAP, I found that the UBA domain is not vital for the cytokinin response and that functionality lies somewhere else in the uncharacterised region of the *PpNOG1* protein. In my final RNA-seq experiment I compared wild type, *nog1dis* and *snog1a*, in the presence and absence of BAP. From these data I could see that genes that were transcriptionally regulated by cytokinin in wild type were often not responsive in *nog1dis*. This aligns with the cytokinin non-responsive phenotype of *nog1dis*. This could suggest that *PpNOG1* lies downstream of the cytokinin signalling that regulates the transition to 3D growth. However, I noticed that most of the genes involved in the 3D transition that are transcriptionally regulated by cytokinin are downstream of *PpNOG1*. The expression of *PpNOG1* and upstream genes such as *PpTPL* and *PpGRAS12* is not

regulated by cytokinin. *PpDEK1* does not have a role in cytokinin signalling as *dek1* mutants have a normal response to cytokinin treatment (Perroud et al., 2014). Also, cytokinin does not regulate the expression of the *PpAPB* genes (Aoyama et al., 2012). All this suggests that in terms of the transcriptional network, cytokinin signalling is occurring downstream of *PpNOG1*, although *PpNOG1* seems to be vital for cytokinin perception in this developmental context.

Finally, I have included PpSNOG1A as a hypothesised repressor of gametophore development, correct division planes and cytokinin response. I have hypothesised that PpNOG1 may repress PpSNOG1A function (although it does not repress transcript expression), based on the fact that overexpression of *PpSNOG1A* in the wild type background does not cause suppression of 3D growth.

In conclusion, my work has resulted in the identification of two novel regulators of the transition to 3D growth in *P. patens*; PpSNOG1A and PpHEC1b. This opens up many exciting avenues for further research. Also, I have further characterised the function of the key early regulator PpNOG1. I have also analysed three RNA-seq datasets which has given further insight into the broad functions and pathways that are perturbed in response to PpNOG1, PpNOG2, PpSNOG1A and the cytokinin analogue BAP.

Bibliography

- Mark D Adams, Susan E Celniker, Robert A Holt, Cheryl A Evans, Jeannine D Gocayne, Peter G Amanatides, Steven E Scherer, Peter W Li, Roger A Hoskins, Richard F Galle, et al. The genome sequence of *Drosophila melanogaster*. *Science*, 287(5461):2185–2195, 2000.
- Victor Aguilar-Hernández, Do-Young Kim, Robert J Stankey, Mark Scalf, Lloyd M Smith, and Richard D Vierstra. Mass spectrometric analyses reveal a central role for ubiquitylation in remodeling the Arabidopsis proteome during photomorphogenesis. *Molecular plant*, 10(6):846–865, 2017.
- Joon-Woo Ahn, Moonil Kim, Jeong Hwa Lim, Gyung-Tae Kim, and Hyun-Sook Pai. Phytocalpain controls the proliferation and differentiation fates of cells in plant organ development. *The Plant Journal*, 38(6):969–981, 2004.
- 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.
- The Amborella Genome Project, Victor A Albert, W Bradley Barbazuk, Claude W dePamphilis, Joshua P Der, James Leebens-Mack, Hong Ma, Jeffrey D Palmer, Steve Rounsley, David Sankoff, et al. The Amborella genome and the evolution of flowering plants. *Science*, 342(6165):1241089, 2013.
- Barbara A Ambrose and Alejandra Vasco. Bringing the multicellular fern meristem into focus. *New Phytologist*, 210(3):790–793, 2016.

- Tsuyoshi Aoyama, Yuji Hiwatashi, Mikao Shigyo, Rumiko Kofuji, Minoru Kubo, Motomi Ito, and Mitsuyasu Hasebe. AP2-type transcription factors determine stem cell identity in the moss *Physcomitrella patens*. *Development*, 139(17): 3120–3129, 2012.
- NW Ashton, NH Grimsley, and DJ Cove. Analysis of gametophytic development in the moss, *Physcomitrella patens*, using auxin and cytokinin resistant mutants. *Planta*, 144(5):427–435, 1979.
- Patricia Ballester, Maria A Martínez-Godoy, Miguel Ezquerro, Marisa Navarrete-Gómez, Marina Trigueros, Manuel Rodríguez-Concepción, and Cristina Ferrándiz. A transcriptional complex of NGATHA and bHLH transcription factors directs stigma development in Arabidopsis. *The Plant Cell*, 33(12):3645–3657, 2021.
- Enrique Balleza, J Mark Kim, and Philippe Cluzel. Systematic characterization of maturation time of fluorescent proteins in living cells. *Nature methods*, 15(1): 47–51, 2018.
- Jo Ann Banks, Tomoaki Nishiyama, Mitsuyasu Hasebe, John L Bowman, Michael Gribskov, Claude DePamphilis, Victor A Albert, Naoki Aono, Tsuyoshi Aoyama, Barbara A Ambrose, et al. The Selaginella genome identifies genetic changes associated with the evolution of vascular plants. *science*, 332(6032):960–963, 2011.
- Hossein Beheshti, Christoph Strotbek, M Asif Arif, Andreas Klingl, Oguz Top, and Wolfgang Frank. *PpGRAS12* acts as a positive regulator of meristem formation in *Physcomitrium patens*. *Plant Molecular Biology*, 107:293–305, 2021.
- Tom A Bennett, Maureen M Liu, Tsuyoshi Aoyama, Nicole M Bierfreund, Marion Braun, Yoan Coudert, Ross J Dennis, Devin O’Connor, Xiao Y Wang, Chris D White, et al. Plasma membrane-targeted PIN proteins drive shoot development in a moss. *Current Biology*, 24(23):2776–2785, 2014.
- Bonnie L Bertolaet, Duncan J Clarke, Meira Wolff, Mark H Watson, Martha Henze,

- Gilles Divita, and Steven I Reed. UBA domains of DNA damage-inducible proteins interact with ubiquitin. *Nature structural biology*, 8(5):417–422, 2001.
- Guillaume Blanc, Irina Agarkova, Jane Grimwood, Alan Kuo, Andrew Brueggeman, David D Dunigan, James Gurnon, Istvan Ladunga, Erika Lindquist, Susan Lucas, et al. The genome of the polar eukaryotic microalga *Coccomyxa subellipsoidea* reveals traits of cold adaptation. *Genome biology*, 13:1–12, 2012.
- Anthony M Bolger, Marc Lohse, and Bjoern Usadel. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15):2114–2120, 2014.
- Martha G Bomar, Sanjay D’Souza, Marzena Bienko, Ivan Dikic, Graham C Walker, and Pei Zhou. Unconventional ubiquitin recognition by the ubiquitin-binding motif within the Y family DNA polymerases ι and Rev1. *Molecular cell*, 37(3):408–417, 2010.
- Thomas Bonacci, Stéphane Audebert, Luc Camoin, Emilie Baudet, Juan-Lucio Iovanna, and Philippe Soubeyran. Regulation of NUB1 activity through non-proteolytic Mdm2-mediated ubiquitination. *PloS one*, 12(1):e0169988, 2017.
- Kim Boutilier, Remko Offringa, Vijay K Sharma, Henk Kieft, Thérèse Ouellet, Lemin Zhang, Jiro Hattori, Chun-Ming Liu, André AM van Lammeren, Brian LA Miki, et al. Ectopic expression of BABY BOOM triggers a conversion from vegetative to embryonic growth. *The Plant Cell*, 14(8):1737–1749, 2002.
- John L Bowman and Yuval Eshed. Formation and maintenance of the shoot apical meristem. *Trends in plant science*, 5(3):110–115, 2000.
- 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.
- Ulrike Brand, Jennifer C Fletcher, Martin Hobe, Elliot M Meyerowitz, and Rudiger

- Simon. Dependence of stem cell fate in *Arabidopsis* on a feedback loop regulated by *CLV3* activity. *Science*, 289(5479):617–619, 2000.
- H Brandes and H Kende. Studies on cytokinin-controlled bud formation in moss protonemata. *Plant Physiology*, 43(5):827–837, 1968.
- Susan H Brawley, Nicolas A Blouin, Elizabeth Ficko-Blean, Glen L Wheeler, Martin Lohr, Holly V Goodson, Jerry W Jenkins, Crysten E Blaby-Haas, Katherine E Helliwell, Cheong Xin Chan, et al. Insights into the red algae and eukaryotic evolution from the genome of *Porphyra umbilicalis* (Bangiophyceae, Rhodophyta). *Proceedings of the National Academy of Sciences*, 114(31):E6361–E6370, 2017.
- Lucie Brejšková, Michal Hála, Anamika Rawat, Hana Soukupová, Fatima Cvrčková, Florence Charlot, Fabien Nogué, Samuel Haluška, and Viktor Žárský. SEC6 exocyst subunit contributes to multiple steps of growth and development of *Physcomitrella* (*Physcomitrium patens*). *The Plant Journal*, 106(3):831–843, 2021.
- Dana E Brown, Aaron M Rashotte, Angus S Murphy, Jennifer Normanly, Brian W Tague, Wendy A Peer, Lincoln Taiz, and Gloria K Muday. Flavonoids act as negative regulators of auxin transport in vivo in *Arabidopsis*. *Plant physiology*, 126(2):524–535, 2001.
- Daniel R Browne, Jerry Jenkins, Jeremy Schmutz, Shengqiang Shu, Kerrie Barry, Jane Grimwood, Jennifer Chiniquy, Aditi Sharma, Thomas D Niehaus, Taylor L Weiss, et al. Draft nuclear genome sequence of the liquid hydrocarbon-accumulating green microalga *Botryococcus braunii* race B (Showa). *Genome announcements*, 5(16):10–1128, 2017.
- Benjamin Buchfink, Chao Xie, and Daniel H Huson. Fast and sensitive protein alignment using DIAMOND. *Nature methods*, 12(1):59–60, 2015.
- Judy Callis, Tami Carpenter, Chih-Wen Sun, and Richard D Vierstra. Structure and evolution of genes encoding polyubiquitin and ubiquitin-like proteins in *Arabidopsis thaliana* ecotype Columbia. *Genetics*, 139(2):921–939, 1995.

- Joseph Cammarata, Adrienne HK Roeder, and Michael J Scanlon. Cytokinin and CLE signaling are highly intertwined developmental regulators across tissues and species. *Current opinion in plant biology*, 51:96–104, 2019.
- Joseph Cammarata, Christopher Morales Farfan, Michael J Scanlon, and Adrienne HK Roeder. Cytokinin-CLAVATA cross-talk is an ancient mechanism regulating shoot meristem homeostasis in land plants. *Proceedings of the National Academy of Sciences*, 119(14):e2116860119, 2022.
- Joseph Cammarata, Adrienne HK Roeder, and Michael J Scanlon. The ratio of auxin to cytokinin controls leaf development and meristem initiation in *Physcomitrium patens*. *Journal of Experimental Botany*, 74(21):6541–6550, 2023.
- Florencia Cano, Radu Rapiteanu, G Sebastiaan Winkler, and Paul J Lehner. A non-proteolytic role for ubiquitin in deadenylation of *MHC-I* mRNA by the RNA-binding E3-ligase MEX-3C. *Nature communications*, 6(1):8670, 2015.
- Guillermo Cardon, Susanne Höhmann, Joachim Klein, Klaus Nettesheim, Heinz Saedler, and Peter Huijser. Molecular characterisation of the Arabidopsis SBP-box genes. *Gene*, 237(1):91–104, 1999.
- Sarah B Carey, Jerry Jenkins, John T Lovell, Florian Maumus, Avinash Sreedasyam, Adam C Payton, Shengqiang Shu, George P Tiley, Noe Fernandez-Pozo, Adam Healey, et al. Gene-rich UV sex chromosomes harbor conserved regulators of sexual development. *Science Advances*, 7(27):eabh2488, 2021.
- Barry Causier, James Lloyd, Laura Stevens, and Brendan Davies. TOPLESS co-repressor interactions and their evolutionary conservation in plants. *Plant signaling & behavior*, 7(3):325–328, 2012.
- Barry Causier, Mary McKay, Tayah Hopes, James Lloyd, Dapeng Wang, C Jill Harrison, and Brendan Davies. The TOPLESS corepressor regulates developmental switches in the bryophyte *Physcomitrium patens* that were critical for plant terrestrialisation. *The Plant Journal*, 115(5):1331–1344, 2023.

- Kong Chen, Guo-Jun Li, Ray A Bressan, Chun-Peng Song, Jian-Kang Zhu, and Yang Zhao. Absciscic acid dynamics, signaling, and functions in plants. *Journal of integrative plant biology*, 62(1):25–54, 2020.
- Shifu Chen, Yanqing Zhou, Yaru Chen, and Jia Gu. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*, 34(17):i884–i890, 2018.
- Zhijian J Chen and Lijun J Sun. Nonproteolytic functions of ubiquitin in cell signaling. *Molecular cell*, 33(3):275–286, 2009.
- Chia-Yi Cheng, Vivek Krishnakumar, Agnes P Chan, Françoise Thibaud-Nissen, Seth Schobel, and Christopher D Town. Araport11: a complete reannotation of the *Arabidopsis thaliana* reference genome. *The Plant Journal*, 89(4):789–804, 2017.
- Shifeng Cheng, Wenfei Xian, Yuan Fu, Birger Marin, Jean Keller, Tian Wu, Wenjing Sun, Xiuli Li, Yan Xu, YU Zhang, et al. Genomes of subaerial Zygnematophyceae provide insights into land plant evolution. *Cell*, 179(5):1057–1067, 2019.
- Cheng-Feng Chiang, David T Okou, Tony B Griffin, C Reynold Verret, and Myron NV Williams. Green fluorescent protein rendered susceptible to proteolysis: positions for protease-sensitive insertions. *Archives of biochemistry and biophysics*, 394(2):229–235, 2001.
- Sung Hyun Cho, Ceyda Coruh, and Michael J Axtell. miR156 and miR390 regulate tasiRNA accumulation and developmental timing in *Physcomitrella patens*. *The plant cell*, 24(12):4837–4849, 2012.
- Michael L Christianson. ABA prevents the second cytokinin-mediated event during the induction of shoot buds in the moss *Funaria hygrometrica*. *American Journal of Botany*, 87(10):1540–1545, 2000.
- Steven E Clark, Steven E Jacobsen, Joshua Z Levin, and Elliot M Meyerowitz. The *CLAVATA* and *SHOOT MERISTEMLESS* loci competitively regulate meristem activity in *Arabidopsis*. *Development*, 122(5):1567–1575, 1996.

- David Cove. The moss *Physcomitrella patens*. *Annual Review of Genetics*, 39: 339–358, 2005.
- David J Cove and Celia D Knight. The moss *physcomitrella patens*, a model system with potential for the study of plant reproduction. *The Plant Cell*, 5(10):1483, 1993.
- David J Cove, Pierre-François Perroud, Audra J Charron, Stuart F McDaniel, Abha Khandelwal, and Ralph S Quatrano. Culturing the moss *Physcomitrella patens*. *Cold Spring Harbor Protocols*, 4(2), 2009a.
- David J Cove, Pierre-François Perroud, Audra J Charron, Stuart F McDaniel, Abha Khandelwal, and Ralph S Quatrano. Transformation of the moss *Physcomitrella patens* using direct DNA uptake by protoplasts. *Cold Spring Harbor Protocols*, 4(2), 2009b.
- Charles Francis Delwiche and Endymion Dante Cooper. The evolutionary origin of a terrestrial flora. *Current Biology*, 25(19):R899–R910, 2015.
- Victor Demko, Tatiana Belova, Maxim Messerer, Torgeir R. Hvidsten, Pierre-François Perroud, Ako Eugene Ako, Wenche Johansen, Klaus F. X. Mayer, Odd-Arne Olsen, and Daniel Lang. Regulation of developmental gatekeeping and cell fate transition by the calpain protease DEK1 in *Physcomitrium patens*. *Communications Biology*, 7(261), 2024.
- Viktor Demko, Pierre-François Perroud, Wenche Johansen, Charles F Delwiche, Endymion D Cooper, Pål Remme, Ako Eugene Ako, Karl G Kugler, Klaus FX Mayer, Ralph Quatrano, et al. Genetic analysis of *DEFECTIVE KERNEL 1* loop function in three-dimensional body patterning in *Physcomitrella patens*. *Plant physiology*, 166(2):903–919, 2014.
- Cynthia DeRenzo and Geraldine Seydoux. A clean start: degradation of maternal proteins at the oocyte-to-embryo transition. *Trends in cell biology*, 14(8):420–426, 2004.

- Yves Deveaux, Claire Toffano-Nioche, Gaelle Claisse, Vincent Thareau, Halima Morin, Patrick Laufs, Hervé Moreau, Martin Kreis, and Alain Lecharny. Genes of the most conserved *wox* clade in plants affect root and flower development in arabidopsis. *BMC Evolutionary Biology*, 8:1–19, 2008.
- Paul R Dohrmann, Warren P Voth, and David J Stillman. Role of negative regulation in promoter specificity of the homologous transcriptional activators Ace2p and Swi5p. *Molecular and cellular biology*, 1996.
- Yanniv Dorone, Steven Boeynaems, Eduardo Flores, Benjamin Jin, Shannon Hateley, Flavia Bossi, Elena Lazarus, Janice G Pennington, Emiel Michiels, Mathias De Decker, et al. A prion-like protein regulator of seed germination undergoes hydration-dependent phase separation. *Cell*, 184(16):4284–4298, 2021.
- Shannon E Dougherty, Austin O Maduka, Toshifumi Inada, and Gustavo M Silva. Expanding role of ubiquitin in translational control. *International journal of molecular sciences*, 21(3):1151, 2020.
- Jim Dover, Jessica Schneider, Mary Anne Tawiah-Boateng, Adam Wood, Kimberly Dean, Mark Johnston, and Ali Shilatifard. Methylation of histone H3 by COMPASS requires ubiquitination of histone H2B by Rad6. *Journal of Biological Chemistry*, 277(32):28368–28371, 2002.
- Mei Duan, Rongxue Zhang, Fugui Zhu, Zhenqian Zhang, Lanming Gou, Jiangqi Wen, Jiangli Dong, and Tao Wang. A lipid-anchored NAC transcription factor is translocated into the nucleus and activates glyoxalase I expression during drought stress. *The Plant Cell*, 29(7):1748–1772, 2017.
- Steffen Durinck, Paul T Spellman, Ewan Birney, and Wolfgang Huber. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nature protocols*, 4(8):1184–1191, 2009.
- Angélique D’hont, France Denoeud, Jean-Marc Aury, Franc-Christophe Baurens, Françoise Carreel, Olivier Garsmeur, Benjamin Noel, Stéphanie Bocs, Gaëtan

- Droc, Mathieu Rouard, et al. The banana (*Musa acuminata*) genome and the evolution of monocotyledonous plants. *Nature*, 488(7410):213–217, 2012.
- Suzanne Elsasser, Devin Chandler-Militello, Britta Muller, John Hanna, and Daniel Finley. Rad23 and Rpn10 serve as alternative ubiquitin receptors for the proteasome. *Journal of Biological Chemistry*, 279(26):26817–26822, 2004.
- David M Emms and Steven Kelly. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome biology*, 20(1):1–14, 2019.
- Eric M Engstrom, Carl M Andersen, Juliann Gumulak-Smith, John Hu, Evguenia Orlova, Rosangela Sozzani, and John L Bowman. Arabidopsis homologs of the *Petunia* HAIRY MERISTEM gene are required for maintenance of shoot and root indeterminacy. *Plant physiology*, 155(2):735–750, 2011.
- Emese Eysholdt-Derzsó, Tilo Renziehausen, Stephanie Frings, Stephanie Frohn, Kira von Bongartz, Clara P Igisch, Justina Mann, Lisa Häger, Julia Macholl, David Leisse, et al. Endoplasmic reticulum-bound ANAC013 factor is cleaved by RHOMBOID-LIKE 2 during the initial response to hypoxia in *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences*, 120(11):e2221308120, 2023.
- Lisa M Farmer, Adam J Book, Kwang-Hee Lee, Ya-Ling Lin, Hongyong Fu, and Richard D Vierstra. The RAD23 family provides an essential connection between the 26S proteasome and ubiquitylated proteins in *Arabidopsis*. *The Plant Cell*, 22(1):124–142, 2010.
- Daniel Finley, Keiji Tanaka, Carl Mann, Horst Feldmann, Mark Hochstrasser, Richard Vierstra, Stephan Johnston, Randolph Hampton, James Haber, John McCusker, et al. Unified nomenclature for subunits of the *Saccharomyces cerevisiae* proteasome regulatory particle. *Trends in biochemical sciences*, 23(7):244–245, 1998.
- Jennifer C Fletcher, Ulrike Brand, Mark P Running, Rudiger Simon, and Elliot M

- Meyerowitz. Signaling of cell fate decisions by *CLAVATA3* in *Arabidopsis* shoot meristems. *Science*, 283(5409):1911–1914, 1999.
- 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.
- Margaret H Frank and Michael J Scanlon. Cell-specific transcriptomic analyses of three-dimensional shoot development in the moss *Physcomitrella patens*. *The Plant Journal*, 83(4):743–751, 2015.
- William E Friedman. One genome, two ontogenies. *Science*, 339(6123):1045–1046, 2013.
- Michael Fuchs and Jan U Lohmann. Aiming for the top: non-cell autonomous control of shoot stem cells in *Arabidopsis*. *Journal of plant research*, 133(3):297–309, 2020.
- Christophe Gaillochet, Thomas Stiehl, Christian Wenzl, Juan-José Ripoll, Lindsay J Bailey-Steinitz, Lanxin Li, Anne Pfeiffer, Andrej Miotk, Jana P Hakenjos, Joachim Forner, et al. Control of plant cell fate transitions by transcriptional and hormonal signals. *Elife*, 6:e30135, 2017.
- Christophe Gaillochet, Suraj Jamge, Froukje van der Wal, Gerco Angenent, Richard Immink, and Jan U Lohmann. A molecular network for functional versatility of HECATE transcription factors. *The Plant Journal*, 95(1):57–70, 2018.
- David Garcias-Morales, V Miguel Palomar, Florence Charlot, Fabien Nogué, Alejandra A Covarrubias, and José L Reyes. N 6-Methyladenosine modification of mRNA contributes to the transition from 2D to 3D growth in the moss *Physcomitrium patens*. *The Plant Journal*, 114(1):7–22, 2023.
- Pierre-Alain Girod, Hongyong Fu, Jean-Pierre Zryd, and Richard D Vierstra. Multiubiquitin chain binding subunit MCB1 (RPN10) of the 26S proteasome is

- essential for developmental progression in *Physcomitrella patens*. *The Plant Cell*, 11(8):1457–1471, 1999.
- James Godwin and Sara Farrona. The importance of networking: Plant polycomb repressive complex 2 and its interactors. *Epigenomes*, 6(1):8, 2022.
- Alfred L Goldberg. Protein degradation and protection against misfolded or damaged proteins. *Nature*, 426(6968):895–899, 2003.
- Emmanuelle Graciet, Franziska Walter, Diarmuid Ó Maoiléidigh, Stephan Pollmann, Elliot M Meyerowitz, Alexander Varshavsky, and Frank Wellmer. The N-end rule pathway controls multiple functions during *Arabidopsis* shoot and leaf development. *Proceedings of the National Academy of Sciences*, 106(32):13618–13623, 2009.
- Linda E Graham, James M Graham, Martha E Cook, and Lee Warren Wilcox. *Algae: Third Edition*. LJLM Press, 2015.
- Thomas Greb, Oliver Clarenz, Elisabeth Schäfer, Dörte Müller, Rubén Herrero, Gregor Schmitz, and Klaus Theres. Molecular analysis of the LATERAL SUPPRESSOR gene in *Arabidopsis* reveals a conserved control mechanism for axillary meristem formation. *Genes & development*, 17(9):1175–1187, 2003.
- Kristina Gremski, Gary Ditta, and Martin F Yanofsky. The *gla* genes regulate female reproductive tract development in *Arabidopsis thaliana*. 2007.
- Ambre Guillory and Sandrine Bonhomme. Phytohormone biosynthesis and signaling pathways of mosses. *Plant Molecular Biology*, pages 1–33, 2021.
- An-Yuan Guo, Qi-Hui Zhu, Xiaocheng Gu, Song Ge, Ji Yang, and Jingchu Luo. Genome-wide identification and evolutionary analysis of the plant specific sbp-box transcription factor family. *Gene*, 418(1-2):1–8, 2008.
- Jieyu Guo, Shuang Guo, Siao Lu, Jun Gong, Long Wang, Liqiong Ding, Qingjie Chen, and Wu Liu. The development of proximity labeling technology and its

- applications in mammals, plants, and microorganisms. *Cell Communication and Signaling*, 21(1):269, 2023.
- C Jill Harrison, Mohi Rezvani, and Jane A Langdale. Growth from two transient apical initials in the meristem of *Selaginella kraussiana*. 2007.
- C Jill Harrison, Adrienne HK Roeder, Elliot M Meyerowitz, and Jane A Langdale. Local cues and asymmetric cell divisions underpin body plan transitions in the moss *Physcomitrella patens*. *Current Biology*, 19(6):461–471, 2009.
- Jill Harrison. Development and genetics in the evolution of land plant body plans. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372(1713): 20150490, 2017.
- Yuki Hata and Junko Kyoizuka. Fundamental mechanisms of the stem cell regulation in land plants: lesson from shoot apical cells in bryophytes. *Plant Molecular Biology*, pages 1–13, 2021.
- Yuki Hata, Juri Ohtsuka, Yuji Hiwatashi, Satoshi Naramoto, and Junko Kyoizuka. Cytokinins specify pluripotent stem cell identity in the moss *Physcomitrium patens*. *bioRxiv*, pages 2023–12, 2023.
- Trevor Hawkins, Elbert Branscomb, Paul Predki, Paul Richardson, Sarah Wenning, Tom Slezak, Norman Doggett, Jan-Fang Cheng, Anne Olsen, Susan Lucas, Christopher Elkin, Edward Uberbacher, Marvin Frazier, Yoshiyuki Sakaki, Asao Fujiyama, Masahira Hattori, Tetsushi Yada, Atsushi Toyoda, Takehiko Itoh, Chiharu Kawagoe, Hidemi Watanabe, Yasushi Totoki, Todd Taylor, Jean Weissenbach, Roland Heilig, William Saurin, Francois Artiguenave, Philippe Brottier, Thomas Bruls, Eric Pelletier, Catherine Robert, Patrick Wincker, André Rosenthal, Matthias Platzer, Gerald Nyakatura, Stefan Taudien, Andreas Rump, Douglas R. Smith, Lynn Doucette-Stamm, Marc Rubenfield, Keith Weinstock, Hong Mei Lee, JoAnn Dubois, Huanming Yang, Jun Yu, Jian Wang, Guyang Huang, Jun Gu, et al. Initial sequencing and analysis of the human genome. *nature*, 409:860–921, 2001.

- Adam L Healey, Bryan Piatkowski, John T Lovell, Avinash Sreedasyam, Sarah B Carey, Sujana Mamidi, Shengqiang Shu, Chris Plott, Jerry Jenkins, Travis Lawrence, et al. Newly identified sex chromosomes in the *Sphagnum* (peat moss) genome alter carbon sequestration and ecosystem dynamics. *Nature Plants*, 9(2): 238–254, 2023.
- Stijn Heessen, Maria G Masucci, and Nico P Dantuma. The UBA2 domain functions as an intrinsic stabilization signal that protects Rad23 from proteasomal degradation. *Molecular cell*, 18(2):225–235, 2005.
- Margaret R Heider and Mary Munson. Exorcising the exocyst complex. *Traffic*, 13(7):898–907, 2012.
- Christian Heinen, Klára Ács, Deborah Hoogstraten, and Nico P Dantuma. C-terminal UBA domains protect ubiquitin receptors by preventing initiation of protein degradation. *Nature communications*, 2(1):191, 2011.
- Linda Hicke, Heidi L Schubert, and Christopher P Hill. Ubiquitin-binding domains. *Nature reviews Molecular cell biology*, 6(8):610–621, 2005.
- Ko Hirano, Masatoshi Nakajima, Kenji Asano, Tomoaki Nishiyama, Hitoshi Sakakibara, Mikiko Kojima, Etsuko Katoh, Hongyu Xiang, Takako Tanahashi, Mitsuyasu Hasebe, et al. The GID1-mediated gibberellin perception mechanism is conserved in the lycophyte *Selaginella moellendorffii* but not in the bryophyte *Physcomitrella patens*. *The Plant Cell*, 19(10):3058–3079, 2007.
- Satoshi Hirano, Masato Kawasaki, Hideaki Ura, Ryuichi Kato, Camilla Raiborg, Harald Stenmark, and Soichi Wakatsuki. Double-sided ubiquitin binding of Hrs-UIM in endosomal protein sorting. *Nature structural & molecular biology*, 13(3):272–277, 2006.
- Candice N Hirsch, Cory D Hirsch, Alex B Brohammer, Megan J Bowman, Ilya Soifer, Omer Barad, Doron Shem-Tov, Kobi Baruch, Fei Lu, Alvaro G Hernandez, et al. Draft assembly of elite inbred line PH207 provides insights into genomic and transcriptome diversity in maize. *The Plant Cell*, 28(11):2700–2714, 2016.

- Diep Thi Hoang, Olga Chernomor, Arndt Von Haeseler, Bui Quang Minh, and Le Sy Vinh. UFBoot2: improving the ultrafast bootstrap approximation. *Molecular biology and evolution*, 35(2):518–522, 2018.
- Sebastian NW Hoernstein, Stefanie J Mueller, Kathrin Fiedler, Marc Schuelke, Jens T Vanselow, Christian Schuessele, Daniel Lang, Roland Nitschke, Gabor L Igloi, Andreas Schlosser, et al. Identification of targets and interaction partners of arginyl-tRNA protein transferase in the moss *Physcomitrella patens*. *Molecular & Cellular Proteomics*, 15(6):1808–1822, 2016.
- Sebastian NW Hoernstein, Buğra Özdemir, Nico van Gessel, Alessandra A Miniera, Bruno Rogalla von Bieberstein, Lars Nilges, Joana Schweikert Farinha, Ramona Komoll, Stella Glauz, Tim Weckerle, et al. A deeply conserved protease, acylamino acid-releasing enzyme (AARE), acts in ageing in *Physcomitrella* and *Arabidopsis*. *Communications Biology*, 6(1):61, 2023.
- Laurent Hoffmann, Sébastien Besseau, Pierrette Geoffroy, Christophe Ritzenthaler, Denise Meyer, Catherine Lapierre, Brigitte Pollet, and Michel Legrand. Silencing of hydroxycinnamoyl-coenzyme a shikimate/quinate hydroxycinnamoyltransferase affects phenylpropanoid biosynthesis. *The Plant Cell*, 16(6):1446–1465, 2004.
- Kay Hofmann and Philipp Bucher. The UBA domain: a sequence motif present in multiple enzyme classes of the ubiquitination pathway. *Trends in biochemical sciences*, 21(5):172–173, 1996.
- Tara J Holman, Peter D Jones, Laurel Russell, Anne Medhurst, Susana Úbeda Tomás, Prabhavathi Talloji, Julietta Marquez, Heike Schmuths, Swee-Ang Tung, Ian Taylor, et al. The N-end rule pathway promotes seed germination and establishment through removal of ABA sensitivity in *Arabidopsis*. *Proceedings of the National Academy of Sciences*, 106(11):4549–4554, 2009.
- Koichi Hori, Fumito Maruyama, Takatomo Fujisawa, Tomoaki Togashi, Nozomi Yamamoto, Mitsunori Seo, Syusei Sato, Takuji Yamada, Hiroshi Mori, Naoyuki

- Tajima, et al. *Klebsormidium flaccidum* genome reveals primary factors for plant terrestrial adaptation. *Nature communications*, 5(1):3978, 2014.
- Yuxin Hu, Qi Xie, and Nam-Hai Chua. The Arabidopsis auxin-inducible gene ARGOS controls lateral organ size. *The Plant Cell*, 15(9):1951–1961, 2003.
- James H Hurley, Sangho Lee, and Gali Prag. Ubiquitin-binding domains. *Biochemical Journal*, 399(3):361–372, 2006.
- Koraljka Husnjak and Ivan Dikic. Ubiquitin-binding proteins: decoders of ubiquitin-mediated cellular functions. *Annual review of biochemistry*, 81:291–322, 2012.
- Sujin Hyoung, Sung Hyun Cho, Joo Hee Chung, Won Mi So, Mei Hua Cui, and Jeong Sheop Shin. Cytokinin oxidase PpCKX1 plays regulatory roles in development and enhances dehydration and salt tolerance in *Physcomitrella patens*. *Plant cell reports*, 39:419–430, 2020.
- The International Brachypodium Initiative. Genome sequencing and analysis of the model grass *Brachypodium distachyon*. *Nature*, 463(7282):763–768, 2010.
- Marta Isasa, Elijah J Katz, Woong Kim, Verónica Yugo, Sheyla González, Donald S Kirkpatrick, Timothy M Thomson, Daniel Finley, Steven P Gygi, and Bernat Crosas. Monoubiquitination of RPN10 regulates substrate recruitment to the proteasome. *Molecular cell*, 38(5):733–745, 2010.
- Kai Ishida, Takafumi Yamashino, Akihiro Yokoyama, and Takeshi Mizuno. Three type-B response regulators, ARR1, ARR10 and ARR12, play essential but redundant roles in cytokinin signal transduction throughout the life cycle of *Arabidopsis thaliana*. *Plant and Cell Physiology*, 49(1):47–57, 2008.
- Yuji Iwata, Makoto Ashida, Chisa Hasegawa, Kazuki Tabara, Kei-ichiro Mishiba, and Nozomu Koizumi. Activation of the Arabidopsis membrane-bound transcription factor bZIP 28 is mediated by site-2 protease, but not site-1 protease. *The Plant Journal*, 91(3):408–415, 2017.

- Olivier Jaillon, Jean Marc Aury, Benjamin Noel, Alberto Policriti, Christian Clepet, Alberto Casagrande, Nathalie Choisne, Sébastien Aubourg, Nicola Vitulo, Claire Jubin, et al. The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. *Nature*, 449(7161):463–467, 2007.
- Chenguang Jiang, Clark K Schommer, Sun Young Kim, and Dae-Yeon Suh. Cloning and characterization of chalcone synthase from the moss, *Physcomitrella patens*. *Phytochemistry*, 67(23):2531–2540, 2006.
- Gerd Jürgens. Apical–basal pattern formation in Arabidopsis embryogenesis. *The EMBO journal*, 20(14):3609–3616, 2001.
- Subha Kalyaanamoorthy, Bui Quang Minh, Thomas KF Wong, Arndt Von Haeseler, and Lars S Jermiin. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature methods*, 14(6):587–589, 2017.
- Maria Kalyna, Craig G Simpson, Naeem H Syed, Dominika Lewandowska, Yamile Marquez, Branislav Kusenda, Jacqueline Marshall, John Fuller, Linda Cardle, Jim McNicol, et al. Alternative splicing and nonsense-mediated decay modulate expression of important regulatory genes in Arabidopsis. *Nucleic acids research*, 40(6):2454–2469, 2012.
- Yasuko Kamisugi, Katja Schlink, Stefan A Rensing, Gabriele Schween, Mark von Stackelberg, Andrew C Cuming, Ralf Reski, and David J Cove. The mechanism of gene targeting in *Physcomitrella patens*: homologous recombination, concatenation and multiple integration. *Nucleic acids research*, 34(21):6205–6214, 2006.
- Mamoru Kato, Naoya Hata, Nilanjana Banerjee, Bruce Futcher, and Michael Q Zhang. Identifying combinatorial regulation of transcription factors and binding motifs. *Genome biology*, 5:1–13, 2004.
- Kazutaka Katoh and Daron M Standley. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular biology and evolution*, 30(4):772–780, 2013.

- David J Katzmann, Greg Odorizzi, and Scott D Emr. Receptor downregulation and multivesicular-body sorting. *Nature reviews Molecular cell biology*, 3(12):893–905, 2002.
- Kensuke Kawade, Gorou Horiguchi, Yuu Hirose, Akira Oikawa, Masami Yokota Hirai, Kazuki Saito, Tomomichi Fujita, and Hirokazu Tsukaya. Metabolic control of gametophore shoot formation through arginine in the moss *Physcomitrium patens*. *Cell Reports*, 32(10), 2020.
- Paul Kenrick. Alternation of generations in land plants: new phylogenetic and palaeobotanical evidence. *Biological Reviews*, 69(3):293–330, 1994.
- Paul Kenrick and Peter R Crane. The origin and early evolution of plants on land. *Nature*, 389(6646):33–39, 1997.
- Erika A Keshishian and Aaron M Rashotte. Plant cytokinin signalling. *Essays in biochemistry*, 58:13–27, 2015.
- Do-Young Kim, Mark Scalf, Lloyd M Smith, and Richard D Vierstra. Advanced proteomic analyses yield a deep catalog of ubiquitylation targets in Arabidopsis. *The Plant Cell*, 25(5):1523–1540, 2013.
- Jeong Hoe Kim and Hirokazu Tsukaya. Regulation of plant growth and development by the GROWTH-REGULATING FACTOR and GRF-INTERACTING FACTOR duo. *Journal of experimental botany*, 66(20):6093–6107, 2015.
- Sang-Gyu Kim, Sangmin Lee, Pil Joon Seo, Soon-Kap Kim, Jeong-Kook Kim, and Chung-Mo Park. Genome-scale screening and molecular characterization of membrane-bound transcription factors in Arabidopsis and rice. *Genomics*, 95(1): 56–65, 2010.
- Tae-Wuk Kim, Chan Ho Park, Chuan-Chih Hsu, Yeong-Woo Kim, Yeong-Woo Ko, Zhenzhen Zhang, Jia-Ying Zhu, Yu-Chun Hsiao, Tess Branon, Krista Kaasik, et al. Mapping the signaling network of BIN2 kinase using TurboID-mediated biotin labeling and phosphoproteomics. *The Plant Cell*, 35(3):975–993, 2023.

- Youn-Sung Kim, Sang-Gyu Kim, Jung-Eun Park, Hye-Young Park, Mi-Hye Lim, Nam-Hai Chua, and Chung-Mo Park. A membrane-bound NAC transcription factor regulates cell division in Arabidopsis. *The Plant Cell*, 18(11):3132–3144, 2006.
- Gary Kleiger and Thibault Mayor. Perilous journey: a tour of the ubiquitin–proteasome system. *Trends in cell biology*, 24(6):352–359, 2014.
- Leopold Kny. *Bau und Entwicklung von Marchantia polymorpha in Botanische wandtafeln mit erläuterndem*. Wiegandt, Hempel Parey, 1890.
- Rumiko Kofuji and Mitsuyasu Hasebe. Eight types of stem cells in the life cycle of the moss *Physcomitrella patens*. *Current opinion in plant biology*, 17:13–21, 2014.
- Claudia Köhler, Lars Hennig, Romaric Bouveret, Jacqueline Gheyselinck, Ueli Grossniklaus, and Wilhelm Gruissem. Arabidopsis MSI1 is a component of the MEA/FIE Polycomb group complex and required for seed development. *The EMBO journal*, 22(18):4804–4814, 2003.
- Ken Kosetsu, Takashi Murata, Moé Yamada, Momoko Nishina, Joanna Boruc, Mitsuyasu Hasebe, Daniël Van Damme, and Gohta Goshima. Cytoplasmic MTOCs control spindle orientation for asymmetric cell division in plants. *Proceedings of the National Academy of Sciences*, 114(42):E8847–E8854, 2017.
- Elena Kozgunova, Mari W Yoshida, Ralf Reski, and Gohta Goshima. Spindle motility skews division site determination during asymmetric cell division in *Physcomitrella*. *Nature Communications*, 13(1):2488, 2022.
- Lucie Kriegshauser, Samuel Knosp, Etienne Grienberger, Kanade Tatsumi, Desirée D Gütle, Iben Sørensen, Laurence Herrgott, Julie Zumsteg, Jocelyn KC Rose, Ralf Reski, et al. Function of the HYDROXYCINNAMOYL-CoA: SHIKIMATE HYDROXYCINNAMOYL TRANSFERASE is evolutionarily conserved in embryophytes. *The Plant Cell*, 33(5):1472–1491, 2021.

- Yong Tae Kwon and Aaron Ciechanover. The ubiquitin code in the ubiquitin-proteasome system and autophagy. *Trends in biochemical sciences*, 42(11):873–886, 2017.
- Yong Tae Kwon, Anna S Kashina, and Alexander Varshavsky. Alternative splicing results in differential expression, activity, and localization of the two forms of arginyl-tRNA-protein transferase, a component of the N-end rule pathway. *Molecular and cellular biology*, 19(1):182–193, 1999.
- Daniel Lang, Kristian K Ullrich, Florent Murat, Jörg Fuchs, Jerry Jenkins, Fabian B Haas, Mathieu Piednoel, Heidrun Gundlach, Michiel Van Bel, Rabea Meyberg, et al. The *Physcomitrella patens* chromosome-scale assembly reveals moss genome structure and evolution. *The Plant Journal*, 93(3):515–533, 2018.
- Michael Lawrence, Wolfgang Huber, Hervé Pagès, Patrick Aboyoun, Marc Carlson, Robert Gentleman, Martin T Morgan, and Vincent J Carey. Software for computing and annotating genomic ranges. *PLoS computational biology*, 9(8):e1003118, 2013.
- Byung Ha Lee, Jae-Heung Ko, Sangman Lee, Yi Lee, Jae-Hong Pak, and Jeong Hoe Kim. The Arabidopsis *GRF-INTERACTING FACTOR* gene family performs an overlapping function in determining organ size as well as multiple developmental properties. *Plant physiology*, 151(2):655–668, 2009.
- Saet Buyl Lee, Sun Ui Yang, Garima Pandey, Myung-Shin Kim, Sujin Hyoun, Doil Choi, Jeong Sheop Shin, and Mi Chung Suh. Occurrence of land-plant-specific glycerol-3-phosphate acyltransferases is essential for cuticle formation and gametophore development in *Physcomitrella patens*. *New Phytologist*, 225(6):2468–2483, 2020.
- Sanghwa Lee, Ling Zhu, and Enamul Huq. An autoregulatory negative feedback loop controls thermomorphogenesis in Arabidopsis. *PLoS Genetics*, 17(6):e1009595, 2021.

- YouJin Lee, Tsui-Fen Chou, Sara K Pittman, Amy L Keith, Babak Razani, and Conrad C Weihl. Keap1/Cullin3 modulates p62/SQSTM1 activity via UBA domain ubiquitination. *Cell reports*, 19(1):188–202, 2017.
- Berbel Lehnert and Martin Bopp. The hormonal regulation of protonema development in mosses i. Auxin-cytokinin interaction. *Zeitschrift für Pflanzenphysiologie*, 110(5):379–391, 1983.
- Andrea Leibfried, Jennifer PC To, Wolfgang Busch, Sandra Stehling, Andreas Kehle, Monika Demar, Joseph J Kieber, and Jan U Lohmann. WUSCHEL controls meristem function by direct regulation of cytokinin-inducible response regulators. *Nature*, 438(7071):1172–1175, 2005.
- Ivica Letunic and Peer Bork. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic acids research*, 49(W1):W293–W296, 2021.
- Fay-Wei Li, Paul Brouwer, Lorenzo Carretero-Paulet, Shifeng Cheng, Jan De Vries, Pierre-Marc Delaux, Ariana Eily, Nils Koppers, Li-Yaung Kuo, Zheng Li, et al. Fern genomes elucidate land plant evolution and cyanobacterial symbioses. *Nature plants*, 4(7):460–472, 2018.
- 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.
- Heng Li. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv preprint arXiv:1303.3997*, 2013.
- Juan Li, Huangwei Chu, Yonghong Zhang, Tongmin Mou, Changyin Wu, Qifa Zhang, and Jian Xu. The rice HGW gene encodes a ubiquitin-associated (UBA) domain protein that regulates heading date and grain weight. *PloS one*, 7(3):e34231, 2012.

- Shanwei Li, Tiantian Sun, and Haiyun Ren. The functions of the cytoskeleton and associated proteins during mitosis and cytokinesis in plant cells. *Frontiers in Plant Science*, 6:282, 2015.
- Yonghua Li-Beisson, Mike Pollard, Vincent Sauveplane, Franck Pinot, John Ohlrogge, and Fred Beisson. Nanoridges that characterize the surface morphology of flowers require the synthesis of cutin polyester. *Proceedings of the National Academy of Sciences*, 106(51):22008–22013, 2009.
- Ivan Liachko, Rachel A Youngblood, Uri Keich, and Maitreya J Dunham. High-resolution mapping, characterization, and optimization of autonomously replicating sequences in yeast. *Genome research*, 23(4):698–704, 2013.
- Zhe Liang, Viktor Demko, Robert C Wilson, Kenneth A Johnson, Rafi Ahmad, Pierre-François Perroud, Ralph Quatrano, Sen Zhao, Kamran Shalchian-Tabrizi, Marisa S Otegui, et al. The catalytic domain CysPc of the DEK 1 calpain is functionally conserved in land plants. *The Plant Journal*, 75(5):742–754, 2013.
- Zhe Liang, Yuke Geng, Changmian Ji, Hai Du, Chui Eng Wong, Qian Zhang, Ye Zhang, Pingxian Zhang, Adeel Riaz, Sadaruddin Chachar, et al. *Mesostigma viride* genome and transcriptome provide insights into the origin and evolution of Streptophyta. *Advanced Science*, 7(1):1901850, 2020.
- Stein Erik Lid, Darren Gruis, Rudolf Jung, Jennifer A Lorentzen, Evgueni Ananiev, Mark Chamberlin, Xiaomu Niu, Robert Meeley, Scott Nichols, and Odd-Arne Olsen. The *defective kernel 1* (*dek1*) gene required for aleurone cell development in the endosperm of maize grains encodes a membrane protein of the calpain gene superfamily. *Proceedings of the National Academy of Sciences*, 99(8):5460–5465, 2002.
- Stein Erik Lid, Lene Olsen, Ragnhild Nestestog, Milo Aukerman, Roy C Brown, Betty Lemmon, Mark Mucha, Hilde-Gunn Opsahl-Sorteberg, and Odd-Arne Olsen. Mutation in the *Arabidopsis thaliana* *DEK1* calpain gene

- perturbs endosperm and embryo development while over-expression affects organ development globally. *Planta*, 221(3):339–351, 2005.
- Qibing Lin, Dan Wang, Hui Dong, Suhai Gu, Zhijun Cheng, Jie Gong, Ruizhen Qin, Ling Jiang, Gang Li, Jiu Lin Wang, et al. Rice APC/C TE controls tillering by mediating the degradation of MONOCULM 1. *Nature communications*, 3(1):752, 2012.
- Hailin Liu, Xiaobo Wang, Guibin Wang, Peng Cui, Shigang Wu, Cheng Ai, Nan Hu, Alun Li, Bing He, Xiujuan Shao, et al. The nearly complete genome of *Ginkgo biloba* illuminates gymnosperm evolution. *Nature Plants*, 7(6):748–756, 2021.
- Yang Liu, Peiyao Li, Li Fan, and Minghua Wu. The nuclear transportation routes of membrane-bound transcription factors. *Cell Communication and Signaling*, 16: 1–9, 2018.
- Jed Long, Thomas P Garner, Maya J Pandya, C Jeremy Craven, Ping Chen, Barry Shaw, Michael P Williamson, Robert Layfield, and Mark S Searle. Dimerisation of the UBA domain of p62 inhibits ubiquitin binding and regulates NF- κ B signalling. *Journal of molecular biology*, 396(1):178–194, 2010.
- 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): 1–21, 2014.
- John T Lovell, Jerry Jenkins, David B Lowry, Sujana Mamidi, Avinash Sreedasyam, Xiaoyu Weng, Kerrie Barry, Jason Bonnette, Brandon Campitelli, Chris Daum, et al. The genomic landscape of molecular responses to natural drought stress in *Panicum hallii*. *Nature communications*, 9(1):5213, 2018.
- Paulina Lozano-Sotomayor, Ricardo A Chávez Montes, Marina Silvestre-Vañó, Humberto Herrera-Ubaldo, Raffaella Greco, Jeanneth Pablo-Villa, Bianca M Galliani, David Diaz-Ramirez, Mieke Weemen, Kim Boutilier, et al. Altered expression of the bZIP transcription factor DRINK ME affects growth and

- reproductive development in *Arabidopsis thaliana*. *The Plant Journal*, 88(3): 437–451, 2016.
- Zhaoxue Ma, Xupeng Hu, Wenjuan Cai, Weihua Huang, Xin Zhou, Qian Luo, Hongquan Yang, Jiawei Wang, and Jirong Huang. Arabidopsis miR171-targeted scarecrow-like proteins bind to GT cis-elements and mediate gibberellin-regulated chlorophyll biosynthesis under light conditions. *PLoS genetics*, 10(8):e1004519, 2014.
- Andrea Mair, Shou-Ling Xu, Tess C Branon, Alice Y Ting, and Dominique C Bergmann. Proximity labeling of protein complexes and cell-type-specific organellar proteomes in Arabidopsis enabled by TurboID. *elife*, 8:e47864, 2019.
- D Blaine Marchant, Guang Chen, Shengguan Cai, Fei Chen, Peter Schafran, Jerry Jenkins, Shengqiang Shu, Chris Plott, Jenell Webber, John T Lovell, et al. Dynamic genome evolution in a model fern. *Nature Plants*, 8(9):1038–1051, 2022.
- Ryan F McCormick, Sandra K Truong, Avinash Sreedasyam, Jerry Jenkins, Shengqiang Shu, David Sims, Megan Kennedy, Mojgan Amirebrahimi, Brock D Weers, Brian McKinley, et al. The *Sorghum bicolor* reference genome: improved assembly, gene annotations, a transcriptome atlas, and signatures of genome organization. *The Plant Journal*, 93(2):338–354, 2018.
- Aaron McKenna, Matthew Hanna, Eric Banks, Andrey Sivachenko, Kristian Cibulskis, Andrew Kernytsky, Kiran Garimella, David Altshuler, Stacey Gabriel, Mark Daly, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome research*, 20(9): 1297–1303, 2010.
- Wen Jing Meng, Zhi Juan Cheng, Ya Lin Sang, Miao Miao Zhang, Xiao Fei Rong, Zhi Wei Wang, Ying Ying Tang, and Xian Sheng Zhang. Type-B Arabidopsis RESPONSE REGULATORS specify the shoot stem cell niche by dual regulation of *WUSCHEL*. *The Plant Cell*, 29(6):1357–1372, 2017.

- Sabeeha S Merchant, Simon E Prochnik, Olivier Vallon, Elizabeth H Harris, Steven J Karpowicz, George B Witman, Astrid Terry, Asaf Salamov, Lillian K Fritz-Laylin, Laurence Maréchal-Drouard, et al. The *Chlamydomonas* genome reveals the evolution of key animal and plant functions. *Science*, 318(5848):245–250, 2007.
- HrW Mewes, K Albermann, M Bähr, D Frishman, A Gleissner, J Hani, K Heumann, K Kleine, A Maierl, SG Oliver, et al. Overview of the yeast genome. *Nature*, 387(6632):7–8, 1997.
- Yoshinobu Mineyuki. The preprophase band of microtubules: its function as a cytokinetic apparatus in higher plants. In *International review of cytology*, volume 187, pages 1–49. Elsevier, 1999.
- Laura A Moody. The 2D to 3D growth transition in the moss *Physcomitrella patens*. *Current opinion in plant biology*, 47:88–95, 2019.
- Laura A Moody. Three-dimensional growth: a developmental innovation that facilitated plant terrestrialization. *Journal of Plant Research*, pages 1–8, 2020.
- Laura A Moody, Steven Kelly, Ester Rabbino-witsch, and Jane A Langdale. Genetic regulation of the 2D to 3D growth transition in the moss *Physcomitrella patens*. *Current Biology*, 28(3):473–478, 2018.
- Laura A Moody, Steven Kelly, Roxaana Clayton, Zoe Weeks, David M Emms, and Jane A Langdale. NO GAMETOPHORES 2 is a novel regulator of the 2D to 3D Growth Transition in the moss *Physcomitrella patens*. *Current Biology*, 31(3):555–563, 2021.
- Assaf Mosquana, Aviva Katz, Eva L Decker, Stefan A Rensing, Ralf Reski, and Nir Ohad. Regulation of stem cell maintenance by the Polycomb protein FIE has been conserved during land plant evolution. 136(14), 2009.
- Iva Mozgova, Claudia Köhler, and Lars Hennig. Keeping the gate closed: functions of the polycomb repressive complex PRC 2 in development. *The Plant Journal*, 83(1):121–132, 2015.

- Ralf Muller, Andrea Bleckmann, and Rudiger Simon. The receptor kinase CORYNE of *Arabidopsis* transmits the stem cell-limiting signal CLAVATA3 independently of CLAVATA1. *The Plant Cell*, 20(4):934–946, 2008.
- Masafumi Muratani and William P Tansey. How the ubiquitin–proteasome system controls transcription. *Nature reviews Molecular cell biology*, 4(3):192–201, 2003.
- Zoe Nemec-Venza, Connor Madden, Amy Stewart, Wei Liu, Ondřej Novák, Aleš Pěňčík, Andrew C Cuming, Yasuko Kamisugi, and C Jill Harrison. *CLAVATA* modulates auxin homeostasis and transport to regulate stem cell identity and plant shape in a moss. *New Phytologist*, 234(1):149–163, 2022.
- Sophia Ng, Aneta Ivanova, Owen Duncan, Simon R Law, Olivier Van Aken, Inge De Clercq, Yan Wang, Chris Carrie, Lin Xu, Beata Kmiec, et al. A membrane-bound NAC transcription factor, ANAC017, mediates mitochondrial retrograde signaling in Arabidopsis. *The Plant Cell*, 25(9):3450–3471, 2013.
- Lam-Tung Nguyen, Heiko A Schmidt, Arndt Von Haeseler, and Bui Quang Minh. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular biology and evolution*, 32(1):268–274, 2015.
- Tomoaki Nishiyama, Hidetoshi Sakayama, Jan De Vries, Henrik Buschmann, Denis Saint-Marcoux, Kristian K Ullrich, Fabian B Haas, Lisa Vanderstraeten, Dirk Becker, Daniel Lang, et al. The Chara genome: secondary complexity and implications for plant terrestrialization. *Cell*, 174(2):448–464, 2018.
- Nir Ohad, Ramin Yadegari, Linda Margossian, Mike Hannon, Daphna Michaeli, John J Harada, Robert B Goldberg, and Robert L Fischer. Mutations in FIE, a WD polycomb group gene, allow endosperm development without fertilization. *The Plant Cell*, 11(3):407–415, 1999.
- Ayako Ohno, JunGoo Jee, Kenichiro Fujiwara, Takeshi Tenno, Natsuko Goda, Hidehito Tochio, Hideki Kobayashi, Hidekazu Hiroaki, and Masahiro Shirakawa.

- Structure of the UBA domain of Dsk2p in complex with ubiquitin: molecular determinants for ubiquitin recognition. *Structure*, 13(4):521–532, 2005.
- Yosuke Okano, Naoki Aono, Yuji Hiwatashi, Takashi Murata, Tomoaki Nishiyama, Takaaki Ishikawa, Minoru Kubo, and Mitsuyasu Hasebe. A polycomb repressive complex 2 gene regulates apogamy and gives evolutionary insights into early land plant evolution. *Proceedings of the National Academy of Sciences*, 106(38):16321–16326, 2009.
- Beatriz Orosa-Puente and Steven H Spoel. Harnessing the ubiquitin code to respond to environmental cues. *Essays in biochemistry*, 66(2):111–121, 2022.
- Shu Ouyang, Wei Zhu, John Hamilton, Haining Lin, Matthew Campbell, Kevin Childs, Françoise Thibaud-Nissen, Renae L Malek, Yuandan Lee, Li Zheng, et al. The TIGR rice genome annotation resource: improvements and new features. *Nucleic acids research*, 35(suppl_1):D883–D887, 2007.
- Brian Palenik, Jane Grimwood, Andrea Aerts, Pierre Rouzé, Asaf Salamov, Nicholas Putnam, Chris Dupont, Richard Jorgensen, Evelyne Derelle, Stephane Rombauts, et al. The tiny eukaryote *Ostreococcus* provides genomic insights into the paradox of plankton speciation. *Proceedings of the National Academy of Sciences*, 104(18):7705–7710, 2007.
- Saurabh Pandey, Amir Bahram Moradi, Oleksandr Dovzhenko, Alisher Touraev, Klaus Palme, and Ralf Welsch. Molecular control of sporophyte-gametophyte ontogeny and transition in plants. *Frontiers in Plant Science*, 12:789789, 2022.
- Andrew H Paterson, Jonathan F Wendel, Heidrun Gundlach, Hui Guo, Jerry Jenkins, Dianchuan Jin, Danny Llewellyn, Kurtis C Showmaker, Shengqiang Shu, Joshua Udall, et al. Repeated polyploidization of gossypium genomes and the evolution of spinnable cotton fibres. *Nature*, 492(7429):423–427, 2012.
- Rob Patro, Geet Duggal, Michael I Love, Rafael A Irizarry, and Carl Kingsford. Salmon provides fast and bias-aware quantification of transcript expression. *Nature methods*, 14(4):417–419, 2017.

Pierre-François Perroud, David J Cove, Ralph S Quatrano, and Stuart F McDaniel.

An experimental method to facilitate the identification of hybrid sporophytes in the moss *Physcomitrella patens* using fluorescent tagged lines. *New Phytologist*, 191(1):301–306, 2011.

Pierre-François Perroud, Viktor Demko, Wenche Johansen, Robert C Wilson, Odd-Arne Olsen, and Ralph S Quatrano. Defective Kernel 1 (DEK 1) is required for three-dimensional growth in *Physcomitrella patens*. *New Phytologist*, 203(3):794–804, 2014.

Pierre-François Perroud, Rabea Meyberg, and Stefan A Rensing. *Physcomitrella patens* Reute mCherry as a tool for efficient crossing within and between ecotypes. *Plant Biology*, 21:143–149, 2019.

Pierre-François Perroud, Rabea Meyberg, Viktor Demko, Ralph S Quatrano, Odd-Arne Olsen, and Stefan A Rensing. DEK1 displays a strong subcellular polarity during *Physcomitrella patens* 3D growth. *New Phytologist*, page 1, 2020.

Jan-Michael Peters. The anaphase promoting complex/cyclosome: a machine designed to destroy. *Nature reviews Molecular cell biology*, 7(9):644–656, 2006.

Cecile M Pickart and David Fushman. Polyubiquitin chains: polymeric protein signals. *Current opinion in chemical biology*, 8(6):610–616, 2004.

Ché S Pillay, Edith Elliott, and Clive Dennison. Endolysosomal proteolysis and its regulation. *Biochemical Journal*, 363(3):417–429, 2002.

Andrew RG Plackett, Verónica S Di Stilio, and Jane A Langdale. Ferns: the missing link in shoot evolution and development. *Frontiers in Plant Science*, 6:166925, 2015.

Alastair Robert Plant, Antoine Larrieu, and Barry Causier. Repressor for hire! the vital roles of TOPLESS-mediated transcriptional repression in plants. *New Phytologist*, 231(3):963–973, 2021.

- Lukas Pluska, Ernst Jarosch, Henrik Zaubner, Andreas Kniss, Anita Waltho, Katrin Bagola, Maximilian von Delbrück, Frank Löhr, Brenda A Schulman, Matthias Selbach, et al. The UBA domain of conjugating enzyme Ubc1/Ube2K facilitates assembly of K48/K63-branched ubiquitin chains. *The EMBO journal*, 40(6): e106094, 2021.
- R Scott Poethig. Small RNAs and developmental timing in plants. *Current opinion in genetics & development*, 19(4):374–378, 2009.
- Juergen EW Polle, Kerrie Barry, John Cushman, Jeremy Schmutz, Duc Tran, Leyla T Hathwaik, Won C Yim, Jerry Jenkins, Zaid McKie-Krisberg, Simon Prochnik, et al. Draft nuclear genome sequence of the halophilic and beta-carotene-accumulating green alga *Dunaliella salina* strain CCAP19/18. *Genome announcements*, 5(43):10–1128, 2017.
- Michael J Prigge and Magdalena Bezanilla. Evolutionary crossroads in developmental biology: *Physcomitrella patens*. *Development*, 137(21):3535–3543, 2010.
- Michael J Prigge, Meirav Lavy, Neil W Ashton, and Mark Estelle. *Physcomitrella patens* auxin-resistant mutants affect conserved elements of an auxin-signaling pathway. *Current Biology*, 20(21):1907–1912, 2010.
- Simon E Prochnik, James Umen, Aurora M Nedelcu, Armin Hallmann, Stephen M Miller, Ichiro Nishii, Patrick Ferris, Alan Kuo, Therese Mitros, Lillian K Fritz-Laylin, et al. Genomic analysis of organismal complexity in the multicellular green alga *Volvox carteri*. *Science*, 329(5988):223–226, 2010.
- Yin-Long Qiu, Libo Li, Bin Wang, Zhiduan Chen, Volker Knoop, Milena Groth-Malonek, Olena Dombrowska, 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.
- R Core Team. *R: A Language and Environment for Statistical Computing*. R

- Foundation for Statistical Computing, Vienna, Austria, 2021. URL <https://www.R-project.org/>.
- Shahri Raasi, Ranjani Varadan, David Fushman, and Cecile M Pickart. Diverse polyubiquitin interaction properties of ubiquitin-associated domains. *Nature structural , molecular biology*, 12(8):708–714, 2005.
- Simin Rahighi, Fumiyo Ikeda, Masato Kawasaki, Masato Akutsu, Nobuhiro Suzuki, Ryuichi Kato, Tobias Kensche, Tamami Uejima, Stuart Bloor, David Komander, et al. Specific recognition of linear ubiquitin chains by NEMO is important for NF- κ B activation. *Cell*, 136(6):1098–1109, 2009.
- Leah Randles and Kylie J Walters. Ubiquitin and its binding domains. *Frontiers in bioscience (Landmark edition)*, 17:2140, 2012.
- Hai Rao and Ashwani Sastry. Recognition of specific ubiquitin conjugates is important for the proteolytic functions of the ubiquitin-associated domain proteins Dsk2 and Rad23. *Journal of Biological Chemistry*, 277(14):11691–11695, 2002.
- Rency J Raquid, Richard Jaeger, and Laura A Moody. CURLY LEAF is required for the auxin-dependent regulation of 3-dimensional growth specification in *Physcomitrium patens*. *Micropublication Biology*, 2023.
- Anamika Rawat, Lucie Brejšková, Michal Hála, Fatima Cvrčková, and Viktor Žárský. The *Physcomitrella patens* exocyst subunit EXO70. 3D has distinct roles in growth and development, and is essential for completion of the moss life cycle. *New Phytologist*, 216(2):438–454, 2017.
- Hugues Renault, Annette Alber, Nelly A Horst, Alexandra Basilio Lopes, Eric A Fich, Lucie Kriegshauser, Gertrud Wiedemann, Pascaline Ullmann, Laurence Herrgott, Mathieu Erhardt, et al. A phenol-enriched cuticle is ancestral to lignin evolution in land plants. *Nature communications*, 8(1):1–8, 2017.
- Stefan A Rensing, Daniel Lang, Andreas D Zimmer, Astrid Terry, Asaf Salamov, Harris Shapiro, Tomoaki Nishiyama, Pierre-François Perroud, Erika A Lindquist,

- Yasuko Kamisugi, et al. The *Physcomitrella* genome reveals evolutionary insights into the conquest of land by plants. *Science*, 319(5859):64–69, 2008.
- Stefan A Rensing, Bernard Goffinet, Rabea Meyberg, Shu-Zon Wu, and Magdalena Bezanilla. The moss *Physcomitrium* (*Physcomitrella*) *patens*: a model organism for non-seed plants. *The Plant Cell*, 32(5):1361–1376, 2020.
- Enrique Rojo, Vijay K Sharma, Valentina Kovaleva, Natasha V Raikhel, and Jennifer C Fletcher. CLV3 is localized to the extracellular space, where it activates the Arabidopsis CLAVATA stem cell signaling pathway. *The Plant Cell*, 14(5):969–977, 2002.
- D Rotin, O Staub, and R Haguenauer-Tsapis. Ubiquitination and endocytosis of plasma membrane proteins: role of Nedd4/Rsp5p family of ubiquitin-protein ligases. *The Journal of membrane biology*, 176:1–17, 2000.
- Ari Sadanandom, Mark Bailey, Richard Ewan, Jack Lee, and Stuart Nelis. The ubiquitin–proteasome system: central modifier of plant signalling. *New Phytologist*, 196(1):13–28, 2012.
- Younousse Saidi, Andrija Finka, Mickhail Chakhporanian, Jean-Pierre Zryd, Didier G Schaefer, and Pierre Goloubinoff. Controlled expression of recombinant proteins in *Physcomitrella patens* by a conditional heat-shock promoter: a tool for plant research and biotechnology. *Plant molecular biology*, 59(5):697–711, 2005.
- Keiko Sakakibara, Pascal Reisewitz, Tsuyoshi Aoyama, Thomas Friedrich, Sayuri Ando, Yoshikatsu Sato, Yosuke Tamada, Tomoaki Nishiyama, Yuji Hiwatashi, Tetsuya Kurata, et al. *WOX13-like* genes are required for reprogramming of leaf and protoplast cells into stem cells in the moss *Physcomitrella patens*. *Development*, 141(8):1660–1670, 2014.
- Yoichi Sakata, Kenji Komatsu, Teruaki Taji, and Shigeo Tanaka. Role of PP2C-mediated ABA signaling in the moss *Physcomitrella patens*. *Plant signaling & behavior*, 4(9):887–889, 2009.

- Shusei Sato, Satoshi Tabata, Hideki Hirakawa, Erika Asamizu, Kenta Shirasawa, Sachiko Isobe, Takakazu Kaneko, Yasukazu Nakamura, Daisuke Shibata, Koh Aoki, et al. The tomato genome sequence provides insights into fleshy fruit evolution. *Nature*, 485(7400):635–641, 2012.
- Estelle Schaefer, Katia Belcram, Magalie Uyttewaal, Yann Duroc, Magali Goussot, David Legland, Elise Laruelle, Marie-Ludivine de Tautzia-Moreau, Martine Pastuglia, and David Bouchez. The preprophase band of microtubules controls the robustness of division orientation in plants. *Science*, 356(6334):186–189, 2017.
- Jeremy Schmutz, Steven B Cannon, Jessica Schlueter, Jianxin Ma, Therese Mitros, William Nelson, David L Hyten, Qijian Song, Jay J Thelen, Jianlin Cheng, et al. Genome sequence of the palaeopolyploid soybean. *nature*, 463(7278):178–183, 2010.
- Heiko Schoof, Michael Lenhard, Achim Haecker, Klaus FX Mayer, Gerd Jürgens, and Thomas Laux. The stem cell population of Arabidopsis shoot meristems is maintained by a regulatory loop between the *CLAVATA* and *WUSCHEL* genes. *Cell*, 100(6):635–644, 2000.
- Christian Schuessele, Sebastian NW Hoernstein, Stefanie J Mueller, Marta Rodriguez-Franco, Timo Lorenz, Daniel Lang, Gabor L Igloi, and Ralf Reski. Spatio-temporal patterning of arginyl-tRNA protein transferase (ATE) contributes to gametophytic development in a moss. *New Phytologist*, 209(3): 1014–1027, 2016.
- Silke Schulze, Barbara Nicole Schäfer, Eneida Abreu Parizotto, Olivier Voinnet, and Klaus Theres. *LOST MERISTEMS* genes regulate cell differentiation of central zone descendants in Arabidopsis shoot meristems. *The Plant Journal*, 64(4): 668–678, 2010.
- Christoph Schuster, Christophe Gaillochet, Anna Medzihradzsky, Wolfgang Busch, Gabor Daum, Melanie Krebs, Andreas Kehle, and Jan U Lohmann. A regulatory

- framework for shoot stem cell control integrating metabolic, transcriptional, and phytohormone signals. *Developmental cell*, 28(4):438–449, 2014.
- Christoph Schuster, Christophe Gaillochet, and Jan U Lohmann. *Arabidopsis HECATE* genes function in phytohormone control during gynoecium development. *Development*, 142(19):3343–3350, 2015.
- Pil Joon Seo, Sang-Gyu Kim, and Chung-Mo Park. Membrane-bound transcription factors in plants. *Trends in plant science*, 13(10):550–556, 2008.
- Tal J Shalev, Omnia Gamal El-Dien, Macaire MS Yuen, Shu Shengqiang, Shaun D Jackman, René L Warren, Lauren Coombe, Lise van der Merwe, Ada Stewart, Lori B Boston, et al. The western redcedar genome reveals low genetic diversity in a self-compatible conifer. *Genome research*, 32(10):1952–1964, 2022.
- Masaki Shimamura. *Marchantia polymorpha*: taxonomy, phylogeny and morphology of a model system. *Plant and Cell Physiology*, 57(2):230–256, 2016.
- Zachery R Smith and Jeff A Long. Control of *Arabidopsis* apical-basal embryo polarity by antagonistic transcription factors. *Nature*, 464(7287):423–426, 2010.
- Marc Somssich, Byoung Il Je, Rüdiger Simon, and David Jackson. CLAVATA-WUSCHEL signaling in the shoot meristem. *Development*, 143(18):3238–3248, 2016.
- Charlotte Soneson, Michael I Love, and Mark D Robinson. Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. *F1000Research*, 4, 2015.
- Victoria Spencer, Zoe Nemec Venza, and Cecily Jill Harrison. What can lycophytes teach us about plant evolution and development? Modern perspectives on an ancient lineage. *Evolution & Development*, 23(3):174–196, 2021.
- Susanne Stary, Xiao-jun Yin, Thomas Potuschak, Peter Schlogelhofer, Victoria Nizhynska, and Andreas Bachmair. PRT1 of *Arabidopsis* is a ubiquitin protein

- ligase of the plant N-end rule pathway with specificity for aromatic amino-terminal residues. *Plant physiology*, 133(3):1360–1366, 2003.
- Olesya V Stepanenko, Olga V Stepanenko, Irina M Kuznetsova, Vladislav V Verkhusha, and Konstantin K Turoverov. Beta-barrel scaffold of fluorescent proteins: folding, stability and role in chromophore formation. *International review of cell and molecular biology*, 302:221–278, 2013.
- Sean R Stevenson, Yasuko Kamisugi, Chi H Trinh, Jeremy Schmutz, Jerry W Jenkins, Jane Grimwood, Wellington Muchero, Gerald A Tuskan, Stefan A Rensing, Daniel Lang, et al. Genetic analysis of *Physcomitrella patens* identifies ABSCISIC ACID NON-RESPONSIVE, a regulator of ABA responses unique to basal land plants and required for desiccation tolerance. *The Plant Cell*, 28(6): 1310–1327, 2016.
- Vivian Su and Alan F Lau. Ubiquitin-like and ubiquitin-associated domain proteins: significance in proteasomal degradation. *Cellular and molecular life sciences*, 66: 2819–2833, 2009.
- Lijun Sun and Zhijian J Chen. The novel functions of ubiquitination in signaling. *Current opinion in cell biology*, 16(2):119–126, 2004.
- Kurt A Swanson, Linda Hicke, and Ishwar Radhakrishnan. Structural basis for monoubiquitin recognition by the Ede1 UBA domain. *Journal of molecular biology*, 358(3):713–724, 2006.
- Heidi Szemenyei, Mike Hannon, and Jeff A Long. TOPLESS mediates auxin-dependent transcriptional repression during Arabidopsis embryogenesis. *Science*, 319(5868):1384–1386, 2008.
- Keiji Tanaka. The proteasome: overview of structure and functions. *Proceedings of the Japan Academy, Series B*, 85(1):12–36, 2009.
- Haibao Tang, Vivek Krishnakumar, Shelby Bidwell, Benjamin Rosen, Agnes Chan, Shiguo Zhou, Laurent Gentzbittel, Kevin L Childs, Mark Yandell, Heidrun

- Gundlach, et al. An improved genome release (version Mt4. 0) for the model legume *Medicago truncatula*. *BMC genomics*, 15(1):1–14, 2014.
- Han Tang, Kilian Duijts, Magdalena Bezanilla, Ben Scheres, Joop EM Vermeer, and Viola Willemsen. Geometric cues forecast the switch from two-to three-dimensional growth in *Physcomitrella patens*. *New Phytologist*, 2019.
- Mattias Thelander, Katarina Landberg, and Eva Sundberg. Minimal auxin sensing levels in vegetative moss stem cells revealed by a ratiometric reporter. *New Phytologist*, 224(2):775–788, 2019.
- Richard T Timms and Itay Koren. Tying up loose ends: the N-degron and C-degron pathways of protein degradation. *Biochemical Society Transactions*, 48(4):1557–1567, 2020.
- Jan Traas and Teva Vernoux. The shoot apical meristem: the dynamics of a stable structure. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 357(1422):737–747, 2002.
- Gerald A Tuskan, Stephen Difazio, Stefan Jansson, Jörg Bohlmann, Igor Grigoriev, Uffe Hellsten, Nicholas Putnam, Steven Ralph, Stephane Rombauts, Asaf Salamov, et al. The genome of black cottonwood, *Populus trichocarpa* (torr. , gray). *science*, 313(5793):1596–1604, 2006.
- Maritza van Dop, Marc Fiedler, Sumanth Mutte, Jeroen de Keijzer, Lisa Olijslager, Catherine Albrecht, Che-Yang Liao, Marcel E Janson, Mariann Bienz, and Dolf Weijers. DIX domain polymerization drives assembly of plant cell polarity complexes. *Cell*, 180(3):427–439, 2020.
- Ruben Vanholme, Barbara De Meester, John Ralph, and Wout Boerjan. Lignin biosynthesis and its integration into metabolism. *Current Opinion in Biotechnology*, 56:230–239, 2019.
- Alexander Varshavsky. N-degron and C-degron pathways of protein degradation. *Proceedings of the National Academy of Sciences*, 116(2):358–366, 2019.

- Liesbeth Vercruyssen, Aurine Verkest, Nathalie Gonzalez, Ken S Heyndrickx, Dominique Eeckhout, Soon-Ki Han, Teddy Jégu, Rafal Archacki, Jelle Van Leene, Megan Andriankaja, et al. ANGUSTIFOLIA3 binds to SWI/SNF chromatin remodeling complexes to regulate transcription during Arabidopsis leaf development. *The Plant Cell*, 26(1):210–229, 2014.
- Tom Viaene, Katarina Landberg, Mattias Thelander, Eva Medvecka, Eric Pederson, Elena Feraru, Endymion D Cooper, Mansour Karimi, Charles F Delwiche, Karin Ljung, et al. Directional auxin transport mechanisms in early diverging land plants. *Current Biology*, 24(23):2786–2791, 2014.
- Luis Vidali, Robert C Augustine, Ken P Kleinman, and Magdalena Bezanilla. Profilin is essential for tip growth in the moss *Physcomitrella patens*. *The Plant Cell*, 19(11):3705–3722, 2007.
- Luis Vidali, Graham M Burkart, Robert C Augustine, Erin Kerdavid, Erkan Tüzel, and Magdalena Bezanilla. Myosin XI is essential for tip growth in *Physcomitrella patens*. *The Plant Cell*, 22(6):1868–1882, 2010.
- Richard D Vierstra. Protein degradation in plants. *Annual review of plant biology*, 44(1):385–410, 1993.
- Klaus von Schwartzberg, Ann-Cathrin Lindner, Njuscha Gruhn, Jan Šimura, Ondřej Novák, Miroslav Strnad, Martine Gonneau, Fabien Nogué, and Alexander Heyl. CHASE domain-containing receptors play an essential role in the cytokinin response of the moss *Physcomitrella patens*. *Journal of experimental botany*, 67(3):667–679, 2016.
- Feng Wang and Xing Wang Deng. Plant ubiquitin-proteasome pathway and its role in gibberellin signaling. *Cell research*, 21(9):1286–1294, 2011.
- Guandong Wang and Weixiong Zhang. A steganalysis-based approach to comprehensive identification and characterization of functional regulatory elements. *Genome biology*, 7:1–16, 2006.

- Jun Wang, Yao-Fu Chang, John I Hamilton, and Miles F Wilkinson. Nonsense-associated altered splicing: a frame-dependent response distinct from nonsense-mediated decay. *Molecular cell*, 10(4):951–957, 2002.
- Sibo Wang, Linzhou Li, Haoyuan Li, Sunil Kumar Sahu, Hongli Wang, Yan Xu, Wenfei Xian, Bo Song, Hongping Liang, Shifeng Cheng, et al. Genomes of early-diverging streptophyte algae shed light on plant terrestrialization. *Nature plants*, 6(2):95–106, 2020.
- Wenkin Wang, G Haberer, H Gundlach, C Gläßer, TCLM Nussbaumer, MC Luo, A Lomsadze, M Borodovsky, RA Kerstetter, J Shanklin, et al. The *Spirodela polyrhiza* genome reveals insights into its neotenuous reduction fast growth and aquatic lifestyle. *Nature communications*, 5(1):3311, 2014.
- Zhishuo Wang, Beatriz Orosa-Puente, Mika Nomoto, Heather Grey, Thomas Potuschak, Takakazu Matsuura, Izumi C Mori, Yasuomi Tada, Pascal Genschik, and Steven H Spoel. Proteasome-associated ubiquitin ligase relays target plant hormone-specific transcriptional activators. *Science Advances*, 8(42):eabn4466, 2022.
- Tomas Werner, Václav Motyka, Valérie Laucou, Rafaël Smets, Harry Van Onckelen, and Thomas Schmulling. Cytokinin-deficient transgenic Arabidopsis plants show multiple developmental alterations indicating opposite functions of cytokinins in the regulation of shoot and root meristem activity. *The Plant Cell*, 15(11):2532–2550, 2003.
- Chris D Whitewoods, Joseph Cammarata, Zoe Nemec Venza, Stephanie Sang, Ashley D Crook, Tsuyoshi Aoyama, Xiao Y Wang, Manuel Waller, Yasuko Kamisugi, Andrew C Cuming, et al. *CLAVATA* was a genetic novelty for the morphological innovation of 3D growth in land plants. *Current Biology*, 28(15):2365–2376, 2018.
- Norman J Wickett, Siavash Mirarab, Nam Nguyen, Tandy Warnow, Eric Carpenter, Naim Matasci, Saravanaraj Ayyampalayam, Michael S Barker, J Gordon Burleigh,

- Matthew A Gitzendanner, et al. Phylotranscriptomic analysis of the origin and early diversification of land plants. *Proceedings of the National Academy of Sciences*, 111(45):E4859–E4868, 2014.
- Debbie Winter, Ben Vinegar, Hardeep Nahal, Ron Ammar, Greg V Wilson, and Nicholas J Provart. An “Electronic Fluorescent Pictograph” browser for exploring and analyzing large-scale biological data sets. *PloS one*, 2(8):e718, 2007.
- Tanja Woelk, Sara Sigismund, Lorenza Penengo, and Simona Polo. The ubiquitination code: a signalling problem. *Cell division*, 2(1):1–12, 2007.
- Hanno Wolters and Gerd Jürgens. Survival of the flexible: hormonal growth control and adaptation in plant development. *Nature Reviews Genetics*, 10(5):305–317, 2009.
- Alexandra Z Worden, Jae-Hyeok Lee, Thomas Mock, Pierre Rouzé, Melinda P Simmons, Andrea L Aerts, Andrew E Allen, Marie L Cuvelier, Evelyne Derelle, Meredith V Everett, et al. Green evolution and dynamic adaptations revealed by genomes of the marine picoeukaryotes *Micromonas*. *Science*, 324(5924):268–272, 2009.
- Gang Wu and R Scott Poethig. Temporal regulation of shoot development in *Arabidopsis thaliana* by miR156 and its target SPL3. *Development*, 133(18):3539–3547, 2006.
- Tianzhi Wu, Erqiang Hu, Shuangbin Xu, Meijun Chen, Pingfan Guo, Zehan Dai, Tingze Feng, Lang Zhou, Wenli Tang, LI Zhan, et al. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *The innovation*, 2(3), 2021.
- Brecht Wybouw and Bert De Rybel. Cytokinin—a developing story. *Trends in plant science*, 24(2):177–185, 2019.
- Lihong Xiao, Abou Yobi, Karen L Koster, Yikun He, and Melvin J Oliver. Desiccation tolerance in *Physcomitrella patens*: rate of dehydration and the

- involvement of endogenous abscisic acid (ABA). *Plant, cell & environment*, 41 (1):275–284, 2018.
- Cao Xu, Yonghong Wang, Yanchun Yu, Jingbo Duan, Zhigang Liao, Guosheng Xiong, Xiangbing Meng, Guifu Liu, Qian Qian, and Jiayang Li. Degradation of MONOCULM 1 by APC/C TAD1 regulates rice tillering. *Nature communications*, 3(1):750, 2012.
- Mingli Xu, Tieqiang Hu, Jianfei Zhao, Mee-Yeon Park, Keith W Earley, Gang Wu, Li Yang, and R Scott Poethig. Developmental functions of miR156-regulated *SQUAMOSA PROMOTER BINDING PROTEIN-LIKE* (*SPL*) genes in *Arabidopsis thaliana*. *PLoS genetics*, 12(8):e1006263, 2016.
- Xiaosa Xu, Praveen Kumar Kathare, Vinh Ngoc Pham, Qingyun Bu, Andrew Nguyen, and Enamul Huq. Reciprocal proteasome-mediated degradation of PIFs and (*hfr1*) underlies photomorphogenic development in *Arabidopsis*. *Development*, 144(10):1831–1840, 2017.
- Ramin Yadegari, Tetsu Kinoshita, Ofra Lotan, Gal Cohen, Anat Katz, Yeonhee Choi, Aviva Katz, Kazuo Nakashima, John J Harada, Robert B Goldberg, et al. Mutations in the FIE and MEA genes that encode interacting polycomb proteins cause parent-of-origin effects on seed development by distinct mechanisms. *The plant cell*, 12(12):2367–2381, 2000.
- Weili Yang, Mike Pollard, Yonghua Li-Beisson, Fred Beisson, Michael Feig, and John Ohlrogge. A distinct type of glycerol-3-phosphate acyltransferase with *sn-2* preference and phosphatase activity producing 2-monoacylglycerol. *Proceedings of the National Academy of Sciences*, 107(26):12040–12045, 2010.
- Weili Yang, Jeffrey P Simpson, Yonghua Li-Beisson, Fred Beisson, Mike Pollard, and John B Ohlrogge. A land-plant-specific glycerol-3-phosphate acyltransferase family in *Arabidopsis*: substrate specificity, *sn-2* preference, and evolution. *Plant Physiology*, 160(2):638–652, 2012.

- Yu Ye, Georg Blaser, Mathew H Horrocks, Maria J Ruedas-Rama, Shehu Ibrahim, Alexander A Zhukov, Angel Orte, David Klenerman, Sophie E Jackson, and David Komander. Ubiquitin chain conformation regulates recognition and activity of interacting proteins. *Nature*, 492(7428):266–270, 2012.
- Mari W Yoshida, Noiri Oguri, and Gohta Goshima. *Physcomitrium patens* SUN2 mediates MTOC association with the nuclear envelope and facilitates chromosome alignment during spindle assembly. *Plant and Cell Physiology*, 64(9):1106–1117, 2023.
- Alexis Zajicek and Wei-Dong Yao. Remodeling without destruction: non-proteolytic ubiquitin chains in neural function and brain disorders. *Molecular psychiatry*, 26(1):247–264, 2021.
- Liechi Zhang, Yuko Sasaki-Sekimoto, Ken Kosetsu, Tsuyoshi Aoyama, Takashi Murata, Yukiko Kabeya, Yoshikatsu Sato, Shizuka Koshimizu, Mie Shimojima, Hiroyuki Ohta, et al. An ABCB transporter regulates anisotropic cell expansion via cuticle deposition in the moss *Physcomitrium patens*. *New Phytologist*, 241(2), 2023.
- Sheng Zhao and Russell D Fernald. Comprehensive algorithm for quantitative real-time polymerase chain reaction. *Journal of computational biology*, 12(8):1047–1064, 2005.
- Xuelian Zheng, Caiyan Qi, Lijia Yang, Quan Quan, Binglin Liu, Zhaohui Zhong, Xu Tang, Tingting Fan, Jianping Zhou, and Yong Zhang. The improvement of CRISPR-Cas9 system with ubiquitin-associated domain fusion for efficient plant genome editing. *Frontiers in plant science*, 11:524429, 2020.
- Ling Zhu, Ruijiao Xin, Qingyun Bu, Hui Shen, Jonathan Dang, and Enamul Huq. A negative feedback loop between PHYTOCHROME INTERACTING FACTORS and HECATE proteins fine-tunes photomorphogenesis in Arabidopsis. *The Plant Cell*, 28(4):855–874, 2016.

Evelyne Zürcher, Deborah Tavor-Deslex, Dmytro Lituiev, Katalin Enkerli, Paul T Tarr, and Bruno Müller. A robust and sensitive synthetic sensor to monitor the transcriptional output of the cytokinin signaling network in planta. *Plant physiology*, 161(3):1066–1075, 2013.

Appendix

Chapter 2: Materials and methods

Name	Target	Sequence
AtNOG1_F_BamHI	AtNOG1 CDS	aaaggatccATGGATTACGATTACAGAAA
AtNOG1_R_XbaI	AtNOG1 CDS	aaatctagaTTAGGAAGAACCATGGAGCA
PpNOG1_F_BamHI	PpNOG1 CDS	aaaggatccATGGAATATGATTATGGACG
PpNOG1_R_XbaI	PpNOG1 CDS	aaatctagaTTAAGAACTAGAGATGATA
PpNOG1noStopXmaI	PpNOG1 CDS	aaacccgggAGAACTAGAGATGATATTG
Pp3c8_19720 int F	Pp3c8_19720 CDS internal fragment	CAAGGCCTACCGTCTCATCC
Pp3c8_19720 int R	Pp3c8_19720 CDS internal fragment	GTGGAGGAGGGACCTCTTGA
Pp3c8_19720 F BamHI	Pp3c8_19720 CDS	aaaGGATCCATGGATCATGTGGGATCC
Pp3c8_19720noSTOP_ R_ XmaI	Pp3c8_19720 CDS	aaaCCCGGGCCGGCCATACCAGC
snog1a_del_genotyping2_F	upstream of SNOG1A locus	GTCCACCAAGACCACGAAAC
snog1a_del_genotyping2_R	within NPTII in SNOG1A del construct	CATCAGAGCAGCCGATTGTC

nog1del_promoter_F	NOG1 promoter fragment	CCATCCATGCACACAACCAA
nog1del_promoter_R	NOG1 promoter fragment	TGCACGTTACTCCCTTGTCT
putSNOG1A_genomic1_F	SNOG1A genomic fragment	CCGAGTTTGTGTAGGCAAGG
putSNOG1A_genomic1_R	SNOG1A genomic fragment	ACCTCCAAATACCACTCACCA
putSNOG1A_genomic2_F	SNOG1A genomic fragment	CAGCTGACAACTTTGGTGCA
putSNOG1A_genomic2_R	SNOG1A genomic fragment	CATCTGCTGCTGAGGAAGTG
putSNOG1A_genomic3_F	SNOG1A genomic fragment	CAAGTCCAGGTTGCGACATC
putSNOG1A_genomic3_R	SNOG1A genomic fragment	AAGCAATCTTCATCACCGGC
SNOG1del_int_F	SNOG1A 5' fragment in construct	CTCTGGTCGTGCTTCAGGTA
SNOG1del_int_R	within NPTII in snog1a del construct	AATACCTGACAGCCCAACCA
DEK1 del genotyping 3' F	within NPTII in DEK1 del construct	AACTAGTTCCTCACACCGG
DEK1 del genotyping 3' R	downstream of DEK1 locus	GCGTTCACATATGCAAGCCT

DEK1 del genotyping 5 prime F2	upstream of DEK1 locus	CACAGATTCCACGCACATGT
DEK1 del genotyping 5 prime R2	within NPTII in DEK1 del construct	TGCTTTGAAGACGTGGTTGG
DEK1del3prime.int_F	within NPTII in DEK1 del construct	TCTACAGCCAGCTTTGGGAA
DEK1del3prime.int_R	SNOG1A 3' fragment in construct	ACTTTGTCAGGCCTCCTTCA
CCS52a_seq_F	CCS52A CDS	ATGGATCATCCTATGGACG
CCS52a_seq_R	CCS52A CDS	TCGAATGTGTGTCCGT
MOC1a_seq_F	MOC1a CDS	ATGCAATTGTTTGTGCCA
MOC1a_seq_R	MOC1a CDS	CATGTTACTTGGATGCCAA
MOC1b_seq_F	MOC1b CDS	ATGCCAGTGGCCC
MOC1b_seq_R	MOC1b CDS	CATGTTATTCGTATTCCAAGAAGA
HEC1a_seq_F	HEC1a CDS	ATGACGCTTAGCTTGAGA
HEC1a_seq_R	HEC1a CDS	CATTCTGCCATCGGAAAA
HEC1b_seq_F	HEC1b CDS	ATGACACTTAGTTTGAGACC
HEC1b_seq_R	HEC1b CDS	CATTGGGGCAATCGTCTTG
DEK1.WT_F	DEK1 genomic fragment	GCCGATTGTCTGCATAACGT
DEK1.WT_R	DEK1 genomic fragment	CCCACAAGATAAGAACCGCG
SNOG1A.dUBA_R	Before SNOG1A UBA domain	AAGCTTGACGCGGTT
SNOG1A.dDS_F	From SNOG1A DS domain	GTCGACATGGCCCC
pGADT7_seq_F	pGADT7 fragment	CCATGGAGTACCCATACGAC

pGBKT7_seq_F	pGBKT7 fragment	CGACTCACTATAGGGCGAG
PptubF	alpha-tubulin CDS	TGTGCTGTTGGACAATGAG
PptubR	alpha-tubulin CDS	ACATCAGATCGAACTTGTG
NOG2_29140_qPCRF2	PpNOG2 CDS fragment 2 (for qPCR)	TGATCTTCACTTGCACCTCCTATT
NOG2_29140_qPCRR2	PpNOG2 CDS fragment 2 (for qPCR)	TTCATCATTCATTCTCGCCAAC
NOG2_F_SalI	PpNOG2 CDS	aaagtcgacATGGCCGCCGCAAGTCAAGTT
NOG2noSTOP_R_HindIII	PpNOG2 CDS	aaaaagcttGAAGGATGCCACTAGTTTGG
EF1alphaqPCRF	EF1alpha CDS fragment (for qPCR)	AATCATACATTTTCACCTCGCC
EF1alphaqPCRR	EF1alpha CDS fragment (for qPCR)	GATCAGTGGGTAGAAGTGAC
PpACT7qPCRF2	PpACT7 CDS fragment 2 (for qPCR)	GGGGAAGGTGAGGATGTCC
PpACT7qPCRR2	PpACT7 CDS fragment 2 (for qPCR)	GATGCTCGGAAACACGGC
AtNOG1_F_SalI	AtNOG1 CDS	aaagtcgacATGGATTACGATTACAGAAA
AtNOG1noSTOP_R_HindIII	AtNOG1 CDS	aaaaagcttGGAAGAACCATGGAGCAAAT
PpNOG1_F_SalI	PpNOG1 CDS	aaagtcgacATGGAATATGATTATGGACG
PpNOG1noSTOP_R_HindIII	PpNOG1 CDS	aaaaagcttAGAACTAGAGATGATATTG
PpNOG1UBA_F_SalI	PpNOG1 from start of UBA domain	aaagtcgacATGGGCCATAATAAAGAAGC
PpNOG1noSTOP_R_BamHI	PpNOG1 CDS	aaagatccAGAACTAGAGATGATATTG

NOG1.5FKpnI	PpNOG1 5 prime flanking sequence	aaaggtaccCCATCCATGCACACAACCAA
NOG1.5RXhoI	PpNOG1 5 prime flanking sequence	aaactcgagCCTCCGCTCCAAACTCCCAC
NOG1.3FNotI	PpNOG1 3 prime flanking sequence	aagcgccgcTAATCTGTGTATGAGTTCAG
NOG1.3RNotI	PpNOG1 3 prime flanking sequence	aagcgccgcGAGTTATCTAGTTTTGTGGA
PpDEK1.5FNotI	PpDEK1 5 prime flanking sequence	aagcgccgcTGGTACAGAGAGCAGTCCAG
PpDEK1.5RXbaI	PpDEK1 5 prime flanking sequence	aaatctagaGGCGCCAATTTACAGCAACT
PpDEK1.3FEcoRV	PpDEK1 3 prime flanking sequence	aaagatatcTGAAAGCCGTGGGTAGTGAT
PpDEK1.3RXmaI	PpDEK1 3 prime flanking sequence	aaacccgggGGAATTGCAGCGGAAAGACT
NOG1.3FSphI	PpNOG1 3 prime flanking sequence	aaagcatgcTAATCTGTGTATGAGTTCAG
NOG1.3RNotI	PpNOG1 3 prime flanking sequence	aagcgccgcGAGTTATCTAGTTTTGTGGA
PpNOG1_F_XmaI	PpNOG1 CDS	aaacccgggATGGAATATGATTATGGACGAAG
PpNOG1_R_XmaI	PpNOG1 CDS	aaacccgggTTAAGAACTAGAGATGATATTGCT
AtUBQ3_F_XmaI	AtUBQ3 CDS	aaacccgggATGCAAATCTTCGTGAAAACA
AtUBQ3_R_XmaI	AtUBQ3 CDS	aaacccgggTTAGAAGCTTCCACCACG

PpUBQ3A_FXmaI	PpUBQ3A CDS	aaaccgaggATGCAGATTTTTGTGAAGAC
PpUBQ3A_R_BamHI	PpUBQ3A CDS	aaaggatccTCAAAATCCTCCTCGAAGGC
PpUBQ3B_F_XmaI	PpUBQ3B CDS	aaaccgaggATGCAGATCTTTGTAAAGAC
PpUBQ3B_R_BamHI	PpUBQ3B CDS	aaaggatccCTACAGACCACCCCGGAGCC
PpUBQ3C_F_XmaI	PpUBQ3C CDS	aaaccgaggATGCAGATCTTCGTGAAAAC
PpUBQ3C_R_BamHI	PpUBQ3C CDS	aaaggatccTCACTGTCCACCTCTGAGGC
AtNOG1_R_BamHI	AtNOG1 CDS	aaaggatccTTAGGAAGAACCATGGAGCAAAT

Table S1: All primer names and sequences used in this work.

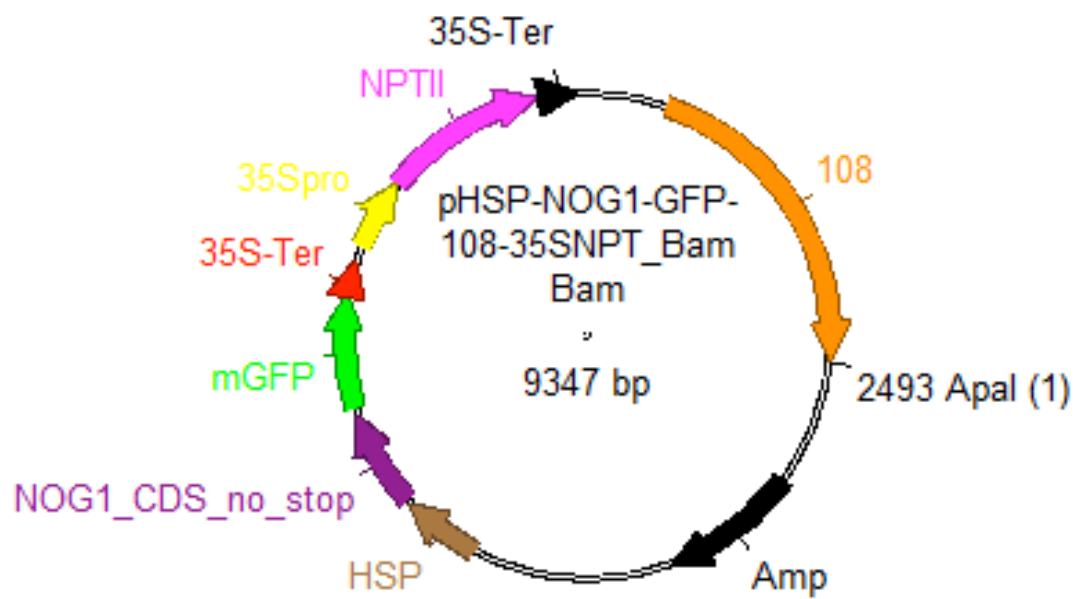


Figure S1: The *pHSP::PpNOG1-GFP* construct

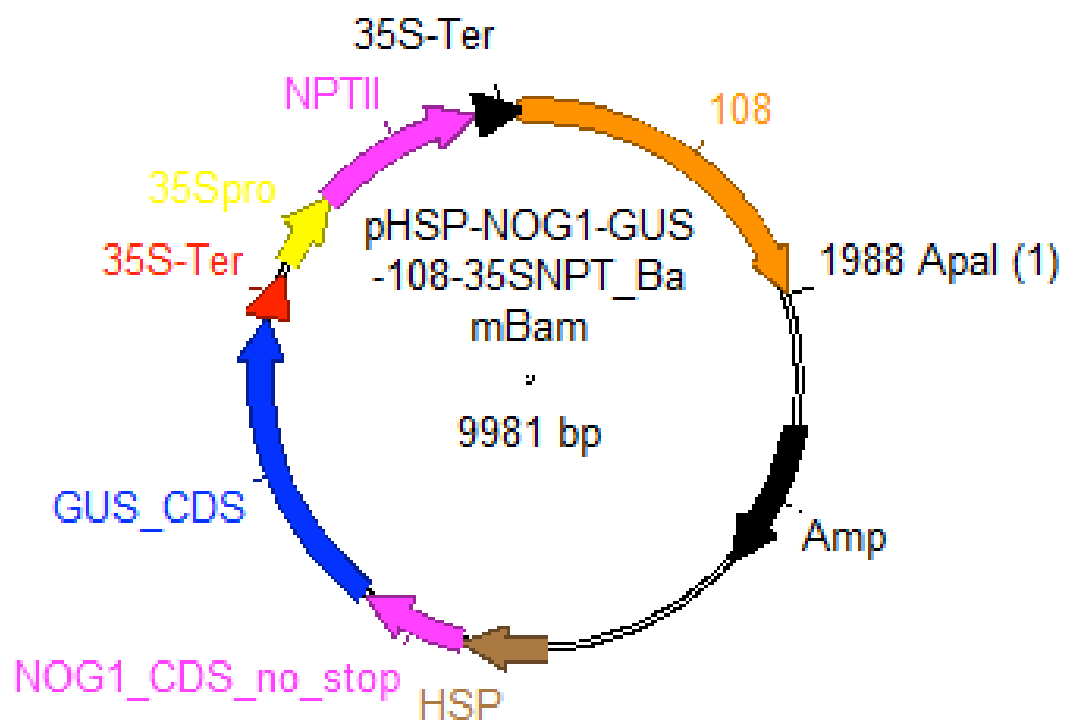


Figure S2: The *pHSP::PpNOG1-GUS* construct

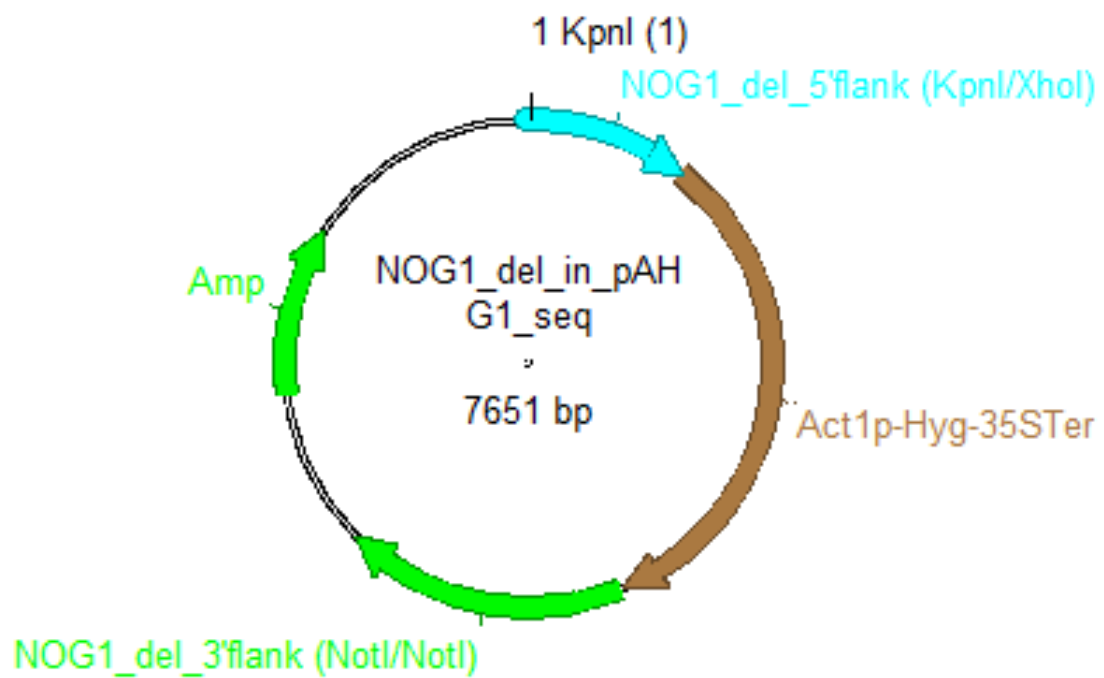


Figure S3: The *pNOG1delH* construct

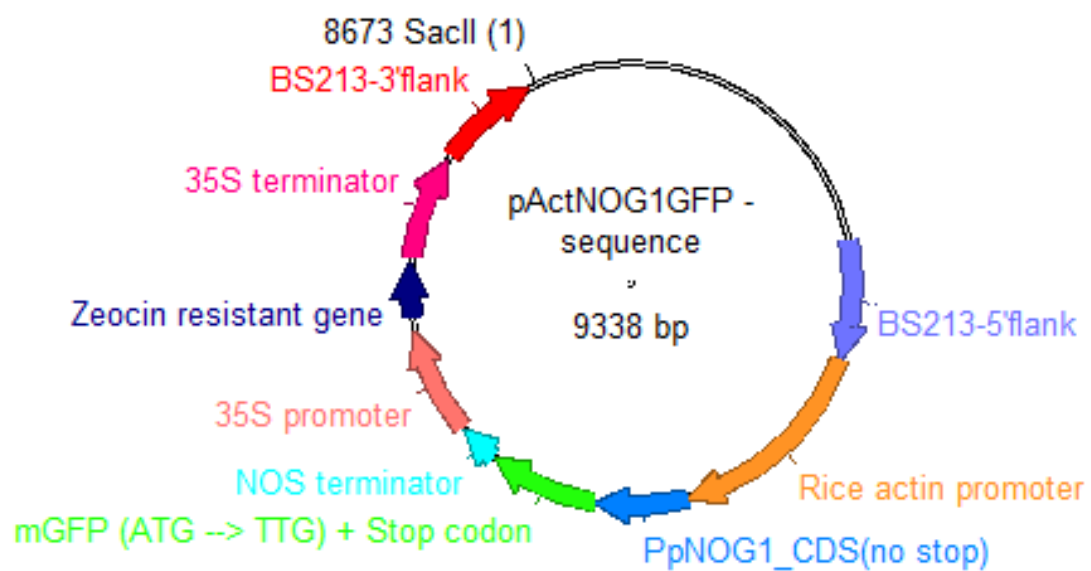


Figure S4: The *pACT::PpNOG1-GFP* construct

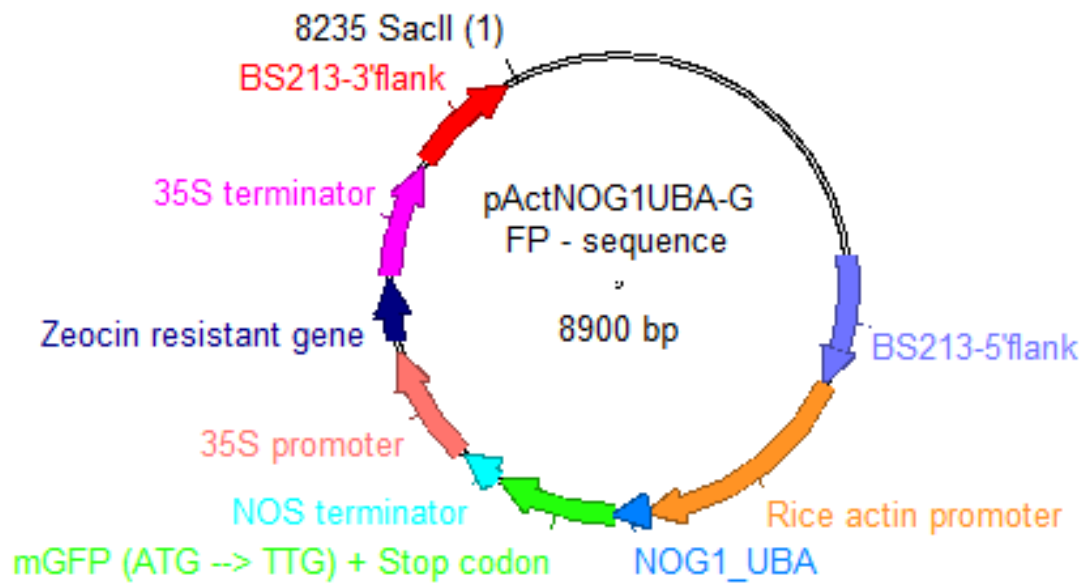


Figure S5: The *pACT::PpNOG1^{UBA}-GFP* construct

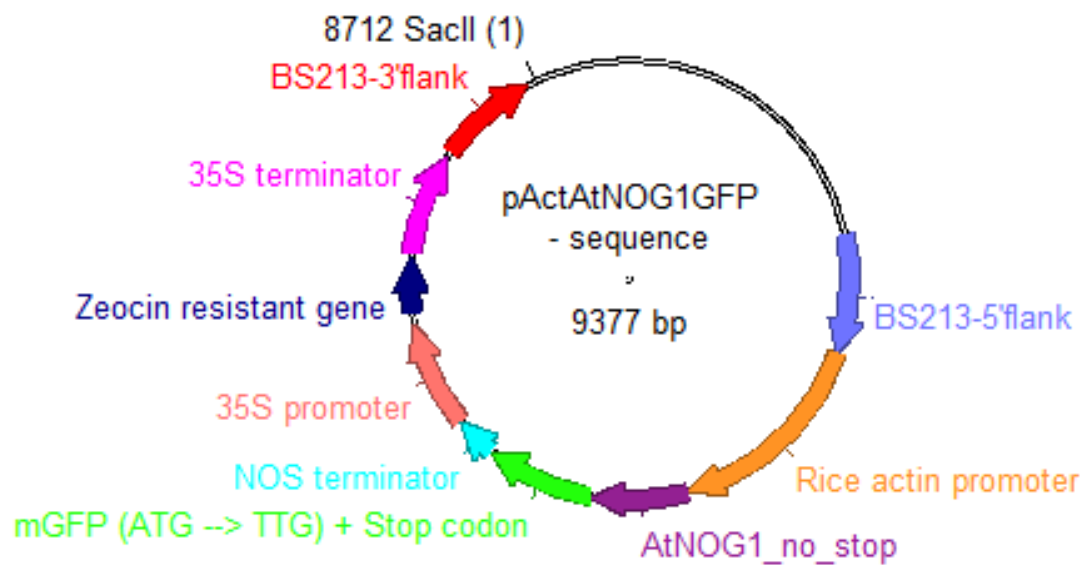


Figure S6: The *pACT::AtNOG1-GFP* construct

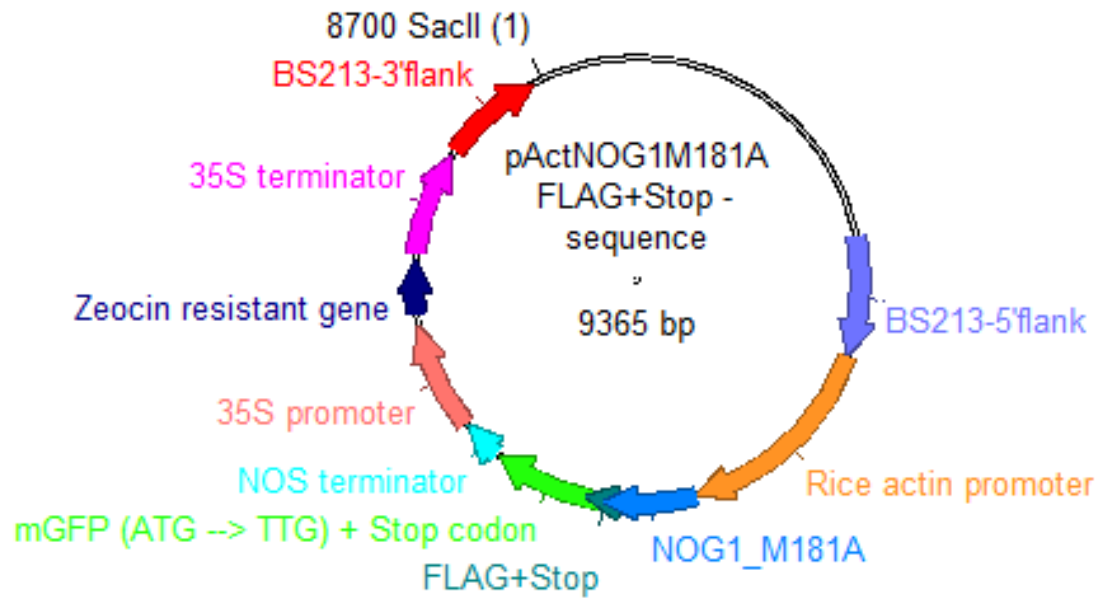


Figure S7: The *pACT::PpNOG1^{M181A}-GFP* construct

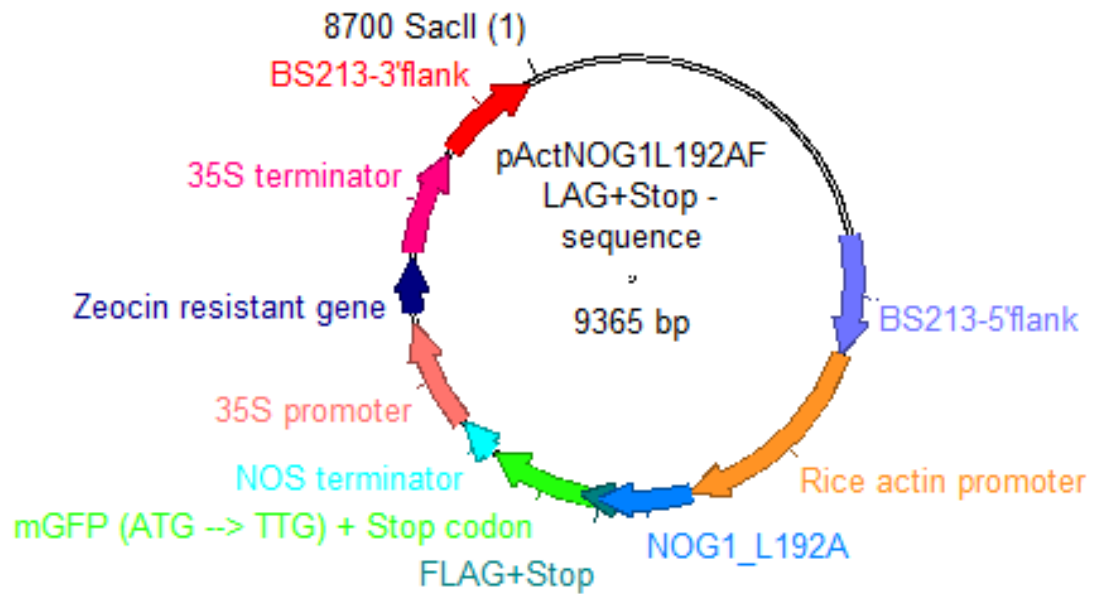


Figure S8: The *pACT::PpNOG1^{L192A}-GFP* construct

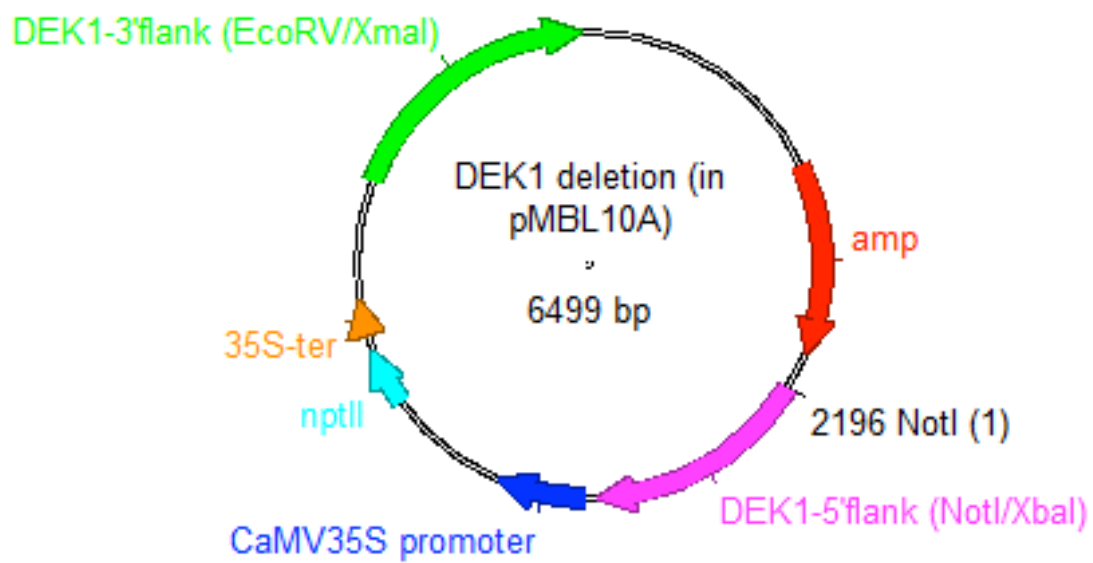


Figure S9: The *PpDEK1* deletion construct

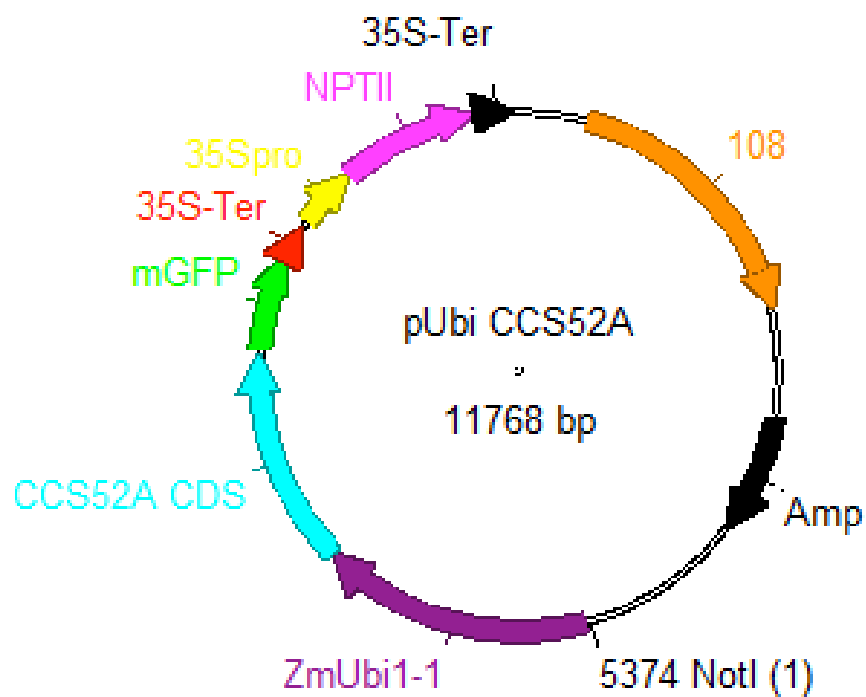


Figure S10: The *pUBI::PpCCS52A-GFP* construct

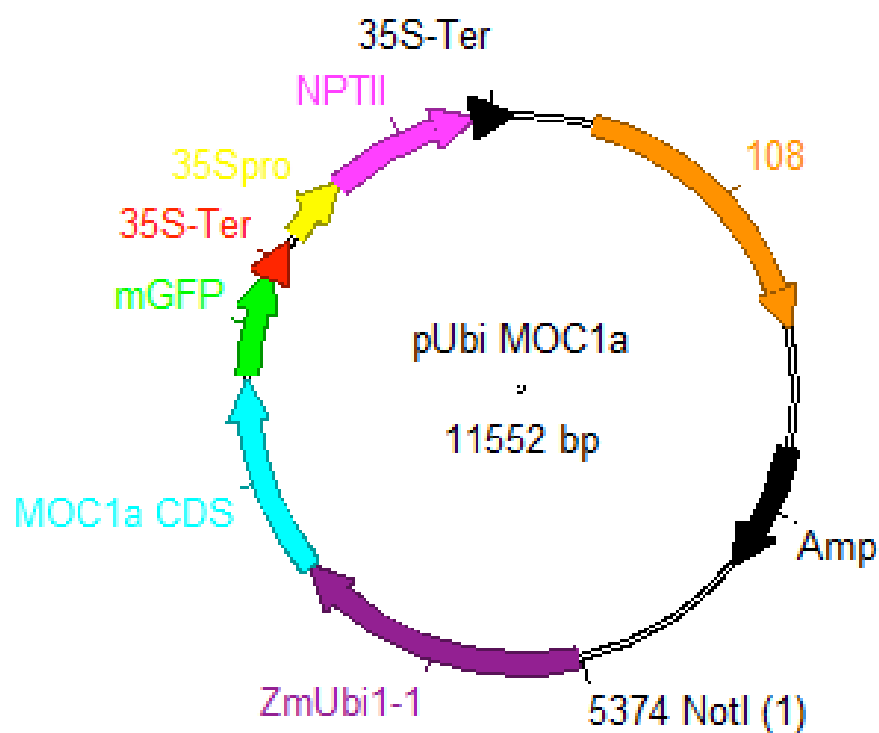


Figure S11: The *pUBI::PpMOC1a-GFP* construct

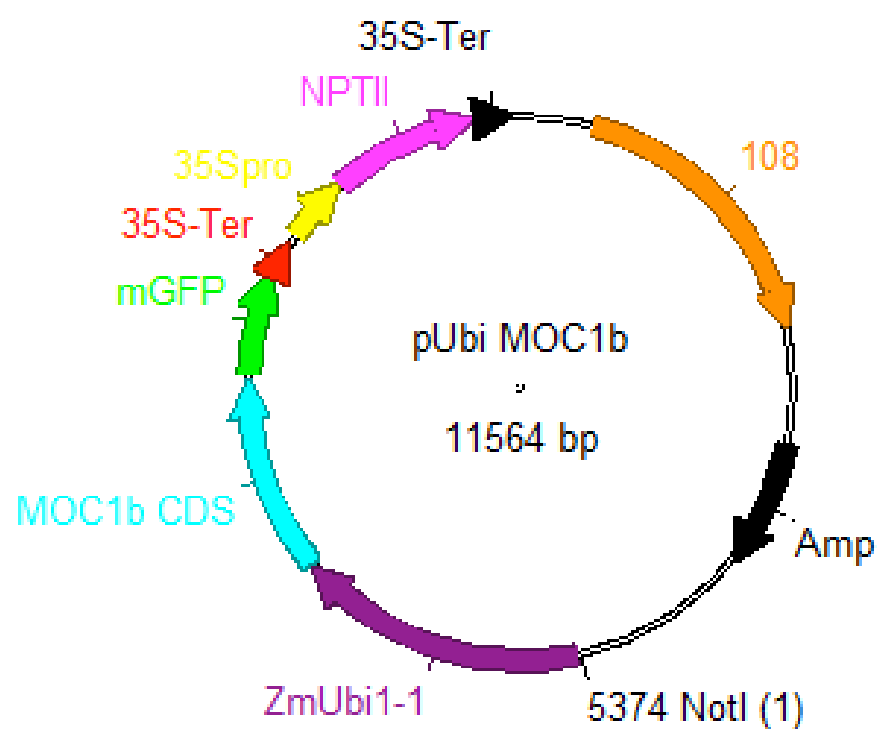


Figure S12: The *pUBI::PpMOC1b-GFP* construct

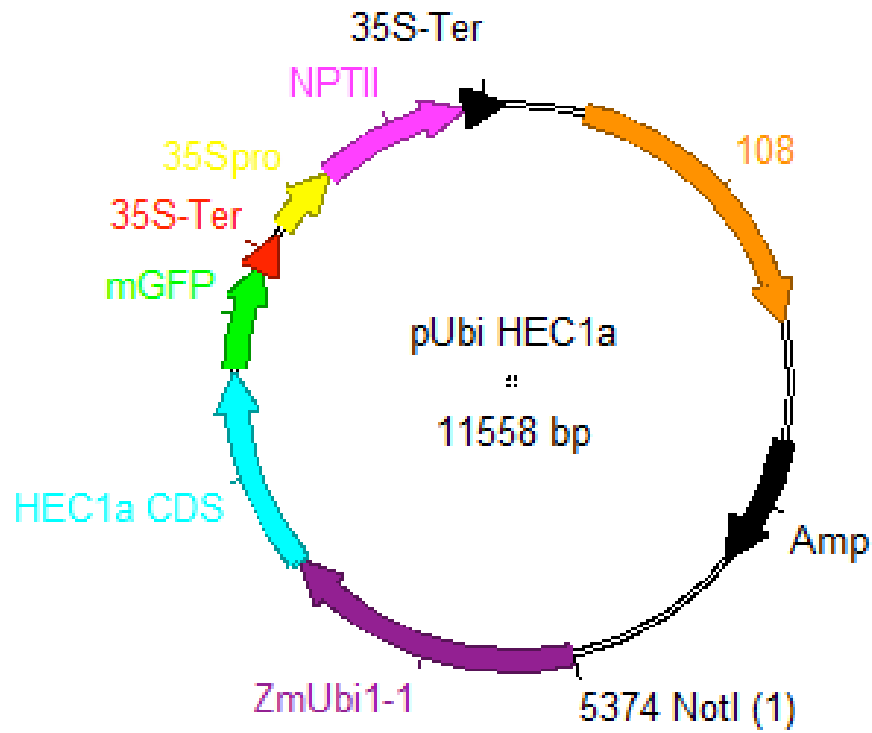


Figure S13: The *pUBI::PpHEC1a-GFP* construct

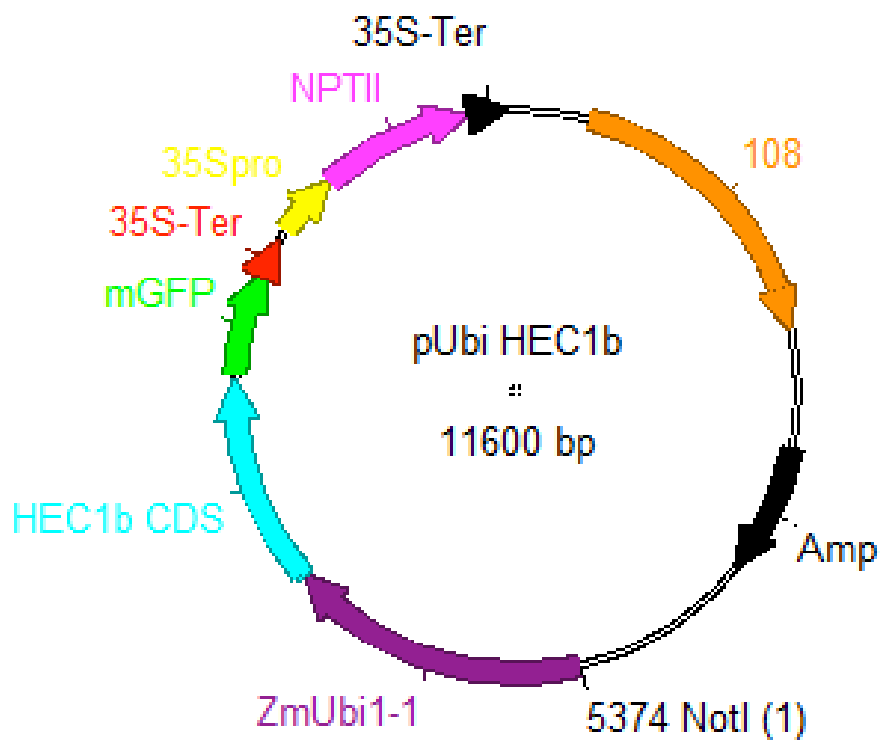


Figure S14: The *pUBI::PpHEC1b-GFP* construct

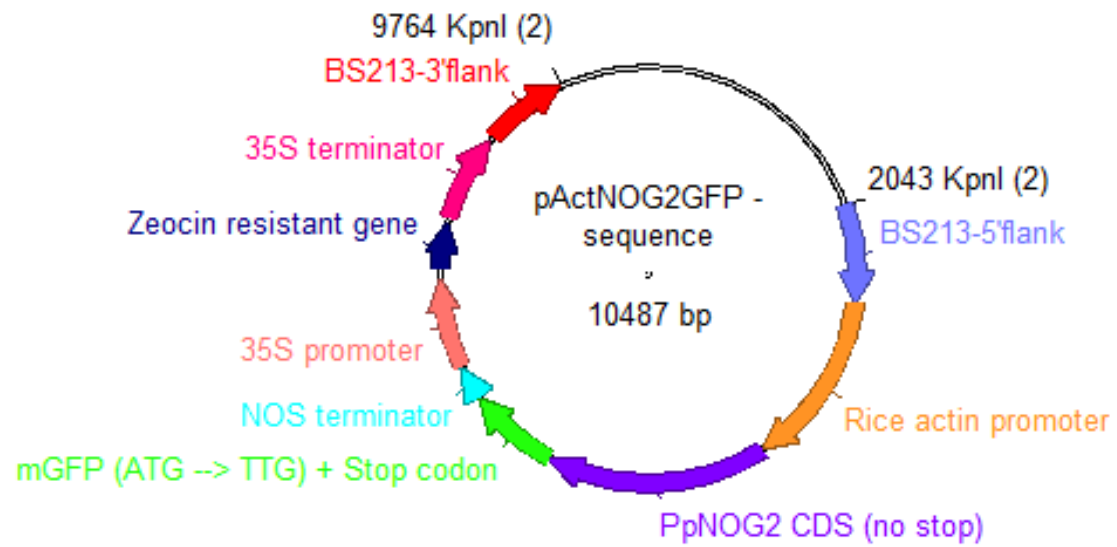


Figure S15: The *pACT::PpNOG2-GFP* construct

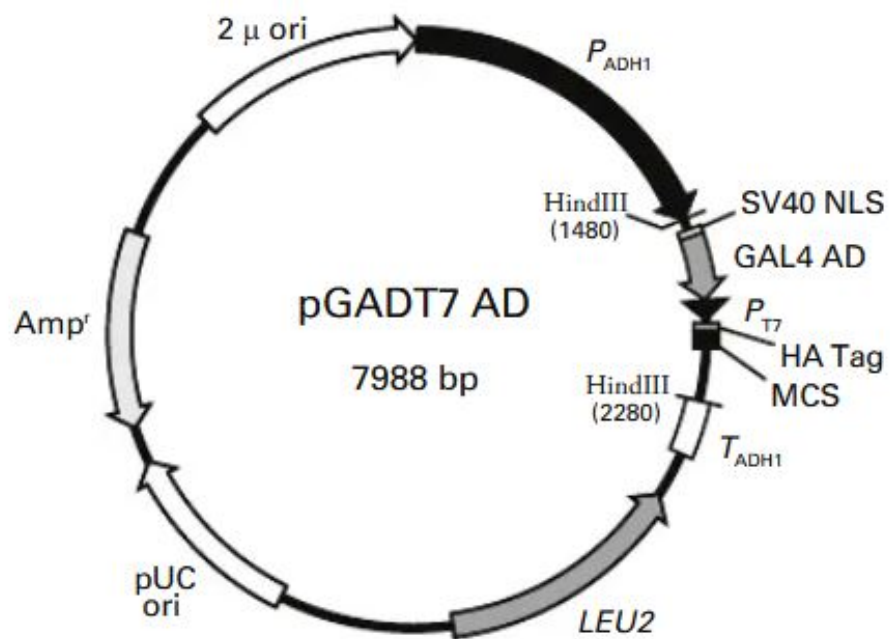


Figure S16: The pGADT7 AD 'prey' yeast two-hybrid vector (Takara Bio; 630442)

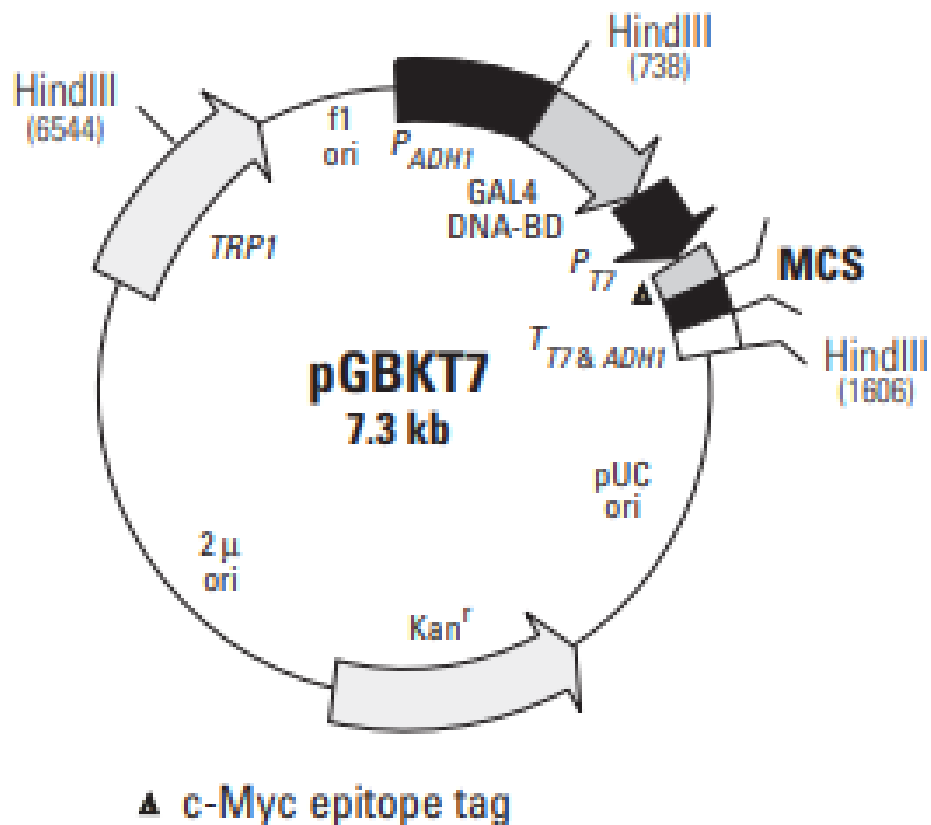


Figure S17: The pGBKT7 ‘bait’ yeast two-hybrid vector (Takara Bio; 630443)

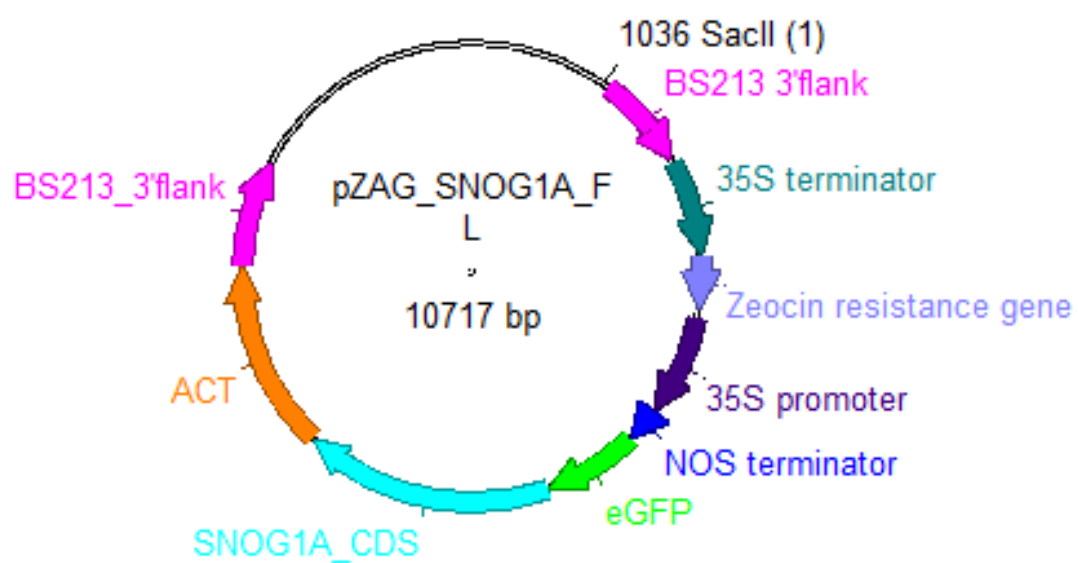


Figure S18: The *pACT::Pp3c8_19720-GFP* construct

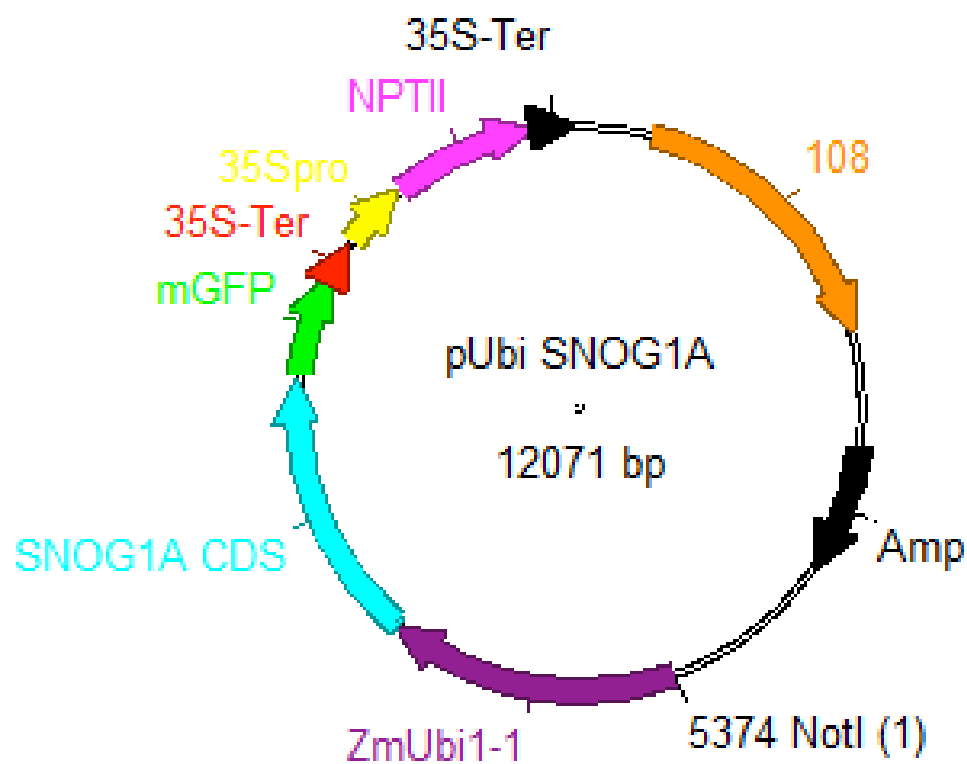


Figure S19: The *pUBI::Pp3c8_19720-GFP* construct

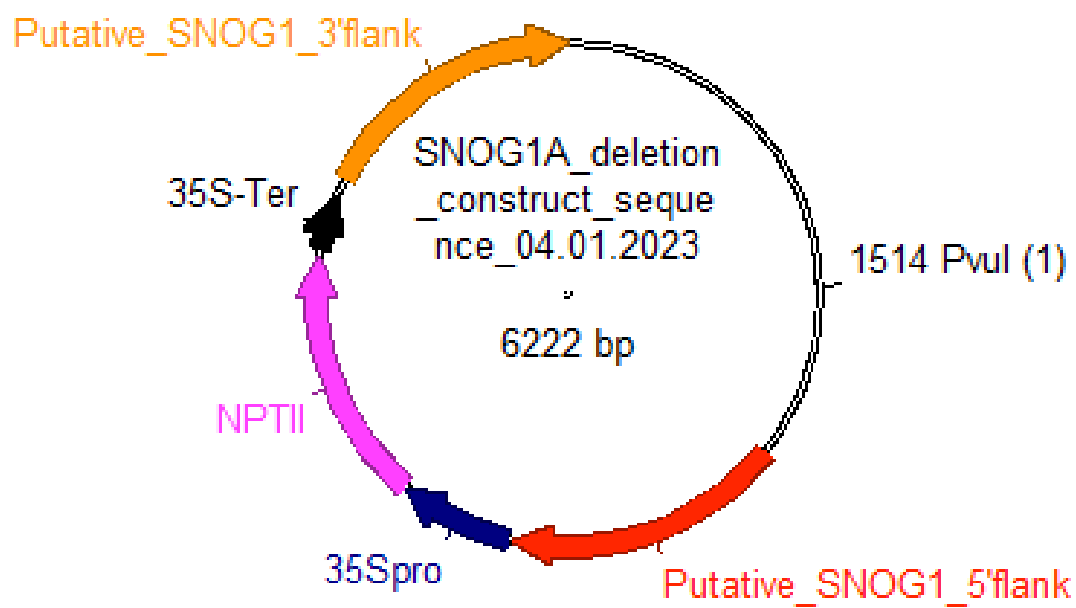


Figure S20: The *Pp3c8_19720* deletion construct

```
#####

#Example code for RNA-seq analysis

#####

#Raw read quality filtering with Fastp (Linux command line)

${samp} = sample ID

fastp \
    -i ${samp}_1.fq.gz \
    -I ${samp}_2.fq.gz \
    -o ${samp}_1_f.fq.gz \
    -O ${samp}_2_f.fq.gz \
    --failed_out ${samp}_failed -h ${samp}_report.html

#####

#Read mapping and quantification with Salmon (Linux command line)

${samp} = sample ID

salmon quant -i Ppatens_index -l A \
    -1 ${samp}_1_f.fq.gz \
    -2 ${samp}_2_f.fq.gz \
    -p 8 --validateMappings -o quants/${samp}_quant

#####

#Differential expression analysis (R)

library(tximport)
library(biomaRt)
library(GenomicFeatures)
library(AnnotationDbi)
library(tidyverse)
library(DESeq2)
library(clusterProfiler)
library(ggrepel)
library(org.At.tair.db)
```

```

##Making TxDb for P.patens

mart <- useMart(biomart="plants_mart", host="plants.ensembl.org")
datasets <- listDatasets(mart)
datasets[ , 1:2]
ppatens_txdb <- makeTxDbFromBiomart(biomart="plants_mart",
dataset="ppatens_eg_gene",
host="plants.ensembl.org")
saveDb(ppatens_txdb, 'ppatens_txdb.PhyPav3.sqlite')

##Creating Tx2gene data.frame

ppatens_txdb = loadDb(file = 'ppatens_txdb.PhyPav3.sqlite')
genes(ppatens_txdb)
k <- keys(ppatens_txdb, keytype = "TXNAME")
tx2gene <- select(ppatens_txdb, k, "GENEID", "TXNAME")

##Converting to matrix with tximport

dir <- "path/to/quant/files"
samples <- c("sample_ID_1", "sample_ID_2", "sample_ID_3", etc.)
files <- file.path(dir, samples, "quant.sf")
names(files) <- c("sample_ID_1", "sample_ID_2", "sample_ID_3", etc.)
all(file.exists(files))#should be yes
txi <- tximport(files, type = "salmon", tx2gene = tx2gene)
names(txi)
count <- txi$counts %>%
  round() %>%
  data.frame()
write.csv(count, "ppatens_counts.csv", row.names = T)
#raw count data (un-nomalised)

#metadata preparation
metadata <- data.frame( "sampleID" = c("sample_ID_1", "sample_ID_2",
"sample_ID_3", etc.), "condition" = factor(c("WT", "WT", "WT", etc)))
rownames(metadata) <- metadata$sampleID

```

```

#check metadata matches the count data
all(colnames(count) %in% rownames(metadata))
all(colnames(count) == rownames(metadata))

#Normalised count data using DESeq
#Experimental design - differences between condition (line)
dds <- DESeqDataSetFromMatrix(count, colData = metadata,
design = ~ condition)
dds <- estimateSizeFactors(dds)
sizeFactors(dds)
normalized_counts <- counts(dds, normalized=TRUE)
dir.create("output", showWarnings=F)
write.table(normalized_counts, file="output/normalized_counts.txt",
sep="\t", quote=F, col.names=NA)

#differential expression analysis with DESeq2
dds <- DESeqDataSetFromMatrix(count, colData = metadata,
design = ~ condition)
dds <- DESeq(dds)

#Contrasts
#Repeat for as many contrasts as needed
#In the contrast the base level condition should be put last.
contrast1 <- c("condition", "mutant", "WT")
res_WTvsmutant = results(dds, contrast = contrast1, alpha=0.05)
#unshrunk data

#Shrink LFC to make it more accurate for genes with
#low counts or high dispersion
res_WTvsmutant_shrunk <- lfcShrink(dds, contrast = contrast1,
type = "normal")
res_WTvsmutant_tb <- res_WTvsmutant_shrunk %>%
  data.frame() %>%
  rownames_to_column(var="gene") %>%
  as_tibble()
write.csv(res_WTvsmutant_tb, "output/WTvsmutant_result.csv")

```

```

#Significant differentially expressed genes filter adjusted p value
sig1 <- res_WTvsmutant_tb %>%
  filter(padj < 0.05)

#### map PPatens gene names to At_ID
#Pp transcriptome blasted against At protein database
#to get top match to provide annotation information
#Clusterprofiler only uses A.thaliana database, not Pp
Arabidopsis_ID <- read.table(file = "C:/Users/ball4695/
OneDrive-Nexus365/Documents/DPhil_project/Transcriptome/WT_nog1_BAP/
data_analysis/annotation/blastx.tsv", sep = '\t', header = FALSE)

Pp_ID_WTvsmutant <- res_WTvsmutant_tb$gene #extract the Pp gene ID
At_ID_WTvsmutant <- c() #Empty vector
gene_num_WTvsmutant <- length(Pp_ID_WTvsmutant) #no. of Pp genes

for (i in 1:gene_num_WTvsmutant){
  match_ID <- Arabidopsis_ID[Arabidopsis_ID$V1 %like%
    paste(Pp_ID_WTvsmutant[i], "V3", sep = ""), ]
  At_match <- as.character(match_ID[1,2])
  At_ID_WTvsmutant <- c(At_ID_WTvsmutant,
    strsplit(At_match, "\\\.")[[1]][1])
}

res_WTvsmutant_tb$AtGeneID <- At_ID_WTvsmutant
geneList_WTvsmutant <- res_WTvsmutant_tb$log2FoldChange
names(geneList_WTvsmutant) <- res_WTvsmutant_tb$AtGeneID
geneList_WTvsmutant <- na.omit(geneList_WTvsmutant)
#Sort list in decreasing order of log2fold change
geneList_WTvsmutant = sort(geneList_WTvsmutant, decreasing = T)
##GSEA with GO
#preparing gseGO object
gse_go_WTvsmutant <- gseGO(geneList = geneList_WTvsmutant,
  ont = "BP",
  keyType = "TAIR",
  minGSSize = 5,
  maxGSSize = 800,

```



```
pvalueCutoff = 0.05,  
verbose = T,  
OrgDb = org.At.tair.db,  
pAdjustMethod = "BH")  
dotplot(gse_go_WTvsmutant, showCategory=8, split=".sign") +  
facet_grid(~.sign)
```

Chapter 3: Genetic manipulation of the *Ppnog1R* mutant allows further characterisation of the function of PpNOG1

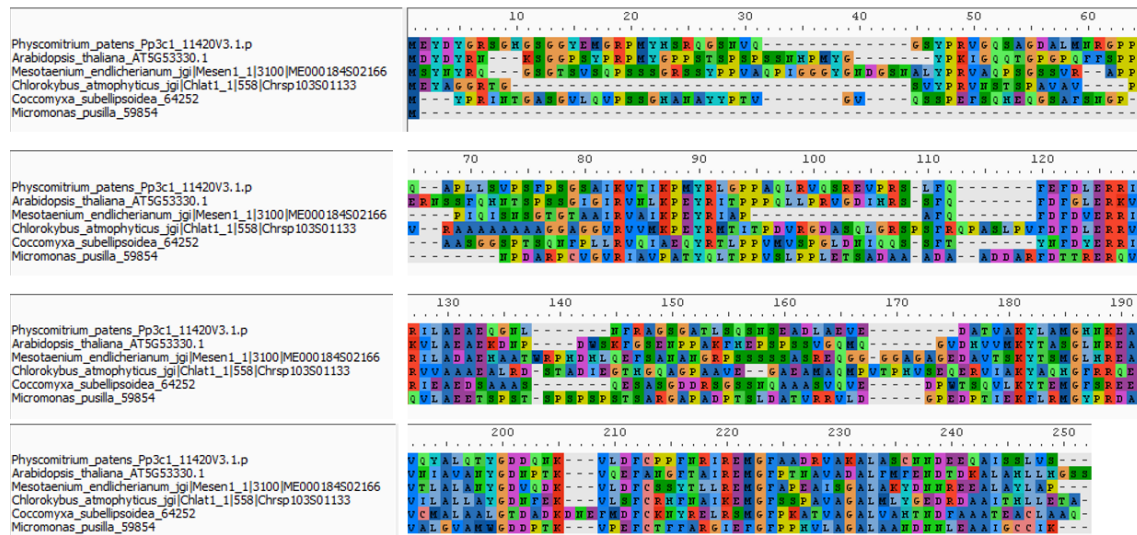


Figure S21: Multiple sequence alignment of algal NOG1-like protein sequences with PpNOG1 and AtNOG1

A.thaliana_NOG1.p	MDYDYRNK---SGGPSYPRPMYGGPPSTSPSPSSNHPMYGYPKIGQQTGPGPQFFSPPERN	57
P.patens_NOG1.p	MEYDYGRSGHSGGYEMGRPMYHSRQGS-----NVQGSYPRVQGSAGDALMNRGPPQA-	53
	*:*** .. *** . ***** . * : .***:***:*. . .***:	
A.thaliana_NOG1.p	SSFQHTSPSSGIGIRVNLKPEYRITPPPQLLRVGDHRRSSFQDFGLERKVLAEAEKD	117
P.patens_NOG1.p	-PLLSVPSPSGSAIKVTIKPMYRLGPPAQLRVQSRVPSLRFQFEFDLERRILAEAEQG	112
	: * ** .***:*** ** : : ** ***:*.***:*****.	
A.thaliana_NOG1.p	NPDWSKFGSENPPAKFHEPSPSSVGQMVGVDHVVVMKYTAGSLNREAVNIIVANYGDNPTK	177
P.patens_NOG1.p	NLNFRAGSGATLS----QSNSEADLAIVEDATVAKYLAMGHNKEAVQYALQTYGDDQNK	167
	* : : .. . * *... : * . * * * * * :***:*. .***: .*	
A.thaliana_NOG1.p	VQEFANGFTAIREMGFPTNAVADALFMFENDTDKALAHLLHGSS*	221
P.patens_NOG1.p	VLDFCPPFNRIREMGFAADRVAKALASCNDEEQAISSLVS*---	208
	* :*. *. ***** : : **.* ** :** :*: : *	

Figure S22: protein alignment of PpNOG1 and AtNOG1

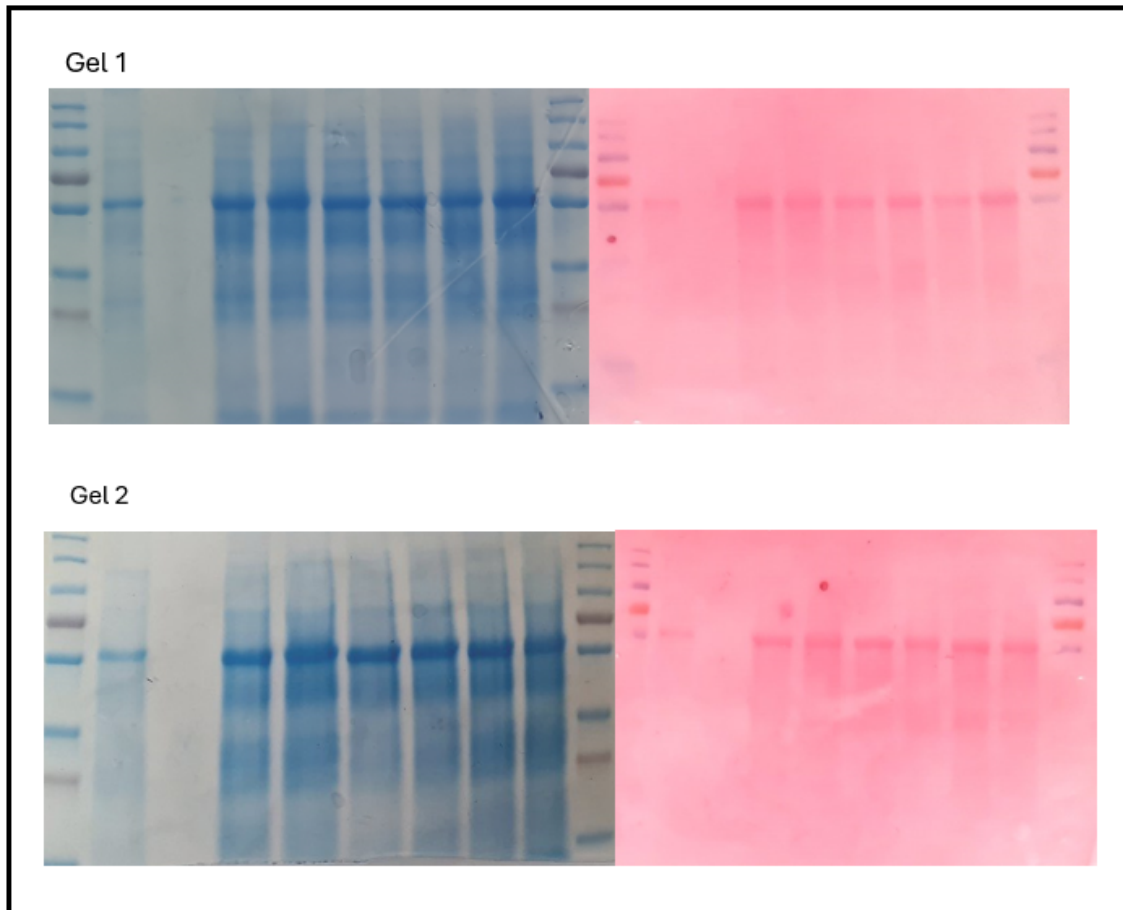


Figure S23: Controls for western blot to detect PpNOG1-GFP. Gel loading control shown with Coomassie blue stain of the SDS-PAGE gel. Effective protein transfer onto the membrane shown with Ponceau S stain.

Chapter 4: A suppressor screen reveals a new regulator of the transition to 3D growth

```

nog1dis      --NNNNNNNNNNNNNNNTANNNNNTTGANNNNNTATTACTCTAATAAATAAATTTATATA
NOG1_pro     CTCTTTATTATGCCTAAATAAGTTGATTTAACATATTACTCTAATAAATAAATTTATATA

                *      *      *****

nog1dis      TACATACACAAAATGCTTGTATAAATATATATATACATATATACATATATAGATCTTA
NOG1_pro     TACATATACAAAATGCTTGTATAAATATAT--ATATATATATATACATATATAGATCTTA

***** ***** ***** ***** *****

nog1dis      CCCTTGCCAAAATGTAGATGACGAAGCATAGTGTGATGTCCACCCTCAGCATATCCGATG
NOG1_pro     CCCTTGCCAAAATGTAGATGACGAAGCATAGTGTGATGTCCACCCTCAGCATATCAGATG

***** *****

nog1dis      TGAGCTTACCAAGTGTGTGTGAAGTCGCAATTCATAATAAATAAATAAAGAAGAAT
NOG1_pro     TGAGCTTACCAAGTGTGTGTGAAGTCGCAATTCATAATAAATAAATAAAGAAGAAT

*****

Nog1dis      TCCAAATTA AAACTTA AACTCGCGGATGGAAATGATTCCATC-----ACCGTTTCTAGGA
NOG1_pro     TCCAAATTA AAACTTA AACTCGCGGATGGAAATGATTCCATCACCAGGACCGTTTCTAGGA

*****

Nog1dis      ACCTTCTTTAAGGCAGTGGTCGGCGAGCGACCTCCAAAGCAAAGCTGATGTACCTATTGA
NOG1_pro     ACCTTCTTTAAGGCAGTGGTCGGCGAGCGACCTCCAAAGCAAAGCTGATGTACCTATTGA

*****

Nog1dis      CCCCCCCCCCCCCCCCCCCCCCNNTTTTTTTTTTNNNNAAATTTAAAATTGAACCCC
NOG1_pro     CCCCCCCCCCCCCCCCCCAACCTCCTTGTCTATGTATGTGAACCTGAATCAGGATCCTC

***** ** * * * * * * * * * *

```

Figure S24: A portion of the *PpNOG1* promoter region that was sequenced from *nog1dis* genomic DNA. A 6 bp deletion that is present in the *nog1dis* sequence is highlighted in green.

```

Pp3c8_19720_genomic      AGCAATCACAGGTTCAAATGCAAGGCCTACCGTCTCATCCTCAGCATCAGCAATTGCATC 2100
nog1dis                  -----NNNNNNNNNN 10
ReCh                     -----NNNNN 5
snog1a                   -----NNNNN 5

Pp3c8_19720_genomic      CGCAAGTCCAGGTTGCGACATCCCTGCATCAACAACCGCAAGGACAGAACCCCTCCACTTC 2160
nog1dis                  CCGNAGTCCNGGTTGCGACATCCCTGCATCAACAACCGCAAGGACAGAACCCCTCCACTTC 70
ReCh                     NTCGNAGNNNNNGTTGCGACATCCCTGCATCAACAACCGCAAGGACAGAACCCCTCCACTTC 65
snog1a                   NTCGNAGNNNNNGTTGCGACATCCCTGCATCAACAACCGCAAGGACAGAACCCCTCCACTTC 65
                        *****

Pp3c8_19720_genomic      CTCAGCAGCAGATGCAAAATGCAAGGCATGCACCCGCCCTCCTCAGCATCAACAACCTGCAGA 2220
nog1dis                  CTCAGCAGCAGATGCAAAATGCAAGGCATGCACCCGCCCTCCTCAGCATCAACAACCTGCAGA 130
ReCh                     CTCAGCAGCAGATGCAAAATGCAAGGCATGCACCCGCCCTCCTCAGCATCAACAACCTGCAGA 125
snog1a                   CTCAGCAGCAGATGCAAAATGCAAGGCATGCACCCGCCCTCCTCAGCATCAACAACCTGCAGA 125
                        *****

Pp3c8_19720_genomic      TTCAAGGTACTCCTCAGGCCCCACAGGTGCCGCCGCAAAATGCTCAGTCCCAGCCTCCTC 2280
nog1dis                  TTCAAGGTACTCCTCAGGCCCCACAGGTGCCGCCGCAAAATGCTCAGTCCCAGCCTCCTC 190
ReCh                     TTCAAGGTACTCCTCAGGCCCCACAGGTGCCGCCGCAAAATGCTCAGTCCCAGCCTCCTC 185
snog1a                   TTCAAGGTACTCCTCAGGCCCCACAGGTGCCGCCGCAAAATGCTCAGTCCCAGCCTCCTC 185
                        *****

Pp3c8_19720_genomic      CGCAACCCCATCTTTCATCACCACCAATTACAGCTGCAGGCGCCAACTGTAAATCAGCCGG 2340
nog1dis                  CGCAACCCCATCTTTCATCACCACCAATTACAGCTGCAGGCGCCAACTGTAAATCAGCCGG 250
ReCh                     CGCAACCCCATCTTTCATCACCACCAATTACAGCTGCAGGCGCCAACTGTAAATCAGCCGG 245|
snog1a                   CGCAACCCCATCTTTCATCACCACCAATTACAGCTGCAGGCGCCAACTGTAAATCAGCCGG 245
                        *****

Pp3c8_19720_genomic      ATTCTCAGTACCAAACCTCAGCAAGCGCCACCGGTTGCTTCTTCACATTCTTCGCAAGTTC 2400
nog1dis                  ATTCTCAGTACCAAACCTCAGCAAGCGCCACCGGTTGCTTCTTCACATTCTTCGCAAGTTC 310
ReCh                     ATTCTCAGTACCAAACCTCAGCAAGCGCCACCGGTTGCTTCTTCACATTCTTCGCAAGTTC 305
snog1a                   ATTCTCAGTACCAAACCTCAGCAAGCGCCACCGGTTGCTTCTTCACATTCTTCGCAAGTTC 305
                        *****

Pp3c8_19720_genomic      CATCTTATTATGCCCGAGCAACAGCAACTGCAGCCTGGACAAACAGGCCCAACTCCAGCCA 2460
nog1dis                  CATCTTATTATGCCCGAGCAACAGCAACTGCAGCCTGGACAAACAGGCCCAACTCCAGCCA 370
ReCh                     CATCTTATTATGCCCGAGCAACAGCAACTGCAGCCTGGACAAACAGGCCCAACTCCAGCCA 365
snog1a                   CATCTTATTATGCCCGAGCAACAGCAACTGCAGCCTGGACAAACAGGCCCAACTCCAGCCA 365
                        *****

Pp3c8_19720_genomic      TGGCGCAACATTATGGCAATCTCAACCAAGACGTTAATCCCTATGGAACAGCACAGGGTA 2520
nog1dis                  TGGCGCAACATTATGGCAATCTCAACCAAGACGTTAATCCCTATGGAACAGCACAGGGTA 430
ReCh                     TGGCGCAACATTATGGCAATCTCAACCAAGACGTTAATCCCTATGGAACAGCACAGGGTA 425
snog1a                   TGGCGCAACATTATGGCAATCTCAACCAAGACGTTAATCCCTATGGAACAGCACAGGGTA 425
                        *****

Pp3c8_19720_genomic      GCTACAATCAAGAGGTCCCTCCTCCACAGGTGCCTGGAGGCTACGAGCTTCGTTTCTTGC 2580
nog1dis                  GCTACAATCAAGAGGTCCCTCCTCCACAA----- 460
ReCh                     GCTACAATCAAGAGGTCCCTNNNNNNNNN----- 454
snog1a                   GCTACAATCAAGAGGTCCCTTCTCCCAA----- 456
                        *****

```

Figure S25: Sequencing of a region of *Pp3c8_19720* from *snog1a*, *nog1dis* and ReCh genomic DNA with *Pp3c8_19720* int F/ *Pp3c8_19720* int R to confirm mutation. Each SNP that produced a nonsense mutation has been highlighted in green

Pp3c8_19720V3.1_CDS	ATGGATCATGTGGGATCCTCTGTCGAGCGCAACCAATATGCGCAGGGCGAGCCTTTCTTT	60
WT_Pp3c8_19720_F	-----NNNNNNNNNCCANATGCGCAGGGCGAGCCTTTCTTT	35
	* *	
Pp3c8_19720V3.1_CDS	GACTTGTGTGCAACAAACCCAAATTCGGGTACTCTGGATCGCAGAAATCTCCGCTGTG	120
WT_Pp3c8_19720_F	GACTTGTGTGCAACAAACCCAAATTCGGGTACTCTGGATCGCAGAAATCTCCGCTGTG	95

Pp3c8_19720V3.1_CDS	TATTTTGGCGATGATGACGTCTTCCGAGCATGGACTTCCAGCCGATGATGGACACAGGC	180
WT_Pp3c8_19720_F	TATTTTGGCGATGATGACGTCTTCCGAGCATGGACTTCCAGCCGATGATGGACACAGGC	155

Pp3c8_19720V3.1_CDS	GTCAGGGGGAAGGTGGATGATAGCAAGCGTGGGTGGCCGTCTCAACATTCAACACTTCA	240
WT_Pp3c8_19720_F	GTCAGGGGGAAGGTGGATGATAGCAAGCGTGGGTGGCCGTCTCAACATTCAACACTTCA	215

Pp3c8_19720V3.1_CDS	AATTCTTGAATCAGCCCCATCTCATGAATTGAATGTGCGCCCCGACAAGGATGCTGTT	300
WT_Pp3c8_19720_F	AATTCTTGAATCAGCCCCATCTCATGAATTGAATGTGCGCCCCGACAAGGATGCTGTT	275

Pp3c8_19720V3.1_CDS	GAGACTGCTACAGTGGCAGCCGTGGAAACGCACAATGAAGAAGTATGCGGACAACTCTGTTG	360
WT_Pp3c8_19720_F	GAGACTGCTACAGTGGCAGCCGTGGAAACGCACAATGAAGAAGTATGCGGACAACTCTGTTG	335

Pp3c8_19720V3.1_CDS	CGGGTTTTGGAGGGCATGAGTGGCGGACTGACTCAGTTAGAAACCAACAGTCTTCGATG	420
WT_Pp3c8_19720_F	CGGGTTTTGGAGGGCATGAGTGGCGGACTGACTCAGTTAGAAACCAACAGTCTTCGATG	395

Pp3c8_19720V3.1_CDS	GAGCACTCAGTCGACAAAATTAAGGTTGATGCACTGACAACTTTGGTGCAACTGATGGA	480
WT_Pp3c8_19720_F	GAGCACTCAGTCGACAAAATTAAGGTTGATGCACTGACAACTTTGGTGCAACTGATGGA	455

Pp3c8_19720V3.1_CDS	AAGTTGAGATCGTTGGAAAATATGATTGAGAGGTTTCCGCGAAGTGTGCAAGTGATTCGT	540
WT_Pp3c8_19720_F	AAGTTGAGATCGTTGGAAAATATGATTGAGAGGTTTCCGCGAAGTGTGCAAGTGATTCGT	515

Pp3c8_19720V3.1_CDS	GATAAGCAGGAATTTGCAAGAAGACAGATGAAGCTTGA AAAA ACTTCAGCAAGTGCTCAA	600
WT_Pp3c8_19720_F	GATAAGCAGGAATTTGCAAGAAGACAGATGAAGCTTGA AAAA ACTTCAGCAAGTGCTCAA	575

Pp3c8_19720V3.1_CDS	CGCGAGAAATCTTCATCTGTTACTTCAACTCCCACTGTGCGAGCCGTGCAATCACCCTC	660
WT_Pp3c8_19720_F	CGCGAGAAATCTTCATCTGTTACTTCAACTCCCACTGTGCGAGCCGTGCAATCACCCTC	635

Pp3c8_19720V3.1_CDS	GCTCTTCCACCTGCAACGTGCGCTGTGAAATCTGTGCGAGGCTCCTCTGCGCAACACGCC	720
WT_Pp3c8_19720_F	GCTCTTCCACCTGCAACGTGCGCTGTGAAATCTGTGCGAGGCTCCTCTGCGCAACACGCC	695

Pp3c8_19720V3.1_CDS	CCAGTTCTACAGCAGCAGTCTCAGCAAGTGGGCGCAGCTTCCCGCGCCAGCCCTCAGATG	780
WT_Pp3c8_19720_F	CCAGTTCTACAGCAGCAGTCTCAGCAAGTGGGCGCAGCTTCCCGCGCCAGCCCTCANATG	755

Pp3c8_19720V3.1_CDS	CAGATGCCTGGTATGGCACCCTCAGCAGCAATCAGAGGTTCAAATGCAAGGCCTACCG	840
WT_Pp3c8_19720_F	CAGATGCCTGGTATGGCACCCTCAGCAGCAATCAGAGGTTCAAATGCAAGGCCTACCG	815

Pp3c8_19720V3.1_CDS	TCTCATCCTCAGCATCAGCAATTGCATCCGCAAGTCCAGGTTGCGACATCCCTGCATCAA	900
WT_Pp3c8_19720_F	TCTCATCCTCAGCATCAGCAATTGCATCCGCAAGTCCAGGTTGCGACATCCCTGCATCAA	875

Pp3c8_19720V3.1_CDS	CAACCGCAAGGACAGAACCTCCACTTCTCAGCAGCAGATGCAAAATGCAAGGCATGCAC	960
WT_Pp3c8_19720_F	CAACCGCAAGGACAGAACCTCCACTTCTCAGCAGCAGATGCAAAATGCAAGGCATGCAC	935

Pp3c8_19720V3.1_CDS	CCGCTCTCTCAGCATCAACAACTGCAGATTCAAGGTA CTCTC AGGCCCCACAGGTG	1017
WT_Pp3c8_19720_F	CCGCTCTCTCAGCATCAACAACTGCAGATTCAAGGTA CTCTC AGGCCCCACAGGTG	269

Pp3c8_19720V3.1_CDS	CCGCGCAAAAATGCTCAGTCCAGGCTCTCCGCAACCCCATCTTCATCACCACAAATTA	1077
WT_Pp3c8_19720_R	CCGCGCAAAAATGCTCAGTCCAGGCTCTCCGCAACCCCATCTTCATCACCACAAATTA	329

Pp3c8_19720V3.1_CDS	CAGCTGCAGGCGCCAACTGTAAATCAGCCGGATTCTCAGTACCAAACTCAGCAAGCGCCA	1137
WT_Pp3c8_19720_R	CAGCTGCAGGCGCCAACTGTAAATCAGCCGGATTCTCAGTACCAAACTCAGCAAGCGCCA	389

Pp3c8_19720V3.1_CDS	CCGGTTGCTTCTTCACATTCTCGCAAGTTCCATCTTATTATGCCAGCAACAGCAACTG	1197
WT_Pp3c8_19720_R	CCGGTTGCTTCTTCACATTCTCGCAAGTTCCATCTTATTATGCCAGCAACAGCAACTG	449

Pp3c8_19720V3.1_CDS	CAGCTGGACAAACAGGCCCAACTCCAGCCATGGCGCAACATTATGGCAATCTCAACCAA	1257
WT_Pp3c8_19720_R	CAGCTGGACAAACAGGCCCAACTCCAGCCATGGCGCAACATTATGGCAATCTCAACCAA	509

Pp3c8_19720V3.1_CDS CGCGAGAAATCTTCATCTGTTACTTCAACTCCCACTGTGAGACCGTCGAATCACCCTTC 660
 snog1a_Fp3c8_19720_F CGCGAGAAATCTTCATCTGTTACTTCAACTCCCACTGTGAGACCGTCGAATCACCCTTC 636

Pp3c8_19720V3.1_CDS GCTCTTCCACCTGCAACGTCGGCTGTGAAATCTGTCGAGGCTCCTCCTGCGCAACCAAGCC 720
 snog1a_Fp3c8_19720_F GCTCTTCCACCTGCAACGTCGGCTGTGAAATCTGTCGAGGCTCCTCCTGCGCAACCAAGCC 696

Pp3c8_19720V3.1_CDS CCAGTTCTACAGCAGCAGTCTCAGCAAGTGGGCCAGCTTCCCGCGCCAGCCCTCAGATG 780
 snog1a_Fp3c8_19720_F CCAGTTCTACAGCAGCAGTCTCAGCAAGTGGGCCAGCTTCCCGCGCCAGCCCTCAGATG 756

Pp3c8_19720V3.1_CDS CAGATGCCTGGTATGGCACCCTCAGCAGCAATCAGAGGTTCAAATGCAAGGCCATACCG 840
 snog1a_Fp3c8_19720_F CAGATGCCTGGTATGGCACCCTCAGCAGCAATCAGAGGTTCAAATGCAAGGCCATACCG 816

Pp3c8_19720V3.1_CDS TCTCATCCTCAGCATCAGCAATTGCAATCCGCAAGTCCAGGTTGCGACATCCCTGCATCAA 900
 snog1a_Fp3c8_19720_F TCTCATCCTCAGCATCAGCAATTGCAATCCGCAAGTCCAGGTTGCGACATCCCTGCATCAA 876

Pp3c8_19720V3.1_CDS CAACCGCAAGGACAGAACCTCCACTTCCCTCAGCAGCAGATGCAATGCAAGGCATGCAC 960
 snog1a_Fp3c8_19720_F CAACCGCAAGGACAGAACCTCCACTTCCCTCAGCAGCAGATGCAATGCAAGGCATGCAC 936

Pp3c8_19720V3.1_CDS CCGCCTCCTCAGCATCAACAATGCAAGTCAAGGTAAGGCTCCTCAGGCCCCAC AGGT 1016
 snog1a_Fp3c8_19720_F CCGCCTCCTCAGCATCAACAATGCAAGTCAAGGTAAGGCTCCTCAGGCCCCAC AGGT 283

Pp3c8_19720V3.1_CDS GCCGCCGCAAAATGCTCAGTCCCAGCCTCCTCCGCAACCCCATCTTCATCACCAACAATT 1076
 snog1a_Fp3c8_19720_R GCCGCCGCAAAATGCTCAGTCCCAGCCTCCTCCGCAACCCCATCTTCATCACCAACAATT 343

Pp3c8_19720V3.1_CDS ACAGCTGCAGGCGCCCACTGTAAATCAGCCGGATTCTCAGTACCAAACTCAGCAAGCGCC 1136
 snog1a_Fp3c8_19720_R ACAGCTGCAGGCGCCCACTGTAAATCAGCCGGATTCTCAGTACCAAACTCAGCAAGCGCC 403

Pp3c8_19720V3.1_CDS ACCGGTTGCTTCTTCACATTCTCGCAAGTCCATCTTATTATGCCAGCAACAGCAACT 1196
 snog1a_Fp3c8_19720_R ACCGGTTGCTTCTTCACATTCTCGCAAGTCCATCTTATTATGCCAGCAACAGCAACT 463

Pp3c8_19720V3.1_CDS GCAGCCTGGACAAACAGGCCCAACTCCAGCCATGGCGCAACATTATGGCAATCTCAACCA 1256
 snog1a_Fp3c8_19720_R GCAGCCTGGACAAACAGGCCCAACTCCAGCCATGGCGCAACATTATGGCAATCTCAACCA 523

Pp3c8_19720V3.1_CDS AGACGTTAATCCCTATGGAACAGCACAGGCTAGCTACAATCAAGAGGTCCCTCCTCCACA 1316
 snog1a_Fp3c8_19720_R AGACGTTAATCCCTATGGAACAGCACAGGCTAGCTACAATCAAGAGGTCCCTCCTCCACA 583

Pp3c8_19720V3.1_CDS GGTGCTGGAGGCTACGAGCTTCGTTTCTTGCCACTGCCTGCACCACTCCTGTAATGGG 1376
 snog1a_Fp3c8_19720_R GGTGCTGGAGGCTACGAGCTTCGTTTCTTGCCACTGCCTGCACCACTCCTGTAATGGG 643

Pp3c8_19720V3.1_CDS GCAAAATGCCTCAGCAAGTGGCCAGCCAGCATCAGCAACATGGACAGAGCGGAGGTCA 1436
 snog1a_Fp3c8_19720_R GCAAAATGCCTCAGCAAGTGGCCAGCCAGCATCAGCAACATGGACAGAGCGGAGGTCA 703

Pp3c8_19720V3.1_CDS GATGTTTGAGCCTCCCCGTCATGCTGTTGGTCCACCTCCACAAGGACCTTATCATCCTCA 1496
 snog1a_Fp3c8_19720_R GATGTTTGAGCCTCCCCGTCATGCTGTTGGTCCACCTCCACAAGGACCTTATCATCCTCA 763

Pp3c8_19720V3.1_CDS AGCCCATCCTTCTCAAGGCCGGGATGGACTCCAAGTATATAACCAAGTATGTCCCCCCCCA 1556
 snog1a_Fp3c8_19720_R AGCCCATCCTTCTCAAGGCCGGGATGGACTCCAAGTATATAACCAAGTATGTCCCCCCCCA 823

Pp3c8_19720V3.1_CDS ATACCGGATGGCCAGCCTGGTAGTCCCTCTCCTGGAGGTTATCCGCACTACCCACAGC 1616
 snog1a_Fp3c8_19720_R ATACCGGATGGCCAGCCTGGTAGTCCCTCTCCTGGAGGTTATCCGCACTACCCACAGC 883

Pp3c8_19720V3.1_CDS CCGGCCGGTTCAACCGTCGATGGGGTCGCAAGGAGGACCAACTCCAGTGGCTGCCAACCG 1676
 snog1a_Fp3c8_19720_R CCGGCCGGTTCAACCGTCGATGGGGTCGCAAGGAGGACCAACTCCAGTGGCTGCCAACCG 943

Pp3c8_19720V3.1_CDS CGTCCCAGTTGACGAAGTGATTGATAAAGTTGCAAGCATGGGGTTTGGGAAGAGATCAGGT 1736
 snog1a_Fp3c8_19720_R CGTCCCAGTTGACGAAGTGATTGATAAAGTTGCAAGCATGGGGTTTGGGAAGAGATCAGGT 1003

Pp3c8_19720V3.1_CDS TCGGGCCACTGTGAGAAATTAATGAAATGGTACATCTGTGGACTTAAACGTTGTCTCT 1796
 snog1a_Fp3c8_19720_R TCGGGCCACTGTGAGAAATTAATGAAATGGTACATCTGTGGACTTAAACGTTGTCTCT 1063

Pp3c8_19720V3.1_CDS TGACAAGTTGATGAATGGTGACTCTAGTGGCAGAAATGAGAGTCAGCAACCTGCCAAAGG 1856
 snog1a_Fp3c8_19720_R TGACAAGTTGATGAATGGTGACTCTAGTGGCAGAAATGAGAGTCAGCAACCTGCCAAAGG 1120

Pp3c8_19720V3.1_CDS CTGGTATGGCCGGTGA 1872
 snog1a_Fp3c8_19720_R ----- 1120

Figure S26: Sequencing of *Pp3c8_19720* transcript from cDNA extracted from wild type and *snog1a* using *Pp3c8_19720* F BamHI/ *Pp3c8_19720* R NO STOP XmaI primers. Nonsense mutations (premature STOP codons) highlighted in yellow. This showed no splicing defects in the *Pp3c8_19720* transcript from *snog1a*.

Pp3c8_19720_genomic
snog1b_frag1 GTTCAAGGTTGGCCGAGTTTGTGTAGGCAAGGTTGGTAGCGGGTTGACGTGGTGGATGCG 600
-----NNNNNNNNNNNNNNNG 16
*

Pp3c8_19720_genomic
snog1b_frag1 TGATGGATCATGTGGGATCCTCGTCGAGCGCGAACCAATATGCGCAGGGCGAGCCTTTCT 660
TGATGGATCATGTGGGATCCTCGTCGAGCGCGAACCAATATGCGCAGGGCGAGCCTTTCT 76

Pp3c8_19720_genomic
snog1b_frag1 TTGACTTGTGCAAAACAAACCCAAATTCGGTTACTCTGGATCGCAGAAATCTCCGCCCTG 720
TTGACTTGTGCAAAACAAACCCAAATTCGGTTACTCTGGATCGCAGAAATCTCCGCCCTG 136

Pp3c8_19720_genomic
snog1b_frag1 TGTATTTTGGCGATGATGACGTCCTTCCGAGCATGGACTTCCAGCCGATGATGGACACAG 780
TGTATTTTGGCGATGATGACGTCCTTCCGAGCATGGACTTCCAGCCGATGATGGACACAG 196

Pp3c8_19720_genomic
snog1b_frag1 GCGTCAAGGGGAAGGTGGATGATAGCAAGCGTGGGTGGCCGCTCTCAACATTCAAACACTT 840
GCGTCAAGGGGAAGGTGGATGATAGCAAGCGTGGGTGGCCGCTCTCAACATTCAAACACTT 256

Pp3c8_19720_genomic
snog1b_frag1 CAAATTCTGTACGTTTTTAGTCCTTCGTCGTTGTGTGAGTTTCAATTGTATTCTGTTGTT 900
CAAATTCTGTACGTTTTTAGTCCTTCGTCGTTGTGTGAGTTTCAATTGTATTCTGTTGTT 316

Pp3c8_19720_genomic
snog1b_frag1 CATACTTCCAAGTAATTGTGGTGGAGTTGCTGGCATGACGACCCCTAATTCAAGTCATCT 960
CATACTTCCAAGTAATTGTGGTGGAGTTGCTGGCATGACGACCCCTAATTCAAGTCATCT 376

Pp3c8_19720_genomic
snog1b_frag1 TCAAACATGAACCTTACGAGGAAACCAGGATAGGTTTTGGAGTGCTGCCACTTTGTACCAA 1020
TCAAACATGAACCTTACGAGGAAACCAGGATAGGTTTTGGAGTGCTGCCACTTTGTACCAA 436

Pp3c8_19720_genomic
snog1b_frag1 TGTGAGTTATGTAGTAGTGAGGAGGTAATGGTTTCAATGATGTTTTCCATGTTTGGCCGC 1080
TGTGAGTTATGTAGTAGTGAGGAGGTAATGGTTTCAATGATGTTTTCCATGTTTGGCCGC 496

Pp3c8_19720_genomic
snog1b_frag1 ATGATCGTGACATTGTGAGATGATGCTGATGGAGCAGTGGAAATCAGCCCCATCTCATGA 1140
ATGATCGTGACATTGTGAGATGATGCTGATGGAGCAGTGGAAATCAGCCCCATCTCATGA 556

Pp3c8_19720_genomic
snog1b_frag1 ATTGAATGTGCGCCCCGACAAGGATGCTGTTGAGACTGCTACAGTGGCAGCCGTGGAAACG 1200
ATTGAATGTGCGCCCCGACAAGGATGCTGTTGAGACTGCTACAGTGGCAGCCGTGGAAACG 616

Pp3c8_19720_genomic
snog1b_frag1 CACAATGAAGAAGTATGCGGACAATCTGTTGCGGGTTTTGGAGGGCATGAGTGGGCGACT 1260
CACAATGAAGAAGTATGCGGACAATCTGTTGCGGGTTTTGGAGGGCATGAGTGGGCGACT 676

Pp3c8_19720_genomic
snog1b_frag1 GACTCAGTTAGAAACCAACAGTCTTCGCATGGAGCACTCAGTCGACAAAATTAGGTTGA 1320
GACTCAGTTAGAAACCAACAGTCTTCGCATGGAGCACTCAGTCGACAAAATTAGGTTGA 736

Pp3c8_19720_genomic
snog1b_frag1 TGCAGCTGACAACTTTGGTGCAACTGATGGAAAGTTGAGATCGTTGAAAATATGATTTCG 1380
TGCAGCTGACAACTTTGGTGCAACTGATGGAAAGTTGAGATCGTTGAAAATATGATTTCG 796

Pp3c8_19720_genomic
snog1b_frag1 AGAGGTAATTTTTTCTTTCTTCATCGACCTTTGATGATTGGTGAGTGGTATTTGGAGGTT 1440
AGAGGTAATTTTTTCTTTCTTCATCGACCTTTGATGATTGGTGAGTGGTATTTGGAGGTT 856

Pp3c8_19720_genomic
snog1b_frag2 TTATAAATTAGTTTATGCAATCAACAGCTATTGTTTGGTTTTTGGATTTCGATCCGTTTA 1500
TTATAAATTAGTTTATGCAATCAACAGCTATTGTTTGGTTTTTGGATTTCGATCCGTTTA 148

Pp3c8_19720_genomic
snog1b_frag2 TTCGATTGCTATCAACAGTCAATGACCAACTTGGTCAAGTTTAGGCCATTTTCGGCGGA 1560
TTCGATTGCTATCAACAGTCAATGACCAACTTGGTCAAGTTTAGGCCATTTTCGGCGGA 208

Pp3c8_19720_genomic
snog1b_frag2 TTTTTTAGATGATACCACTCTGAACCATGAAATTTCAAATTTTCTTTCTATTCTCAGC 1620
TTTTTTAGATGATACCACTCTGAACCATGAAATTTCAAATTTTCTTTCTATTCTCAGC 268

Pp3c8_19720_genomic
snog1b_frag2 GTTGTGTCGTTGATTATCTTGTGGACAGAAGCTGTGATTGCACTTGACGTCATTCTATC 1680
GTTGTGTCGTTGATTATCTTGTGGACAGAAGCTGTGATTGCACTTGACGTCATTCTATC 328

Pp3c8_19720_genomic
snog1b_frag2 ATACAAATTATTATTTTTTGTGTTTTTGGAACTAAATTTAATGAATTTTATCGTTTTT 1740
ATACAAATTATTATTTTTTGTGTTTTTGGAACTAAATTTAATGAATTTTATCGTTTTT 388

Pp3c8_19720_genomic
snog1b_frag2 TGCAGGTTGAGCGAAGTGTGAGGTTGATTGATGATAAGCAGGAAATTCAGAGAAGACAGA 1800
TGCAGGTTGAGCGAAGTGTGAGGTTGATTGATGATAAGCAGGAAATTCAGAGAAGACAGA 448

Pp3c8_19720_genomic
snog1b_frag2 TGAAGCTTGAAAACTTCAGCAAAGTGCTCAACGCGAGAAATCTTCATCTGTTACTTCAA 1860
TGAAGCTTGAAAACTTCAGCAAAGTGCTCAACGCGAGAAATCTTCATCTGTTACTTCAA 508

Pp3c8_19720_genomic
snog1b_frag2 CTCCTACTGTGAGACCGTCAATACCCCTCGCTCTTCCACCTGCAACGTCGGCTGTGA 1920
CTCCTACTGTGAGACCGTCAATACCCCTCGCTCTTCCACCTGCAACGTCGGCTGTGA 568

Pp3c8_19720_genomic
snog1b_frag2 AATCTGTGAGGCTCCTCCTGCGCAACCGCCAGTTCTACAGCAGCAGTCTCAGCAAG 1980
AATCTGTGAGGCTCCTCCTGCGCAACCGCCAGTTCTACAGCAGCAGTCTCAGCAAG 628

Pp3c8_19720_genomic
snog1b_frag2 TGGGCCAGCTTCCCGCGCCAGCCCTCAGATGCAGATGCTGGTATGGCAGCCCTCAGC 2040
TGGGCCAGCTTCCCGCGCCAGCCCTCAGATGCAGATGCTGGTATGGCAGCCCTCAGC 688

Pp3c8_19720_genomic
snog1b_frag2 AGCAATCAGAGTTCAAATGCAAGGCTACCGTCTCATCTCAGCATCAGCAATTGCATC 2100
AGCAATCAGAGTTCAAATGCAAGGCTACCGTCTCATCTCAGCATCAGCAATTGCATC 748

Pp3c8_19720_genomic
snog1b_frag2 CGCAAGTCCAGGTTGCGACATCCCTGCATCAACAACCGCAAGGACAGAACCTCCACTTC 2160
CGCAAGTCCAGGTTGCGACATCCCTGCATCAACAACCGCAAGGACAGAACCTCCACTTC 808

Pp3c8_19720_genomic
snog1b_frag3 CTCAGCAGCAGATGCAAAATGCAAGGCTGCAACCGCCCTCCTCAGCATCAACAATGCAGA 2220
CTCAGCAGCAGATGCAAAATGCAAGGCTGCAACCGCCCTCCTCAGCATCAACAATGCAGA 80

Pp3c8_19720_genomic
snog1b_frag3 TTCAAGGTACTCCTCAGGCCCCACAGGTGCGCGCGCAAAATGCTCAGTCCCAGCCTCCTC 2280
TTCAAGGTACTCCTCAGGCCCCACAGGTGCGCGCGCAAAATGCTCAGTCCCAGCCTCCTC 140

Pp3c8_19720_genomic
snog1b_frag3 CGCAACCCCATCTTTCATCACCACAAATTACAGCTGCAGGCGCCAACTGTAAATCAGCCGG 2340
CGCAACCCCATCTTTCATCACCACAAATTACAGCTGCAGGCGCCAACTGTAAATCAGCCGG 200

Pp3c8_19720_genomic
snog1b_frag3 ATTCTCAGTACCAAATCAGCAAGCGCCACCGGTTGCTTCTTCACATTCCTCGCAAGTTC 2400
ATTCTCAGTACCAAATCAGCAAGCGCCACCGGTTGCTTCTTCACATTCCTCGCAAGTTC 260

Pp3c8_19720_genomic
snog1b_frag3 CATCTTATTATGCCAGCAACAGCAACTGCAGCCTGGACAAACAGGCCCACTCCAGCCA 2460
CATCTTATTATGCCAGCAACAGCAACTGCAGCCTGGACAAACAGGCCCACTCCAGCCA 320

Pp3c8_19720_genomic
snog1b_frag3 TGGCGCAACATTATGGCAATCTCAACCAAGACGTTAATCCCTATGGAACAGCACAGGGTA 2520
TGGCGCAACATTATGGCAATCTCAACCAAGACGTTAATCCCTATGGAACAGCACAGGGTA 380

Pp3c8_19720_genomic
snog1b_frag3 GCTACAATCAAGAGGTCCCTCCTCCACAGGTGCCTGGAGGCTACGAGCTTCGTTTCTTGC 2580
GCTACAATCAAGAGGTCCCTCCTCCACAGGTGCCTGGAGGCTACGAGCTTCGTTTCTTGC 440

Pp3c8_19720_genomic
snog1b_frag3 CACTGCCTGCACAGCTCCTGTAATGGGGCAAATGCCTCAGCAAGTGCCGACCAGCCAGC 2640
CACTGCCTGCACAGCTCCTGTAATGGGGCAAATGCCTCAGCAAGTGCCGACCAGCCAGC 500

Pp3c8_19720_genomic
snog1b_frag3 ATCAGCAACATGGACAGAGCGGAGGTCAGATGTTTGAGCCTCCCGTCATGCTGTTGGTC 2700
ATCAGCAACATGGACAGAGCGGAGGTCAGATGTTTGAGCCTCCCGTCATGCTGTTGGTC 560

Pp3c8_19720_genomic
snog1b_frag3 CACCTCCACAAGGACCTTATCATCTCAAGCCCATCCTTCTCAGGCGGGGATGGACTCC 2760
CACCTCCACAAGGACCTTATCATCTCAAGCCCATCCTTCTCAGGCGGGGATGGACTCC 620

Pp3c8_19720_genomic
snog1b_frag3 AAGTATATAACCAAGTATGTCCCCCCCCAATACCGGATGGCCAGCCTGGTAGTCCCTCTC 2820
AAGTATATAACCAAGTATGTCCCCCCCCAATACCGGATGGCCAGCCTGGTAGTCCCTCTC 680

```

Pp3c8_19720_genomic      CTGGAGGTTATCCGCAGCTACCCACAGCCCGGCCGGTTCAACCGTCGATGGGGTCGCAAG 2880
snog1b_frag3             CTGGAGGTTATCCGCAGCTACCCACAGCCCGGCCGGTTCAACCGTCGATGGGGTCGCAAG 740
*****

Pp3c8_19720_genomic      GAGGACCAACTCCAGTGGCTGCCAACCGCGTCCCAGTTGACGAAGTGATTGATAAAGTTG 2940
snog1b_frag3             GAGGACCAACTCCAGTGGCTGCCAACCGCGTCCCAGTTGACGAAGTGATTGATAAAGTTG 800
*****

Pp3c8_19720_genomic      CAAGCATGGGGTTTGGAAAGAGATCAGGTTTCGGGCCACTGTCAAGAAATTAAGTAAAATG 3000
snog1b_frag3             CAAGCATGGGGTTTGGAAAGAGATCAGGTTTCGGGCCACTGTCAAGAAATTAAGTAAAATG 860
*****

Pp3c8_19720_genomic      GTACATCTGTGGACTTAAACGTTGTCCTTGACAAGTTGATGAATGGTGACTCTAGTGGCA 3060
snog1b_frag3             GTACATCTGTGGACTTAAACGTTGTCCTTGACAAGTTGATGAATGGTGACTCTAGTGGCA 920
*****

Pp3c8_19720_genomic      GAAATGAGAGTCAGCAACCTGCGAAAGGCTGGTATGGCCGGTCAATGAAGATTGCTTTTCGT 3120
snog1b_frag3             GAAATGAGAGTCAGCAACCTGCGAAAGGCTGGTATGGCCGGTCAATGAAGATTGCTTTAANA 980
*****

Pp3c8_19720_genomic      GATTCTTCTCTGGTCGTGCTTCAGGTACCTTAGGTCATGCTGTAATATTCAATAACATCT 3180
snog1b_frag3             A----- 981

```

Figure S27: Sequencing of *Pp3c8_19720* from *snog1b* using three pairs of overlapping primers (putSNOG1A_genomic1_F/ putSNOG1A_genomic1_R, putSNOG1A_genomic2_F/ putSNOG1A_genomic2_R, putSNOG1A_genomic3_F/ putSNOG1A_genomic3_R). The START codon is highlighted in yellow. The STOP codon is highlighted in green. An intronic SNP is highlighted in red.

281

nogidis snogla del-6

snogla del genotyping predicted amplicon	GTCCACCAAGACCAAGAAACGGTCTTACGGTTGAGATCAACGACGCCAGGGACTACGGTC	60
no_6_amplicon_sequencing_F	-----NNNNNNNNNNNNNCAGCCNCGGACTACGGTC	33
snogla del genotyping predicted amplicon	GTGGTTAAGAGACCCCATACCGTGTCCCAAGTCCATCGTCGTCGTCCTCTCTCG	120
no_6_amplicon_sequencing_F	GTGGTTAAGAGACCCCATACCGTGTCCCAAGTCCATCGTCGTCGTCCTCTCTCG	93
snogla del genotyping predicted amplicon	ATAGCGCCCCGCGCTTTCGTCGCTTCTCTCTCTCGGCGTCGATGCTTCTTGCCCAAG	180
no_6_amplicon_sequencing_F	ATAGCGCCCCGCGCTTTCGTCGCTTCTCTCTCTCGGCGTCGATGCTTCTTGCCCAAG	153
snogla del genotyping predicted amplicon	CTCATTGATTATTCATCATATTTATCATCTCGCGAGCCTCGCCCAATTGTTGTCCCAT	240
no_6_amplicon_sequencing_F	CTCATTGATTATTCATCATATTTATCATCTCGCGAGCCTCGCCCAATTGTTGTCCCAT	213
snogla del genotyping predicted amplicon	TGAGCCCATTCGCGGTCAAACTATTGCTTTAGTCAATTGAGACTAGTTTGTTCCTCG	300
no_6_amplicon_sequencing_F	TGAGCCCATTCGCGGTCAAACTATTGCTTTAGTCAATTGAGACTAGTTTGTTCCTCG	273
snogla del genotyping predicted amplicon	TGGTGATTGGGTGTGGGTAGGGTTTGTGCATGAATGTGAACCTTGGAGTCTGTGTTAT	360
no_6_amplicon_sequencing_F	TGGTGATTGGGTGTGGGTAGGGTTTGTGCATGAATGTGAACCTTGGAGTCTGTGTTAT	333
snogla del genotyping predicted amplicon	CTCTCGAATTGGGTCATAGTTTCTTAAGGTTGGAAGTTAGTAACGTAGCGTCAACTGA	420
no_6_amplicon_sequencing_F	GTTCTGGAATTGGGTCATAGTTTCTTAAGGTTGGAAGTTAGTAACGTAGCGTCAACTGA	393
snogla del genotyping predicted amplicon	TCAATTGGGTATGATTGATAGGTTTGTGAAGTTATGTTGTTTATTTTCATTTTGGGG	480
no_6_amplicon_sequencing_F	TCAATTGGGTATGATTGATAGGTTTGTGAAGTTATGTTGTTTATTTTCATTTTGGGG	453
...		
snogla del genotyping predicted amplicon	TTGGCCGAGATCGTGGACATGTGAGATGATGCTGATGGACCAATTAGCGGCCGGGGGA	1256
no_6_amplicon_sequencing_R	TTGGCCGAGATCGTGGACATGTGAGATGATGCTGATGGACCAATTAGCGGCCGGGGGA	296
snogla del genotyping predicted amplicon	AATTCCCATGGAGTCAAGATTCAAATAGAGGACCTAACAGAACTCGCCGTAAAGACTGG	1316
no_6_amplicon_sequencing_R	AATTCCCATGGAGTCAAGATTCAAATAGAGGACCTAACAGAACTCGCCGTAAAGACTGG	356
snogla del genotyping predicted amplicon	CGAACAGTTTATACAGAGTCTCTTACGACTCAATGACAAGAAGAAATCTTCTGCAACAT	1376
no_6_amplicon_sequencing_R	CGAACAGTTTATACAGAGTCTCTTACGACTCAATGACAAGAAGAAATCTTCTGCAACAT	416
snogla del genotyping predicted amplicon	GGTGGAGCAGCAGCAGCTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCA	1436
no_6_amplicon_sequencing_R	GGTGGAGCAGCAGCAGCTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCA	476
snogla del genotyping predicted amplicon	AAGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTG	1496
no_6_amplicon_sequencing_R	AAGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTG	536
snogla del genotyping predicted amplicon	CCCAGCTATCTGTCACTTTATTGTGAAGATAGTGGAAAAGGAAGTGGCTCCTACAAATG	1556
no_6_amplicon_sequencing_R	CCCAGCTATCTGTCACTTTATTGTGAAGATAGTGGAAAAGGAAGTGGCTCCTACAAATG	596
snogla del genotyping predicted amplicon	CCATCATTTGCGATAAAGGAAAGGCCATCGTTGAAGATGCTCTGCCACAGTGGTCCCAA	1616
no_6_amplicon_sequencing_R	CCATCATTTGCGATAAAGGAAAGGCCATCGTTGAAGATGCTCTGCCACAGTGGTCCCAA	656
snogla del genotyping predicted amplicon	AGATGGACCCCCACCCACGAGGAGCATCGTGGAAAAAGAGACGTTCCAAACCAGTCTTC	1676
no_6_amplicon_sequencing_R	AGATGGACCCCCACCCACGAGGAGCATCGTGGAAAAAGAGACGTTCCAAACCAGTCTTC	716
snogla del genotyping predicted amplicon	AAAGCAAGTGGATTGATGTGATATCTCCACTGACGTAAGGGATGACGCACATCCCACTA	1736
no_6_amplicon_sequencing_R	AAAGCAAGTGGATTGATGTGATATCTCCACTGACGTAAGGGATGACGCACATCCCACTA	776

snog1a_del_genotyping_predicted_amplicon	TCCTTCGCAAGACCCCTTCCTCTATATAAGGAAGTTCATTTCATTGGAGAGGACAGGGTA	1796
no_6_amplicon_sequencing_R	TCCTTCGCAAGACCCCTTCCTCTATATAAGGAAGTTCATTTCATTGGAGAGGACAGGGTA	836
snog1a_del_genotyping_predicted_amplicon	CCCATCAAAAGGCTACTGCTATTATAACACATCTCTGGAGACTGAGAAATCAGACCTC	1856
no_6_amplicon_sequencing_R	CCCATCAAAAGGCTACTGCTATTATAACACATCTCTGGAGACTGAGAAATCAGACCTC	896
snog1a_del_genotyping_predicted_amplicon	CAAGCATGGAGAACATAGAAAACCTCTCATGCAAGGAAAATCTAATGCTAGAGCTCG	1916
no_6_amplicon_sequencing_R	CAACATGGAGAACATAGAAAACCTCTCATGCAAGGAAAATCTAATGCTAGAGCTCG	956
snog1a_del_genotyping_predicted_amplicon	ATCTAGTAAGAGCAAAAATAAGCTTGGATGGATTGCACGCAGGTTCTCCGGCCGCTTGGG	1976
no_6_amplicon_sequencing_R	ATCTAGTAAGAGCAAAAATAAGCTTGGATGGATTGCACGCAGGTTCTCCGGCCGCTTGGG	1016
snog1a_del_genotyping_predicted_amplicon	TGGAGAGGCTATTTCGGCTATGACTGGGCACACACAGCAATCGGCTGCTCTGATG	2030
no_6_amplicon_sequencing_R	NGGAGAGGCTATCGNN	1047

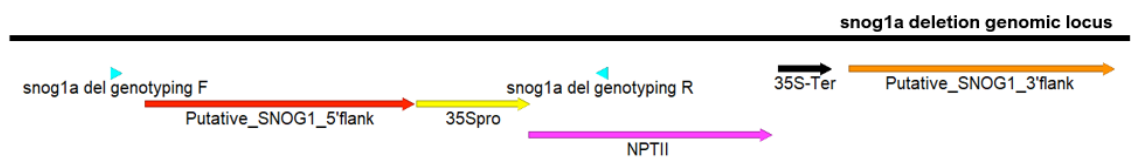


Figure S28: Sequencing of the amplicons acquired with snog1a_del_genotyping2_F/ snog1a_del_genotyping2_R primers from *nog1dis snog1a del-4* and *nog1dis snog1a del-6*. This shows that the deletion construct has integrated into the targeted locus as the amplicons align to the endogenous genomic sequence at one end and the introduced NPTII sequence at the other end. Green highlights the endogenous genomic sequence upstream of 5' flank region. Red highlights the *PpSNOG1A* 5' flanking region included in the deletion construct. Yellow highlights the 35S promoter. Pink highlights the NPTII sequence in the deletion construct. Schematic shows construct integrated at the correct locus in the genome and the position of the genotyping primers.

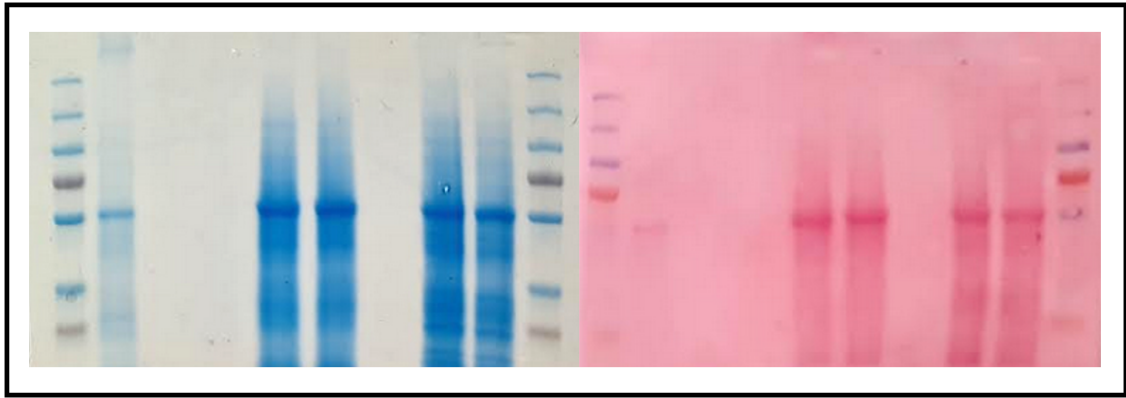


Figure S29: Controls for western blot to detect PpSNOG1A-GFP. Gel loading control shown with Coomassie blue stain of the SDS-PAGE gel. Effective protein transfer onto the membrane shown with Ponceau S stain.

Chapter 5: RNA-seq experiments yield insights into gene expression in 3D-defective mutants

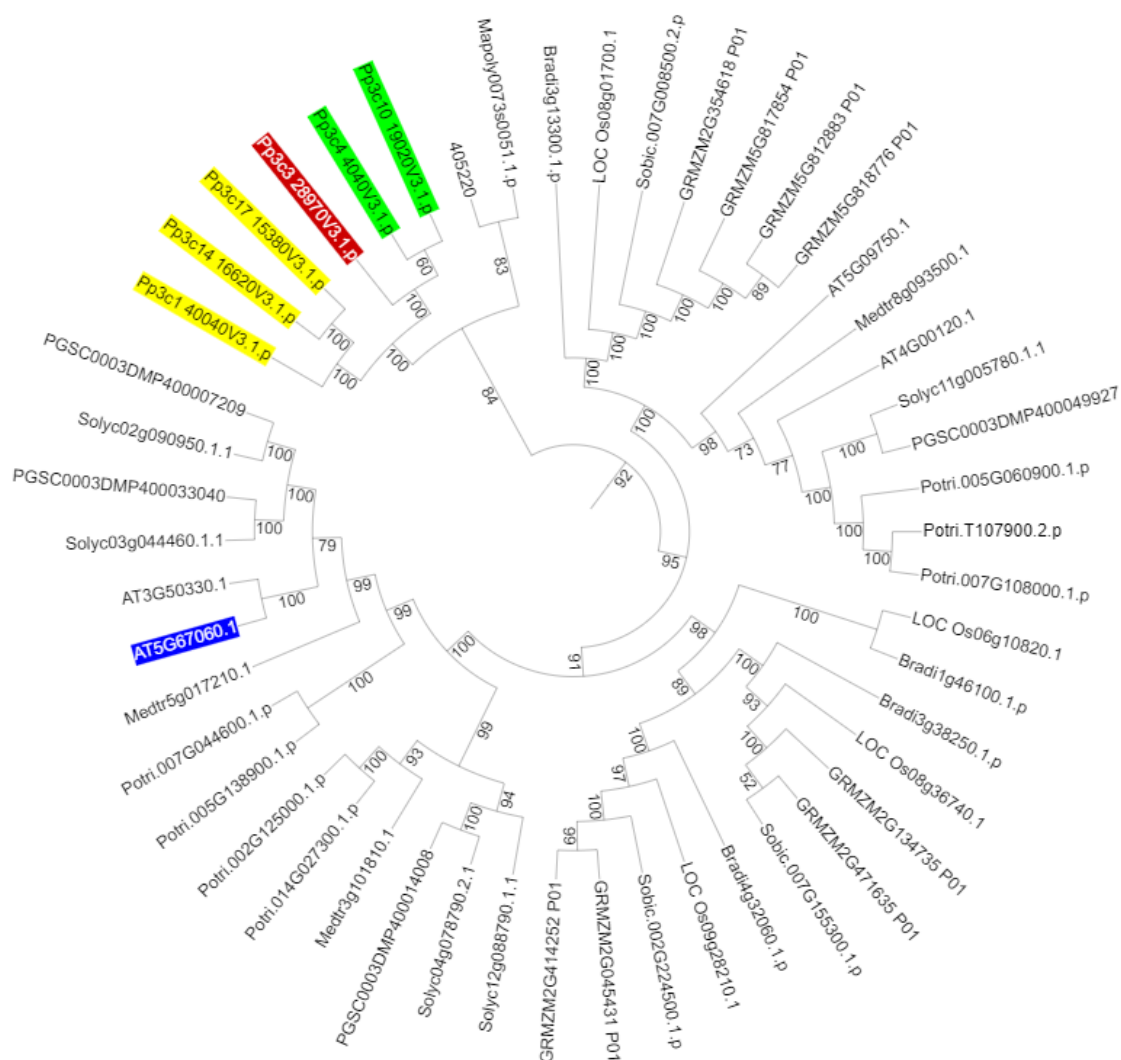


Figure S30: Phylogenetic tree based on multiple sequence alignment of HECATE protein homologues. Blue is AtHEC1. Green is PpHEC1a/PpHEC1b. Red is homologue for which no expression was detected in protonemal tissues. Yellow are three other PpHEC1 homologues. Branch numbers are bootstrap values.

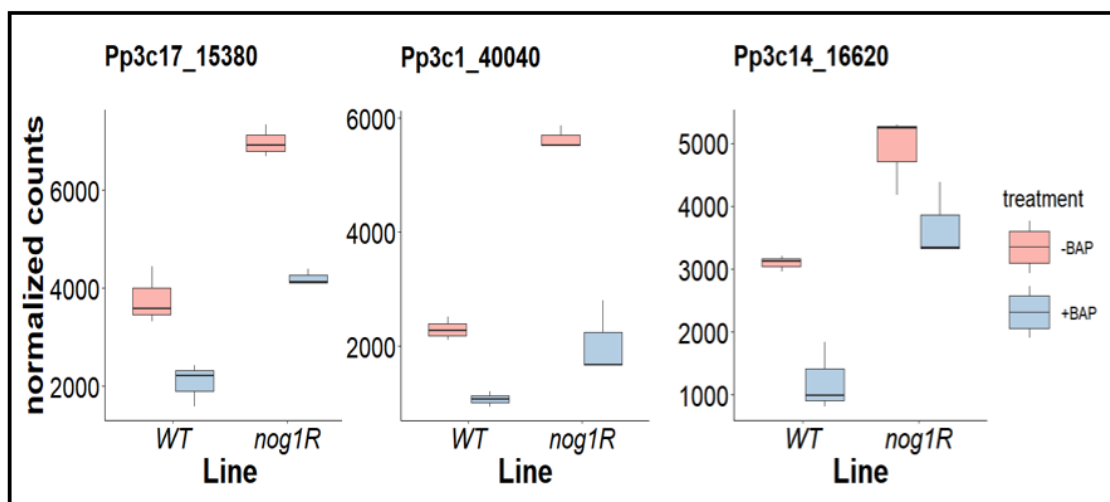


Figure S31: The normalised read count data showing expression levels of three *PpHEC1* genes in wild type, wild type + BAP, *Ppnog1R* and *Ppnog1R* + BAP.

Accession no.	Description
<i>Pp3c1.1080</i>	
<i>Pp3c1.26200</i>	12-oxophytodienoate reductase 2
<i>Pp3c1.36390</i>	ADP-ribosylation factor A1E
<i>Pp3c1.36940</i>	homeobox protein 16
<i>Pp3c1.8860</i>	
<i>Pp3c10.10040</i>	
<i>Pp3c10.12560</i>	PLATZ transcription factor family protein
<i>Pp3c10.14840</i>	plastid transcriptionally active 4
<i>Pp3c10.1490</i>	Galactose oxidase/kelch repeat superfamily protein
<i>Pp3c10.15520</i>	phosphoglycerate kinase
<i>Pp3c10.1580</i>	
<i>Pp3c10.18930</i>	
<i>Pp3c10.3520</i>	diaminopimelate epimerase family protein
<i>Pp3c10.8570</i>	glyoxal oxidase-related protein
<i>Pp3c11.14690</i>	Integrase-type DNA-binding superfamily protein
<i>Pp3c11.23650</i>	Transducin family protein / WD-40 repeat family protein
<i>Pp3c11.2770</i>	Transducin/WD40 repeat-like superfamily protein
<i>Pp3c12.11230</i>	ATP binding cassette subfamily B4
<i>Pp3c12.12140</i>	multidrug resistance-associated protein 2
<i>Pp3c12.17180</i>	myb domain protein 103
<i>Pp3c12.19420</i>	NAD(P)-binding Rossmann-fold superfamily protein
<i>Pp3c12.23100</i>	Molybdenum cofactor sulfurase family protein
<i>Pp3c12.5390</i>	
<i>Pp3c12.7878</i>	
<i>Pp3c13.1670</i>	glycine-rich protein
<i>Pp3c13.2080</i>	basic helix-loop-helix (bHLH) DNA-binding superfamily protein
<i>Pp3c13.24590</i>	MADS-box transcription factor family protein
<i>Pp3c13.4760</i>	PLATZ transcription factor family protein
<i>Pp3c13.5660</i>	
<i>Pp3c13.6760</i>	basic helix-loop-helix (bHLH) DNA-binding superfamily protein
<i>Pp3c14.11830</i>	pleiotropic drug resistance 7
<i>Pp3c14.12710</i>	
<i>Pp3c14.23490</i>	
<i>Pp3c14.5440</i>	SPFH/Band 7/PHB domain-containing membrane-associated protein family

<i>Pp3c14.5560</i>	
<i>Pp3c14.6160</i>	
<i>Pp3c15.15470</i>	
<i>Pp3c15.22000</i>	
<i>Pp3c15.2490</i>	Pyruvate kinase family protein
<i>Pp3c15.3810</i>	
<i>Pp3c15.5890</i>	C2 domain-containing protein / GRAM domain-containing protein
<i>Pp3c15.7120</i>	Cellulose synthase family protein
<i>Pp3c15.7150</i>	cellulose synthase 1
<i>Pp3c15.830</i>	
<i>Pp3c16.1070</i>	Glycosyl hydrolase family protein with chitinase insertion domain
<i>Pp3c16.10920</i>	
<i>Pp3c16.17550</i>	Integrase-type DNA-binding superfamily protein
<i>Pp3c16.2310</i>	pectin-related
<i>Pp3c16.3180</i>	Early-responsive to dehydration stress protein (ERD4)
<i>Pp3c17.10170</i>	erf domain protein 9
<i>Pp3c17.17300</i>	Tetratricopeptide repeat (TPR)-like superfamily protein
<i>Pp3c17.21660</i>	GDSL-like Lipase/Acylhydrolase superfamily protein
<i>Pp3c17.21830</i>	Ribosomal protein S5/Elongation factor G/III/V family protein
<i>Pp3c17.887</i>	
<i>Pp3c17.8870</i>	
<i>Pp3c17.9890</i>	serine carboxypeptidase-like 48
<i>Pp3c18.14540</i>	
<i>Pp3c18.16700</i>	RING/U-box superfamily protein
<i>Pp3c18.20200</i>	Aldolase-type TIM barrel family protein
<i>Pp3c18.4620</i>	histidine kinase 2
<i>Pp3c18.5730</i>	Leucine-rich repeat transmembrane protein kinase
<i>Pp3c18.8480</i>	BTB/POZ domain-containing protein
<i>Pp3c19.14690</i>	
<i>Pp3c19.19330</i>	Thiamine pyrophosphate dependent pyruvate decarboxylase family protein
<i>Pp3c19.19870</i>	
<i>Pp3c19.21290</i>	
<i>Pp3c19.2820</i>	Transducin/WD40 repeat-like superfamily protein
<i>Pp3c19.3900</i>	
<i>Pp3c19.4520</i>	P-loop containing nucleoside triphosphate hydrolases superfamily protein

<i>Pp3c19.9271</i>	
<i>Pp3c2.12240</i>	mitochondrial HSO70 2
<i>Pp3c2.12680</i>	alternative oxidase 1A
<i>Pp3c2.1670</i>	Phosphoglycerate mutase family protein
<i>Pp3c2.16830</i>	decapping 2
<i>Pp3c2.20450</i>	
<i>Pp3c2.26770</i>	acyl-CoA binding protein 2
<i>Pp3c2.27000</i>	APRATAXIN-like
<i>Pp3c2.34670</i>	myb domain protein 105
<i>Pp3c20.11540</i>	proteasome activating protein 200
<i>Pp3c20.12200</i>	WRKY family transcription factor
<i>Pp3c20.124</i>	
<i>Pp3c20.16210</i>	
<i>Pp3c20.3140</i>	
<i>Pp3c20.8330</i>	Esterase/lipase/thioesterase family protein
<i>Pp3c21.11360</i>	Lactate/malate dehydrogenase family protein
<i>Pp3c21.6800</i>	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein
<i>Pp3c21.9280</i>	adenosine kinase
<i>Pp3c22.12260</i>	Leucine-rich repeat (LRR) family protein
<i>Pp3c22.12540</i>	glutamate-cysteine ligase
<i>Pp3c22.14060</i>	
<i>Pp3c22.15320</i>	
<i>Pp3c22.19000</i>	
<i>Pp3c22.8510</i>	tryptophan synthase beta-subunit 2
<i>Pp3c23.12980</i>	
<i>Pp3c23.15050</i>	RHOMBOID-like 1
<i>Pp3c23.18540</i>	
<i>Pp3c23.18780</i>	Plant protein of unknown function (DUF869)
<i>Pp3c23.2880</i>	
<i>Pp3c23.3290</i>	basic helix-loop-helix (bHLH) DNA-binding superfamily protein
<i>Pp3c23.7740</i>	plastidic type i signal peptidase 1
<i>Pp3c23.800</i>	C2H2 and C2HC zinc fingers superfamily protein
<i>Pp3c24.10130</i>	CBS domain-containing protein with a domain of unknown function (DUF21)
<i>Pp3c24.13910</i>	Translation initiation factor IF6
<i>Pp3c24.13920</i>	DNA directed RNA polymerase, 7 kDa subunit

<i>Pp3c25.2290</i>	Peptidase C78, ubiquitin fold modifier-specific peptidase 1/ 2
<i>Pp3c25.950</i>	WWE protein-protein interaction domain protein family
<i>Pp3c26.1820</i>	
<i>Pp3c26.290</i>	
<i>Pp3c26.30</i>	
<i>Pp3c26.3410</i>	Galactose oxidase/kelch repeat superfamily protein
<i>Pp3c27.2480</i>	
<i>Pp3c27.3260</i>	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein
<i>Pp3c27.4280</i>	AAA-type ATPase family protein
<i>Pp3c3.12290</i>	Poly (ADP-ribose) glycohydrolase (PARG)
<i>Pp3c3.15470</i>	Galactose oxidase/kelch repeat superfamily protein
<i>Pp3c3.18850</i>	
<i>Pp3c3.21830</i>	2-isopropylmalate synthase 1
<i>Pp3c3.23790</i>	phytochrome B
<i>Pp3c3.30710</i>	Homeodomain-like superfamily protein
<i>Pp3c3.32100</i>	alpha/beta-Hydrolases superfamily protein
<i>Pp3c3.37060</i>	TraB family protein
<i>Pp3c3.9650</i>	STELAR K ⁺ outward rectifier
<i>Pp3c4.12510</i>	myb domain protein 103
<i>Pp3c4.13790</i>	myb domain protein 105
<i>Pp3c4.21490</i>	sodium:hydrogen antiporter 1
<i>Pp3c4.24710</i>	Integrase-type DNA-binding superfamily protein
<i>Pp3c4.2680</i>	Integrase-type DNA-binding superfamily protein
<i>Pp3c4.29450</i>	
<i>Pp3c4.29810</i>	SET domain group 4
<i>Pp3c4.30510</i>	alpha/beta-Hydrolases superfamily protein
<i>Pp3c5.14260</i>	Xanthine/uracil permease family protein
<i>Pp3c5.19220</i>	aldehyde oxidase 4
<i>Pp3c5.3340</i>	ABC-2 and Plant PDR ABC-type transporter family protein
<i>Pp3c5.5590</i>	
<i>Pp3c5.8660</i>	zinc finger (C3HC4-type RING finger) family protein
<i>Pp3c6.14680</i>	ER lumen protein retaining receptor family protein
<i>Pp3c6.15020</i>	CHASE domain containing histidine kinase protein
<i>Pp3c6.23710</i>	EF hand calcium-binding protein family
<i>Pp3c6.24500</i>	Mov34/MPN/PAD-1 family protein

<i>Pp3c6.3710</i>	HSP20-like chaperones superfamily protein
<i>Pp3c6.4290</i>	
<i>Pp3c6.7260</i>	ATP binding cassette subfamily B19
<i>Pp3c7.10960</i>	D-3-phosphoglycerate dehydrogenase
<i>Pp3c7.12220</i>	NAD(P)-binding Rossmann-fold superfamily protein
<i>Pp3c7.13080</i>	glutathione S-transferase PHI 10
<i>Pp3c7.20200</i>	ethylene responsive element binding factor 2
<i>Pp3c7.20870</i>	myb-like HTH transcriptional regulator family protein
<i>Pp3c7.23000</i>	tripeptidyl peptidase ii
<i>Pp3c7.26870</i>	hydroxycinnamoyl-CoA shikimate/quinate hydroxycinnamoyl transferase
<i>Pp3c7.3360</i>	cold, circadian rhythm, and RNA binding 1
<i>Pp3c8.1820</i>	
<i>Pp3c8.1940</i>	photosynthetic electron transfer C
<i>Pp3c8.19560</i>	gamete expressed protein 1
<i>Pp3c8.230</i>	zinc finger (C2H2 type, AN1-like) family protein
<i>Pp3c8.25270</i>	receptor-like kinase in in flowers 3
<i>Pp3c8.25400</i>	zinc finger (C2H2 type, AN1-like) family protein
<i>Pp3c9.18930</i>	
<i>Pp3c9.19810</i>	
<i>Pp3c9.3190</i>	Early-responsive to dehydration stress protein (ERD4)
<i>Pp3c9.760</i>	

Table S2: List of genes that are deregulated in both *Ppnog1R* and *nog1dis* and also not deregulated in *snog1a* compared to wild type

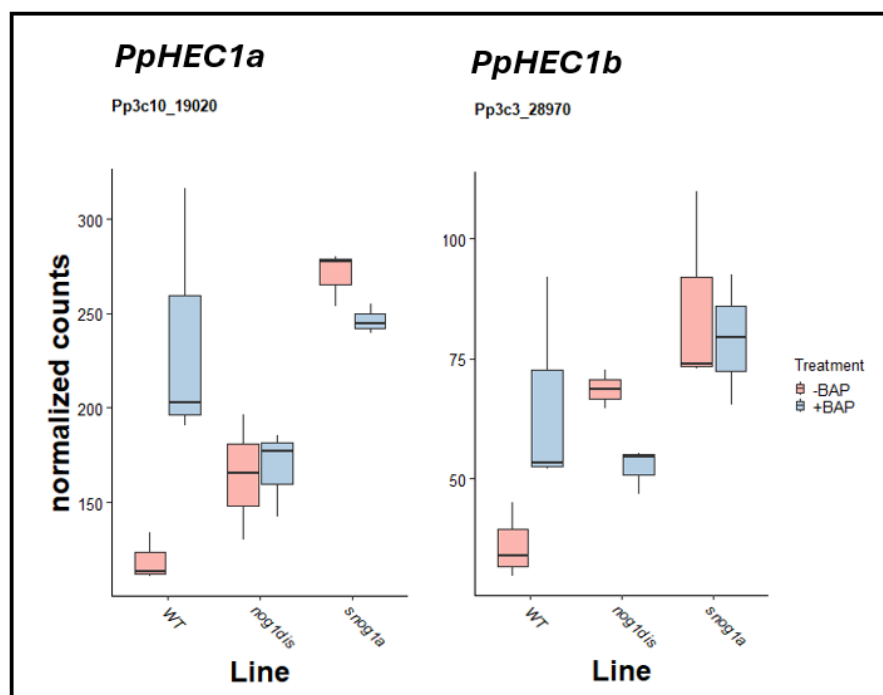


Figure S32: The normalised read count data showing expression levels of *PpHEC1a* and *PpHEC1b* in wild type, wild type + BAP, *nog1dis*, *nog1dis* + BAP, *snog1a* and *snog1a* + BAP.