The Proteins Found at Plasmodesmata and the Interactions Between Them

Matthew Graham Johnston

Thesis submitted for the degree of Doctor of Philosophy

University of East Anglia

John Innes Centre

January 2021

i Abstract

Plasmodesmata are specialised membrane-lined pores that create cell-to-cell connections through the cell wall. This cytoplasmic and membrane continuity allows for communication and co-ordination between cells, a prerequisite for multicellularity. All land plants contain plasmodesmata derived from a single evolutionary event.

Plasmodesmata are not passive conduits. Instead, the aperture of plasmodesmata is dynamically regulated by the deposition of a complex carbohydrate, callose. This controls the cell-to-cell flux through plasmodesmata.

In this thesis, I explore the protein composition of plasmodesmata. I developed a plasmodesmata extraction protocol for mature plant tissue. This protocol is used to biochemically localise transiently expressed proteins to plasmodesmata in *Nicotiana benthamiana*. This technique is then extended to define the native plasmodesmal proteome of the bryophyte *Physcomitrella patens*. I used a comparative phyloproteomic approach to identify conserved protein families at plasmodesmata. This approach identified two classes of structural proteins, C2 lipid-binding proteins and tetraspanins, which may have been present in the plasmodesmata of the last common ancestor between land plants and algae.

Secondly, I investigate interactions between PLASMODESMATA-LOCALISED PROTEINs (PDLPs) and other plasmodesmata proteins. PDLP overexpression leads to the misregulation of callose deposition, ultimately dwarfing *Arabidopsis thaliana* plants. I exploited this phenotype to find novel components in the PDLP-callose deposition pathway. Ultimately, I propose a common pathway downstream of PDLPs which is required for plasmodesmata callose deposition.

Overall, the results herein offer candidate proteins that may be ancient components of plasmodesmata. These may have both structural and biochemical functions. I characterised one PDLP interactor, NDR1/HIN1-LIKE 3, and produced a list of other likely interactors, by comparing interaction data with *A. thaliana* plasmodesmal proteomes. An additional putative genetic interactor with PDLP1, KISS ME DEADLY 2, was identified by a forward genetic screen. This will guide the direction of future investigations.

ii   Contents

iii List of Figures

iv   List of Tables

v   List of Equations

vi   List of Supplementary Tables

Table S1 Unpublished level 0 plasmids used in Golden Gate

Table S2 Mature tissue plasmodesmata proteomes generated in this thesis

Table S3 *Arabidopsis thaliana* datasets used in this thesis

vii   Publications arising from this thesis

1. A bootstrap approach is a superior statistical method for the comparison of non-normal data with differing variances. **Johnston MG**, Faulkner C. *New Phytologist* **230**(*1*):23-6 (2021). DOI: 10.1111/nph.17159

2. Chitin perception in plasmodesmata characterises submembrane immune signalling specificity in plants. Cheval C*, Samwald S*, **Johnston MG***, de Keijzer J, Breakspear A, Liu X, Bellandi A, Kadota Y, Zipfel C, Faulkner C. *Proceedings of the National Academy of Sciences of the United States of America* **117**(*17*):9621-29 (2020). DOI: 10.1073/pnas.1907799117

    *These authors contributed equally to this work.

viii Acknowledgements

1 Introduction: the intracellular network of plants

Multicellularity has evolved many times within Eukarya, allowing organisms to grow larger and specialise tissues. A consequence of multicellularity is a requirement for communication and exchange between cells. This allows cells to co-ordinate growth and development, as well as provide the resources which tissues require.

Plants can communicate between cells in numerous ways. The main two routes for cell-to-cell communication is apoplastically, in the cell wall, and symplastically, the movement of cytoplasm between cells. The classic example of apoplastic signalling is the movement of the hormone auxin, potentiated by the acidic environment of the apoplast. Plant cells also communicate symplastically, despite being embedded in a cell wall which isolates plant cells from one another. Plasmodesmata are membrane-lined channels that traverse the cell wall connecting neighbouring cells, allowing cytoplasmic continuity between cells. This forms a connected symplasm throughout a plant permitting direct cell-to-cell communication.

Plasmodesmata are dynamic structures that can modulate the degree of cell-to-cell connectivity. This means that plants can isolate tissues, such as dormant tree buds or even generate isolated domains within tissues. Plasmodesmata also respond to the environment over shorter timescales. Cellular stress can lead to a rapid reduction in flux of molecules between cells. This response is an important part of plant defence, required for plant immunity from pathogens. The signalling pathways that lead to plasmodesmal closure are still poorly understood. This work aims to better characterise the components of plasmodesmata, and how they interact when signalling in plant plasmodesmal closure.
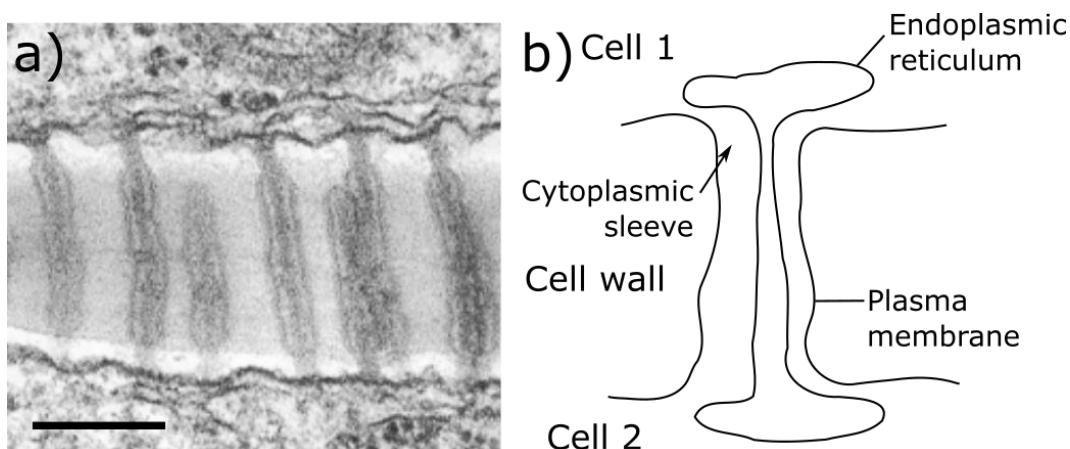
Figure 1-1 Ultrastructure of a plasmodesmata
a) Electron micrograph of a pit field containing plasmodesmata in a mature sugarcane leaf (*Saccharum sp.*). Plasmodesmata traverse the cell wall between two cells. Scale bar = 200 nm. Reproduced with permission from Robinson-Beers and Evert (1991). b) Cartoon of a single plasmodesma from the micrograph, demonstrating membrane and cytoplasmic continuity between the cells.

1.1   Ultrastructure of a plasmodesmata

Plasmodesmata provide both membrane and cytoplasmic continuity between plant cells (Figure 1-1). Due to the nanoscopic size of plasmodesmata, their structural features can only be imaged with electron microscopy techniques. Transmission electron micrographs show that the plasma membrane is continuous from cell to cell (Overall *et al.*, 1982). Within the centre of the plasmodesmata a central core is observed that is continuous with the endoplasmic reticulum, termed the desmotubule (López-Sáez *et al.*, 1966). The cytoplasmic sleeve is thought to be 20 – 40 nm wide (Ehlers & Kollmann, 2001), and the desmotubule to have a diameter of about 8 nm (Grison *et al.*, 2015a).

The desmotubule has been observed to have spokes radiating out to the cytoplasm (Ding *et al.*, 1992b). Interestingly, spokes have been observed in basal algal plasmodesmata, as well as in the higher land plants (Cook *et al.*, 1997; Brecknock *et al.*, 2011). Plasmodesmata can be separated into two classes, Type I and Type II, depending on the presence of these spokes and size of the cytoplasmic sleeve (Nicolas *et al.*, 2017). Type II plasmodesmata are found in mature walls and look "textbook", containing a cytoplasmic sleeve and spoke elements, whereas Type I plasmodesmata have no cytoplasmic sleeve or spokes, and are found in nascent post-cytokinesis walls. The maturation of plasmodesmata from Type I to Type II has been shown to be prevented in a sphingolipid biosynthesis mutant (Yan *et al.*, 2019) implicating membrane composition in plasmodesmal maturation. Spokes are not the only component to have been observed with plasmodesmata, with extracellular sphincters and extracellular spirals also observed (Badelt *et al.*, 1994).

1.2   Genesis and form of plasmodesmata

Plasmodesmata are thought to be formed primarily during cytokinesis in higher plants, when the endoplasmic reticulum is entrapped by the cell plate (Porter & Palade, 1957; Hepler, 1982). Plasmodesmata formed in this manner are termed primary plasmodesmata. Secondary plasmodesmata are formed *de novo* across existing cell walls as cells expand. They were first observed between plant-plant junctions, such as synthetic grafts, as well as those between parasitic plants and their hosts (see Jones (1976) for a review). Secondary plasmodesmata have since be shown to be a natural part of cell development as the cell expands (Faulkner *et al.*, 2008). Secondary plasmodesmata can also be induced within a cell, either endogenously, such as by the floral transition and concomitant increase in cytokinin in *Sinapis alba* (Ormenese *et al.*, 2006) or exogenously, as in the case of nematode invasion (*Meloidogyne sp.*) (Jones & Dropkin, 1976). Both primary and secondary formation have been

observed within the basal streptophytic algae *Chara corallina* and *Chara zeylanic,* suggesting that the mechanisms of plasmodesmata development have an ancient evolutionary origin (Franceschi *et al.*, 1994; Cook *et al.*, 1997)*.*
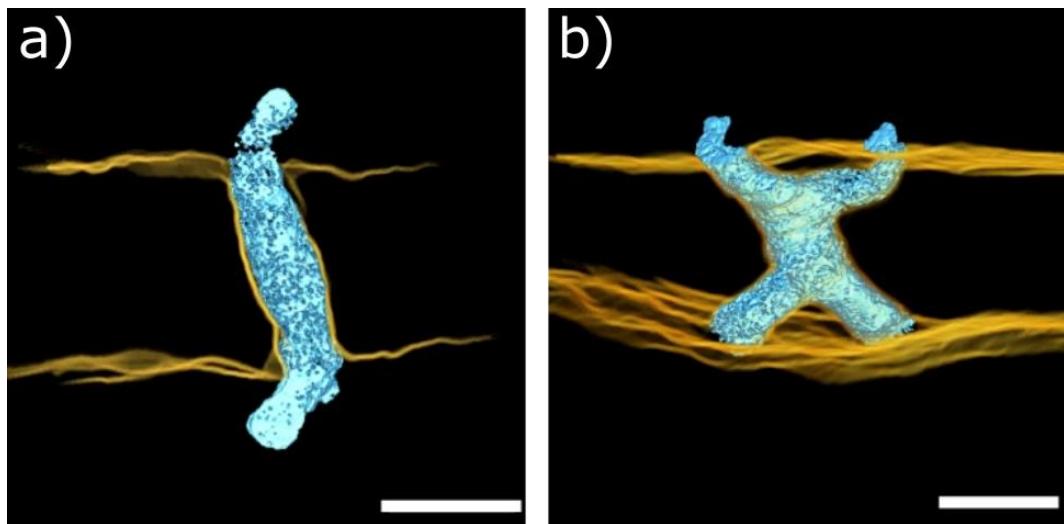
Figure 1-2 Simple and complex plasmodesmata
3D reconstructions from 2D images of simple (a) and complex (b) plasmodesmata in *Arabidopsis thaliana*. The plasma membrane is rendered in yellow and the desmotubule in light blue. Scale bar = 50 nm. Reproduced with permission from Yan et al. (2019).

Primary plasmodesmata are initially formed as simple plasmodesmata. Simple plasmodesmata are single channels that cross the cell wall with a single opening on either side of the cell wall, whereas complex plasmodesmata have multiple branches and openings that often meet in the middle lamella to form a central cavity (Figure 1-2). Plasmodesmata tend to increase in complexity as the tissue matures (Oparka *et al.*, 1999; Roberts *et al.*, 2001; Kwiatkowska, 2003). The maturation of plasmodesmata from simple to complex is impaired in a choline transporter mutant, indicating that plasmodesmata-specific or plasmodesmata-localised lipid synthesis is required for the maturation of plasmodesmata (Kraner *et al.*, 2017a). Branched plasmodesmata (or twinned secondary plasmodesmata) may form from simple plasmodesmata by actively inserting new endoplasmic reticulum strands into the plasmodesma as the wall grows (Faulkner *et al.*, 2008). Alternatively, cell wall loosening may allow a plasmodesma to expand and new cell wall material to deposit in the middle of the plasmodesma (Ehlers & van Bel, 2010). The formation of twinned plasmodesmata help explain the presence of pit fields, as it suggests that secondary plasmodesmata are formed adjacent to existing plasmodesmata leading to clusters of plasmodesmata (or pit fields) in mature walls, where many plasmodesmata reside together within the cell wall (Faulkner *et al.*, 2008).

Recently, serial block-face scanning electron microscopy (as used to produce the images in Figure 1-2) has been used to demonstrate the diversity between plasmodesmata within a single cell: the phloem-pole pericycle (Ross-Elliott *et al.*, 2017). Ross-Elliott *et al.* discovered that the plasmodesmata between protophloem sieve elements and phloem-pole pericycle are funnel shaped, with much larger apertures on the protophloem sieve element face. On other cell interfaces, there are 'pore' plasmodesmata with multiple branches on one side of the cell wall filtering into a single opening on the phloem-pole pericycle side. In addition, the phloem-pole pericycle cell also has simple plasmodesmata and branched plasmodesmata (Ross-Elliott *et al.*, 2017). The localised diversity of plasmodesmal structure between cell-cell interfaces within a single cell suggests a diversity of function within plasmodesmata. Features such as asymmetrical branching or pore opening might allow selective or directional trafficking of molecules. Mathematical modelling may be able to shed light on how geometry affects plasmodesmal function. This work has been started for simple plasmodesmata, but has yet to be applied to funnel plasmodesmata (Deinum *et al.*, 2019)

1.3 Transport through plasmodesmata

1.3.1 The size exclusion limit of cytoplasmic trafficking

The cytoplasmic sleeve provides a conduit for molecules to move between two cells, allowing the trafficking of a diverse range of substrates from metabolites to proteins. A contentious question in the field has been the size of molecules that can fit through the cytoplasmic sleeve, known as the size exclusion limit (SEL). Original studies focussed on microinjection of dyes and their conjugates to determine the SEL, yielding a SEL of ≈1 kDa (Erwee & Goodwin, 1983; Burnell, 1988; Goodwin *et al.*, 1990).

This view has been challenged more recently with evidence that proteins also move through plasmodesmata. An updated SEL of roughly 50 kDa was found using GREEN FLUORESCENT PROTEIN (GFP) in *Nicotiana tabacum* var *Samsun* epidermal cells (Imlau *et al.*, 1999; Crawford & Zambryski, 2000). The difference in the SEL is thought to be due to the method of dye introduction: invasive (microinjection) compared to less invasive (bombardment and transformation). This is underscored by the observation that high pressure bombardment or detachment of leaves can prevent GFP movement, whereas GFP movement is observed with low-pressure bombardment (Crawford & Zambryski, 2000). Indeed, it has been shown that microinjection itself lowers the SEL in *A. thaliana* (Radford & White, 2001).

The modulation of the SEL, dependent on sample treatment, highlights the idea that plasmodesmata are dynamic structures. An early example of this is the response of

plasmodesmata to plasmolysis. When a cell is plasmolysed and subsequently deplasmolysed, the SEL is significantly decreased (Jarvis & House, 1970; Oparka & Prior, 1987; Weiner *et al.*, 1988). Physiologically relevant stimuli can also alter the SEL, such as light (Epel & Erlanger, 1991; Tylewicz *et al.*, 2018) and signalling molecules (e.g. IP$_3$ (Tucker, 1988) and Ca^{2+} (Erwee & Goodwin, 1983)), circadian rhythms (Liarzi & Epel, 2005), and biotic stresses (Lee *et al.*, 2011b; Faulkner *et al.*, 2013). Plasmodesmal permeability can also be modulated in response to longer-term stimuli, such as reducing cell-to-cell flux in winter (Shepherd & Goodwin, 1992; Tylewicz *et al.*, 2018).

The SEL also varies developmentally, changing with age and within tissues. For example, symplastic domains are formed within meristems and embryos where there is free movement between some cells but not others (Rinne & Van der Schoot, 1998; Wróbel-Marek *et al.*, 2017). As tissue matures, the SEL also changes. However, in some contexts it has been observed to increase with age as in *Setcreasea purpurea*, whereas in tobacco leaves the SEL decreases with the sink-to-source transition (Shijie *et al.*, 1995; Oparka *et al.*, 1999).

There is no simple relationship between plasmodesmata SEL and the complexity of a plasmodesmata. Naively, one would assume that branched plasmodesmata would allow a greater cell-to-cell flux than simple plasmodesmata, as they are larger. Indeed, the mutants *increased size exclusion limit 1* (*ise1*) and *ise2*, which have an increased SEL, also have a great proportion of branched and twinned plasmodesmata (Stonebloom *et al.*, 2009). However, the opposite relationship was observed in tobacco leaves (Oparka *et al.*, 1999). Roberts *et al.* observed that the sink-to-source maturation is accompanied by a simple-to-complex transition of plasmodesmata (Roberts *et al.*, 2001). However, free GFP could only move cell-to-cell in sink tissue with simple plasmodesmata (Oparka *et al.*, 1999). Thus, the control of the SEL is more complex than simply the type of plasmodesmata.

It was originally proposed that plasmodesmata may exist in three conformations (closed, open or dilated) to dynamically control the SEL (Crawford & Zambryski, 2000). This idea has recently been overturned by the finding that Type I plasmodesmata (with no cytoplasmic sleeve) enable a faster diffusion of molecules that Type II plasmodesmata (Nicolas *et al.*, 2017; Yan *et al.*, 2019). Recent data from our research group has found no evidence of a SEL, rather a exponentially decreasing probability of movement with protein size (Ohtsu *et al.*, 2021). The implication of this is that while cell-to-cell mobility is dependent on molecular size, that there is no hard limit at which it is impossible for a molecule to pass through plasmodesmata.

1.3.2 Lipid trafficking at plasmodesmata

Lipids have also been shown to move via plasmodesmata. Fluorescent phospholipid analogues have been used to demonstrate the membrane continuity between plant cells (Baron-Epel *et al.*, 1988; Grabski *et al.*, 1993). Grabski *et al.* elegantly showed that there is movement of phospholipid analogues through the endoplasmic reticulum, but not the plasma membrane (Grabski *et al.*, 1993). This refutes the earlier work of Baron-Epel *et al.* (1988) who undertook similar experiments, with the same analogues. However, Grabski *et al.* noted that the plasma membrane analogue (NBD-PC) degraded into the endoplasmic reticulum analogue (NBD-DAG), and so showed that the lipids in the plasma membrane had erroneously been found to be mobile. This had the further consequence of demonstrating the movement of physiological lipid signalling molecules (i.e. diacylglycerol (DAG)) (Grabski *et al.*, 1993). Therefore, plasmodesmata provide continuity of both membranes and the cytosol between cells. Presumably, the transport of lipids between cells is not hindered by the occlusion of the cytoplasmic sleeve, unlike cytoplasmic constituents. Thus, lipid signalling between cells may continue even when symplastic transport has been reduced.

1.3.3 The mechanisms of symplastic transport

It is clear that molecules of varying size, from small dyes the size of monomeric sugars (e.g. fluorescein) to large proteins, can move cell-to-cell. However, how molecules of different sizes move from cell to cell has not been well established. Goodwin *et al.* found the rate of dye movement to be proportional to mass, suggesting translocation by simple diffusion (Goodwin *et al.*, 1990). Indeed, modelling movement as diffusion also matched experimental data for Tyree and Tammes, who observed movement proportional to the square root of time, and at a rate slower than in water (Tyree & Tammes, 1975). Interestingly, the rate of diffusion was modelled to the number of cells, not the physical distance (Tyree & Tammes, 1975). This suggests two things: that the rate-limiting step of movement is the flow through plasmodesmata, and that there is near-perfect mixing within a single cell before movement (Tyree, 1970). Therefore, it is thought there may be a passive movement of small molecules through plasmodesmata, aided by the active process of cytoplasmic streaming within each cell. While this may be the case for small molecules, it appears unlikely that large molecules such as proteins and mRNA, could translocate through plasmodesmata without the aid of chaperones, transporters, or unravelling making it an active process.

There is strong evidence for the translocation of some proteins being active, as movement can be directional and selective and a passive process is neither of these. Unidirectional movement is exemplified by KNOTTED1 (KN1), an endogenous transcription factor, which

shows movement only from the mesophyll to the epidermis and not vice versa (Kim *et al.*, 2003). Furthermore, it has been shown that KN1 translocation requires chaperones, and so at least the unfolding of KN1 in preparation for transport is an active process (Xu *et al.*, 2011). Selective (and directional) movement was shown more recently in the case of bHLH transcription factors (Lu *et al.*, 2018). Lu *et al*. compared the transport within a family of bHLH transcription factors, and demonstrated that the protein sequence, not size, was crucial for symplastic transport (Lu *et al.*, 2018). A passive process could not be sequence selective.

However, unidirectional movement is not the norm. Tyree and Tammes showed that in *Tradescantia* staminal hairs, the number of cells traversed by fluorescein (a small dye of similar mass to glucose) is the same "whether transport is from the base or the tip" (Tyree & Tammes, 1975). Moreover, most observations of GFP movement in model organisms are non-directional (Imlau *et al.*, 1999; Crawford & Zambryski, 2000; Dashevskaya *et al.*, 2008; Guseman *et al.*, 2010; Faulkner *et al.*, 2013). Even when GFP is coupled to functional proteins, such as *Colletotrichum higginsianum* effectors, non-directional movement was observed in *Nicotiana benthamiana* leaf tissue (Ohtsu *et al.*, 2021). On the other hand, Kitagawa and Fujita found that DENDRA2 (a fluorescent protein) moved in a directional and ATP-dependent manner in *Physcomitrella patens* (Kitagawa & Fujita, 2015). This suggests a possible difference in plasmodesmata translocation between species, or it could be due to directional bulk flow towards the growing tip in *P. patens* that is absent in *N. benthamiana*.

It has been suggested that Stokes radius (or the hydrodynamic radius, which is the radius of a hard non-ionic sphere that diffuses at the same rate as the molecule when in solution) is more important for movement than mass of small molecules as it defines the radius that must fit into the cytoplasmic sleeve (Terry & Robards, 1987). The same result was obtained for proteins with the exogenous expression of GFP (Dashevskaya *et al.*, 2008). Furthermore, this result can be extended to pathogen effector proteins using the result of Ohtsu *et al.* (2021), if mass is assumed to be proportional to Stokes radius. This would appear to confirm the idea of passive movement through plasmodesmata. However, passive movement of proteins still seems unlikely due to their large size, making it possible that there is non-selective active trafficking with a rate dependent on size.

Therefore, there appears to be multiple mechanisms of movement at plasmodesmata. In addition to passive movement of small molecules, there may be a non-selective size-dependent active mechanism of protein movement. This bears the hallmarks of a passive process but may require the active translocation or chaperoning of large molecules. On top

of these non-selective mechanisms, there is evidence of selective and directional movement at plasmodesmata, as well, which must be an active process. The development of new tools, e.g. tagged sucrose analogues, may also find the selective transport of small molecules in the future (De Moliner *et al.*, 2021).

In addition to mobility being dependent on size of a molecule, Crawford *et al.* showed that the availability of a protein, via its subcellular localisation, was also relevant to its mobility (Crawford & Zambryski, 2000). It was shown that both cytosolic- and nuclear-localised GFP were mobile, but ER- and actin-localised GFP were not (Crawford & Zambryski, 2000). Furthermore, a small amount of movement was detected of cytosolic 2×GFP, but not of nuclear-targeted 2×GFP. Thus, if a molecule is anchored or restricted in a cellular domain it is likely immobile and thus cell-autonomous.

In addition to the native mechanisms of symplastic transport, pathogens can manipulate plasmodesmata to increase cell-to-cell mobility (Waigmann *et al.*, 1994; Aung *et al.*, 2020; Tomczynska *et al.*, 2020; Ohtsu *et al.*, 2021). The classic example of this is the cell-to-cell spread of viruses. *Tobacco mosaic virus* MOVEMENT PROTEIN increases the SEL of plant cells to allow the passage of virus replication complexes (Waigmann *et al.*, 1994; Kawakami *et al.*, 2004). Ohtsu *et al*. (2021) demonstrated that *C. higginsianum* can alter cell-to-cell spread by at least three different mechanisms. One effector generically opened plasmodesmata for itself and another marker protein in a way that is similar to *Tobacco mosaic virus* MOVEMENT PROTEIN. However, other effectors only increased cell-to-cell flux for other proteins, or for themselves alone. Future work will follow up exactly how these proteins are altering plasmodesmal function.

1.3.4   Endogenous symplastic trafficking

The discussion so far has centred mainly around the movement of exogenous and presumably inert molecules. However, a wide range of endogenous molecules with physiological relevance move cell to cell. The first endogenous protein observed to have non-cell autonomous action was KN1, where the knotted phenotype could be reproduced from expression of *KN1* within just the middle mesophyll-bundle sheath layer of the leaf (Sinha & Hake, 1990). Moreover, it was then shown that the protein KN1 moved beyond the expression profile of its mRNA (Jackson *et al.*, 1994; Lucas *et al.*, 1995). In addition, KN1 also increases the SEL, allowing the movement of 20 kDa dextrans up from 1 kDa (Lucas *et al.*, 1995). Physiological importance of non-cell autonomous action has been elegantly shown for the transcription factor SHORTROOT (SHR) (Nakajima *et al.*, 2001). SHR is essential for

specifying the single cell layer of endodermis in the root. SHR mRNA is found solely in the stele, internal to the endodermis, whereas SHR-GFP is seen in the endodermis. SHR was shown to induce the endodermis-specifying transcription factor SCARECROW, and so acts in two ways: giving positional information, and determining cell fate (Nakajima *et al.*, 2001).

The ability of transcription factors to give positional information reiterates that symplastic movement of endogenous proteins can be directional. Another example of this is DEFICIENS, which is required for the correct patterning of *Antirrhinum majus* flowers. When expressed in the epidermis DEFICIENS cannot move inward towards the meristem, but can move in the opposite direction (Perbal *et al.*, 1996). Movement is not only directional, but can be selective; the observation that only a subset of bHLH transcription factors are translocated from cell to cell in the root (Lu *et al.*, 2018) identifies that mobility is specific and thus likely selective.

Endogenous movement is not limited to proteins: RNAs also move from cell to cell. The first example of this was again KN1, which can selectively transport its own mRNA through plasmodesmata (Lucas *et al.*, 1995). However, there is also a more general RNA movement pathway, as grafting experiments have shown that thousands of mRNA are mobile, including against the phloem from root-to-shoot (Thieme *et al.*, 2015). It is still a contentious issue as to whether mRNA movement is selective or not (Calderwood *et al.*, 2016; Zhang *et al.*, 2016). As well as mRNAs, small RNAs have been shown to move, such as micro RNAs in the phloem and small-interfering RNAs from the tapetum to the male meiocyte (Bari, 2006; Long *et al.*, 2021).

In addition to the movement of macromolecules such as nucleic acids and proteins, the movement of small molecules via the symplast has also been demonstrated. The movement of carbohydrates has been shown multiple times, in different ways to be symplastic. One example of this is the movement of ^{14}C- sucrose from the stele to the cortex in pea (*Pisum sativum*). ^{14}C- sucrose could not be washed out of the roots by high concentrations of either glucose, fructose or sucrose (Dick & ap Rees, 1975). Thus, the sucrose was being unloaded from the stele symplastically, otherwise the apoplastic washes would have removed the ^{14}C signal. Circumstantial evidence for the movement of photoassimilates symplastically come from a comparison of plasmodesmata in C3 and C4 plants. C4 photosynthesis requires high metabolic fluxes between bundle sheath and mesophyll cells. This requirement is correlated with a five-fold increase in pit field area in C4 plants at this interface, and a doubling of plasmodesmata density within the pit field, strongly implying that photoassimilate

movement is through the plasmodesmata (Danila *et al.*, 2016). The cell-to-cell movement of sugars has also been directly shown through tagged sucrose analogues (De Moliner *et al.*, 2021). As well as the movement of sugars, hormones move via plasmodesmata. This can be illustrated by auxin, which classically was thought to move solely apoplastically via PIN transporters (Barbez *et al.*, 2017). Recent work has taken both genetic and mathematical modelling approaches to show that the symplastic movement of auxin is required to explain the observe distributions of the hormone (Mellor *et al.*, 2020; Sager *et al.*, 2020). Beyond sugars and hormones, plasmolysis has also been shown to reduce the rate of ion export in maize (*Zea mays*) (Jarvis & House, 1970) and symplastic transport of ATP has also been shown plausible, using the fluorescent ATP analogue TNP-ADP (Cleland *et al.*, 1994).

Overall, therefore, symplastic movement has been observed of most biological classes of molecules, with few having no members which are able to move cell-to-cell, and so intercellular movement appears to be more of a rule than an exception. The physiological role of lipid transport has yet to be observed, but lipid transport may continue even when cytoplasm cannot move from cell to cell. The precise nature of the movement of larger molecules is yet to be established, with multiple signatures of movement observed. There is complex regulation of cell-to-cell movement, which is not simply explained by ultrastructure. Whether small molecules can be prevented from moving, while proteins can continue to transport is unknown. Experiments down this route may provide a handle to explore the nature of protein transport at plasmodesmata.

1.4 Composition of plasmodesmata

Both the observations made by electron microscopy and the complex regulation of movement through plasmodesmata imply that plasmodesmata are not simple structures. To fully understand how plasmodesmata function and are regulated, the protein and lipid constituents of plasmodesmata have been closely examined. Proteomics has led to an exponential increase in the number of proteins identified at plasmodesmata over the last 15 years, most notably the plasmodesmata proteome (Fernandez-Calvino *et al.*, 2011). A wide range of proteins have now been identified at plasmodesmata, from receptor-like kinases to structural proteins. Further, plasmodesmata have been recognised to have a unique lipid composition (Grison *et al.*, 2015a). Pharmacological or genetic disruption of plasmodesmal lipids prevents the targeting of proteins to plasmodesmata identifying lipids as critical to plasmodesmal function and structure. Finally, other macromolecules, such as callose, are also known to be associated with plasmodesmata, and are required for dynamic plasmodesmal responses.

Callose is a polysaccharide synthesised from glucose by a callose synthase complex. The constituents of the callose synthase complex remain a matter of debate, but potentially includes phragmoplastins and annexins (De Storme & Geelen, 2014). There is a constitutive basal plasmodesmata-associated callose level (Radford et al., 1998; Faulkner et al., 2009) and an increase in callose from this baseline is associated with reduced symplastic transport (Levy et al., 2007; Vatén et al., 2011).

1.4.1   Protein composition of plasmodesmata

1.4.1.1   Structural proteins

Several structural proteins have been localised to plasmodesmata. The first of these were actin and myosin, localised by protein immunolocalisation (Blackman & Overall, 1998). This is corroborated by *in vivo* experiments, where the depolymerisation of actin increased the SEL of plasmodesmata (White *et al.*, 1994). Correspondingly, actin-binding proteins have been subsequently immunolocalised to plasmodesmata (Van Gestel *et al.*, 2003; Faulkner *et al.*, 2009). Moving away from the cytoskeleton, several families of 'scaffolding' membrane associated proteins have been localised to plasmodesmata, namely: remorins, tetraspanins, centrins, C2-lipid binding proteins (also known as MCTPs) and reticulons.

Remorins are marker proteins for membrane nanodomains, with an unknown molecular function. Lipid rafts were originally defined biochemically by being resistant to non-ionic detergents (Simons & Ikonen, 1997; Pike, 2006; Jarsch & Ott, 2011). Remorins were one of the first proteins to be associated with lipid rafts, being used to label lipid rafts in potato (Raffaele *et al.*, 2007). Remorins cluster to negatively charged phospholipids and can produce nanodomains *in vitro* (Gronnier *et al.*, 2017; Legrand, 2020)*.* When expressed as a GFP translation fusion, a remorin localised to plasma membrane nanodomains, as well as plasmodesmata (Raffaele *et al.*, 2007). Moreover, multiple remorins have been shown to have a physiological role at plasmodesmata, either by hampering virus movement (Raffaele *et al.*, 2009; Perraki *et al.*, 2014) or altering grain setting in rice (*Oryza sativa*) (Gui *et al.*, 2014). S-acylation is required for remorin function in rice (Gui *et al.*, 2015), and a virus (rice strip virus) actively removes S-acylation from NbREM1, strengthening the idea S-acylation is key for function (Fu *et al.*, 2018). Further, calcium-dependent phosphorylation of remorins may be required to restrict viral movement (Perraki *et al.*, 2017). These data imply that remorins are recruited to and maintain lipid rafts, which are essential to plasmodesmal function. This ties in with the unique lipid composition of plasmodesmata, which is required for the recruitment of plasmodesmal proteins (see 1.4.2 for a discussion on plasmodesmata lipids).

Tetraspanins are a family of structurally-related proteins defined by having four transmembrane domains and have been localised to plasmodesmata via proteomics and subsequent GFP-tagging in *Arabidopsis thaliana* (Fernandez-Calvino *et al.*, 2011; Boavida *et al.*, 2013). Tetraspanins are thought to create or maintain nanodomains within membranes, termed tetraspanins webs (Zuidscherwoude *et al.*, 2015). In animals, tetraspanins webs have been found to be important for clustering of membrane proteins (Vogt *et al.*, 2002; Płóciennikowska *et al.*, 2015) and are required for the transition from unicellular to multicellular life (Huang *et al.*, 2005). No role has yet been ascribed to tetraspanins at plasmodesmata, but it is plausible that they are required for the recruitment or activity of signalling components at plasmodesmata.

A centrin-like protein was localised to plasmodesmata by immunolocalization (Blackman *et al.*, 1999). Centrins bind calcium and are contractile, leading to the suggestion that centrin might link the plasmodesmata plasma membrane to the desmotubule and control SEL in response to calcium (Erwee & Goodwin, 1983; Overall & Blackman, 1996). Thus, centrin-like proteins may provide a callose-independent mechanism for plasmodesmata closure.

The spokes observed at plasmodesmata were hypothesized to be a combination of myosin and centrin proteins, which were anchored to a spiral of actin around the desmotubule (Overall & Blackman, 1996; Baluška *et al.*, 2001). However, pharmacological treatments destabilising F-actin did not alter the spokes, meaning that growing actin is not required for the spokes (Nicolas *et al.*, 2017). Nonetheless, the spokes could still be made of stable actin, or arranged in an non-canonical configuration (Tilsner *et al.*, 2011).

Recently, C2 lipid-binding proteins have been proposed to be these spokes instead (Brault *et al.*, 2019). C2 lipid-binding proteins bridge the cytoplasmic sleeve: being tethered into the endoplasmic reticulum by transmembrane domains and attached to plasma membrane by C2 domains. Hence, removing the requirement for actin to tether the protein to the desmotubule, and so being consistent with the F-actin result above. However, C2 lipid-binding proteins were significantly enriched in Type I (no cytoplasmic sleeve) over Type II plasmodesmata, whereas the spokes were only observed in Type II plasmodesmata (Nicolas *et al.*, 2017; Brault *et al.*, 2019), reducing the likelihood of plasmodesmal spokes being constructed from C2 lipid-binding proteins. While the presence of C2 lipid-binding proteins in Type I plasmodesmata is hard to explain, the fact that there are fewer endoplasmic reticulum-plasma membrane tethering proteins in Type II plasmodesmata may explain why the cytoplasmic sleeve becomes looser in these plasmodesmata. C2 lipid-binding proteins

may not be specific to plasmodesmata, as their localisation to endoplasmic reticulum-plasma membrane contact sites is conserved in unicellular yeast (Brault *et al.*, 2019). This data concurs with the prevailing view that plasmodesmata are specialised endoplasmic reticulum-plasma membrane contact sites (Tilsner *et al.*, 2011; Pérez-Sancho *et al.*, 2016; Chen *et al.*, 2021). Membrane contacts sites occur in all eukaryotic cells and are sites of organelle cross-talk.

Unlike the other structural proteins mentioned so far, reticulons are thought to help the desmotubule form (Knox *et al.*, 2015). Reticulon proteins localise at curved membranes (Sparkes *et al.*, 2010), and themselves have the ability to reshape membranes (Tolley *et al.*, 2008, 2010). This may explain their localisation to the desmotubule, as it is one of the most constricted membranes in nature (Tilsner *et al.*, 2011). Knox *et al.* followed up the localisation of two reticulons identified at plasmodesmata by proteomics (Fernandez-Calvino *et al.*, 2011) and found that both reticulons localised to the cell plate in dividing cells, and to the endoplasmic reticulum and plasmodesmata when the cells were in interphase (Knox *et al.*, 2015).

1.4.1.2 *Signalling proteins*

A wide array of signalling proteins have been identified at plasmodesmata, leading some authors to term them 'a signalling hub' (Lee, 2015). Plasmodesmata-localised receptor proteins have been shown to respond to multiple signalling pathways including plant growth and pathogen defence (for a recent review see Vu et al., (2020)).

There has been in depth investigation into defence signalling at plasmodesmata. Receptors for both bacterial and fungal pathogens have been localised to plasmodesmata: FLAGELLIN SENSING 2 (FLS2) and LYSIN MOTIF (LysM) DOMAIN-CONTAINING GLYCOSYLPHOSPHATIDYLINOSITOL-ANCHORED PROTEIN 2 (LYM2), respectively (Faulkner *et al.*, 2013). The response of plasmodesmata to chitin is independent from canonical chitin responses, and so does not require CHITIN ELICITOR RECEPTOR KINASE 1 (CERK1). Thus, the specificity of plasmodesmata responses is generated at the receptor level. LYM2 has no intracellular signalling domain, as it is a glycosylphosphatidylinositol (GPI) anchored protein, and so requires two co-receptor proteins for fungal perception (Cheval *et al.*, 2020). I found that only one of these two proteins localise to plasmodesmata: LysM RECEPTOR-LIKE KINASE 4 (see Chapter 2).

Plasmodesmal specificity to bacterial perception, unlike chitin, does not rest with receptor specificity. FLS2 is required for all known cellular responses to flg22. CALMODULIN-LIKE 41

(CML41), a peripheral plasmodesmata protein (i.e. without a transmembrane domain), was shown to act downstream of FLS2 for plasmodesmata-specific responses only. CML41 is required for full immunity in *A. thaliana* to bacteria (Xu *et al.*, 2017). How these proteins, CML41 and LYM2/LYK4, signal downstream to elicit callose deposition is still unknown, but both fungal and bacterial responses result in plasmodesmal callose deposition. This suggests there may be integration of these two signalling cascades at some point upstream of callose synthesis.

PLASMODESMATA-LOCALISED PROTEIN 5 (PDLP5) is also localised at plasmodesmata and required for full immunity from bacteria in *A. thaliana*. However, it is not yet clear how PDLP5 signalling relates to FLS2 and CML41 signalling, and if they lie in the same pathway. It appears likely as they both have the same endpoint: callose synthesis. The downstream signalling of PDLPs is explored in greater depth in Chapter 4.

In addition to defence signalling, several developmental signalling pathways have been shown to act at plasmodesmata. For example, the receptor kinases CLAVATA 1 (CLV1) and ARABIDOPSIS CRINKLY 4 (ACR4) localise at plasmodesmata in a heteromeric complex within the root meristem and control meristem maintenance and differentiation (Stahl *et al.*, 2013). CLV1 and ACR4 also interact in the plasma membrane, but in a complex with a different composition to that of the complex formed at plasmodesmata. Another example, from the root meristem are receptor-like kinase STRUBBELIG and downstream C2 lipid-binding protein QUIRKY (Vaddepalli *et al.*, 2014). Both proteins localise to plasmodesmata and are required for root hair patterning, as well as having a flower defect when mutated. However, despite this clear role in development, the benefit from plasmodesmata-specific complexes and perception is not yet clear. It may be that plasmodesmata control is important for defining the meristem niche.

There is currently only one example of plasmodesmata proteins involved in abiotic stress tolerance. *A. thaliana* plants are more tolerant to salt stress when a DUF26 protein CYSTINE-RICH RECEPTOR-LIKE KINASE 2 (CRK2) is overexpressed (Hunter *et al.*, 2019). CRK2 responds dynamically to osmotic stress relocalising from an even plasma membrane distribution to a punctate, plasmodesmal one (Hunter *et al.*, 2019). Interestingly, this relocalisation phenomenon upon osmotic stress has been shown for several other proteins, which have no described role in abiotic stress (Grison *et al.*, 2019). As LYM2 also accumulates at plasmodesmata in response to chitin (Cheval *et al.*, 2020), these collective data might

indicate a general accumulation of signalling machinery at plasmodesmata in response to stress.

1.4.1.3 Enzymes associated to callose turnover at plasmodesmata

Callose is deposited around the neck of a plasmodesmata to reduce cell-to-cell flux (Radford *et al.*, 1998). Callose, a β-1,3-glucan of UDP-glucose, is synthesised by CALLOSE SYNTHASEs (CalS) of which 12 have been described in *A. thaliana* (Hong *et al.*, 2001). Six CalS have been determined to localise at plasmodesmata (Guseman *et al.*, 2010; Vatén *et al.*, 2011; Xie *et al.*, 2011; Cui & Lee, 2016; Saatian *et al.*, 2018). Callose is maintained at a basal level at plasmodesmata, through at least Cals8 and Cals10 (Chen *et al.*, 2009; Cui & Lee, 2016). When plants perceive a pathogen, the levels of callose dramatically increase within 30 minutes closing the plasmodesmata and isolating the cell (Xu et al., 2017).

In the other direction, callose is removed by β-1,3-glucanases (glycosyl hydrolase family 17), of which there are at least 50 grouped into three clades in *A. thaliana* (Doxey *et al.*, 2007). Several β-1,3-glucanases have been localised to plasmodesmata via proteomics (Levy *et al.*, 2007; Fernandez-Calvino *et al.*, 2011), and confirmed by GFP-tagged protein fusions (Levy *et al.*, 2007; Gaudioso-Pedraza & Benitez-Alfonso, 2014). Gaudioso-Pedraza *et al*. combined these two sets of data and, taking a phylogenetic approach, showed that all plasmodesmata localised β-1,3-glucanases (PDBG) were in the α clade (Gaudioso-Pedraza & Benitez-Alfonso, 2014). So far, four PDBGs have been identified: AtBG_PPAP (Levy *et al.*, 2007), and PDBG1-3 (Benitez-Alfonso *et al.*, 2013).

In addition to callose related enzymes, another class of proteins have been discovered to interact with callose at plasmodesmata: PLASMODESMATA CALLOSE-BINDING PROTEINs (PDCBs). PDCBs are a small three-member family, within the larger X8 domain family (Simpson *et al.*, 2009). They have no catalytic domain, and yet have been shown to alter callose levels. It is possible their mode of action is to bind and stabilise callose itself (Simpson *et al.*, 2009) but their specific role in plasmodesmal function is not yet known.

As of now, there is no known mechanistic pathway completely linking a stimulus to callose deposition (or prevention of degradation). However, another group of proteins, the PDLPs, has been found to upregulate callose deposits (Lee *et al.*, 2011b; Caillaud *et al.*, 2014). The PDLPs are an eight-member family in *A. thaliana*, containing two DOMAIN OF UNKNOWN FUNCTION 26 (DUF26) domains (Thomas *et al.*, 2008). These domains have canonical disulphide bridges that were initially linked to ROS responsiveness (Wrzaczek *et al.*, 2010b). Recent work has elucidated the crystal structure of the PDLP5 DUF26 domain, and concluded

that the disulphide bridges are more important for structure than ROS sensing (Vaattovaara *et al.*, 2019). PDLP overexpression lines show dramatic increases in callose, reduced symplastic flux and severe developmental phenotypes: for example PDLP1 overexpressing plants are dwarfed (Thomas *et al.*, 2008). These phenotypes demonstrate that misregulation of plasmodesmal function causes significant perturbation of growth.

How PDLPs enhance plasmodesmal callose is unknown. The transmembrane domain of PDLP1 was originally shown to be sufficient for plasmodesmata-localisation, as shown by YFP tagging (Thomas *et al.*, 2008). However, a recent study has demonstrated that the transmembrane domain does not target proteins to plasmodesmata, but rather is the site of interaction with other PDLP proteins and so helps retain PDLPs at plasmodesmata (Wang *et al.*, 2020). The transmembrane domain alone is sufficient to increase the deposition of callose (Caillaud *et al.*, 2014). This implies that the transmembrane domain is also responsible for downstream interactions or activates other PDLPs through interaction. As the PDLPs act highly redundantly molecular characterisation of their functions is difficult (Bricchi *et al.*, 2013).

1.4.2   Lipid composition of plasmodesmata

Plasmodesmal membranes have been shown to be significantly different in lipid composition to plasma membrane membranes, with an enrichment of sterols, sphingolipids, and a greater degree of saturation of phospholipids in *A. thaliana* (Grison *et al.*, 2015a). How specific membrane composition is generated and maintained at plasmodesmata is still an open question but it is known to be important for function.

Perturbation of the lipid composition of plasmodesmata has been observed to impair plasmodesmata development and function. Reduced plasmodesmal permeability and increased callose deposits were observed when suppressing a sterol carrier gene in cotton (*Gossypium hirsutum*) fibres, thus reducing the sterol content at plasmodesmata (Zhang *et al.*, 2017). This was validated by the addition of a sterol synthesis inhibitor, lovastatin, leading to a similar reduction in sterols and concomitant increase in callose deposition (Zhang *et al.*, 2017). Zhang *et al.* argued the increase in callose was due to reduced expression and targeting of a plasmodesmata-targeted β-1,3-glucanase (Zhang *et al.*, 2017). This follows from a similar finding by Grison *et al.* that PD-LOCALIZED CALLOSE BINDING PROTEIN 1 (PDCB1) and PD-LOCATED B-1,3-GLUCANASE 2 (PDBG2) are mistargeted upon the addition of sterol synthesis inhibitors in *A. thaliana* (Grison *et al.*, 2015a). All three of these plasmodesmal proteins have a C-terminal glycosylphosphatidylinositol (GPI) modification,

which has been shown to be important for plasmodesmal-targeting (Zavaliev *et al.*, 2016). The conclusion is that the sterol composition at plasmodesmata is essential for of GPI-targeting of proteins to plasmodesmata.

In addition to being important for GPI targeting to plasmodesmata, sterol enrichment is also thought to recruit remorins to plasmodesmata (Gronnier *et al.*, 2017). As aforementioned, remorins associate with lipid rafts (Raffaele *et al.*, 2007), which are known to be rich in sphingolipids and sterols. These lipid rafts can be assembled in response to salicylic acid (a defence hormone) in a remorin-dependent fashion (Huang *et al.*, 2019). This may provide a mechanism for the relocalisation of GPI-anchored proteins to plasmodesmata in response to stress (Cheval *et al.*, 2020). Interestingly, PDCB1 also co-extracted with lipid rafts, whereas other plasmodesmata-associated proteins such as PDLP1 localised with detergent soluble membranes (Grison *et al.*, 2015a).

While PDLPs did not co-elute with lipid rafts, their recruitment to plasmodesmata has been shown to dependent on sphingolipids within the plasmodesmata-plasma membrane (Liu *et al.*, 2020). Moreover, callose accumulation driven by PDLPs was shown to be sphingolipid-dependent (using the *sld1 sld2* double mutant, which cannot make the plasmodesmata-enriched t18:0-based sphingolipid) (Liu *et al.*, 2020). Sphingolipid biosynthesis has also been shown to be important for the maturation of plasmodesmata from Type I to Type II plasmodesmata with the *plm* mutant (which cannot make very-long-chain fatty acid sphingolipids) (Yan *et al.*, 2019). In both sphingolipid synthesis mutant lines (*sld1 sld2* and *plm*), cell-to-cell flux was greater. However, in the *sld1 sld2* line this was due to decreased callose levels from reduced PDLP targeting (Liu *et al.*, 2020). By contrast, in the *plm* mutant cell-to-cell flux was callose insensitive and increased flux is thought to be due to the altered plasmodesmata architecture (Yan *et al.*, 2019). Thus, the function of lipids at plasmodesmata are pleiotropic and as yet are poorly understood.

1.4.3   Cell wall composition around plasmodesmata

Plasmodesmata are membranous channels and so naturally have lipids that make up the membrane and proteins within the membrane. However, they do not exist in isolation and the cell wall environment around the channel is of great importance. The most notable example of a plasmodesmal cell wall component is callose, which is dynamically regulated at the neck of the pore (Rinne & Van der Schoot, 1998; Radford & White, 2001; Levy *et al.*, 2007; Vatén *et al.*, 2011; Wu *et al.*, 2018). Callose accumulation has been correlated with the reduced cell-to-cell flux between cells (Levy *et al.*, 2007) and direct callose induction has been

shown to prevent the symplastic trafficking of transcription factors (Vatén *et al.*, 2011). Callose interacts with cellulose (the major cell wall constituent) and deceases cell wall stiffness *in vitro*, perhaps aiding in the closure of plasmodesmata (Abou-Saleh *et al.*, 2018).

The cell wall environment around plasmodesmata, like lipids, is also differentially regulated (Knox & Benitez-Alfonso, 2014). There is more unesterified pectin and less cellulose in the cell wall in plasmodesmal pit fields(Casero & Knox, 1995; Faulkner et al., 2008). Pectin methylesterases, required for pectin de-esterification, have been immunolocalised to the plasmodesmata in flax and were found in the plasmodesmal proteome (Morvan *et al.*, 1998; Fernandez-Calvino *et al.*, 2011). Pectin itself is a group of complex polysaccharides, with the two major constituents being homogalacturonan and rhamnogalacturonan-I (RG-I). The pectin at plasmodesmata is enriched for RG-I with arabinan side chains, whereas galactan RG-I is specifically excluded (Roy *et al.*, 1997; Orfila & Knox, 2000; Giannoutsou *et al.*, 2013). In general, pectins modified in this manner are less rigid than cellulose, suggesting the wall is more flexible at plasmodesmata (Burton *et al.*, 2010).

Many other cell wall synthesis and remodelling proteins were proteomically identified at plasmodesmata (Fernandez-Calvino *et al.*, 2011) such as proteins controlling xyloglucans. This has not yet been followed up, but immunogold labelling has suggested that xylan is deposited at plasmodesmata in response to viral infection (Otulak-Kozieł *et al.*, 2018). There is plenty of scope for further research into the unique cell wall environment of plasmodesmata.

1.5   Diversity of plasmodesmata

Membrane-lined pores are not exclusive to the Viridiplantae (true plants) (Figure 1-3). Structures similar to plasmodesmata have been observed in fungi (Beckett, 1981; Bauer *et al.*, 2006; Van Peer *et al.*, 2009). In fact, the first reports of pores in filamentous fungi were made before that of plasmodesmata in higher plants (de Bary, 1866; Tangl, 1880). Phylogenetically closer, within Plantae, some brown algae have membrane-lined pores, with much larger pit plugs compared to a desmotubule (Lee, 1971; Schmitz & Srivastava, 1974; Terauchi *et al.*, 2012), suggesting that direct connections between the cytoplasm of adjacent cells is an common feature of multicellular organisms.

Within the Viridiplantae numerous types of plasmodesmata have been observed, both with and without a desmotubule (Fraser & Gunning, 1969; Floyd *et al.*, 1971; Stewart *et al.*, 1973a). The diversity of plasmodesmata and the phylogenetic pattern of their occurrence has

led to the suggestion that there are at least two, and up to five, independent evolutionary origins of plasmodesmata within Viridiplantae (Raven, 1997, 2005).

Importantly, however, a single origin of plasmodesmata is postulated for all Embryophytes (land plants). In line with this, bryophytes such as the liverworts *Asterella wilmsii* and *Monoclea gottschei*, hornwort *Notothylas orbicularis*, and moss *Sphagnum fimbriatum* all contain plasmodesmata with desmotubules (Ligrone & Duckett, 1994, 1998; Cook *et al.*, 1997). Moreover, even the basal streptophytic algae *Chara corallina* is thought to contain a desmotubule (Brecknock *et al.*, 2011). This conservation of structure substantiates the hypothesis that all Embryophytic plasmodesmata are related. Consequently, this implies that plasmodesmata have a core, shared composition that defines their ontogeny and function.
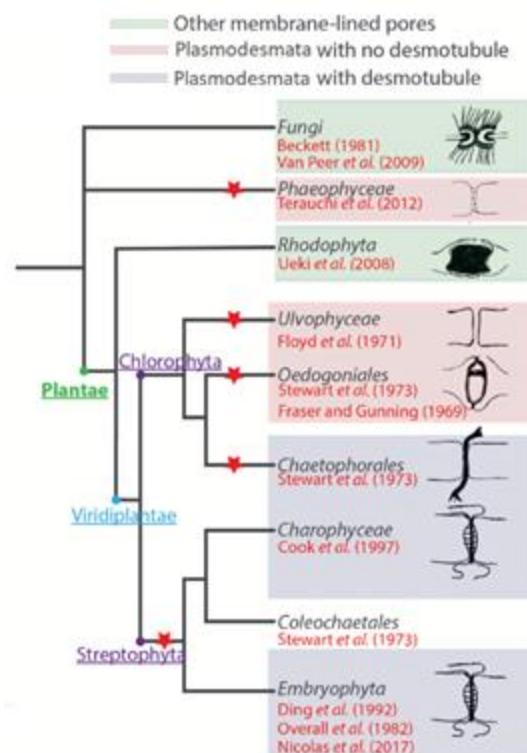
Figure 1-3 Multiple origins of plasmodesmata within Viridiplantae
Phylogenetic tree and suggested independent evolutionary events (red stars) drawn after Raven *et al.* (2005). Shading indicates whether observed plasmodesmata have desmotubules (blue) or not (red). Green shading indicates another type of membrane-lined pore: septal pores in fungi and pit-plugs in red algae (Rhodophyta). Cartoons of the ultrastructure of plasmodesmata as viewed by electron microscopy illustrate the presence or absence of a desmotubule. Reproduced and modified with permission from Nicolas et al. (2018).

1.6 Aims and objectives of this thesis

1. Plasmodesmata are nanoscopic and embedded within the cell wall making them intractable to work with. Strides have been made in identifying components of plasmodesmata over the last decade, especially with the aid of proteomic techniques.

The first aim of this study is to extend the utility of plasmodesmata purifications, by developing a method of plasmodesmata extraction for mature tissue. This will allow the biochemical localisation of transiently expressed proteins.

2. As mentioned above, I hypothesise that plasmodesmata have a core, evolutionarily conserved machinery. I aim to investigate this by comparison of *Arabidopsis thaliana* and *Physcomitrella patens* plasmodesmal proteomes. I will use my method of plasmodesmata extraction to generate novel plasmodesmata proteomes from mature tissue. The proteomes will be used to refine current plasmodesmal proteomes, as well as to find conservation of plasmodesmata proteins across Viridiplantae.

3. PDLP proteins have remained enigmatic in the signalling pathway to callose deposition, despite having a clear role in plant defence. I will exploit them to identify novel components that regulate callose synthesis. I will take a multi-faceted approach to find novel interactors of PDLP proteins. This study will use genetic and cell biology techniques to characterise a downstream interacting protein of PDLP1.

Ultimately, the aim of this thesis is to identify novel components of plasmodesmata and build further understanding of plasmodesmal function.

2 Developing a method for the extraction of mature plasmodesmata

2.1 Introduction

2.1.1 Plasmodesmata extractions

Attempts to purify plasmodesmata began two decades before the first plasmodesmal proteome became available (Monzer & Kloth, 1991; Yahalom *et al.*, 1991). Initial techniques sought to purify plasmodesmata away from other membranes and organelles, whilst leaving the plasmodesmata within the cell wall (Kotlizky *et al.*, 1992; Turner *et al.*, 1994; Ritzenthaler *et al.*, 2000). This technique was successfully used to find some of the first plasmodesmata proteins (Epel *et al.*, 1996; Kishi-Kaboshi *et al.*, 2005; Lee *et al.*, 2005; Levy *et al.*, 2007). Moreover, proteomics on this cell wall fraction (Bayer *et al.*, 2006) aided in the discovery of the canonical PLASMODESMATA LOCALISED PROTEINs (PDLPs) and PLASMODESMATA CALLOSE BINDING PROTEINs (PDCBs) (Thomas *et al.*, 2008; Simpson *et al.*, 2009).

This technique was extended to generate "wall-free plasmodesmata" by the use of cellulase (Epel *et al.*, 1995). The cell wall purification separates all other cellular membranes from the cell wall and embedded plasmodesmata. The cellulase digest liberates the plasmodesmal membranes from the cell wall, creating a plasmodesmata-specific membrane fraction in the sample. Epel and colleagues (1995) note that "the selection of the biological material is of utmost importance in the successful isolation of a wall-free plasmodesmatal fraction", and so explain that etiolated maize mesocotyls were chosen for two reasons: the lack of secondary walls and the lack of chloroplasts; both of which interfere with plasmodesmata extraction. For the same reasons, *Arabidopsis thaliana* suspension cells were explored as a source of plasmodesmata material, as they have uniform, large cells which are easily disrupted with, an assumed single-type of, simple plasmodesmata (Bayer *et al.*, 2004).

Suspension cell cultures were then used for the first plasmodesmal proteome, released in 2011 (Fernandez-Calvino *et al.*, 2011), and the method used has been described several times since, with no significant changes (Grison *et al.*, 2015b; Faulkner & Bayer, 2017). The same technique has been used multiple times in the literature, namely to generate a plasmodesmata proteome for *Populus trichocarpa* and to produce a refined list of plasmodesmal proteins using semi-quantitative proteomics (Leijon *et al.*, 2018; Brault *et al.*, 2019) (see 3.1.4 for more details on attempted to characterise the protein composition of

plasmodesmata). Notably, however, there were no substantial alterations to the method described by Fernandez-Calvino *et al.* (2011).

More recently, a group successfully isolated plasmodesmata from mature 4-week-old *A. thaliana* rosette leaves (Liu *et al.*, 2020). Liu *et al.* (2020) used this method to compare the sphingolipid profiles of the plasmodesmata and plasma membrane, concluding that phytosphinganine modifies the levels of PDLP5 at plasmodesmata. Liu *et al.* (2020) used the same buffers and conditions as described in Grison *et al.* (2015), with the only modification to the grinding of tissue. As mature rosette tissue was used instead of suspension cells, leaves were ground in liquid nitrogen and then ultrasonnicated instead of being passed through a liquid nitrogen bomb.

Following this, our group published a different method for the isolation of plasmodesmata from mature tissue (*Nicotiana benthamiana* leaves) (Cheval *et al.*, 2020), which was developed in this chapter. Similarly, to Liu *et al.* (2020) a liquid nitrogen bomb could not be used, so tissue was ground and sonicated. However, in addition to the steps Liu *et al.* (2020), I passed the crude extract through a cell disrupter to break apart the cell walls more fully. The method described in Cheval *et al.* (2020) also uses simplified buffers compared to prior work, for example the crude extraction buffer does not cotain glycerol, EDTA, nor manitol. Further, I used detergents to aid in the removal of chloroplast contaminants from the cell walls. The essence of the extraction stays true to that of Epel *et al.* (1995): the purification of cell walls followed by the enzymatic release of plasmodesmal membranes.

2.1.2   Aims

This chapter describes the development of a plasmodesmata purification protocol from mature leaf tissue and an example a test case, both published in Cheval *et al.* (2020).

I aimed to optimise the established plasmodesmata extraction protocol to that of mature *Nicotiana benthamiana* tissue (Grison *et al.*, 2015b). To do this, I considered the use of different protein tags to find the most stable throughout the protocol. Further, I aimed to utilise the protocol to biochemically localise proteins to plasmodesmata. This is particularly useful when proteins have an even plasma membrane localisation by fluorescence microscopy, as this is indistinguishable from a plasma membrane and plasmodesmata dual localisation.

The work in this chapter was technically assisted by a Master's student, Dan Zhang, from the Lozano-Duran research group (Shanghai Center for Plant Stress Biology, Chinese Academy of

Sciences Center for Excellence in Molecular Plant Sciences, Shanghai, China). Specifically, she helped me to produce: Figure 2-5, Figure 2-6, Figure 2-8 and Figure 2-9.

2.2   Materials and Methods

2.2.1   Plant material

Nicotiana benthamiana plants were grown on soil in long days (16h light/8h dark) at 22°C and 85% humidity. The light intensity was 200 μmol/m^2/s.

2.2.2   Gene synthesis and cloning

2.2.2.1   Golden Gate cloning

All cloning carried out in this thesis was by Golden Gate cloning, using the principals outlined by Engler *et al.* (2008, 2009). The overhangs used follow the standard as set by Engler *et al.* (2014). Level 0 modules contain gene parts (promoters, coding sequences, terminators, etc). All level 1 and level 2 modules are binary vectors suitable for expression in plant tissue by *Agrobacteria tumifaceins*. Level 0 modules, level 1 acceptors, level 1 components, and level 2 acceptors were provided by the TSL SynBio platform (Norwich, UK).

Level 1 and 2 modules were produced as recommended by TSL SynBio. Level 1 components were made in a one-pot 15 μL reaction containing: 50 ng level 1 acceptor and 100 ng of each level 0 module, 1.5 μL 10 × Bovine Serum Albumin (New England Biolabs), 1.5 μL 10 × T4 DNA ligase buffer (New England Biolabs), 1 U *BsaI*-HFV2 (New England Biolabs) and 200 U T4 DNA ligase (New England Biolabs). Level 2 components were made in a one-pot 15 μL reaction containing: 50 ng level 2 acceptor and 100 ng of each level 1 module, 1.5 μL 10 × Bovine Serum Albumin (New England Biolabs), 1.5 μL 10 × T4 DNA ligase buffer (New England Biolabs), 5 U *BbsI* (Thermo Scientific) and 200 U T4 DNA ligase (New England Biolabs). The assembly reaction was carried out in a Nexus G2 (Eppendorf) thermocycler with the following programme: 20 seconds at 37°C; followed by 30 cycles of 3 minutes at 37°C and 4 minutes at 16°C; finished by 5 minutes at 50°C and 5 minutes at 80°C.

2.2.2.2   Bacterial transformation

Escherichia coli (DH5α) and *Agrobacteria tumifaceins* (GV3101 (pMP90)) were transformed by electroporation (200 Ω, 25 μF, 2.5 kV) with 100 ng of plasmid or 1 μL of Golden Gate product. *E. coli* were then allowed to recover for 1 h at 37°C and *A. tumifaceins* for 2 h at 28°C in Super Optimal broth with Catabolite repression (SOC) medium (Hanahan, 1983). Bacteria were plated on to lysogeny broth (LB) agar containing the appropriate antibiotics (Bertani, 1951). The antibiotic concentrations used were: kanamycin (50 μg/mL), rifampicin (50 μg/mL), spectinomycin (50 μg/mL), carbenicillin (100 μg/mL). *E. coli* were grown for 12 h at 37°C and *A. tumifaceins* for 36 h at 28°C.

E. coli positive colonies were grown in liquid culture overnight at 37°C in LB containing the appropriate antibiotics. Plasmids were extracted using the QIAprep Spin Miniprep Kit (Qiagen) following the manufacturer's instructions. The plasmid was verified by restriction digests using *EcoRI* and *HindIII*. This was done in a 10 µL reaction containing: 100 ng plasmid, 5 U enzyme, 1 µL 10 × buffer (as provided by the manufacture). The reaction was incubated at 37°C for one hour before being run on a 1% agarose tris-borate EDTA gel containing 1% ethidium bromide (Melford and Severn Biotech).

2.2.2.3   Synthesis and cloning of CML41

The coding sequence of CALMODULIN-LIKE41 (AT3G50770, CML41) was domesticated by synonymous substitutions of *BsaI* and *BbsI* restriction sites. The domesticated sequence was synthesised as a level 0 module with overhangs AATG and TTCG by ENSA (now Crop Engineering Consortium) as EC00475 in the pTwist30 backbone with kanamycin resistance.

Two level 1 constructs were constructed: 35s::CML41-GFP and 35s::CML41-HA. Both level 1 reaction contained a level 1 position 3 acceptor (pICH47751), a short 35s promoter and 5' UTR (pICH51277), CML41 (EC00475), and a 35s terminator (pICH41414). The CML41-GFP reaction also included a C-terminal eGFP tag module (overhangs TTCG and GCTT). The CML41-HA reaction also included a C-terminal 6×HA tag module (pICSL50009). These plasmids are available from Addgene deposited by Sylvestre Marillonnet & Nicola Patron (Weber *et al.*, 2011; Engler *et al.*, 2014). Published plasmids used in this thesis and their Addgene reference is given in Table 2-1. The sequences for unpublished modules, including CML41 and eGFP used in this chapter, are provided in Table S1.

Table 2-1 Published Golden Gate plasmids used in this thesis

Name	Addgene ID	Level	Resistance	Overhangs		Description
pICH51277	50267	0	Spectinomycin	GGAG	AATG	Short 35s promoter and 5' UTR
pICH41414	50337	0	Spectinomycin	GCTT	CGCT	35s 3' UTR and terminator
pICSL50004	50316	0	Spectinomycin	TTCG	GCTT	mCherry
pICSL30003	50304	0	Spectinomycin	CCAT	AATG	mCherry
pICSL50009	50309	0	Spectinomycin	TTCG	GCTT	6×HA
pICSL13002	50266	0	Spectinomycin	GGAG	CCAT	Short 35s promoter and 5' UTR
pICSL01003	47989	0	Spectinomycin	TTCG	GCTT	Acceptor for C-terminal tag modules
pICSL01002	47986	0	Spectinomycin	CCAT	AATG	Acceptor for N-terminal tag modules
pICSL01005	47996	0	Spectinomycin	AATG	TTCG	Acceptor for CDS (no stop) modules
pICH41308	47998	0	Spectinomycin	AATG	GCTT	Acceptor for CDS (stop) modules
pICSL11024	51144	1	Carbenicillin	TGCC	GCAA	Acceptor Forward P1
pICH47751	48002	1	Carbenicillin	ACTA	TTAC	Acceptor Forward P3
pICH47822	48009	1	Carbenicillin	ACTA	TTAC	Acceptor Reverse P3
pICH54022	48066	1	Carbenicillin	GCAA	ACTA	Dummy P2
pICH41766	48018	1	Carbenicillin	TTAC	GGGA	End P3
pICSL4723	86173	2	Kanamycin	NA	NA	Acceptor

2.2.3   Protein expression

Agrobacteria tumifaceins (GV3101 (pMP90)) were grown at 28°C shaking at 180 rpm in appropriate antibiotics overnight (Koncz *et al.*, 1994). Once turbid, cells were pelleted and washed in $MgCl_2$ (10 mM) twice. Washed cells were resuspended in $MgCl_2$ (10 mM) and acetosyringone (100 nM) in the same volume as grown in, so that OD_{600} > 1. Appropriate co-cultures of equal volumes of agrobacteria were prepared for protein expression, including the genes of interest and P19 (a silencing suppressor). These cultures were infiltrated with a blunt 1 mL syringe into the abaxial side of third and fourth oldest leaves of four- to five-week-old *N. benthamiana* leaves. Leaves were harvested 2 days post infiltration.

2.2.4   Nuclei Extraction

Nuclei were extracted from expanded *Nicotiana benthamiana* leaves in a modified protocol from (Folta & Kaufman, 2006; Sikorskaite *et al.*, 2013; Pendle & Shaw, 2017), as detailed (Figure 2-3A). Leaf tissue (1.5 g) was cut into 1 cm^2 squares, and homogenised using gentleMACs™ Octo Dissociator (MACS) in 3 mL NIB (10 mM MES pH 5.5, 0.2 M sucrose, 2.5 mM EDTA, 0.1 mM spermine, 0.5 mM spermidine, 2.5 mM DTT, 10 mM NaCl, 10 mM KCl, cOmplete™ ULTRA protease inhibitors (Sigma)). The lysate was filtered through a 30 µm mesh. Triton X-100 (10% v/v) was added dropwise to the filtrate to a final concentration of 0.5% (150 µL). The solution was agitated at 4°C for 10 mins. 1 mL of the solution was layered on top of 1 mL of a 2 M sucrose cushion. This was spun for 10 mins at 2000g at 4°C. The pellet was washed twice in 600 µL NIB and spun as before for 5 mins. Finally, the pellets from all the cushions were combined and resuspended in a final volume of 150 µL.

2.2.5   Plasmodesmata Extraction

Plasmodesmata were extracted from adult tissue in a modified protocol from (Grison *et al.*, 2015b), as detailed in Figure 2-1. Four expanded 5-week-old *N. benthamiana* leaves, transiently expressing the desired construct(s), were ground in liquid nitrogen to a fine powder and suspended in 15 mL extraction buffer (EB: 50 mM Tris-HCl pH 7.5, 150 mM NaCl, 1 × cOmplete™ ULTRA protease inhibitors (Sigma), 1 mM PMSF, 1% (w/v) PVP-40kDa (Sigma)). The sample was ultrasonicated for 1 minute in six 10-second pulses with a five second pause between each pulse (Soniprep 150 Plus, MSE). The sample was passed twice through a high-pressure homogenizer (EmulsiFlex™-B15, Avestin) at 80 PSI. Triton X-100 (10% v/v) was added dropwise to the resultant homogenate to a final concentration of 0.5% (v/v). Cell walls were aggregated by a 5-minute spin at 400*g*. The cell wall pellet was washed three times with EB (15 mL) and spun as before. The cleaned cell wall pellet was incubated in an equal volume of cellulase buffer (CB: 20 mM MES-KOH pH 5.5, 100 mM NaCl, 2% w/v

Cellulase R-10 (Yakult Pharmaceutical Co., Ltd., Japan), 1 × cOmplete™ ULTRA protease inhibitors (Sigma), 1 mM PMSF) for 1 h at 37°C, 200 rpm. Undigested cell wall was removed with a 15-minute spin at 5,000g. The supernatant was kept. The pellet was washed with CB (3 mL) and spun again as before. The two supernatants were combined and ultracentrifuged at 135,000g for 1 h. The membrane pellet (containing plasmodesmal membranes) was resuspended in buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 5 mM DTT, 1×cOmplete™ ULTRA EDTA-free protease inhibitors (Sigma), 1 mM PMSF, 0.2% (v/v) IPEGAL®CA-630 (Sigma)).

2.2.6 Western Blots

Samples were prepared for SDS-PAGE by addition of 6 × loading dye (120 mM Tris-HCl pH 6.8, 50% glycerol, 6% SDS, 0.05% bromophenol blue) (Laemmi, 1970). Samples were then heated at 95°C for 5 mins or 70°C for 15 mins and centrifuged at 10,000g for 5 mins before loading.

The Bio-Rad Mini-PROTEAN II apparatus (Bio-Rad Laboratories) was used for casting and running gels. Gels were 1 mm thick, 8 cm wide and 7 cm long, with 6 cm separating gel and 1 cm stacking gel. The separating gel was made of 10% (w/v) acrylamide:bis-acrylamide (37.5:1), 0.1% SDS (w/v), and 375 mM Tris-HCl pH 8.8. The stacking gel was made of 4% (w/v) acrylamide:bis-acrylamide (37.5:1), 0.1% SDS (w/v), and 125 mM Tris-HCl pH 6.8. Both gels were polymerised by the addition of 0.05% (w/v) ammonium persulfate and 0.1% (v/v) TEMED per gel. The stacking and separating gels were allowed to polymerise for a minimum time of 30 min, with the separating gel polymerising under a layer of deionised water.

Gels were run in running buffer (192 mM glycine, 25 mM Tris and 0.1% (w/v) SDS) at a constant 120V, until the dye front ran to the bottom of the gel. The proteins were transferred to Immun-Blot® PVDF 0.2 μm (Bio-Rad Laboratories) with a Mini Trans-Blot Electrophoretic Transfer Cell (Bio-Rad Laboratories) in 25 mM Tris, 192 mM glycine, 20% (v/v) methanol at a constant 150 mA for 90 mins. Membranes were blocked for 1 hour in 5% milk (w/v) in TBST at room temperature or overnight at 4°C.

Blots were incubated with diluted 1:5,000-20,000 primary antibodies at room temperature for one hour. For non-conjugated antibodies, blots were washed three times with 5% milk (w/v) in TBST and incubated for 1 h with 1:5,000 secondary antibody. Blots were washed with TBST three times for a minimum of 5 mins. Chemiluminescence of SuperSignal™ West Femto Maximum Sensitivity Substrate (Thermo Scientific) was imaged with ImageQuant LAS 500 (GE Healthcare Life Sciences). Alkaline phosphatase signal was detected by equilibrising the blot

in buffer (150 mM Tris-HCl pH 9.5, 50 mM MgCl$_2$, 100 mM NaCl) and then incubating the blot with 1:1 NBT:BCIP at room temperature.

2.2.7   Calcofluor staining

Calcofluor white M2R (F3543, Sigma) stock was made by dissolving 100 mg powder in 10 mL H$_2$O with a base. Calcofluor white M2R stock was added in a 1:10 ratio to samples to be imaged. 20 µL of sample was loaded onto a slide inside a chamber made of vacuum grease. Confocal microscopy was carried out on a ZEISS LSM800 with a 20 × objective (PLAN APOCHROMAT 20 ×/0.8). Calcofluor white was excited at 405 nm with a UV laser and collected at 430 – 470 nm.

2.2.8   Plasmids used in Chapter 2

The plasmids used in this chapter for transient expression in Nicotiana benthamiana are outline in Table 2-2. pICSL11016 was made by Mark Youles using: pICH47742 (Addgene ID #48001), pICH45173 (#50259) pAGM5355 (#50296), pICH41531 (#50321), pICH41421 (#50339).

Table 2-2 Plasmids used for transient expression in *Nicotiana benthamiana* in Chapter 2

Name	Backbone	Construct	Reference
P19	pCB301	35s::P19	(Win & Kamoun, 2004)
Chloroplastic GFP (pICSL11016)	pICH47742	AtRbcs1B::RbcS-GFP	Mark Youles
H2A-GFP	pMDC7	OLexA::H2A-GFP	Yvonne Stahl (Schatlowski *et al.*, 2010)
C4-GFP	pGWB5	35s::C4-GFP	(Rosas-Diaz *et al.*, 2018)
CML41-GFP	pICH47751	35s::CML41-eGFP	This thesis
LYK4-GFP	pB7FWG2.0	35s::LYK4-eGFP	(Cheval *et al.*, 2020)
PDLP1-HA	pEarleygate301	35s::PDLP1-HA	(Thomas *et al.*, 2008)
PDLP5-HA	pEarleygate301	35s::PDLP5-HA	Christine Faulkner
C4-HA	pGWB514	35s::C4-3×HA	(Rosas-Diaz *et al.*, 2018)
CML41-HA	pICH47751	35s::CML41-6×HA	This thesis
LYK4-HA	pGWB14	35s::LYK4-3×HA	(Cheval *et al.*, 2020)
LYK5-myc	pGWB17	35s::LYK5-4×myc	Xiaokun Liu
LYK4-myc	pGWB17	35s::LYK4-4×myc	Xiaokun Liu
LYK5-HA	pGWB14	35s::LYK5-3×HA	(Cheval *et al.*, 2020)

2.3   Results

Plasmodesmata have previously been extracted from *Arabidopsis thaliana* cultured cells (Grison *et al.*, 2015b). I modified this protocol to work with mature *Nicotiana benthamiana* tissue. The final protocol is outlined in Figure 2-1. This chapter moves systematically through the different stages of extraction, and what was learnt at each step. Broadly, there are four key parts to the extraction: homogenisation of the tissue, purification of the cell wall, digestion of the cell wall and collection of the plasmodesmata membranes. Each had to be optimised and adapted to a mature tissue protocol.
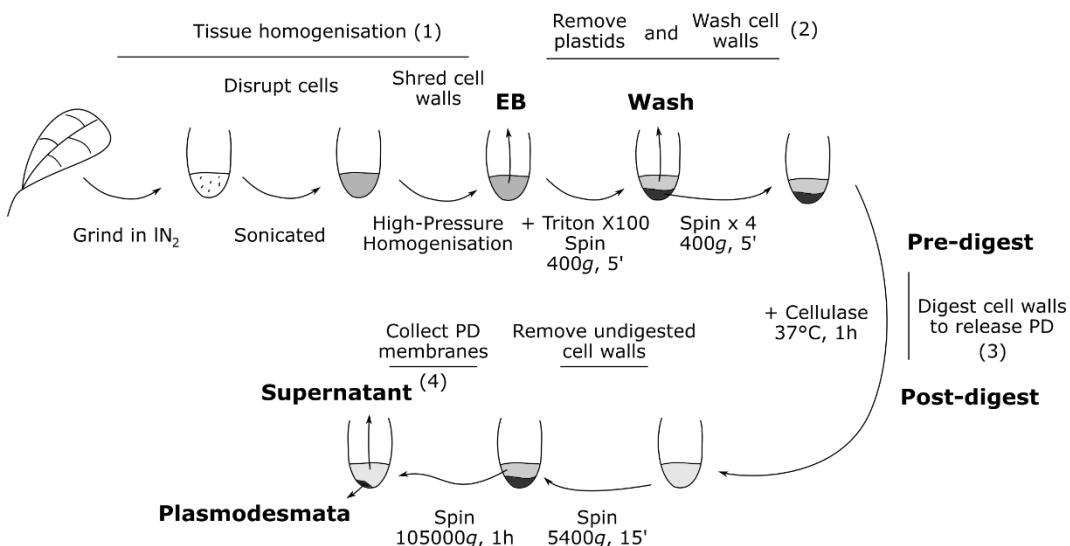
Figure 2-1 Schematic for the extraction of plasmodesmata from mature tissue.
Plasmodesmata extraction can be grouped into four stages: tissue homogenisation (1), cell wall purification (2), cell wall digestion (3), and collection of plasmodesmata membranes (4). Tissue is homogenised using by grinding in liquid nitrogen (lN$_2$) followed by sonication and two passes through a cell disrupter to give uniform small wall fragments and disrupt nearly all cells. Cell walls are washed with Triton X-100 to remove chloroplast contaminants and low-speed centrifugation is used to collect the cell wall fraction. Walls are digested at 37°C for one hour with Cellulase Onozuka R10 (Yakult Pharmaceutical Co., Ltd., Japan). The supernatant is collected from the digestion containing plasmodesmata membranes. The plasmodesmata membranes are collected by high-speed ultracentrifugation. Diagnostic fractions are collected throughout the protocol. EB – Extraction Buffer – a total extract of all proteins; Wash – cellular contents not in the cell wall fraction; Pre-digest – purified cell walls; Post-digest – purified cell walls after treatment with cellulase; Supernatant – proteins released by cellulase treatment not attached to plasmodesmal membranes; Plasmodesmata – purified plasmodesmata membranes. The protocol is modified from Grison *et al.* (2015) and published in Cheval *et al.* (2020).

2.3.1   Homogenisation of tissue

As Epel *et al.* noted in 1995, "the selection of the biological material is of utmost importance" and so selected tissue without secondary cell walls nor chloroplasts. For a similar reason, Grison *et al.* (2015) chose to use *A. thaliana* suspension culture cells, as they are readily friable, unlike mature tissue which is "often resistant to breakage" (Bayer *et al.*, 2004; Grison *et al.*, 2015b). However, mature *N. benthamiana* leaf tissue has both secondary cell walls and chloroplasts (for chloroplast removal see 2.3.2).

To fully homogenise mature leaf tissue, I devised a three-step approach. I modified the published plasmodesmata extraction protocol to no longer include starting homogenisation with a liquid nitrogen bomb, as this equipment is only suitable for liquids (Grison *et al.*, 2015b). First, frozen mature tissue was ground into a powder in liquid nitrogen. Then the tissue was made into a viscous suspension with extraction buffer ("EB"; see 2.2.5 for the buffer composition; see 2.3.4 for a consideration of extraction buffers) and sonicated. Finally, the tissue was passed twice through a cell disruptor. The rationale for this is that the grinding breaks the tissue down into a powder which can be suspended. Sonication disrupts any remaining cells which are still joined into clusters. Finally, the single cells are disrupted by the cell homogeniser. This rationale was validated by calcofluor staining of cell walls at the different stages of fractionation (Figure 2-2). Leaf tissue after grinding in liquid nitrogen contains broken cells, as well as small clusters of cells. The small clusters of cells are broken apart by sonication. The resulting individual cells and large fragments are broken down into small cell wall fragments by homogenisation (Figure 2-2). The size of cell wall fragments are similar to those derived from *A. thaliana* suspension cells ($30 - 100$ µm) (Grison *et al.*, 2015b).
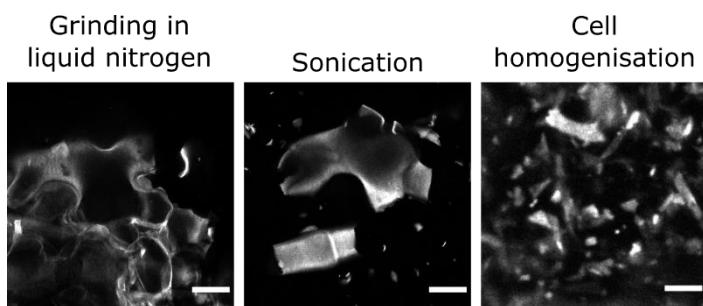
Figure 2-2 Mature tissue requires three stages of homogenisation to give cell wall fragments
Cell walls of *N. benthamiana* were stained with calcofluor white M2R (F3543, Sigma) after each step of mature tissue homogenisation. Grinding in liquid nitrogen produced clusters of cells, which were disrupted into single cells or large fragments by sonication. Small cell wall fragments resulted after passing the solution through a cell homogeniser. Scale bar = 25 µm.

2.3.2   Purification of cell walls

The second hurdle to plasmodesmata extraction from mature tissue is chloroplast contamination, as chloroplasts take up roughly 5% cell volume. Moreover, chloroplastic RuBisCO is one of the most predominant proteins in plant cells with up to five orders of magnitude more protein molecules than other protein constituents (Malcevschi & Marmiroli, 2012). This equates to about 50% of soluble protein, or roughly 5% of leaf total dry mass (Rowland-Bamford *et al.*, 1991; Bar-On & Milo, 2019). When developing a different protocol, nuclear purification, I found that detergents can be used to selectively remove chloroplasts. First, I will outline that technique, and then show how it was cross applied to plasmodesmata purifications.

2.3.2.1 Triton X-100 can be used to selectively remove chloroplasts

CALMODULIN-LIKE 41 (CML41) has been published to localise to plasmodesmata (Xu *et al.*, 2017). However, when expressed transiently in *N. benthamiana* or stably in *A. thaliana* I observed that CML41-GFP also had a nuclear-cytoplasmic localisation. I was interested in the nuclear fraction of CML41 specifically and so I aimed to purify intact nuclei from mature *N. benthamiana* leaves. As with most extractions from plant tissue, contamination from chloroplasts had to be overcome. While developing this protocol, I came across the idea of using detergents to selectively remove chloroplasts, leveraging their relatively thin membranes. Here, I describe how this was done in the context of nuclei purification and in 2.3.2.2 how it was then applied to the plasmodesmata purification from mature leaf tissue.

I started with the protocol described by Pendle and Shaw (2017), which uses a Percoll gradient to purify nuclei. I made two modifications to this protocol. The first, was to selectively remove chloroplasts from nuclei by exploiting their differing membrane properies. A low concentration of detergent (0.5% Trition X-100) is sufficient to permeabilise chloroplastic membranes, but not nuclear membranes (Folta & Kaufman, 2006). Note that the concentration of detergent required is species dependent (Sikorskaite *et al.*, 2013). Pendle and Shaw (2017) recommend using root tips or protoplasts, so as to avoid chloroplastic contamination, which was not possible in this case.

The second alteration was during the purification of nuclei, by centrifuging nuclei through a 2 M sucrose cushion. This was recommended by a Sigma technical bulletin (Product code: CELLYTPN1). Purification through a cushion offers a compromise between purity (as given by Percoll gradient extraction) and yield (offered by differential centrifugation). However, by combining the sucrose cushion with the selective disruption of chloroplast membranes a greater yield can be obtained with similar purity. This protocol yielded intact nuclei with no other fluorescent contaminants (i.e. chloroplasts). However, there was non-fluorescent contamination that is likely be starch granules (Figure 2-3).
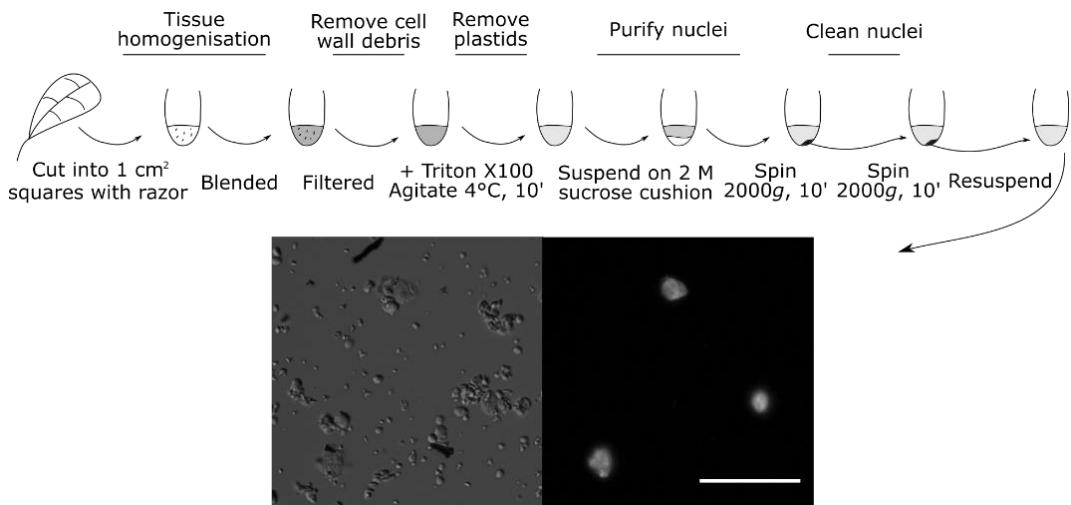
Figure 2-3 Schematic of a nuclear extraction protocol for intact nuclei.
Intact nuclei were extracted from mature *N. benthamiana* tissue following the shown protocol based off of Pendle and Shaw (2017). Nuclei were stained with DAPI and imaged (right) showing whole nuclei. Non-fluorescent contamination is observed with transmitted light (left). Scale bar = 50 μm.

To demonstrate that this was not proteinaceous contamination, and that chloroplasts were successfully removed I tracked a nuclear and chloroplast protein through the extraction. I transiently transformed *N. benthamiana* to express HISTONE 2A (H2A), a nuclear protein, and chloroplast-targeted GFP (GFP N-terminally tagged with the RbcS chloroplast targeting sequence). The nuclear extract was shown to be enriched for nuclear protein (H2A) and depleted in plastid proteins (chloroplast-targeted GFP) (Figure 2-4). The same blot was stripped and reprobed with α-H3-Ac(lys9) to confirm there were similar amounts of nuclear protein in both samples (Figure 2-4).
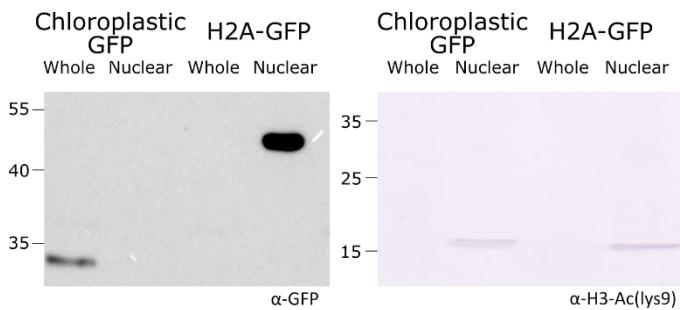
Figure 2-4 Nuclei fractions were free of chloroplastic contamination.
Whole and nuclear fractions were extracted from tissue transformed with GFP targeted to chloroplasts (pICSL11016) and nuclei (H2A-GFP). Chloroplastic GFP was only found in the whole fraction and histone-tagged GFP was only found in the nuclear fraction. The same blot was stripped and re-probed for H3-Ac(lys9), demonstrating that nuclei were efficiently extracted and protein loaded on the 'chloroplastic GFP – nuclear lane'. The calculated weights from the amino acid sequences are: RbcS-GFP – 33 kDa, H2A-GFP – 41 kDa. Blots were probed with α-GFP-HRP (Miltenyi Biotec, 130-091-833) or α-H3-Ac(lys9) (Millipore, 06-942) followed by α-rabbit-Alkaline Phosphatase (Sigma, A3687). The location of size markers are noted in kDa.

2.3.2.2   Application to plasmodesmata extractions

Detergents should only be added once the homogenisation is complete. This prevents bubbles from forming during sonication (plasmodesmata extraction) or blending (nuclei extraction). Therefore, I opted to add Triton X-100 after cell wall disruption by the cell disruptor (Figure 2-1). To demonstrate that chloroplasts were efficiently lysed by the addition of detergent, I monitored the amount of chlorophyll in the "Wash" from two extractions with and without Triton X-100 addition. Chlorophyll concentration within samples was measured by the absorbance at λ_{450} and λ_{680} and normalised to the absorbance before detergent addition (Table 2-3). At both wavelengths, the absorption was 10% greater in the wash with detergent addition, than without. Thus, there was a greater concentration of chlorophyll in the wash fraction upon addition of detergent. This demonstrates that chloroplasts are being lysed in the total extract and removed from the cell walls, and so a greater amount of chlorophyll is being released into the wash fraction.

Table 2-3 Chlorophyll absorbance in two extractions with and without Triton X-100 addition.

	- Triton X-100		+ Triton X-100	
	λ_{450}	λ_{680}	λ_{450}	λ_{680}
EB	1	1	1	1
Wash	0.74	0.77	0.82	0.84

2.3.3 Digestion of cell walls

Once purified cell walls have been obtained, the plasmodesmata can be released from the wall fraction by cellulases digesting away the cell wall. Across the literature the same cellulase is used, Cellulase Onozuka R10 (Yakult Pharmaceutical Co., Ltd., Japan), which was first used for protoplast isolation (Takebe *et al.*, 1968). Onozuka R10 is derived from the fungus *Trichoderma viride*, and contains a mixture of enzymes, containing: cellulases, hemicellulases, pectinases and proteases. The protease activity is advertised at 1% of the cellulase activity. Experimentally, the protease activity was found to have little effect on rice straw silage protein levels (Hidayat *et al.*, 2005).

I expressed plasmodesmal proteins to be extracted transiently in *N. benthamiana*. Originally, I used GFP tagged proteins, so as to monitor the expression using epifluorescence microscopy, before extracting the proteins. I used a variety of plasmodesmal membrane and peripheral (membrane-associated but with no known membrane anchorage) proteins: CML41 (peripheral), C4 (peripheral; from *Tomato yellow leaf curl virus*), and LYK4 (membrane; LysM RECEPTOR-LIKE KINASE 4) (Figure 2-5). In addition, I attempted to use PDLP1 and PDLP5 (membrane; PLASMODESMATA LOCALISED PROTEIN) tagged with GFP, however they did not express well in my hands. In all cases, GFP-tagged proteins were fully degraded by the cellulase treatment of the cell wall fraction. The GFP tag was labile, and was released as free cytosolic GFP (27 kDa).
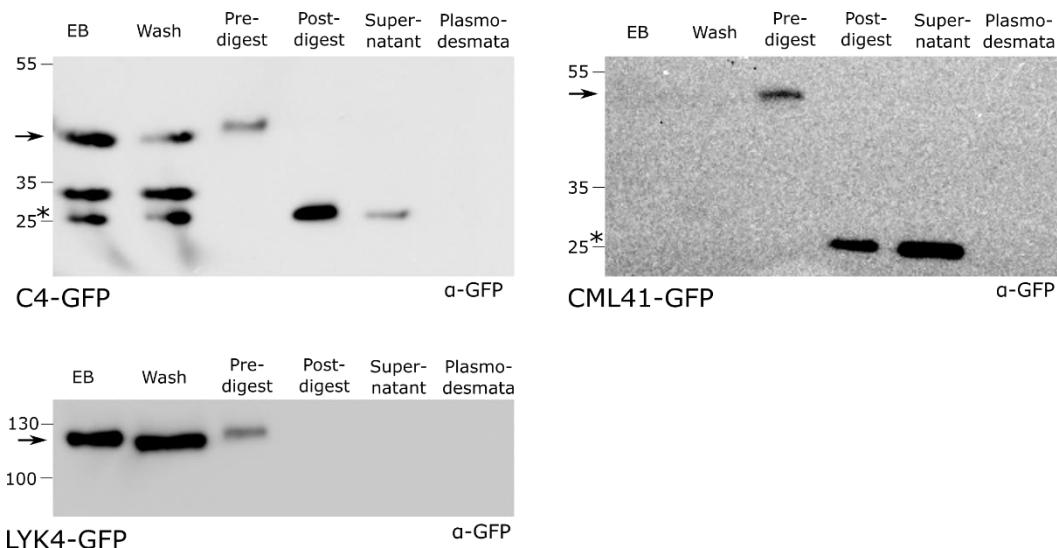
Figure 2-5 GFP tagged proteins are cleaved during cell wall digestion.
Western blot analyses of sequential extraction steps of protein-GFP fusions. In all cases, full length protein was observed in the first steps of extraction ('EB', 'Wash' and 'Pre-digest). GFP was then cleaved from the proteins during the digestion, as only a GFP-weight band can be observed after digestion (*) and no full-length protein (arrow). The calculated weights from the amino acid sequences are: C4-GFP – 41 kDa, CML41-GFP – 50 kDa, LYK4-GFP – 95 kDa, GFP – 27 kDa. Proteins were transiently expressed for two days in *N. benthamiana* co-expressed with the silencing suppressor P19. Blots were probed with α-GFP-HRP (Miltenyi Biotec, 130-091-833). The location of size markers are noted in kDa.

GFP may be particularly labile as it is a large tag that folds into a discrete separate structure from the protein it is tagged to (a β-barrel) and is usually connected by a linker sequence. Thus, an epitope tag, such as human influenza hemagglutinin (HA) (9 amino acids), may be less susceptible to cleavage due to its small size.

I trialled the plasmodesmata purification with HA tagged PDLP1 and PDLP5 proteins (Figure 2-6). With these two proteins, full length protein could be detected in the plasmodesmata fraction. Degradation was still apparent with HA tagged proteins, as observed by comparison of pre-digest to post-digest band strength (Figure 2-6). However, the degradation was to a lesser degree compared to the GFP tag, as some full-length protein remained after digestion (compare Figure 2-5 and Figure 2-6). Therefore, HA tags are suitable for use with the plasmodesmata extraction protocol, while GFP tags are not as they are more readily cleaved during cell wall digestion.
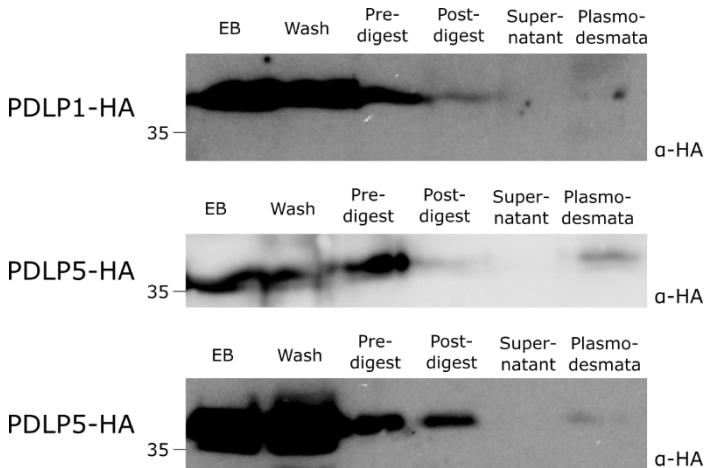
Figure 2-6 PDLP1 and PDLP5 tagged with HA are observed in the plasmodesmal fraction.
Western blot analyses of sequential extraction steps of PDLP-HA fusions. Degradation of PDLP proteins was still observed during cell wall digestion by comparison of pre-digest to post-digest band strength, but to a lesser extent than that of GFP-tagged proteins. Full length protein was observed in the plasmodesmata fraction. The calculated weights from the amino acid sequences are: PDLP1-HA – 38 kDa, PDLP5-HA – 38 kDa. Proteins were transiently expressed for two days in *N. benthamiana* co-expressed with the silencing suppressor P19. Blots were probed with α-HA-HRP (abcam, ab173826). The location of size markers are noted in kDa. Two replicated of PDLP5-HA are presented.

I wanted to confirm the conclusion that HA tags are degraded less readily than GFP tags, and so I repeated the experiment in Figure 2-5 with HA tagged constructs. This experiment demonstrates that it was not the proteins themselves (CML41, C4 and LYK4) that led to increased tag cleavage. When these proteins were tagged with HA, full length protein could be observed in the plasmodesmata fraction, in direct contrast to when they are tagged with GFP (Figure 2-7). This reiterates the finding that GFP tags are more labile than HA tags during this plasmodesmata extraction protocol.
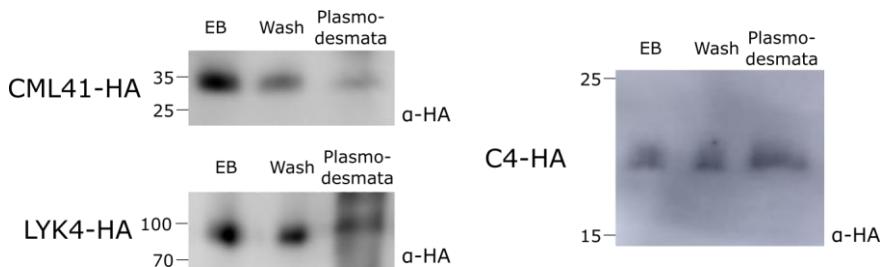
Figure 2-7 HA tagged proteins are stable through the plasmodesmata extraction.
Western blot analyses of sequential extraction steps of protein-HA fusions. When tagged with HA instead of GFP, CML41, LYK4, and C4 are not fully degraded in the cell wall digestion, as they are present in the plasmodesmata fraction. The calculated weights from the amino acid sequences are: CML41-HA – 30 kDa, LYK4-HA – 73 kDa, C4-HA – 18 kDa. Proteins were transiently expressed for two days in *N. benthamiana* co-expressed with the silencing suppressor P19. Blots were probed with α-HA-HRP (abcam, ab173826). The location of size markers are noted in kDa.

Degradation of proteins was observed during cell wall digestion irrespective of the tag used (Figure 2-5, Figure 2-6). I attempted to reduce the degree of degradation in the protocol.

There are three plausible factors that may be influencing the degradation of proteins during cell wall digestion: the 37°C temperature, the acidic buffer, and/or the Cellulase Onozuka R10. Thus, I incubated C4-HA protein at 37°C with and without Cellulase Onozuka R10 in the same buffer (Figure 2-8). C4-HA was only degraded in the presence of cellulase. Thus, the proteins are being degraded by protease activity of the Cellulase Onozuka R10 enzyme cocktail, and not the 37°C temperature or the acidic buffer.
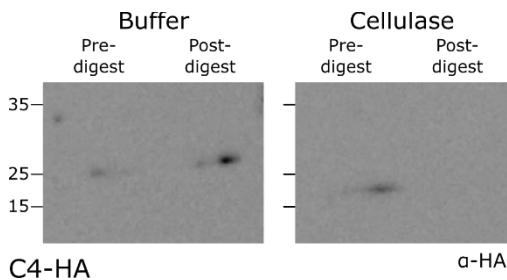
Figure 2-8 C4-HA was only degraded when Cellulase Onozuka R10 was added to the cell wall digestion step.
Western blot analysis of C4-HA stability during cell wall digestion. To locate the source of degradation C4-HA was incubated at 37°C in cellulase buffer for 1 h with and without Cellulase Onozuka R10 addition. C4-HA was only degraded when the enzyme was added. C4-HA was transiently expressed for two days in *N. benthamiana* co-expressed with the silencing suppressor P19. The calculated weight from the amino acid sequences is: C4-HA – 18 kDa. Blots were probed with α-HA-HRP (abcam, ab173826). The location of size markers are noted in kDa.

Having concluded that Cellulase Onozuka R10 contained proteolytic activity, I attempted to alter the protocol to reduce this. Thus, I reduced the temperature at which the digestion occurs from 37°C to 4°C. Cellulases have a Q_{10} temperature coefficient of ≈2.6 (Sørensen *et al.*, 2015), therefore I increased the incubation time to 24 h from 1 h ($1\,h \times 2.6^{\frac{37°C-4°C}{10°C}} = 23.4\,h$) to maintain the same amount of cellulase activity. While this change in temperature did not significantly alter the amount of protein in the plasmodesmata fraction, neither did it reduce the amount of protein degradation (Figure 2-9).
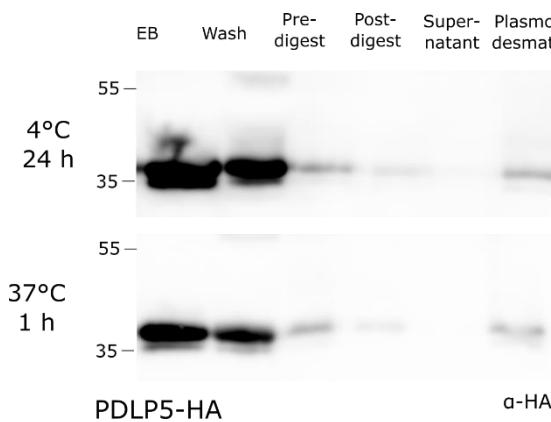
Figure 2-9 Protein degradation was not reduced by a longer colder incubation
Western blot analyses of sequential extraction steps of PDLP5-HA, with the cell wall digestion at two different temperatures and durations. To see whether temperature had an effect on protein degradation during cell wall digestion, PDLP5-HA was extracted in the same conditions with differing temperatures of digestion. The digestion of cell walls was carried out at 4°C for 24 h and 37°C for 1 h. The amount of degradation was similar in both cases.

PDLP5-HA was transiently expressed for two days in *N. benthamiana* co-expressed with the silencing suppressor P19. The calculated weight from the amino acid sequences is: PDLP5-HA – 38 kDa. Blots were probed with α-HA-HRP (abcam, ab173826). The location of size markers are noted in kDa.

Combining the above results, it can be concluded that both GFP and HA tagged proteins are degraded by Cellulase Onozuka R10 enzyme cocktail. Though, the GFP tag is more labile than the smaller HA tag. Consequently, full length protein of both membrane-bound and membrane-associated proteins could only be localised to plasmodesmata when tagged with HA. The amount of degradation was not altered by varying the incubation temperature.

2.3.4 Collection of the plasmodesmata fraction

Different proteins are extracted by using different buffer compositions (Ignatoski & Verderame, 1996). I attempted to do a total protein extraction of PDLP1-HA, using two buffers with the main difference being the presence of glycerol in Buffer 2 (Buffer 1: 50 mM Tris-HCl pH 6.8, 150 mM NaCl, 1% Triton X-100, protease inhibitor cocktail; Buffer 2: 150 mM Tris-HCl pH 7.5, 150 mM NaCl, 10% glycerol, 5 mM DTT, 1% IGEPAL CA-630, protease inhibitor cocktail). Glycerol is commonly added to extraction buffers of membrane proteins to mimic the hydrophobic environment of the cell wall, and so theoretically stabilise membrane proteins in the buffer. PDLP1-HA extracted with much greater efficiency without glycerol, but still extracted to a lesser extent when glycerol was added to the buffer (Figure 2-10). Therefore, unless stated otherwise, I used buffers without glycerol throughout the plasmodesmata extraction protocol, unlike Grison *et al.* (2015).
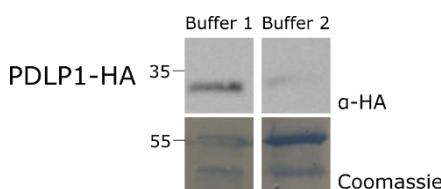
Figure 2-10 Total protein extraction of PDLP1 is more efficient without glycerol.
Western blot analysis of PDLP1-HA extraction efficiency in a total protein extraction in differing buffers. Total protein was extracted in two buffers with (Buffer 2) and without (Buffer 1) glycerol. PDLP1-HA was extracted more efficiently in Buffer 1, i.e. without glycerol. Total loading is shown by the RuBisCO Coomassie Brilliant Blue staining (Zess & Kamoun, 2019). PDLP1-HA was transiently expressed for two days in *N. benthamiana* co-expressed with the silencing suppressor P19. The calculated weight from the amino acid sequences is: PDLP1-HA – 38 kDa. Blots were probed with α-HA-HRP (abcam, ab173826). The location of size markers are noted in kDa.

I found that PDLP1, PDLP5, CML41, C4 and LYK4 extracted well from mature tissue in a buffer without glycerol (Figure 2-6, Figure 2-7). However, I also found that H$^+$ ATPases (AHAs) were not stable without glycerol (Figure 2-11). AHAs are used as a plasma membrane control, as they have been shown to be actively excluded from plasmodesmata (Fleurat-Lessard *et al.*, 1995; Grison *et al.*, 2015b,a).
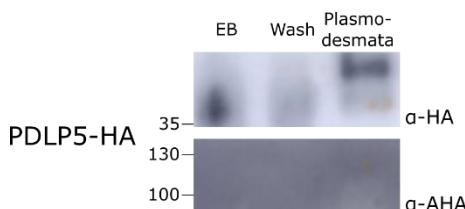
Figure 2-11 The plasmodesmata extraction protocol is unsuitable for the extraction of H⁺ ATPases
Western blot analysis of plasmodesmata extraction of PDLP5-HA, probed for endogenous AHAs. Plasmodesmata were extracted with buffers not containing glycerol. PDLP5-HA was observed at in the input (EB) and plasmodesmata fractions, indicating the plasmodesmata extraction worked. AHA could not be observed at any stage, including the input (EB) where endogenous AHAs must be present. Native H⁺ ATPases were detected by a polyclonal antibody. PDLP5-HA was transiently expressed for two days in *N. benthamiana* co-expressed with the silencing suppressor P19. The calculated weight from the amino acid sequences is: PDLP1-HA – 38 kDa. Blots were probed with α-HA-HRP (abcam, ab173826) or α-H⁺ ATPase (AS07 260, Agrisera) followed by α-rabbit-HRP (A0545, Sigma). The location of size markers are noted in kDa.

Thus, I reintroduced glycerol to all buffers to determine whether AHAs were extracted into the plasmodesmal fraction. In addition, as per the manufacture's suggestion (Agrisera), samples were denatured at 70°C instead of 95°C. In this manner, AHAs were extracted into the total 'EB' fraction, but were not present at plasmodesmata (Figure 2-12). Therefore, the plasmodesmata extraction produces a fraction containing plasmodesmata proteins and devoid of plasma membrane proteins.
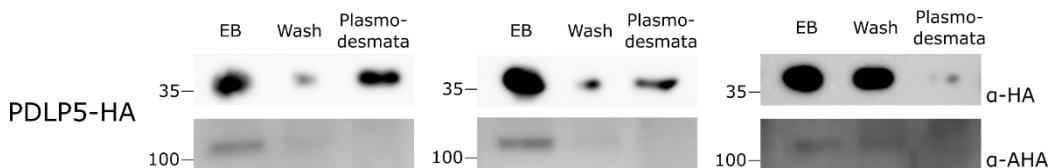
Figure 2-12 H⁺ ATPases could only be observed when extracted with buffers containing glycerol
Western blot analysis of plasmodesmata extraction of PDLP5-HA, probed for endogenous AHAs. Plasmodesmata were extracted with buffers containing glycerol (10% v/v). Native AHAs were stabilised by the glycerol and could be detected in the input (EB) (detected by a polyclonal antibody). AHA was not present in the plasmodesmata fraction, whereas PDLP5-HA was, indicating that the plasmodesmata fraction is pure from plasma membrane contaminants. PDLP5-HA was transiently expressed for two days in *N. benthamiana* co-expressed with the silencing suppressor P19. The calculated weight from the amino acid sequences is: PDLP5-HA – 38 kDa. Blots were probed with α-HA-HRP (abcam, ab173826) or α-H⁺ ATPase (Agrisera, AS07 260) followed by α-rabbit-HRP (Sigma, A0545). The location of size markers are noted in kDa. Three replicates are presented.

Previously, I showed that PDLP1 extracts less efficiently with glycerol in the extraction buffer (Figure 2-10). Despite this inefficiency, there was still sufficient PDLP5 to detect in the plasmodesmata fraction of a plasmodesmata extraction done with glycerol added. Thus, I attempted to repeat the extraction of other membrane proteins, such as LYK4, with glycerol to have all the work done in the same setting. However, LYK4 was not stable in buffers containing 10% glycerol (Figure 2-13). Therefore, the composition of extraction buffers must be optimised for each protein extracted: no buffer suitable for all proteins tested was found.
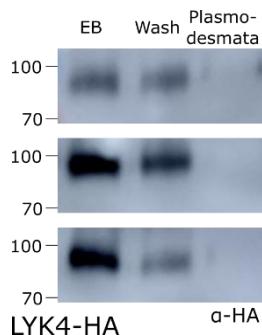
Figure 2-13 LYK4-HA is not stable in buffers containing 10% glycerol
Western blot analysis of plasmodesmata extraction of LYK4-HA extracted with buffers containing glycerol (10% v/v). LYK4-HA could not be extracted efficiently in plasmodesmata fractions when glycerol was added to the extraction buffers. LYK4-HA was transiently expressed for two days in *N. benthamiana* co-expressed with the silencing suppressor P19. The calculated weight from the amino acid sequences is: LYK4-HA – 73 kDa. Blots were probed with α-HA-HRP (abcam, ab173826). The location of size markers are noted in kDa. This extraction was carried out with 10% glycerol added to all buffers. Three replicates are presented.

2.3.5   Practical application

Having concluded that the extraction is enriched for plasmodesmata proteins and that epitope tags were suitable for plasmodesmata extraction, I investigated the localisation of two LysM RECEPTOR-LIKE KINASES (LYKs). LYK4 and LYK5 (AT2G23770 and AT2G23580, respectively) were implicated in plasmodesmal responses by Cecilia Cheval (Cheval *et al.*, 2020). However, both LYKs were observed to have an even plasma membrane distribution by confocal microscopy when transiently expressed in *N. benthamiana*. This raised the question of whether the LYK proteins were acting directly at plasmodesmata (i.e. are they present at plasmodesmata)?

Thus, I co-expressed LYK4-HA and LYK5-myc in *N. benthamiana* and probed the plasmodesmal localisation of the proteins. The co-expression allowed direct comparison of the two proteins within the same extraction. However, as the proteins had similar predicted sizes (LYK4-HA – 73 kDa, LYK5-HA – 78 kDa), the proteins required different tags. LYK4-HA was localised to plasmodesmata, whereas LYK5-myc was not (Figure 2-14).
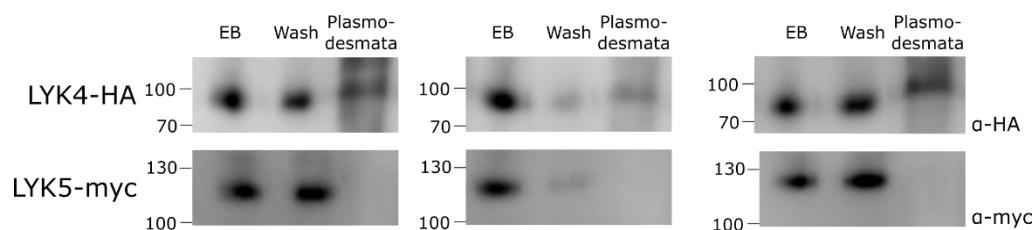
Figure 2-14 LYK4-HA was localised to plasmodesmata, whereas LYK5-myc was not
Western blot analyses of plasmodesmal extracts of mature leaf tissue co-expressing LYK4-HA and LYK5-myc. LYK4-HA was repeatedly extracted into the plasmodesmal fraction, whereas LYK5-myc was never observed in the plasmodesmal fraction. Proteins were transiently expressed for two days in *N. benthamiana* co-expressed with the silencing suppressor P19. The calculated weight from the amino acid sequences were: LYK4-HA – 73 kDa, LYK5-myc – 77 kDa. Blots were probed with α-HA-HRP (abcam, ab173826) or α-myc-HRP (abcam, ab622928). The

location of size markers are noted in kDa. Please note that the first replicate of LYK4-HA was reused in Figure 2-7. Three replicates are presented.

Having seen different degradation effects from using GFP and HA tags, I complemented the previous experiment by switching the tags, and so co-expressed LYK4-myc and LYK5-HA in *N. benthamiana*. However, in this case neither LYK4-myc nor LYK5-HA purified efficiently to plasmodesmata (Figure 2-15).

There are two conclusions to make from these data. First, not all eptiope tags are effective for monitoring plasmodesmata purification. HA is much more stable than myc through the extraction. Secondly, combining the two results using only the HA tagged proteins, to conclude that LYK4 is present at plasmodesmata but LYK5 is not (Cheval *et al.*, 2020).
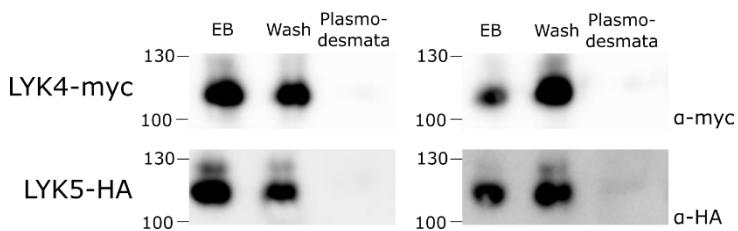
Figure 2-15 Neither LYK4-myc nor LYK5-HA purified to plasmodesmata
Western blot analyses of plasmodesmal extracts of mature leaf tissue co-expressing LYK4-myc and LYK5-HA. Neither LYK5-myc nor LYK5-HA were present in plasmodesmata fractions, despite strong expression in the input (EB). This contrasts with the observation of LYK4-HA in plasmodesmata in Figure 2-14. Proteins were transiently expressed for two days in *N. benthamiana* co-expressed with the silencing suppressor P19. The calculated weight from the amino acid sequences were: LYK4-myc – 71 kDa, LYK5-HA – 78 kDa. Blots were probed with α-HA-HRP (abcam, ab173826) or α-myc-HRP (abcam, ab622928). The location of size markers are noted in kDa. Two replicates are presented.

2.4   Discussion

In this chapter, I developed a plasmodesmata extraction protocol for mature tissue that is free from plasma membrane contaminants. I have established that buffer composition, especially the presence of glycerol, must be optimised for each membrane protein tested. I have demonstrated that the HA epitope tag is suitable for plasmodesmata extractions, whereas GFP and myc tags are not. Also, in contrast to Tilney *et al.* (1991) and Turner *et al.* (1994) I have shown that detergents can be combined with plasmodesmal extractions to remove selectively chloroplastic contaminants while preserving plasmodesmata proteins. I have used the developed method to characterise differential association of proteins with plasmodesmata-related functions to plasmodesmata.

2.4.1   The use of detergents

Detergents were effectively used to reduce chloroplastic contamination from the crude cell extract (Table 2-3). However, it has been suggested that detergents also damage plasmodesmata in the cell wall extracts (Tilney *et al.*, 1991; Turner *et al.*, 1994). Both groups

found that incubating with 1-2% Triton X-100, the plasma membrane of the plasmodesmata could no longer be observed by electron microscopy. However, this was after a 30 minute – 120 minute incubation, opposed to the 10 minutes at 1% Triton X-100 in this protocol. Experimentally, I found that plasmodesmal proteins were extracted from cell walls incubated with detergents for a short duration (Figure 2-12). Therefore, the plasmodesmata membranes cannot have been removed entirely in my set up.

This raises the question as to whether the plasmodesmata observed by Tilney *et al.* (1991) and Turner *et al.* (1994) were truly degraded. Our knowledge of plasmodesmata ultrastructure has advanced with the advent of serial block-face scanning electron microscopy (Nicolas *et al.*, 2017). The plasmodesmata observed where desmotubule cannot be distinguished from plasma membrane after Triton X-100 treatment in Tilney *et al.* (1991) could be Type I plasmodesmata, where the two membranes are firmly appressed. Turner *et al.* (1994) observe that after Triton X-100 treatment "annular structures" are removed. However, Waigmann *et al.* (1997) found that trichome plasmodesma have a synonymous "neck region", whereas mesophyll plasmodesma do not. Thus, there is precedent for plasmodesmata from within the same leaf to have different annular structures. This question, as to whether plasmodesmata are truly disrupted by detergents, could be answered experimentally. With the mild Triton X-100 treatment in the protocol developed here, plasmodesmata proteins could still be observed biochemically. The same more harsh detergent conditions as Tilney *et al.* (1991) and Turner *et al.* (1994) could be trialled in my protocol to observe whether this prevented or reduced the identification of plasmodesmata proteins in the plasmodesmata fraction.

Detergent disruption of plasmodesmata raises the possibility that plasmodesmata could be selectively extracted from cell walls using prolonged detergent incubation opposed to cellulases, as currently used. This may change the protein profiles extracted from plasmodesmata by avoiding Cellulase R10 degradation (Figure 2-5, Figure 2-8). However, this method would not be compatible with co-immunoprecipitation (see Chapter 4), as the detergent would also break apart interacting protein complexes.

Tilney *et al.* (1991) also found that incubation of fern gametophytes with a protease, papain, selectively removed the desmotubule from TEM images, yet left the plasmodesmata plasma membrane intact. Conversely, Turner *et al.* (1994) found that protease treatment of the purified cell walls from maize root tips removed the plasmodesmal plasma membrane and a surrounding collar (of unknown nature). I argue that the Cellulase R10 treatment during cell

wall digestion also contains protease activity (Figure 2-8). However, plasmodesmal proteins are still recovered (Figure 2-12), and so plasmodesmata must not be fully digested. It is unlikely that the desmotubule is fully digested as suggested by Tilney *et al.* (1991) in my experiments. This is because a member of the only known desmotubule-localised proteins, the reticulons, was identified in a proteome generated from my protocol (Table S3). Thus, I reason that whole plasmodesmata are being extracted by this method. It would be an exciting experiment to compare electron micrographs of plasmodesmata preparations from the protocol presented here with detergents (Figure 2-1) to that of the original proteome or *ex vivo* images (Tilney *et al.*, 1991; Turner *et al.*, 1994; Waigmann *et al.*, 1997; Fernandez-Calvino *et al.*, 2011; Nicolas *et al.*, 2017).

2.4.2 Test case: LYK4 is present at plasmodesmata while LYK5 is not

LYK4 and LYK5 are integral membrane proteins from the same family, with seemingly the same plasma membrane localisation when observed with fluorescent microscopy (Cheval *et al.*, 2020). However, when plasmodesmata were biochemically purified from LYK4-HA and LYK5-HA expressing plants a difference emerged. Only LYK4, and not LYK5, is present at plasmodesmata (Figure 2-14, Figure 2-15).

The hypothesis that LYK5 is not present at plasmodesmata is corroborated by two orthogonal pieces of evidence, confirming the results of the biochemical method. First, LYK4-GFP and LYK5-RFP were shown to interact at the plasma membrane by FLIM microscopy (Cheval *et al.*, 2020). However, LYK4-GFP and LYK5-RFP did not interact in plasmodesmal regions of interest (as defined by the localisation of Citrine-LYM2) (Cheval *et al.*, 2020). These data suggest that either one of LYK4/LYK5 are not present at plasmodesmata or that they dissociate at plasmodesmata. Secondly, LYK5 prevents the accumulation of LYM2 accumulation at plasmodesmata upon the addition of chitin, whereas LYK4 does not. Given that all three proteins have been shown to interact, these data are congruent with the idea that LYK5 prevents the movement of LYM2 to plasmodesmata because LYK5 is excluded from plasmodesmata and so retains LYM2 in the plasma membrane. Taken together, these two experiments substantiate the results of the plasmodesmata extraction. An ideal next experiment would be to use electron microscopy and immunogold staining of LYK4-HA and LYK5-HA. The exclusion of LYK5 would then be validated in the same way as the gold-standard AHA2 (Fleurat-Lessard *et al.*, 1995; Grison *et al.*, 2015a).

2.4.3   Conclusions

The chapter describes a modified plasmodesmata extraction protocol which is suitable for the extraction of plasmodesmata from mature leaf tissue. Moreover, it allows for the biochemical localisation of transiently expressed proteins. This includes integral membrane proteins, such as PDLPs (Thomas *et al.*, 2008) and LYK4 (Cheval *et al.*, 2020), as well as peripheral membrane proteins, including C4 (Rosas-Diaz *et al.*, 2018) and CML41 (Xu *et al.*, 2017). I found that heterologous proteins could only be extracted with an HA tag, as other tags were removed during cell wall digestion by the commercial enzyme cocktail in Cellulase R10. These preparations were shown to be pure of plasma membrane contaminants by the lack of H^+ ATPases. H^+ ATPases highlighted the need for different buffer conditions for different protein familys, as they were only being extracted successfully when glycrerol was added to the buffer. Finally, the protocol was used to differentiate between LYK4 and LYK5, which displayed indistinguishable localisations when observed with confocal microscopy. I idientified that only LYK4 resides at plasmodesmata. This led to the generation of a novel model of how LYM2, LYK4 and LYK5 signal at plasmodesmata.

3 A phyloproteomic approach to identify evolutionarily conserved proteins at plasmodesmata

3.1 Introduction

The first evidence of plant life on land comes from the fossil record of spore microfossils and tissue fragments from the Middle Ordovician around 470 million years ago (Wellman *et al.*, 2003; Steemans *et al.*, 2009; Rubinstein *et al.*, 2010). The rise of photosynthetic organisms caused one of the most transformative periods in Earth's history, dramatically altering the carbon cycle (Kenrick *et al.*, 2012; Taylor *et al.*, 2012) and ultimately paving the way for terrestrial colonisation by animals (Miller *et al.*, 1996). Though, it is important to note that plants were not the first photosynthetic organisms on land, rather following algae and microbes (Wellman & Strother, 2015).
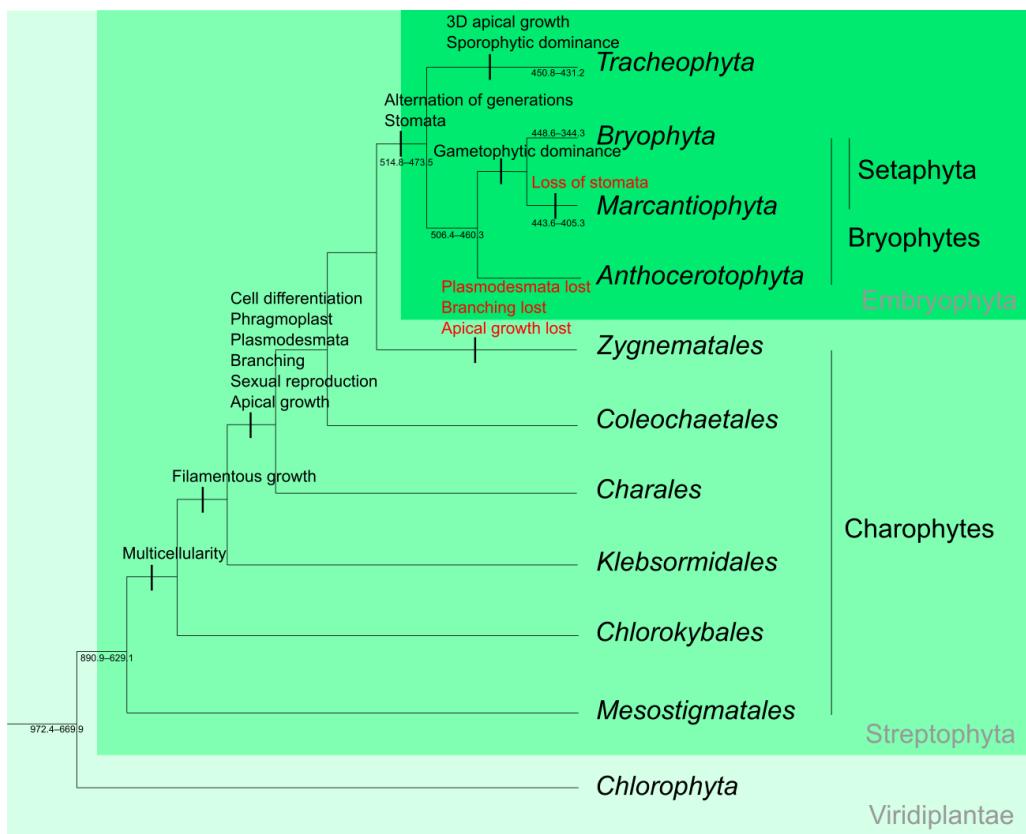
Figure 3-1 A cladogram of the leading theory of Viridiplantae evolution.
Selected feature gain (black) and loss (red) are annotated on the tree (Umen, 2014; Moody, 2020). Some nodes are annotated with 95% highest posterior density age estimates (millions of years ago) (Morris *et al.*, 2018). Green shading denote monophyletic clades within plant evolution, labelled in the bottom right of each shading.

It is well accepted that embryophytes (Embryophyta, land plants) arose from streptophytic green algae (together forming the monophyletic Streptophyta) (Manhart & Palmer, 1990; Devereux *et al.*, 1990) and that this was a single event (Karol *et al.*, 2001). However, a great

deal of debate is ongoing over two key phylogenetic issues: which streptophytic algae is sister to Embryophyta, and the relationship among the earliest lineages of embryophytes (bryophytes). The topology of the Viridiplantae phylogenetic tree affects the conclusions drawn by the comparison of extant traits between living ancestors. Traits which have been retained in multiple lineages can be thought to be important, and comparison between the biological functioning of these traits (physiologically, genetically or otherwise) can elucidate the traits fundamental characteristics through conservation.

3.1.1   From whom did Embryophyta arise?

The first contention is how to untangle the relationship between Embryophyta and three lineages of algae (members of the paraphyletic clade 'charophytes'): Charales, Coleochaetales, and Zygnematales. Previously, the Charales were thought to be a sister group to Embryophyta, due to morphological characteristics and preliminary DNA evidence (Bremer & Wanntorp, 1981; Mattox & Stewart, 1984; Karol *et al.*, 2001; Lewis & McCourt, 2004; Becker & Marin, 2009; Cocquyt *et al.*, 2010; Wickett *et al.*, 2014). This is appealing as it allows the simple hypothesis of progressive evolution of cellular complexity from Zygnematales to Charales to Embryophyta. Moreover, *Chara sp.*, the group's eponym, contain plasmodesmata (Cook *et al.*, 1997), meaning that all embryophytic plasmodesmata are related.

However, more recent analyses have instead found that it is more likely that the Zygnematales alone (or the Coleochaetales and Zygnematales together, which cannot be excluded as the Coleochaetales are poorly sampled in sequencing data) are sister to Embryophyta, with Charales the most distantly related (Turmel *et al.*, 2006; Lemieux *et al.*, 2007; Wodniok *et al.*, 2011; Timme *et al.*, 2012; Laurin-Lemay *et al.*, 2012; Ruhfel *et al.*, 2014; Puttick *et al.*, 2018). The change is due to a wealth of new sequencing data for both plants and the Zygnematales (Leebens-Mack *et al.*, 2019; Cheng *et al.*, 2019), as well as improved phylogenetic algorithms (Cooper, 2014; Puttick *et al.*, 2018). This proposed relationship among the Charophytes implies significant simplification within the Zygnematales, such as the loss flagella and plasmodesmata (Figure 3-1).

Placing Zygnematophyceae sister to Embryophyta suggests that the ancestral state of the more recently diverged charophyte lineages would be a multicellular and branched form. The unicellularity and simplicity seen in the Zygnematales (Becker & Marin, 2009) would then be a derived state. It is interesting to postulate that the highly complex forms of land plants and the severely reduced forms of the Zygnematales are both adaptions to the same

intermittently wet habitats (Delwiche & Cooper, 2015). Each has developed radically different strategies to survive drought, with the former building cuticles to prevent desiccation, and the latter becoming smaller to facilitate dormancy during desiccation. In addition, a small size allows greater utilisation of transient hydration, as a thinner water film is required to cover the organism, as well as having a shorter division time when there is water present. While morphologically the sisters have diverged a long way, hormone signalling in response to desiccation has remained remarkably similar between *Spirogyra pratensis (*Zygnematales*), Physcomitrella patens* (Bryophyta), and *Rumex palustris (Tracheophyta,* vascular plants*).* Ethylene is used in all three cases to regulate growth of submerged tissue, causing cell-elongation (Voesenek *et al.*, 2003; Yasumura *et al.*, 2012; Ju *et al.*, 2015). This is proposed to help the organism escape submergence, as ethylene is incredibly insoluble in water and so only diffuses away when the atmosphere is reached (Davis & Mcketta, 1960). This highlights the potential for using incredibly ancient ancestors to understand fundamental plant processes within Tracheophyta.

Sister to the Charales are another charophytic lineage the Klebsormidiales. *Klebsormidium flaccidum*, an alga within this order, is multicellular with filamentous growth. It has the genes to produce 'plant hormones' such as ethylene but whether or not it has hormone responses has yet to be tested (Hori *et al.*, 2014). Plasmodesmata are lacking in *K. flaccidum* (Stewart *et al.*, 1973b)*,* which then places the evolution of plasmodesmata at the last common ancestor of the Charales: roughly 600 million years ago during the Ediacaran (Morris *et al.*, 2018).

Thus, it is thought that, irrespective of the exact resolution of the Charophytes and Embryophyta, plasmodesmata had evolved prior to the evolution of land plants. Only the timing of the evolution of plasmodesmata is affected. This is supported by the fact that all embryophytes have plasmodesmata (Cook & Graham, 1999a) and that land plants arose from a single event. This leads to the most parsimonious explanation that the last common ancestor of embryophyte had plasmodesmata.

3.1.2   What about the bryophytes?

Once inside the embryophytes, the issues do not become less contentious. This is in part due to poor sampling of early land plants, with the first hornwort genome only published in 2020 (Zhang *et al.*, 2020). The question entails how to place the three bryophyte lineages (Anthocerotophyta, hornworts; Marcantiophya, liverworts; and Bryophyta, mosses) among themselves and in respect to the tracheophytes (Tracheophyta, vascular plants).

The current prevailing view is that bryophytes form a monophyletic clade, which is sister to Tracheophyta. Inside of which, the hornworts are sister to a second monophyletic clade formed of mosses and liverworts, termed Setaphyta (Nishiyama *et al.*, 2004; Puttick *et al.*, 2018; de Sousa *et al.*, 2019; Sousa *et al.*, 2020a; Harris *et al.*, 2020). However, bryophyte monophylly is controversial, and hornworts being sister to Setaphyta and the closest ancestor to Embryophyta cannot be ruled out (Nishiyama *et al.*, 2004; Wickett *et al.*, 2014; Puttick *et al.*, 2018; Zhang *et al.*, 2020; Sousa *et al.*, 2020b).

This runs contrary to the traditional view of liverworts being an extant lineage to all other land plants (Qiu *et al.*, 1998, 2006; Groth-Malonek *et al.*, 2005; Chang & Graham, 2011; Ligrone *et al.*, 2012). This view was supported by the supposed liverwort nature of the ancient spore record (lamellae in the spore walls) (Gray, 1993). However, this evidence is uncertain as no megafossils have been found, which produced the spores. Moreover, similar spore tetrads are made by mosses (Gray, 1993).

3.1.3   The implications of a monophyletic bryophyte clade

3.1.3.1   Alternation of generations

Bryophytes forming a monophyletic clade dramatically alters our perspective on land plant evolution, as it means that tracheophytes are not derived from bryophytes. This largely has an impact on theories surrounding the alternation of generations. The gametophytic dominance of charophyte algal ancestors implies the land plant ancestor had a multicellular gametophyte. However, the unbranched sporophyte of the bryophytes may be a reduction from the ancestral state. Indeed, it is possible there were near monomorphic generations in the ancestor, and bryophytes and tracheophytes reduced different life stages.

Ancestral monomorphic generations seem less likely, as the predominately gametophytic charophyte algae nurture the sporophyte generation, similar to bryophytes. However, this supposed synapomorphy may be coincidence, and simply a case of convergent evolution. On the other hand, *Anthoceros fusiformis (*Anthocerotophyta, hornworts) has almost free-living sporophytes, which when excised onto sterile soil can survive three months independently (Campbell, 1924), suggesting that the ancestral sporophyte may have been self-sufficient. Potential evidence for a branched sporophyte (polysporangiophytes) may come from the fossil plants *Horneophyton* and *Aglaophyton*, from the Rhynie chert flora (Kenrick & Crane, 1991). Traditionally, these fossils were considered early tracheophytes (Kenrick & Crane, 1997) but they could instead pre-date the bryophyte-tracheophyte split, and perhaps have retained traits from the land plant ancestor: a less-reduced sporophyte.

3.1.3.2 Comparative phylogenetics

Beyond early plant life cycles, the monophylly of bryophytes alters our view of how comparative phylogenetics should be used between bryophytes and tracheophytes to find ancestral states. Research has to be phylogenetically constrained to carefully consider whether traits are basal or secondarily lost or gained. The dogma was to compare extant Tracheophytic traits against liverworts, as they were thought to be the most basal extant lineage. However, in light of the monophylly of bryophytes, any bryophyte may be equally as insightful.

An example of this from developmental biology is the presence of stomata within Anthocerotophyta and Bryophyta, but not Marcantiophya. Originally, two independent evolutionary events were postulated (homoplasy) (Field *et al.*, 2015). However, it is now thought that stomata were an ancestral feature that has been lost in Marcantiophya (Puttick *et al.*, 2018; Harris *et al.*, 2020). In short, I agree with Puttick *et al.* (2018) that "no one lineage can be considered a developmental, genetic, or physiological surrogate for the ancestral embryophyte; all are a mélange of shared primitive and unshared derived characters specific to their respective lineages."

While this phylogenetic distance is problematic for finding basal traits, it can be leveraged to our advantage when known basal traits are being compared, such as plasmodesmata. It is important to remember that all extant lineages have gone through an equal number of millions of years of evolution, but when they diverged from each other changes. For example, when comparing Brassicaceae with Setaphyta both lineages have undergone a separate ~450 million years of evolution. This means it is incredibly difficult to derive what a basal plasmodesmata may have looked like. However, if we want to find the most essential parts of a plasmodesma, this situation allows us to find which proteins were conserved (or independently co-opted) in two independent lines over this time from the same starting point. Therefore, the ancient split between Bryophyta and Tracheophyta allows us to probe which parts of plasmodesmata are similar after millions of years of evolution, and thus which may be the most important.

As mentioned above, this does not mean that the extant plasmodesmata are similar to each other or to their common ancestor. However, overlapping proteins in them are likely to be critical to plasmodesmata structure or function. The proteins may have been present in the ancestor, or they may have been predisposed to recruitment into plasmodesmata. In both cases, it highlights the proteins as prime candidates. This logic can be seen in action in the

evolution of C_4 photosynthesis, where predisposed recruitment is observed when the same enzyme, such as NAD-dependent malic enzyme, is independently co-opted in multiple independent events (Brown *et al.*, 2011). The importance of NAD-dependent malic enzyme in C_4 photosynthesis is seen by 95% reduction in CO_2 assimilation when it is mutated (Dever *et al.*, 1998).

3.1.4 Attempts to characterise the protein composition of plasmodesmata

To be able to compare the protein composition of plasmodesmata from different species, high-throughput identification is required: protein mass spectrometry. This first requires the biochemical purification of plasmodesmata. The range of attempts and methods to do this was described previously in section 2.1.1. Ultimately, a robust method for plasmodesmata extraction from cell suspension cultures was developed (Faulkner & Bayer, 2017), which has been used for all published plasmodesmal proteomes (Fernandez-Calvino *et al.*, 2011; Leijon *et al.*, 2018; Brault *et al.*, 2019).

The first plasmodesmal proteome provided the plasmodesmal community with a resource to target plasmodesmal proteins directly for the first time. For example, this aided in the discovery of novel plasmodesmal receptor-like kinases (Faulkner *et al.*, 2013), components of cell-to-cell signalling (Vaddepalli *et al.*, 2014), and structural components of plasmodesmata (Knox *et al.*, 2015; Wang *et al.*, 2015). The authors validated the proteome, recovering 39 previously validated plasmodesmal proteins and discovering five new plasmodesmal proteins (33% of the 15 tested). However, the authors estimated about 35% of the proteome was contaminants (Fernandez-Calvino *et al.*, 2011).

To reduce the degree of contamination in the proteome, the same plasmodesmal purification technique has been used with a more advanced quantitate proteomic technique (tandem mass tagging (TMT)) on multiple cell fractions (Brault *et al.*, 2019). The abundance of proteins within plasmodesmata was compared to contaminant fractions, such as the cell wall and plasma membrane. These ratios were used to define a filtered proteome, where only 115 proteins highly enriched at plasmodesmata were included. This simultaneously produced a plasmodesmal enrichment factor, comparing the abundance of protein within the plasmodesmata and plasma membrane, which is referred to in this thesis as the Bayer enrichment factor (Brault *et al.*, 2019).

A proteome has also been generated from *Populus trichocarpa* suspension cell cultures (Leijon *et al.*, 2018). To verify the plasmodesmal nature of the extractions, Leijon *et al*. demonstrated an increased callose synthase activity of plasmodesmal fractions compared to

control microsomal fractions. Unfortunately, this proteome was not compared to the prior *A. thaliana* proteome (Fernandez-Calvino *et al.*, 2011).

Notably, all plasmodesmata proteomes to date have been performed on suspension cell culture extractions. However, plasmodesmata-enriched proteomes, i.e. cell wall fractions without cellulase digestion, have been generated on other tissues. For example, using rice callus cultured cells to identify receptor like kinases in the cell wall fraction, and cloning them to identify plasmodesmal members (Jo *et al.*, 2011). This is the same method as employed to identify the PDLPs and PDCBs (Bayer *et al.*, 2006; Thomas *et al.*, 2008; Simpson *et al.*, 2009).

Plasmodesmata-enriched proteomes have also been used to derive plasmodesmal proteomes, by comparison between plasmodesmata-replete and plasmodesmata-deplete cell wall fractions. CHOLINE TRANSPORTER-LIKE1 (CHER1) was first implicated in sieve pore formation, and is associated with the phragmoplast during cytokinesis (Dettmer *et al.*, 2014). Later, it was found that *cher1* mutants had reduced numbers of total plasmodesmata, especially complex plasmodesmata (Kraner *et al.*, 2017a). Thus, an *A. thaliana* mature leaf tissue plasmodesmal proteome was generated by comparing two plasmodesmata-enriched proteomes one from each of Col-0 (wild type) and *cher1*, yielding a list of 61 depleted proteins in *cher1* (Kraner *et al.*, 2017b). These proteins are likely to be enriched in complex plasmodesmata. In the same comparative manner, protein spot patterns generated by two-dimensional gel electrophoresis from plasmodesmata-rich nodal complexes and plasmodesmata-replete internodal cell walls of *Chara corallina* were compared (Faulkner *et al.*, 2005). This approached identified several putative plasmodesmal proteins by probing plasmodesmata-specific protein spots (Faulkner *et al.*, 2005).

3.1.5   Aims

I wanted to leverage the ancient phylogenetic origin of plasmodesmata and the synapomorphy of the trait among Embryophyta to uncover highly conserved plasmodesmal proteins. First, I considered whether multiple proteomes of the same species can be used to find conserved plasmodesmal components. I used the existing two proteomes and a third I generated to define a list of proteome-recurrent proteins. Secondly, I extend the proteome comparisons to phylogenetically distant species, including the *P. trichocarpa* proteome and a generated *P. patens* proteome. This required finding homology between extant proteins in the proteomes, and defining phylogenetically conserved orthogroups. Thirdly, I validated the list of orthogroups, by examining the localisations of five orthogroups. Of those examined, 3/5 groups had protein members which localised to plasmodesmata when heterologously

expressed in a non-native system. These findings underline the conservation of the plasmodesmata-localisations of these proteins over 450 million years of evolution, and indicate these proteins may have been present at plasmodesmata in an ancient ancestor.

3.2   Materials and Methods

3.2.1   Extraction of plasmodesmata

Plasmodesmata were extracted from fully expanded leaves of five-week-old *Arabidopsis thaliana* plants grown on soil in short day conditions (10 h light / 14 h dark) at 22°C. Plasmodesmata were extracted from a mix of *Physcomitrella patens* protonemal and gametophore tissue (Figure 3-2). *P. patens* was grown for four weeks on BCDAT medium in long day conditions (16 h light / 8 h dark) at 20°C. Protonemal tissue was grown on top of nitrocellulose membrane, whereas gametophore tissue was grown directly on the medium.
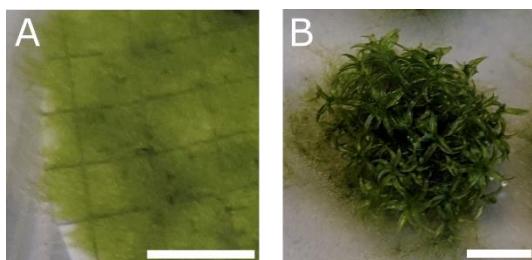
Figure 3-2 *Physcomitrella patens* tissue used for plasmodesmal proteome.
A) Protonemal tissue B) Gametophytic tissue. Scale bar = 0.5 cm.

Tissue was extracted as described in section 2.2.5. I confirmed the homogenisation of *P. patens* tissue by following the extraction with calcofluor white staining, as described in section 2.2.7, as was done in Figure 2-2. In addition to homogenising the tissue with an EmulsiFlex™-B15 cell disruptor, I tested the homogenisation of tissue using a handheld Potter-Elvehjem homogeniser (Fisherscientific, 15381321) (Figure 3-3). *P. patens* tissue was suitably homogenised into small cell wall fragments by the method established in Chapter 2. Homogenisation with the handheld Potter-Elvehjem homogeniser was less effective than by the EmulsiFlex™-B15 cell disruptor, as evidenced by large cell wall fragments only remaining in the handheld-homogenised sample.

The only modification to section 2.2.5, was that *P. patens* tissue required an additional wash with extraction buffer. The final cell wall fraction for *P. patens* contained a brown precipitate, which could not enter a polyacrylamide gel.
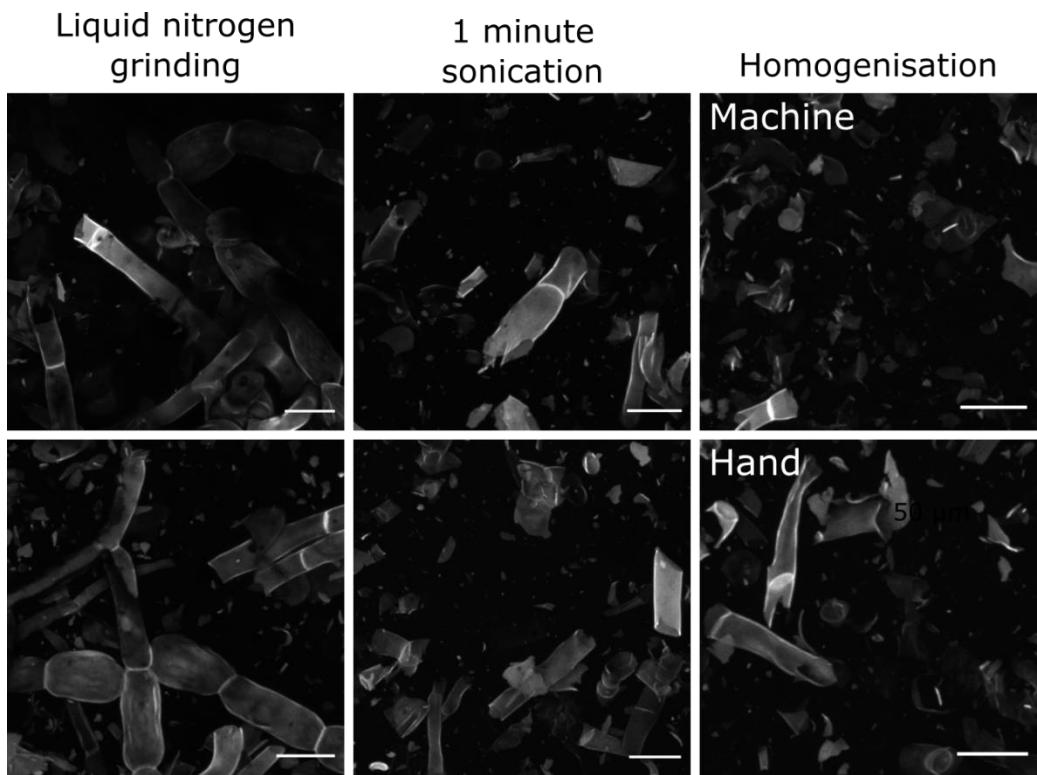
Figure 3-3 Homogenisation of *Physcomitrella patens* tissue stained with calcofluor white
Physcomitrella patens tissue was broken into short chains of cells by grinding in liquid nitrogen. The short chains were broken into single cells and fragments by sonication. Homogenisation by machine (EmulsiFlex™-B15) reduced the wall tissue to uniform small fragments and plasmodesmata-rich septum (protonemal cross walls) could be observed. Homogenisation by hand (Potter-Elvehjem homogeniser) was less effective than by machine, as indicated by the presence large cell wall fragments. Scale bar = 50 μm.

Plasmodesmal samples were run 5 mm into a 1.5 mm thick 10% TRIS resolving gel (containing 0.1% SDS) without a stacking gel, ran in a glycine 0.1% SDS running buffer. The gel was washed in dH_2O and then the band was exercised. The bands were washed four times in 20% acetonitrile at 40°C for 15 minutes to remove detergents, and then stored at 4°C in a 1.5 mL Eppendorf with 100 μL of dH_2O. These gel bands were provided to the Cambridge Centre of Proteomics.

3.2.2   Mass Spectrometry

Mike Deery of the Cambridge Centre of Proteomics carried out the sample preparation and mass spectrometry of the plasmodesmata samples and provided the materials and methods used.

1D gel bands were cut into 1 mm^2 pieces, destained, reduced (DTT) and alkylated (iodoacetamide) and subjected to enzymatic digestion with trypsin overnight at 37°C. The resultant supernatant from each band was pipetted into a sample vial and loaded into an autosampler for automated LC-MS/MS analysis.

All LC-MS/MS experiments were performed using a Dionex Ultimate 3000 RSLC nanoUPLC (Thermo Fisher Scientific Inc, Waltham, MA, USA) system and a Q Exactive Orbitrap mass spectrometer (Thermo Fisher Scientific Inc, Waltham, MA, USA). Separation of peptides was performed by reverse-phase chromatography at a flow rate of 300 nL/min and a Thermo Scientific reverse-phase nano-Easy-spray column (Thermo Scientific PepMap C18, 2 μm particle size, 100A pore size, 75 μm i.d. x 50 cm length). Peptides were loaded onto a pre-column (Thermo Scientific PepMap 100 C18, 5 μm particle size, 100 A pore size, 300 μm i.d. x 5 mm length) from the Ultimate 3000 autosampler with 0.1% formic acid for 3 minutes at a flow rate of 15 μL/min. After this period, the column valve was switched to allow elution of peptides from the pre-column onto the analytical column. Solvent A was water + 0.1% formic acid and solvent B was 80% acetonitrile, 20% water + 0.1% formic acid. The linear gradient employed was 2-40% B in 90 minutes (the total run time including column washing and re-equilibration was 120 minutes).

The LC eluant was sprayed into the mass spectrometer by means of an Easy-spray source (Thermo Fisher Scientific Inc.). All m/z values of eluting ions were measured in an Orbitrap mass analyzer, set at a resolution of 70,000 and scanned between m/z 380 - 1,500 Data dependent scans (Top 20) were employed to automatically isolate and generate fragment ions by higher energy collisional dissociation (HCD, Normalised collision energy (NCE):25%) in the HCD collision cell and measurement of the resulting fragment ions was performed in the Orbitrap analyser, set at a resolution of 17,500. Singly charged ions and ions with unassigned charge states were excluded from being selected for MS/MS and a dynamic exclusion of 20 seconds was employed.

Post-run, all MS/MS data were converted to mgf files and the files were then submitted to the Mascot search algorithm (Matrix Science, London UK, version 2.6.0) and searched against the Cambridge Centre of Proteomics database, including common contaminant sequences containing non-specific proteins such as keratins and trypsin. Variable modifications of oxidation (M) and deamidation (NQ) were applied as well a fixed modification of carbamidomethyl (C). The peptide and fragment mass tolerances were set to 20 ppm and 0.1 Da, respectively. A significance threshold value of $p < 0.05$ and a peptide cut-off score of 20 were also applied. All data (DAT files) were then imported into the Scaffold program (Version_4.10.0, Proteome Software Inc, Portland, OR).

Identified proteins are listed in Table S2.

3.2.3   Gene Ontology (GO) Analysis

Gene ontology (GO) was used to test gene lists for cellular localisation enrichment (Ashburner *et al.*, 2000; Carbon *et al.*, 2019). Cellular localisation GO overrepresentation test, release 20200728 was performed (Mi *et al.*, 2019), using the Panther database (Thomas *et al.*, 2003, 2006) and GO Ontology database released 2020-09-10 (Carbon & Mungall, 2020) with a Fisher's exact test and FDR reported. *P. patens* genes were annotated bioinformatically using phylogenetic backpropagation of GO terms (Gaudet *et al.*, 2011). Graphs were drawn using ggplot2 in R (v4.0.0) (Wickham, 2011).

3.2.4   Protein property prediction

Representative peptide models were downloaded from TAIR (Araport11) for *Arabidopsis thaliana* and from Phytozome v12.1 (*Populus trichocarpa* v3.1) for *Populus trichocarpa*. Protein sequences were analysed using SignalP v5.0 (Almagro Armenteros *et al.*, 2019) and TMHMM v2.0 (Sonnhammer *et al.*, 1998; Krogh *et al.*, 2001). Significant deviation from a control microsome fraction was tested for using bootstrap hypothesis tests, adjusted for multiple comparisons (Johnston & Faulkner, 2021). Analysis was carried out in R (v4.00) (R Core Team, 2020). GPI anchors were predicted using PredGPI (Pierleoni *et al.*, 2008). Plasmodesmata localisation was predicted using PDloc (Li *et al.*, 2020).

3.2.5   HMMER

HMMER v3.3 was used (Eddy, 1998, 2020). The *P. patens* plasmodesmal proteome was downloaded as peptide sequences from UniProt and used as the reference database for a 'phmmer' search against which the *A. thaliana* UniProt proteome was ran (UP000006548, accessed 24/04/2020) (Cheng *et al.*, 2017). Protein matches were filtered at either $E < 1 \times 10^{-100}$ or $E < 1 \times 10^{-50}$ as stated in the text.

3.2.6   OrthoFinder

Orthofinder (v2.2.6) was used to create *de novo* orthogroups (Emms & Kelly, 2015, 2019). Plasmodesmal proteome protein sequences were downloaded using UniProt, TAIR (Araport11), and Phytozome v12.1 (*Populus trichocarpa* v3.1). Orthofinder was run on these sequences with default settings. Venn diagrams were drawn using Venny (Oliveros, 2007) and the University of Ghent Venn diagrams tools (http://bioinformatics.psb.ugent.be/webtools/Venn/).

3.2.7   Unrooted Phylogeny of Orthogroup Proteins

A peptide sequence was downloaded from UniProt for each protein within the orthogroup. The protein FASTA sequences were aligned with Clustal Omega (v1.2.4, (Sievers *et al.*, 2011;

Madeira *et al.*, 2019)) to build a consensus sequence. The consensus sequence, in Stockholm format, was used as the basis for a hmmsearch (EBI, HmmerWeb version 2.41.1, (Potter *et al.*, 2018)). A search was conducted against the EMBL Reference Proteomes database restricted to *A. thaliana* (taxon id: 3702), *P. patens* (taxon id: 3218), and *P. trichocarpa* (taxon id: 3694) species sequences with a sequence E-value cut off of 1×10^{-100}, unless otherwise stated. Protein sequences were manually deduplicated for each gene.

The FASTA sequences for all identified homologues, from the hmmsearch, in all three species were downloaded and a bootstrapped non-rooted phylogenetic was generated using the 'standard_trimmed_phyml_bootstrap' ete (v3.1.1, (Huerta-Cepas *et al.*, 2016)) workflow through xvfb (Wiggins, 2020) on Ubuntu 20.04. In this workflow, sequences are aligned with Clustal Omega (Sievers *et al.*, 2011), trimmed with TrimAI (Capella-Gutiérrez *et al.*, 2009), and a phylogeny determined with 100 bootstraps using PhyML (Guindon *et al.*, 2010). Trees were drawn using ggtree in R (v4.0.0) (Yu *et al.*, 2017).

3.2.8   Cloning of conserved plasmodesmal proteins

Golden Gate cloning was used to create constructs to test the localisation of proteins of interest. Golden Gate was carried out as described in section 2.2.2.1. The construction of the plasmids is described below.

3.2.8.1   Level 0

Protein coding sequences of selected protein candidates were retrieved from UniProt. Sequences were domesticated for Golden Gate cloning by removing all *BbsI*, *BsaI*, *DraIII* restriction cut sites with synonymous mutations. The proteins were *de novo* synthesised (GENEWIZ, China) in a kanamycin level 0 module with overhangs (AATG – ggTTCG).

MULTIPLE C2 DOMAINS AND TRANSMEMBRANE REGION PROTEINS (MCTPs) were synthesised in two parts: an extracellular region, and transmembrane and intracellular region (overhangs AATG – AGTC, AGTC – TTCG). β-1,3-GLUCANASE (Pp3c10_5480V3.1 and AT5G42100) signal peptides, as predicted by SignalP 5.0 and Phobius, were amplified with a high-fidelity enzyme and attached to citrine to make an level 0 module with overhangs CCAT – AATG. The rest of the protein was amplified into a separate level 0 module with overhangs AATG – TTCG.

A full list of sequences is provided in Table S1.

3.2.8.2 Level 1

The synthesised sequences were placed in a level 1 position 3 reverse module (pICH47822), under a 35s short promoter (pICH51277) and 35s terminator (pICH41414). The protein coding sequence was C-tagged with eGFP or mCherry (pICSL50004) and N-tagged with mCherry (pICSL30003) and a C-terminator (pICSL50028).

3.2.8.3 Level 2

Tagged coding sequences were placed into a level 2 module (pICSL4723) with kanamycin (NptII) resistance in forward position one (pICSL11024), a dummy position two (pICH54022), and a position three end-linker (pICH41766). Ultimately, this produced 20 constructs that could be floral dipped in *Arabidopsis thaliana* (see 3.2.9) or transiently expressed in *Nicotiana benthamiana* (see 2.2.3).

3.2.9 Stable transformation of *Arabidopsis thaliana*

Arabidopsis thaliana was transformed by floral dip (Clough & Bent, 1998). *A. thaliana* plants were grown on soil with nine plants in a 9 cm × 9 cm square pot in long day conditions (16 h light / 8 h dark). The first bolt was removed to encourage a greater number of flowers.

A single colony of *Agrobacterium tumefaciens* transformed with the construct of interest (see 2.2.2.2) was grown overnight in 10 mL LB liquid culture with appropriate antibiotics. Cultures were centrifuged at 5,000*g* for 10 minutes to pellet the cells. The cells were resuspended in 20 mL infiltration medium (0.5 × MS medium with 5% (w/v) sucrose). Just before dipping, 5 µl of Silwet L-77 was added to the bacterial suspension.

The bacterial suspension was decanted into a 10 cm × 10 cm square plate. The *A. thaliana* inflorescences were submerged in the suspension for 60 seconds with gentile agitation. Plants were left overnight covered by a opaque plastic bag. T1 seed was collected from the plants after approximately six weeks.

Successful T1 transformants were selected by the kanamycin selection cassette. Seeds were surface sterilised by a 5 minute incubation in sterilization solution (100 mL of sterilisation solution contained 95 mL water, 5 mL bleach and 20 µL TWEEN-20 (Sigma-Aldrich)). Seeds were then washed three times with sterile water. Sterilised seeds were sown on 1× MS agar (0.8%) plates supplemented with 1% sucrose and 50 µg/mL kanamycin. After two weeks, surviving seedlings were transferred to soil.

3.2.10 Microscopy

Leaves were cut into 1 cm^2 samples and mounted adaxially on a slide. Samples were imaged on a ZEISS LSM800 confocal microscope with a 63×/1.2 water immersion objective lens (C-APOCHROMAT 63×/1.2 water). GFP and citrine were excited at 488 nm with an argon laser and collected at 500 − 545 nm. mCherry was excited at 561 nm with a DPSS laser and collected at 590 − 620 nm. Aniline blue (0.1% w/v in 1× PBS pH 7.4 (Sigma-Aldrich)) was infiltrated adaxially and excited at 405 nm with a UV laser and collected at 430 − 470 nm.

3.3   Results

3.3.1   Published *Arabidopsis thaliana* plasmodesmata proteomes show significant overlap

There are currently two published plasmodesmata proteomes of *Arabidopsis thaliana* (Fernandez-Calvino *et al.*, 2011; Brault *et al.*, 2019). Both studies followed very similar extraction protocols, generating proteomes from suspension culture cells, followed by differing mass spectrometry techniques. Therefore, you could expect significant overlap between the proteins identified by both papers.

However, the Brault proteome is more stringent than the Fernandez-Calvino proteome, as it is refined by several plasmodesmal enrichment factors. These factors are the are the ratio of protein abundance between the plasmodesmata and other cell fractions, e.g. the plasma membrane (PD/PM) or cell wall. These filters, whilst removing likely contaminants, may limit comparisons between the two proteomes. The reasons are two-fold: first, the final Brault list is very short as only high confidence hits were included; secondly, the list may be biased and by the selection criteria using ratios.

Thus, I defined a list more suitable for comparison from the Brault raw data (Bayer, *pers. comms.*) of the 1,000 most abundant proteins in the plasmodesmal fraction. This list is of similar size and methodology (abundance) to that drawn up by Fernandez-Calvino *et al*. (2011). Henceforth, to avoid confusion, the proteomes shall be referred to by their species, extraction method, and identifier, as set out in Table 3-1. I defined a filtered proteome as a proteome which had had a inter-organelle abundance ratio cut-off applied.

I wanted to define a core plasmodesmal proteome. I reasoned that proteins occurring in both unfiltered gene lists are plasmodesmal candidates, as they are conserved members across experimental set-ups. I compared the gene lists of AtCells1 and AtCells2, and found they overlapped significantly (Figure 3-4a, 425 proteins, $p < 1 \times 10^{-300}$; hypergeometric test). To test whether the overlapping proteins were good plasmodesmal candidates, I compared the intersection of AtCells1 and AtCells2 (AtCells1∩AtCells2) with the filtered list AtCells2_filtered. These two sets of proteins also overlapped significantly (Figure 3-4b, 56 proteins, $p < 9.14 \times 10^{-71}$; hypergeometric test). Further, I considered whether the recurrent proteins had a higher median Bayer Plasmodesmal Enrichment factor than the proteomes from which they were derived: AtCells1∩AtCells2 proteins were more enriched at plasmodesmata than AtCells1 (6.7 vs 5.3, $p < 0.01$), but not AtCells2 (6.7 vs 6.2, $p > 0.05$).

3.3.1 Published Arabidopsis thaliana plasmodesmata proteomes show significant overlap

Thus, one can overlap independent experimental protein lists to find high-confidence plasmodesmal proteins, and filter out contaminants.

Table 3-1 Summary of current plasmodesmal proteomes.

Name	Species	Extraction Method	Number of Proteins	Reference
AtCells1	*Arabidopsis thaliana*	Cell suspension culture	887	Fernandez-Calvino 2011
AtCells2	*Arabidopsis thaliana*	Cell suspension culture	1,000	This thesis; Bayer *pers. comms.*
AtCells2_filtered	*Arabidopsis thaliana*	Cell suspension culture	115	Brault 2019
AtPlant1	*Arabidopsis thaliana*	Whole plant	238	This thesis
PtCells1	*Populus trichocarpa*	Cell suspension culture	1,113	Leijon 2018
PtCells1_filtered	*Populus trichocarpa*	Cell suspension culture	201	Leijon 2018
PpPlant1	*Physcomitrella patens*	Whole plant	215	This thesis
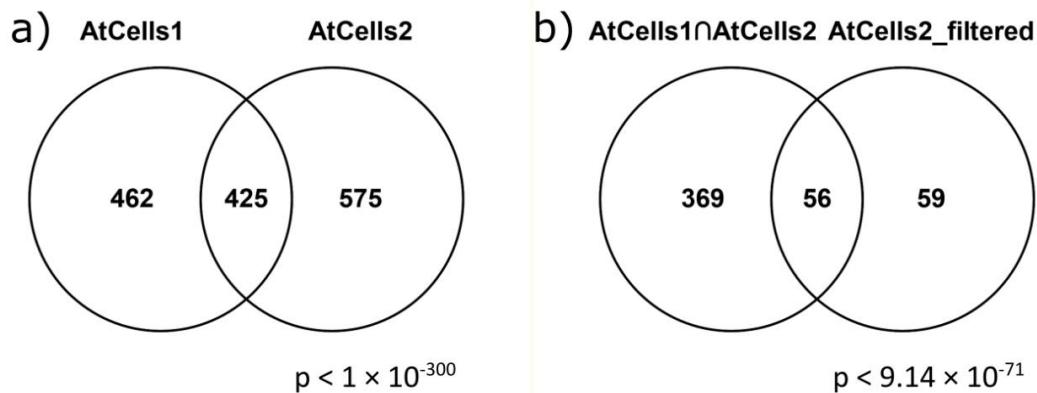
Figure 3-4 Significant overlap between existing A. thaliana proteomes.
a) The overlap between the AtCells1 and AtCells2 proteomes. b) The overlap between the two published proteomes: AtCells1 and AtCells2_filtered. The probability of the observed between proteomes was calculated with a hypergeometric test.

Using a mark-release-recapture method, it can be estimated there are between 2,087 ($\frac{887 \times 1000}{425}$) plasmodesmal proteins (Petersen, 1896; Lincoln, 1930). However, this is likely to be an underestimate, as m will be artificially inflated (Equation 3-1). m is inflated as not all

proteins have an equal likelihood of being identified: more abundant proteins are more likely to be detected by mass spectrometry. Thus, marked proteins in population one are more likely than random to be detected in population two. Therefore, a lower bound on plasmodesmal proteins can be estimated at ≈2,000 proteins.

$$\widehat{N} = \frac{c_1 c_2}{m}$$

Equation 3-1 Petersen–Lincoln method for estimating populations.

Where $\widehat{N}$ is the estimated population, c_i is the number of individuals captured on the ith visit and m is the number of marked recaptured individuals.

3.3.2   Defining a novel *Arabidopsis thaliana* proteome from mature leaf tissue

As AtCells1 and AtCells2 are were derived from a common tissue, it is plausible they contain common contaminants. Thus, to identifies core plasmodesmal proteins I wanted to compare these proteomes to a novel *Arabidopsis thaliana* plasmodesmata proteome derived from a different tissue. Comparing to plasmodesmata from a new tissue, should mean that contaminants will differ and not be conserved, while the plasmodesmal proteins are constant. However, there is potential that plasmodesmata protein composition may vary between tissues, and so this method will only capture 'core' plasmodesmal proteins.

To define a novel *Arabidopsis thaliana* plasmodesmata proteome, I extracted plasmodesmata from mature 5-week-old plants and characterised the proteome of the rosette leaves. Proteins were considered positively identified in the same manner as Leijon 2018; if the protein (95% certainty; Searle, 2010) was present in at least two of the three samples by at least one peptide (95% certainty; Keller et al., 2002). Under these conditions, 238 proteins were in the fraction.

To assess if the method established in Chapter 2 has sufficient purity to define a plasmodesmal proteome, cellular localisation GO enrichment of identified proteins was carried out. AtPlant1 was benchmarked against AtCells1, AtCells2 and AtCells2_filtered (Figure 3-5). All four proteomes were significant enriched for plasmodesmata labelled proteins. Moreover, all proteomes were significantly enriched for "cell wall" and "plasma membrane" proteins. These categories contain both likely contaminants, as well as potentially undiscovered plasmodesmal proteins. The enrichment factor filtering of AtCells2_filtered worked extremely well, with all other likely contaminant categories (e.g. "Golgi apparatus" or "chloroplast") not being over-represented, unlike the unfiltered proteomes (Figure 3-5). The degree of over-representation of plausible contaminants is similar between the unfiltered proteomes, despite AtPlant1 being derived from tissue with a

much greater abundance of chloroplasts. This analysis suggests that my novel mature plasmodesmal proteome (AtPlant1) is of comparable quality to that of those published before it.
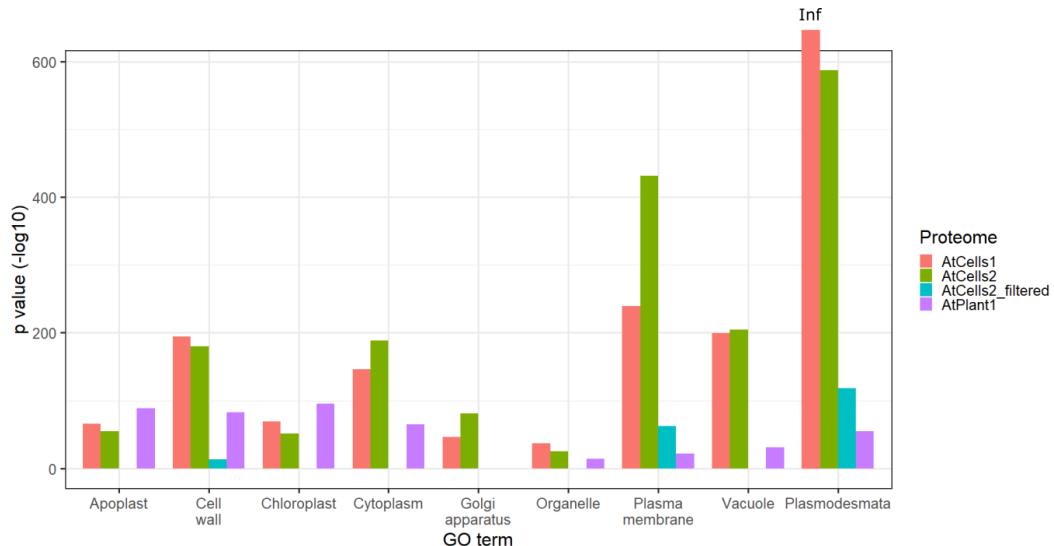
Figure 3-5 AtPlant1 is over-represented for the "plasmodesmata" cellular localisation GO term.
AtCells1, AtCells2, AtCells2_filtered and AtPlant1 proteomes were tested for over representation of selected cellular compartment GO terms ($p < 0.05$). The significance of over representation is given by the p value, which combines the size of the proteome, the size of the GO term, and the number of proteins matched in the proteome by the GO term. p values are plotted as $-log_{10}(p)$ so that a larger number is more significant. Note, AtCells1 was over represented with "plasmodesmata" GO terms to such a degree that $p \approx 0$, thus the $-log_{10}(p)$ is infinite (Inf).

Having shown that AtPlant1 is enriched in plasmodesmal proteins by GO analysis. The next step is to cross-compare AtPlant1 to AtCells1 and AtCells2 to find core plasmodesmal proteins. Between the three proteomes there were 1,576 unique proteins, of which 64 were shared by all three proteomes (Figure 3-6a). AtPlant1 significantly overlapped with the union of AtCells1 and AtCells2 (Figure 3-6a, 64/425, p = 5.76 × 10⁻⁶¹, hypergeometric test). This list of 64 proteins significantly overlapped with AtCells2_filtered (Figure 3-6b, 3/115, p = 0.00243, hypergeometric test).

However, this was a significant reduction ($p < 0.05$) from the 56/115 when comparing AtCells1∩AtCells2 and AtCells2_filtered (Figure 3-4b), even when accounting for the reduced size (425 vs 64 proteins). Naïvely, one would expect 8.4 proteins overlapping ($56 \times \frac{64}{425}$), and a 95% confidence interval can be constructed by repeatedly sampling 64 objects without replacement from a pool of 425 with 56 positives which gives a lower bound 4 and upper bound of 14 proteins. Assuming the Brault proteome as a gold-standard, this suggests two things. This list of 64 proteins is still significantly enriched for *bona fide* plasmodesmal proteins. Secondly, the significant reduction in overlap implies that the comparison to AtPlant1 is removing plasmodesmal proteins. Perhaps, these proteins are unique to simple

plasmodesmata (Bayer *et al.*, 2004). However, the remaining proteins must be core plasmodesmal proteins in both primary and secondary plasmodesmata (Oparka *et al.*, 1999).

The 64 overlapping genes between AtPlant1, AtCells1 andAtCells2 are outlined in Table 3-2. Therefore, the new *A. thaliana* proteome significantly overlaps with the existing proteomes, and, by comparison to other proteomes, identifies a set of proteome-recurrent proteins.
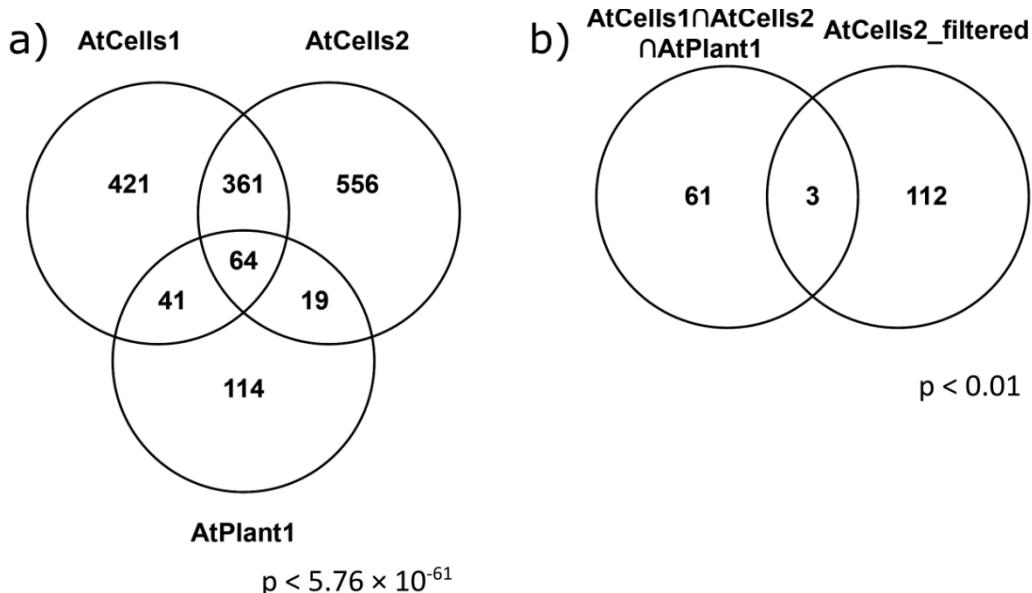
Figure 3-6 Defining a core proteome-recurrent subset of proteins at the intersection of three A. thaliana proteomes.
a) The overlap between AtCells1, AtCells2 and AtPlant1. b) The overlap between AtCells1∩AtCells2∩AtPlant1 and AtCells2_filtered

Table 3-2 Proteome-recurrent *A. thaliana* proteins.
Gene list and TAIR name (capitals) or description (italics, Araport 11) from AtCells1∩AtCells2∩AtPlant1. Genes marked with an asterisk and in bold are also in AtCells2_filtered. The table is divided by chromosome.

Identifier	Name
AT1G04270	CYTOSOLIC RIBOSOMAL PROTEIN S15 (RPS15)
AT1G07660	*histone superfamily protein*
AT1G20620	CATALASE 3 (CAT3)
AT1G22780	POINTED FIRST LEAVES (PFL)
AT1G31330	PHOTOSYSTEM I SUBUNIT F (PSAF)
AT1G42970	GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE B SUBUNIT (GAPB)
AT1G51570*	**MULTIPLE C2 DOMAIN AND TRANSMEMBRANE REGION PROTEIN 4 (MCTP4)**
AT1G56070	LOW EXPRESSION OF OSMOTICALLY RESPONSIVE GENES 1 (LOS1)
AT1G59610	DYNAMIN-LIKE 3 (DL3)
AT1G61520	PHOTOSYSTEM I LIGHT HARVESTING COMPLEX GENE 3 (LHCA3)
AT1G67090	RIBULOSE BISPHOSPHATE CARBOXYLASE SMALL CHAIN 1A (RBCS1A)
AT1G71695	*peroxidase superfamily protein*
AT1G74470	*geranylgeranyl reductase*
AT1G78830	*curculin-like (mannose-binding) lectin family protein*
AT1G78900	VACUOLAR ATP SYNTHASE SUBUNIT A (VHA-A)
AT2G04390	*ribosomal S17 family protein*
AT2G15620	NITRITE REDUCTASE 1 (NIR1)
AT2G19730	*ribosomal L28e protein family*
AT2G21330	FRUCTOSE-BISPHOSPHATE ALDOLASE 1 (FBA1)
AT2G21660	COLD, CIRCADIAN RHYTHM, AND RNA BINDING 2 (CCR2)
AT2G28000	CHAPERONIN-60ALPHA (CPN60A)
AT2G33210	HEAT SHOCK PROTEIN 60-2 (HSP60-2)
AT2G36620	RIBOSOMAL PROTEIN L24 (RPL24A)
AT2G37270	RIBOSOMAL PROTEIN 5B (RPS5B)
AT2G39730	RUBISCO ACTIVASE (RCA)
AT2G40590	*ribosomal protein S26e family protein*
AT3G02880	KINASE 7 (KIN7)
AT3G08580	ADP/ATP CARRIER 1 (AAC1)
AT3G08600	*transmembrane protein*
AT3G11820	SYNTAXIN OF PLANTS 121 (SYP121)

AT3G12780	PHOSPHOGLYCERATE KINASE 1 (PGK1)
AT3G47520	MALATE DEHYDROGENASE (MDH)
AT3G58610	*ketol-acid reductoisomerase*
AT3G60750	TRANSKETOLASE 1 (TKL1)
AT4G04640	*gamma subunit of Arabidopsis chloroplast ATP synthase*
AT4G10340	LIGHT HARVESTING COMPLEX OF PHOTOSYSTEM II 5 (LHCB5)
AT4G13340	LEUCINE-RICH REPEAT/EXTENSIN 3 (LRX3)
AT4G13930	SERINE HYDROXYMETHYLTRANSFERASE 4 (SHM4)
AT4G13940	HOMOLOGY-DEPENDENT GENE SILENCING 1 (HOG1)
AT4G18670	LEUCINE-RICH REPEAT EXTENSIN 5 (LRX5)
AT4G20360	RAB GTPASE HOMOLOG E1B (ATRABE1B)
AT4G24190	SHEPHERD (SHD)
AT4G34150*	***calcium-dependent lipid-binding (CaLB domain) family protein***
AT4G34670	*ribosomal protein S3Ae*
AT5G01530	LIGHT HARVESTING COMPLEX PHOTOSYSTEM II (LHCB4.1)
AT5G02500	HEAT SHOCK COGNATE PROTEIN 70-1 (HSC70-1)
AT5G06870	POLYGALACTURONASE INHIBITING PROTEIN 2 (PGIP2)
AT5G13450	DELTA SUBUNIT OF MT ATP SYNTHASE (ATP5)
AT5G16590	LEUCINE RICH REPEAT PROTEIN 1 (LRR1)
AT5G17920	METHIONINE SYNTHESIS 1 (ATMS1)
AT5G20950	*beta-glucosidase involved in xyloglucan metabolism*
AT5G22880	HISTONE B2 (HTB2)
AT5G25460	DUF642 L-GALL RESPONSIVE GENE 2 (DGR2)
AT5G26742	DEAD-BOX RNA HELICASE 3 (RH3)
AT5G42100*	**BETA-1,3-GLUCANASE_PUTATIVE (BG_PPAP)**
AT5G44130	FASCICLIN-LIKE ARABINOGALACTAN PROTEIN 13 PRECURSOR (FLA13)
AT5G44340	TUBULIN BETA CHAIN 4 (TUB4)
AT5G66570	PS II OXYGEN-EVOLVING COMPLEX 1 (PSBO1)
ATCG00120	ATP SYNTHASE SUBUNIT ALPHA (ATPA)
ATCG00270	PHOTOSYSTEM II REACTION CENTER PROTEIN D (PSBD)
ATCG00280	PHOTOSYSTEM II REACTION CENTER PROTEIN C (PSBC)
ATCG00480	ATP SYNTHASE SUBUNIT BETA (PB)
ATCG00490	RUBISCO LARGE SUBUNIT (RBCL)
ATCG00680	PHOTOSYSTEM II REACTION CENTER PROTEIN B (PSBB)

3.3.3 Proteome-recurrent proteins are more enriched at plasmodesmata than expected by chance

It has been previously shown in *Populus trichocarpa* that plasmodesmata enriched proteins have different properties to the general membrane proteins (Leijon *et al.*, 2018). For example, the authors found that enriched plasmodesmal proteins were more likely to have transmembrane domains compared to a microsomal fraction. Moreover, they showed that the number of amino acids in the transmembrane was increased in plasmodesmal proteins, speculating this could be due to a thicker plasmodesmal lipid bilayer (Leijon *et al.*, 2018). I wanted to consider whether the same was true for *A. thaliana* plasmodesmata enriched proteins (i.e. AtCells2_filtered) and for proteome-recurrent proteins (Table 3-2). I compared these *A. thaliana* gene lists to a microsomal fraction from the literature (Mitra *et al.*, 2007).

Table 3-3 Predicted properties of plasmodesmal proteins.
Protein properties were compared to the microsome fraction within species. Significance was determined using a bootstrap method. * = p < 0.05, † = p < 0.1. TMD = transmembrane domain. The 'mean TMD count' only includes proteins with a TMD.

| | *A. thaliana* | | | *P. trichocarpa* | |
	Microsome	AtCells2_ filtered	Proteome- recurrent	Microsome	PtCells1_ filtered
Number of proteins	630	115	64	1393	201
Contains TMD (%)	26	**65***	25	20	**61***
Mean TMD count	4.3	4.0	**2.2***	3.1	**4.3***
Mean TMD length (amino acids)	23.2	24.3	23.8	22.5	**24.8***
Contains a signal peptide (%)	12	**32***	20†	9	**38***
Median Bayer Plasmodesmal Enrichment Factor	3.8	**87.1***	**6.7***	NA	NA

First, I could replicate the results of Leijon *et al.* (2018) and found that proteins that are enriched at plasmodesmata (PtCells1_filtered) are statistically more likely to contain signal peptides and at least one transmembrane domain. Proteins with transmembrane domains are more likely to have more transmembrane domains than microsome proteins, and each domain is on average 2.3 amino acids longer (Table 3-3).

3.3.3 Proteome-recurrent proteins are more enriched at plasmodesmata than expected by chance

Secondly, this work can be directly compared to the AtCells2_filtered list, as both are defined by enrichment of proteins are plasmodesmata. Similarly, to *P. trichocarpa*, plasmodesmata enriched proteins in *A. thaliana* (AtCells2_filtered) are more likely to contain at least one transmembrane domain and a signal peptide. However, the number of transmembrane passes nor the length of each transmembrane domain significantly differs from microsome transmembrane proteins (Table 3-3). This difference could have been due to differences in the microsome controls between *A. thaliana* and *P. trichocarpa*: if the AtCells2_filtered list were compared against the *P. trichocarpa* control, then both metrics would be significant. However, the microsome control does not vary far from the average transmembrane length of all predicted transmembrane domains in the *A. thaliana* proteome (0.3 amino acids difference, 22.9 on average). Moreover, the reciprocal is not true, if the transmembrane proteins of PtCells1_filtered were compared against the *A. thaliana* microsomal proteins, the domain length still remains larger ($p < 0.05$). Thus, the difference in the controls are not the cause in the difference in significance.

Thirdly, the same set of comparisons can be made for the proteome-recurrent list of proteins (Table 3-2) against generic membrane proteins (microsomal) (Mitra *et al.*, 2007). Unlike plasmodesmata-enriched proteins, the proteome-recurrent proteins were not more likely to contain transmembrane domains, and those that did had significantly fewer transmembrane passes. The only observed similarity to plasmodesmata-enriched proteins was that there was a weak signal of an increased proportion of signal peptides ($p = 0.086$). Naturally, the AtCells2_filtered which was, in part, defined by plasmodesmal enrichment had a significantly higher plasmodesmal enrichment factor. Unsurprisingly, the proteome-recurrent proteins also had an increased plasmodesmal enrichment factor, perhaps as they came from a plasmodesmal proteome. Notably, though, this was significantly higher than AtPlant1 (4.54, $p < 0.05$).

Overall, proteins that are enriched at plasmodesmata are more likely to contain a transmembrane domain than general membrane proteins and are more likely to contain a signal peptide. However, the transmembrane properties are not constant across species. Proteins which are repeatedly found in proteomic studies are also enriched at plasmodesmata, but are not more likely to be transmembrane proteins. Thus, transmembrane proteins appear to be enriched at plasmodesmata, but transmembrane domains do not define which proteins are present at plasmodesmata. Moreover, transmembrane domains are not required for plasmodesmal protein enrichment.

3.3.4   Defining a novel *Physcomitrella patens* plasmodesmata proteome

To define a novel *Physcomitrella patens* plasmodesmata proteome, I extracted plasmodesmata from five-week-old plants and characterised the proteome of an equal mix of protonema and gametophore tissue (Figure 3-2). Proteins were considered positively identified in the same manner as AtPlant1 and PtCells1. Under these conditions, 215 proteins were identified in the plasmodesmal purified extracts, which shall be referred to as PpPlant1. This plasmodesmal proteome derived from a Bryophyte, expands the sampling of plasmodesmata proteins across the phylogenetic tree.

Well established plasmodesmata marker proteins, such as the PDLPs, are not present in the *P. patens* genome (Lee, 2014; Vaattovaara *et al.*, 2019). As there are no known markers of plasmodesmata in *P. patens*, the efficacy of plasmodesmata extraction in *P. patens* could not be verified by Western blot before mass spectrometry. Therefore, to test whether the established mature plasmodesmata extraction protocol works in *P. patens*, a GO term analysis was conducted on PpPlant1.

Using Panther, *P. patens* identifiers can be directly tested for cellular localisation over-representation of a gene list. 185 (86%) of the UniProt identifiers were mapped to the database. Plasmodesma-annotated genes were significantly over-represented (7 proteins, *p* = 3.19 x 10^{-5}, 0.51 proteins expected) in PpPlant1. The degree of over-representation of different cellular components by GO analysis was reduced, but in largely overlapping terms AtPlant1 (7/9 categories, Figure 3-7). The values are likely reduced due to the smaller component sizes in the Panther database, due to poor annotation via phylogenetic backpropagation of GO terms (Gaudet *et al.*, 2011). Therefore, I concluded that the plasmodesmata extraction protocol enables plasmodesmal purification in *P. patens* and PpPlant1 is a *bona fide* bryophyte plasmodesmal proteome.

Figure 3-7 PpPlant1 is over represented for plasmodesmal proteins.
PpPlant1 and AtPlant1 proteomes were tested for over representation of the selected cellular compartment GO terms as Figure 3-5 ($p < 0.05$). The significance of over representation is given by the p value, which combines the size of the proteome, the size of the GO term, and the number of proteins matched in the proteome by the GO term. p values are plotted as $-log_{10}(p)$ so that a larger number is more significant. AtPlant1 and PpPlant1 were over represented for all the same GO terms, bar plasma membrane and vacuole for which PpPlant1 was not over represented.

3.3.5 Finding *A. thaliana* homologues of the PpPlant1

Previously, in 3.3.1 – 3.3.3, I demonstrated how the overlapping of independent proteomes from the same species can be used to select proteins that are enriched at plasmodesmata (Figure 3-4, Figure 3-6, Table 3-1). Here, I aim to extend this technique to proteomes of differing species to find phylogenetically conserved plasmodesmal proteins. I hypothesise these proteins are critical plasmodesmal components, as they have been either conserved at or repeatedly recruited to plasmodesmata. However, to find conservation between *A. thaliana* and *P. patens* proteomes, a shared naming convention is required that spans the evolutionary gap. I explore three methods of finding homologous proteins, either by converting all proteins to their closest *A. thaliana* protein or creating orthogroups.

3.3.5.1 *Homologue database searches: one-to-one*

Initially, I translated *P. patens* proteins to *A. thaliana* orthologues using database lookup tools. I used two different databases: InParanoid 8.0 and MetaPhOrs. InParanoid defines orthogroups (all proteins from an ancestral protein) for *A. thaliana* seed proteins using a pairwise BLAST approach, whereas MetaPhOrs creates orthogroups from a meta-analysis of many homologue databases (including InParanoid). For each protein in PpPlant1, I converted the *P. patens* identifier to the seed *A. thaliana* identifier (Table 3-4). Note, as multiple *P. patens* proteins can map to the same orthogroup, the number of unique loci is lower than the number of proteins matched.

Of those proteins successfully matched in each search, there were 41 PpPlant1 members identified by both methods. Only 20 of these 41, were mapped to the same *A. thaliana* locus (18 unique TAIR loci). We can have more confidence in the homology of proteins independently matched to the same *A. thaliana* locus. However, this only represents 9% of PpPlant1.

As with the AtPlant1 and PpPlant1, I ran a GO enrichment analysis for plasmodesmata cellular localisation with the identified *A. thaliana* homologues. All the tested lists of *A. thaliana* homologues were significantly enriched for plasmodesmal proteins (Table 3-4).

Table 3-4 GO enrichment for *A. thaliana* homologues from the PpPlant1 proteome.

Database	PpPlant1 proteins matched to *A. thaliana* proteins (%)	Unique *A. thaliana* loci	Plasmodesmata GO term over representation (FDR-corrected *p* value)
InParanoid	62 (29%)	56	7.11×10^{-16}
MetaPhOrs	50 (23%)	43	2.82×10^{-8}
Conserved Intersection	20 (9%)	18	1.20×10^{-4}

These data further confirmed that the plasmodesmata extraction protocol works for *P. patens*. However, the poor consistency between databases, and low conversion rate prompted me to find another way to compare inter-species gene lists.

3.3.5.2   HMMER sequence homology: one-to-many

Instead of relying on databases to convert *P. patens* proteins to *A. thaliana* homologues, HMMER (v3.3) can be used to find the closest homologue for any set of proteins in any other set. Therefore, I ran the protein sequences of PpPlant1 from UniProt against the *A. thaliana* proteome (Araport11).

HMMER returns an E-value (the number of hits expected to have a sequence bit score as high if the database contained only nonhomologous random sequences) with a one-to-many mapping of each *P. patens* protein above a set threshold. Using two arbitrary thresholds of E < 1×10^{-50} and E < 1×10^{-100}, HMMER matched 147 (68%) and 80 (37%) *P. patens* proteins, respectively. Even at these conservative values, a HMMER search matched more proteins than database lookup tools.

A one-to-many mapping makes it difficult to translate PpPlant1 to *A. thaliana* proteins. One approach would be to take the most significant (i.e. most likely) homologue for each protein.

However, taking *P. patens* A0A2K1JXU2 ("X8 domain-containing protein") as an example there are two almost indistinguishable top hits in *A. thaliana*: O49737 (E = 4.2 x 10^{-101}) and Q8L837 (E = 6.3 x 10^{-101}). This suggests it is likely the ancestral protein of A0A2K1JXU2 has undergone a duplication event in *A. thaliana* giving two equally likely homologues. Therefore, it is unwise to compare proteome lists this way.

Instead of trying to define an exact match from *P. patens* to *A. thaliana,* it can be accepted that all hits above a threshold are equally probable homologues. In this way one *P. patens* protein, can translate to several similar *A. thaliana* proteins. Thus, for A0A2K1JXU2, both O49737 and Q8L837 would be considered homologues. In essence, building orthogroups which are restricted to one *P. patens* member. Thus, the 147 and 80 *P. patens* proteins which were found to have a homologue (E < 1 x 10^{-50} and E < 1 x 10^{-100}, respectively) translate to 1,126 and 326 *A. thaliana* proteins, respectively.

These enlarged lists can be used to compare PpPlant1 with previous *A. thaliana* proteomes. The number of proteins in each *A. thaliana* proteome which are matched by *P. patens* homologues can be obtained (Table 3-5). The percentage of the proteome matched by *P. patens* homologues are directly comparable, as PpPlant1 is constant. AtPlant1 has many more proteins in common with PpPlant1, than either of the AtCell proteomes. This could be due to the mass spectrometry being done on mature tissue or being done in the same laboratory, and so, in both cases, the results are more similar. It is also worth noting that a single *A. thaliana* protein could match multiple PpPlant1 (Table 3-7).

In addition, the number of individual *P. patens* proteins which are being matched by an *A. thaliana* proteome can be computed. These numbers will be confounded by the size of the search list, i.e. it is expected that AtCells2 will match more of PpPlant1 than AtCells2_filtered as it is a larger list. That said, a majority of PpPlant1 is matched by the unfiltered proteomes, suggesting a large conservation of plasmodesmal proteins across evolutionary time (Table 3-6).

Table 3-5 *A. thaliana* proteins matched by PpPlant1 homologues defined by HMMER at two stringencies.
The number of *A. thaliana* proteins in each proteome, matched by PpPlant1 homologues. HMMER was run on PpPlant1 sequences against the *A. thaliana* proteome and filtered at two stringencies to give candidate homologues. Note that the number of proteome members differs from Table 3-1 as each locus maps to multiple UniProt identifiers.

Proteome	E < 1 x 10^{-50}		E < 1 x 10^{-100}	
	Matched	%	Matched	%
AtCells1	197	16	108	9
AtCells2	197	11	92	5
AtCells2_filtered	20	11	4	2
AtPlant1	121	35	65	19
Proteome-recurrent (Table 3-2)	30	34	20	22

Table 3-6 Number of PpPlant1 proteins matched by their *A. thaliana* HMMER homologues and *A. thaliana* proteomes.
Note that the number of proteome members differs from Table 3-1 as PpPlant1 was filtered for proteins with at least one putative homologue (147 at E < 1 x 10^{-50} and 80 at E < 1 x 10^{-100}).

Proteome	E < 1 x 10^{-50}		E < 1 x 10^{-100}	
	Matched	%	Matched	Matched
AtCells1	101	69	57	71
AtCells2	88	60	43	54
AtCells2_filtered	9	6	2	3
AtPlant1	93	63	44	55
Proteome-recurrent (Table 3-2)	43	29	23	29

Table 3-7 Non-redundant one-to-many mapping of PpPlant1 proteins to *A. thaliana* proteins leads to individual *A. thaliana* proteins matching numerous PpPlant1 members

PpPlant1 proteins matched by single *A. thaliana* protein	E < 1 x 10^{-50}	E < 1 x 10^{-100}
1	555	241
2	440	78
3	32	7
4	38	0
5	5	0
6	33	0
7	23	0

While this is encouraging, the HMMER method of assigning homologues has a major drawback. To find phylogenetically conserved proteins, i.e. to concurrently compare several lists among several species, one list would have to be chosen as the reference frame. For the *A. thaliana* comparison, I have exemplified this with the PpPlant1 as the reference (as in Table 3-6) meaning that AtCells1, AtCells2_filtered, and AtPlant1 are compared to it (Figure 3-8). The distribution of *P. patens* hits across these proteomes can be compared, but any nuance between the *A. thaliana* lists are lost. There are two proteins identified in all four proteomes: A0A2K1JXU2, and A0A2K1J8R8; both are β-1,3-glucanases.
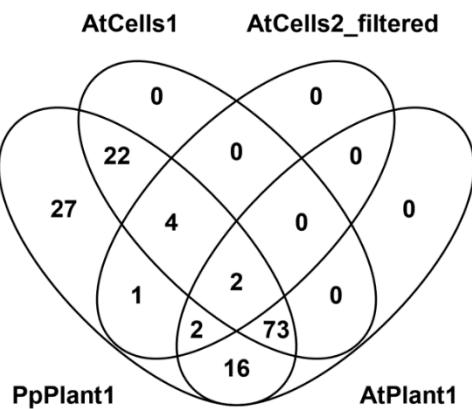
Figure 3-8 Venn diagram of HMMER PpPlant1 homologues.
PpCells1 used as a reference database for a phmmer search of *A. thaliana* proteins from AtCells1, AtCells2_filtered, and AtPlant1. An *A. thaliana* protein was deemed to be a *P. patens* homologue, if any *P. patens* protein matched the *A. thaliana* protein with $E < 1 \times 10^{-50}$. This method allows the comparison of *A. thaliana* and *P. patens* proteomes, but prevents the comparison of *A. thaliana* proteins not present in PpPlant1.
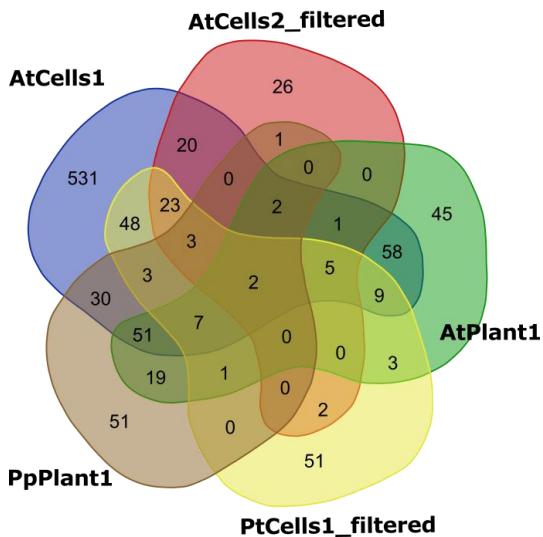
3.3.5.3   OrthoFinder: many-to-many

So far, a one-to-one and a one-to-many approach has been attempted to translate between protein homologues. I tried a third many-to-many method approach by forming *de novo* orthogroups using the OrthoFinder software (Emms & Kelly, 2019). OrthoFinder uses a pairwise BLAST approach to build orthogroups from an input set of protein sequences. I used OrthoFinder (v2.2.6) to define orthogroups between the five defined proteomes: AtCells1, AtCells2_filtered, AtPlant1, PtCells1_filtered, and PpPlant1.

OrthoFinder returned 992 orthogroups, of which 289 had more than one member and 288 contained proteins from multiple proteomes (Figure 3-9). Two orthogroups had members from all proteomes, and 17 had members from four of the five proteomes. These 19 orthogroups contain proteins I define as 'phylogenetically conserved plasmodesmal proteins' (Table 3-8).

Table 3-8 Phylogenetically conserved protein orthogroups within plasmodesmata.

Orthogroup	Protein Class	Number of Proteomes in orthogroup	Number of Proteins in orthogroup	Focus
OG000	β-1,3-glucanase	5	27	Yes
OG001	Peroxidase	4	22	No
OG003	C2 lipid-binding	4	16	Yes
OG004	SKU5	4	13	Yes
OG005	GDSL esterase/lipase	4	13	Yes
OG006	Tetraspanin	4	12	Yes
OG007	ATP-binding cassette	4	11	No
OG008	Aspartyl protease	4	10	No
OG009	Leucine-rich repeat receptor-like kinase	4	10	No
OG010	Leucine-rich repeat extensin-like	4	10	Yes
OG013	Histone H2B	4	9	No
OG014	Tubulin beta-7	4	9	No
OG016	RNA-binding glycine-rich protein	4	8	Yes
OG018	Inflorescence meristem receptor-like kinase 2	5	7	Yes
OG019	DUF26 containing protein	4	7	Yes
OG028	Eukaryotic translation initiation factor 4A	4	6	No
OG040	Subtilisin-like protease	4	5	No
OG050	Calcium-dependent lipid-binding	4	4	Yes
OG063	Ribosomal protein	4	4	No

Figure 3-9 Venn diagram representing proteome constituency of 992 orthogroups
Orthogroups of proteins from AtCells1, AtCells2_filtered, AtPlant1, PpPlant1, and PtCells1_filtered were made *de novo* using Orthofinder. Orthogroups are sorted by their representation within each proteome.

As aforementioned the AtCells2_filtered and PtCells1_filtered proteomes are helpful at filtering out common contaminants, as they are defined by a plasmodesmal enrichment factor. Using this, and a Bayesian idea that the probability of an event is dependent on prior knowledge, it is a worthwhile exercise to consider which orthogroups are identified when the protein and peptide identification level is set to a lower threshold in the 'Plant' proteomes, risking misidentifying proteins. In doing so, contaminants are still removed using the 4/5 proteome threshold by the two filtered proteomes and proteins which we would otherwise think are misidentified, are more likely to be true as there is prior evidence they are plasmodesmal from the other proteomes. Therefore, I reran the Orthofinder software on peptide lists generated from AtCells1, AtCells2_filtered, PtCells1_filtered, and the loose proteomes of AtPlant1_loose and PpPlant1_loose: if the protein (50% certainty) was present in at least one of the three samples by at least one peptide (50% certainty).

OrthoFinder returned 1,280 orthogroups, of which 426 had more than one member and 424 contained proteins from multiple proteomes (Figure 3-10). Seven orthogroups had members from all proteomes, and 30 had members from four of the five proteomes. This list contains proteins which may merit further investigation due to their high phylogenetic conservation, but may have been misidentified in AtPlant1_loose and PpPlant1_loose due to the lower thresholds (Table 3-9).

Table 3-9 Phylogenetically conserved protein orthogroups within plasmodesmata made from loose AtPlant1 and PpPlant1 proteomes.

Orthogroup	Protein Class	Number of Proteomes in orthogroup	Number of Proteins in orthogroup	In Table 3-8
OG000	Lectin receptor-like kinase	5	56	No
OG001	Peroxidase	4	25	Yes
OG002	C2 lipid-binding	4	20	Yes
OG003	GDSL esterase/lipase	4	17	Yes
OG004	Callose synthase	4	15	No
OG005	Tubulin beta-7	4	14	Yes
OG006	β-1,3-glucanase	5	14	Yes
OG007	SKU5	4	14	Yes
OG008	Glycine-rich RNA-binding	4	13	Yes
OG010	Leucine-rich repeat extensin-like	4	12	Yes
OG012	ATP-binding cassette	5	11	Yes
OG013	Aspartyl protease	4	10	Yes
OG015	Leucine-rich repeat receptor-like kinase	4	10	Yes
OG018	NDR1/HIN1-like protein	5	10	No
OG019	Transmembrane protein	5	10	No
OG020	Histone H2B	4	9	Yes
OG023	Heavy metal associated isoprenylated plant protein	4	9	No
OG026	DUF26 containing protein	4	8	Yes
OG030	Eukaryotic initiation factor 4A	4	7	Yes
OG033	CSC1-like protein ERD4	5	7	No
OG037	Xyloglucan endotransglucosylase/ hydrolase	4	7	No
OG042	Subtilisin-like protease	4	6	Yes
OG045	Serine carboxypeptidase-like	4	6	No
OG071	Polyubiquitin	4	5	No

OG073	Prohibitin	4	5	No
OG076	Calcium-dependent lipid-binding	5	5	Yes
OG085	Tetraspanin	4	5	Yes
OG094	Transmembrane protein	4	4	No
OG106	Ribosomal protein	4	4	Yes
OG107	β-glucosidase	4	4	No
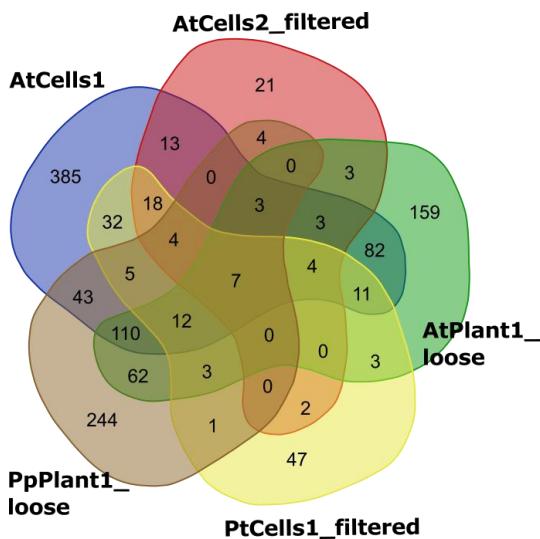
Figure 3-10 Venn diagram of the proteome constituency of 1,280 orthogroups generated using loose AtPlant1 and PpPlant1 proteomes.
Orthogroups of proteins from AtCells1, AtCells2_filtered, AtPlant1_loose, PpPlant1_loose, and PtCells1_filtered were made *de novo* using Orthofinder. Orthogroups are sorted by their representation within each proteome.

3.3.6   Investigating phylogenetically conserved plasmodesmal orthogroups

Ultimately, I wanted to compare the localisation of members of phylogenetically conserved plasmodesmal proteins from both *A. thaliana* and *P. patens*. I refined the list of orthogroups from Table 3-8 for further phylogenetic analysis. I chose not to proceed with orthogroups that had a well-known localisation and function outside of plasmodesmata (e.g. histone H2B) or were exceptionally large multi-gene families where conserved orthology could be due to chance (e.g. leucine-rich repeat receptor-like kinase). In addition, I did not analyse the canonical PDLP clade (OG019, DUF26 containing protein), as it well established that there are no *P. patens* homologues (Vaattovaara *et al.*, 2019). Having selected 9 groups for further study, I investigated the phylogenetic pattern of plasmodesmal proteins within the wider family of proteins. Doing so, I was able to select members for localisation experiments.

To examine the pattern of plasmodesmal localisation in each orthogroup, I first retrieved non-plasmodesmal proteins missing from the constructed orthogroup. To do so, I aligned all of the protein sequences within the orthogroup to make a search pattern. I searched (hmmsearch) the UniProt Reference Proteomes database for similar proteins (E < 1 x 10^{-100}, unless otherwise stated) within the three species (*A. thaliana*, *P. patens*, *P. trichocarpa*). With the enlarged group, I drew unrooted phylogenetic trees with 100 bootstraps. Nodes with weak support (<70) have been marked by white circles (Figures 3-11 – 3-20).

I compared the tree topologies with the location of proteins identified by mass spectrometry. By doing so, I identified four orthogroups where plasmodesmal proteins clearly grouped to

one part of the tree (OG003, 005, 016, 018); one orthogroup where plasmodesmal proteins appeared evenly distributed among the tree (OG006); and four orthogroups where no pattern was distinguishable (OG000, 004, 010, 050). Orthogroups with a clear phylogenetic signal clearly indicate which members to localise, as plasmodesmal members are clustering. Moreover, it may suggest an ancestral plasmodesmal protein which can been retained at plasmodesmata. The same rings true for groups where plasmodesmata proteins are ubiquitous – the founding members may have been plasmodesmal. Lastly, groups without a clear signal, do not lend themselves to localisation studies as divergent members are likely to have been recruited independently to plasmodesmata.

3.3.6.1 Orthogroups with a clear phylogenetic grouping

3.3.6.1.1 OG003 – C2 lipid-binding
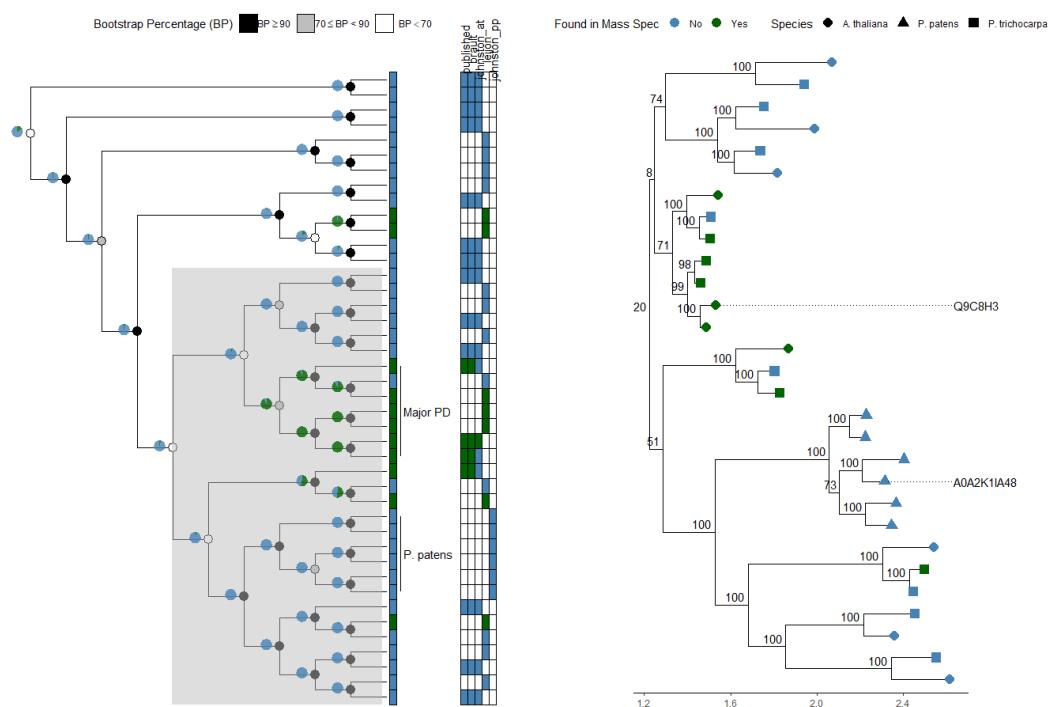
Figure 3-11 Phylogenetic tree of C2 lipid-binding proteins and selected homologues to clone.
Left) Unrooted cladogram of homologues of OG003 from *A. thaliana*, *P. trichocarpa*, and *P. patens* as defined by a hmmsearch with a threshold of $E < 1 \times 10^{-100}$. Proteome hits are marked in green. Pie charts estimate the likely ancestral plasmodesmal localisation by phylogenetic backpropagation. Node support is indicated by a greyscale circles. Centre) A heatmap of proteome matches for each protein. Right) Enlarged phylogenetic tree of grey box (left) with cloned proteins marked. Bootstrap values from 100 bootstraps are marked. Distance is measured in substitutions per base.

There is a clear monophyletic plasmodesmal clade (labelled "Major PD") within the C2 lipid-binding orthogroup, which contains *A. thaliana* and *P. trichocarpa* proteins. There is a sister expansion of *P. patens* C2 lipid-binding proteins, which contains no plasmodesmal proteomic hits (labelled "*P. patens*") (Figure 3-11).

As there are no *P. patens* proteomic hits, I would not have normally chosen to clone proteins from this group to localise. However, C2 lipid-binding proteins have been postulated to be required for desmotubule formation (Brault *et al.*, 2019) and *P. patens* has a desmotubule (Cook *et al.*, 1997). Therefore, it is an interesting hypothesis as to whether P. patens C2 lipid-binding proteins localise to plasmodesmata.

Thus, I chose MCTP4 (Q9C8H3, At1g51570) to clone as it has the most proteome support of all the *A. thaliana* members (3/3, Table 3-2) and has been well characterised in the literature (Brault *et al.*, 2019). I then selected Pp3c10_5480 (A0A2K1IA48) to clone from *P. patens* as it has the closest homology to MCTP4 in a HMMER phmmer search (E = 2.5 × 10^{-308}).

3.3.6.1.2   OG005 – GDSL esterase/lipase
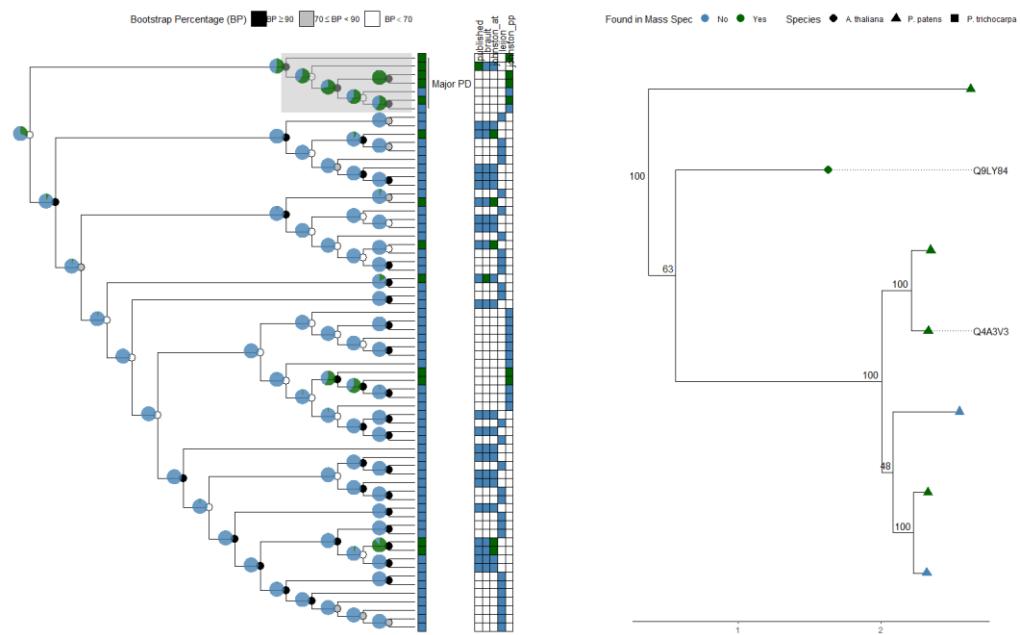
Figure 3-12 Phylogenetic tree of GDSL esterase/lipase proteins and selected homologues to clone.
Left) Unrooted cladogram of homologues of OG005 from *A. thaliana*, *P. trichocarpa*, and *P. patens* as defined by a hmmsearch with a threshold of E < 1 x 10^{-100}. Proteome hits are marked in green. Pie charts estimate the likely ancestral plasmodesmal localisation by phylogenetic backpropagation. Node support is indicated by a greyscale circles. Centre) A heatmap of proteome matches for each protein. Right) Enlarged phylogenetic tree of grey box (left) with cloned proteins marked. Bootstrap values from 100 bootstraps are marked. Distance is measured in substitutions per base. Note that A0A2K1YJK0 was added back manually as it fell under the HMMER threshold (E = 1.0 × 10^{-89}).

GDSL esterase/lipases are a non-canonical class of lipolytic enzymes with a GDSL motif compared to the canonical GxSxG motif (Upton & Buckley, 1995). There are at least 105 GDSL-type esterase/lipase genes in *A. thaliana*, which have been shown to have diverse functions from the biochemical to physiological level (Lai *et al.*, 2017). They accept a wide variety of substrates, with mutants having a subsequent range of phenotypes from aberrant seed germination to changes in biotic and abiotic stress responses (Ding *et al.*, 2019).

Within the GDSL esterase/lipase tree, there was a monophyletic clade containing mostly plasmodesmal proteins (labelled "Major PD", Figure 3-12). Within this clade, there was only one *A. thaliana* protein, At5g14450 (Q9LY84, AtGELP91), which I chose to clone, as it was identified at plasmodesmata by mass spectrometry.

Of the six *P. patens* proteins, within the same clade, four were detected by mass spectrometry. I selected Pp3c18_1550 (Q4A3V3) to clone, as it has the closest homology to At5g14450 in a HMMER phmmer search (E = 9.1×10^{-41}) of the four proteins. In addition, it had the highest unique peptide count of all four proteins (10, 3, 1 in each replicate).

3.3.6.1.3   OG016 – RNA-binding glycine-rich protein
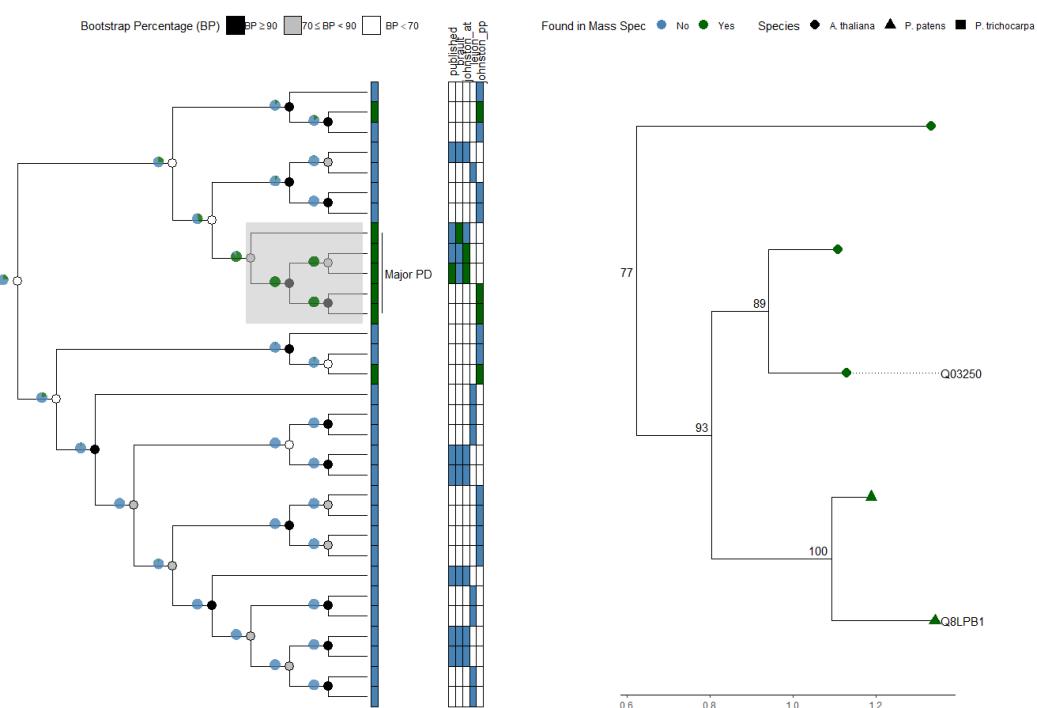
Figure 3-13 Phylogenetic tree of RNA-binding glycine-rich proteins and selected homologues to clone.
Left) Unrooted cladogram of homologues of OG016 from *A. thaliana*, *P. trichocarpa*, and *P. patens* as defined by a hmmsearch with a threshold of E < 1×10^{-50}. Proteome hits are marked in green. Pie charts estimate the likely ancestral plasmodesmal localisation by phylogenetic backpropagation. Node support is indicated by a greyscale circles. Centre) A heatmap of proteome matches for each protein. Right) Enlarged phylogenetic tree of grey box (left) with cloned proteins marked. Bootstrap values from 100 bootstraps are marked. Distance is measured in substitutions per base.

Within the monophyletic clade which contained the most plasmodesmal proteins (labelled "Major PD", Figure 3-13), there were three *A. thaliana* proteins. I selected to clone GRP7 (At2g21660, Q03250), as it was present in the most proteomes (2/3). In the same clade, there were two *P. patens* proteins both of which were detected by mass spectrometry. I decided to clone GRP2 (Q8LPB1, Pp3c11_19620), as it had the most unique peptides in the proteome (0, 2, 2) compared to A0A2K1KA61 (0, 1, 2). However, A0A2K1KA61 had closer homology to

GRP7 than Pp3c11_19620, in a phmmer search (E = 1.2 × 10^{-43} opposed to E = 1.2 × 10^{-42}). Note that HMMER threshold was increased to E < 1.0 × 10^{-50}, to increase the number of proteins included in the tree from two at E < 1.0 × 10^{-100}.

3.3.6.1.4   OG018 – INFLORESCENCE MERISTEM RECEPTOR-LIKE KINASE 2
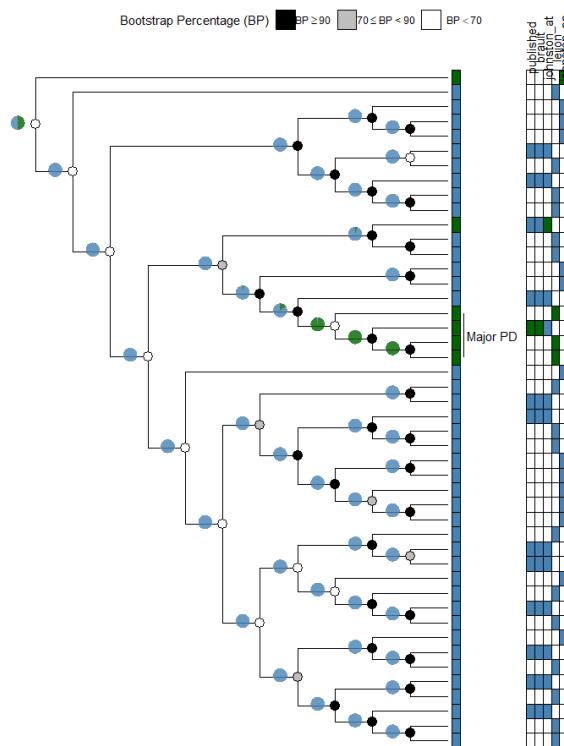
Figure 3-14 Phylogenetic tree of INFLORESCENCE MERISTEM RECEPTOR-LIKE KINASE 2 proteins.
Left) Unrooted cladogram of homologues of OG018 from *A. thaliana*, *P. trichocarpa*, and *P. patens* as defined by a hmmsearch with a threshold of E < 1 x 10^{-230}. Proteome hits are marked in green. Pie charts estimate the likely ancestral plasmodesmal localisation by phylogenetic backpropagation. Node support is indicated by a greyscale circles. Right) A heatmap of proteome matches for each protein.

As receptor-like kinase proteins are highly homologous, I decreased the threshold to E < 1 × 10^{-230} to reduce the number of proteins from >700. However, this would have excluded A0A2K1KXD6_PHYPA which was manually added back (E = 5.6 × 10^{-198}).

I did not clone any proteins from OG018, as the monophyletic clade which contained the most plasmodesmal proteins (labelled "Major PD", Figure 3-14), contained no *P. patens* proteins. The *P. patens* protein in OG018 that was detected by mass spectrometry (A0A2K1KXD6_PHYPA) acted as an outgroup and is likely not phylogenetically related.

3.3.6.2   Ubiquitous throughout the phylogenetic tree

3.3.6.2.1   OG006 – Tetraspanin
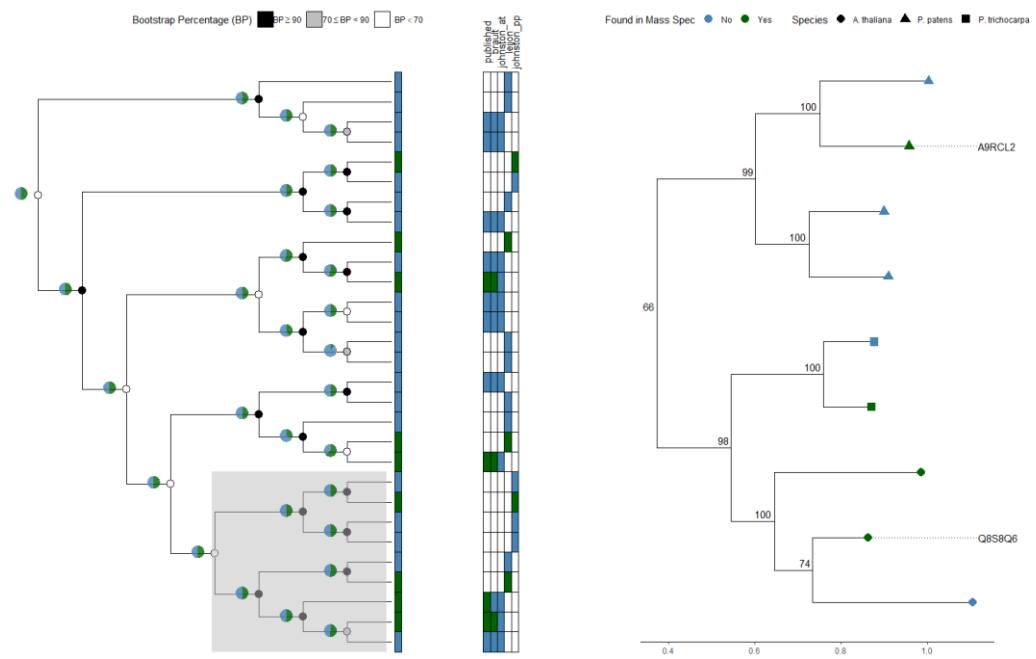
Figure 3-15 Phylogenetic tree of tetraspanin proteins and selected homologues to clone.
Left) Unrooted cladogram of homologues of OG006 from *A. thaliana*, *P. trichocarpa*, and *P. patens* as defined by a hmmsearch with a threshold of E < 1 x 10^{-50}. Proteome hits are marked in green. Pie charts estimate the likely ancestral plasmodesmal localisation by phylogenetic backpropagation. Node support is indicated by a greyscale circles. Centre) A heatmap of proteome matches for each protein. Right) Enlarged phylogenetic tree of grey box (left) with cloned proteins marked. Bootstrap values from 100 bootstraps are marked. Distance is measured in substitutions per base.

Phylogenetic backpropagation of the plasmodesmal state found that all tetraspanin ancestors were almost equally likely to have been located at plasmodesmata (Figure 3-15). Therefore, instead of choosing a monophyletic plasmodesmal clade, I chose to clone the *A. thaliana* and *P. patens* proteins which most closely clustered and were detected by mass spectrometry. I chose the *P. patens* protein Pp3c7_23740 (A9RCL2). Three *A. thaliana* proteins are in a sister clade, of which two were identified by mass spectrometry. Of these two, TET8 (At2g23810, Q8S8Q6) was recovered in more proteomes than the other (2/3). Note that HMMER threshold was increased to E < 1.0 × 10^{-50}.

3.3.6.3   No clear phylogenetic signal

3.3.6.3.1   OG000 – β-1,3-glucanase
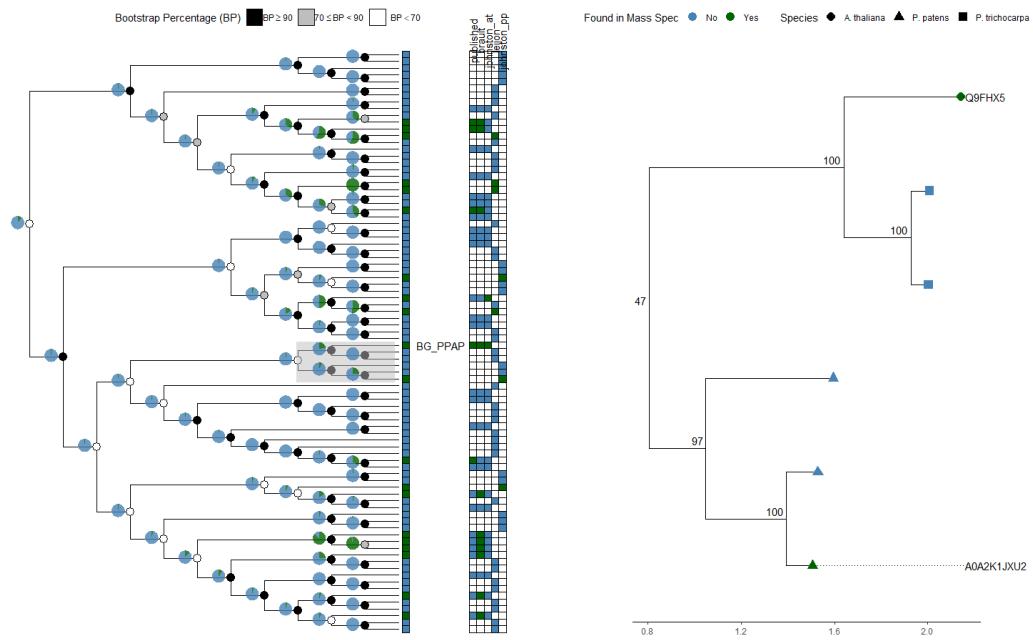
Figure 3-16 Phylogenetic tree of β-1,3-glucanase proteins and selected homologues to clone.
Left) Unrooted cladogram of homologues of OG000 from *A. thaliana*, *P. trichocarpa*, and *P. patens* as defined by a hmmsearch with a threshold of E < 1 x 10^{-100}. Proteome hits are marked in green. Pie charts estimate the likely ancestral plasmodesmal localisation by phylogenetic backpropagation. Node support is indicated by a greyscale circles. Centre) A heatmap of proteome matches for each protein. Right) Enlarged phylogenetic tree of grey box (left) with cloned proteins marked. Bootstrap values from 100 bootstraps are marked. Distance is measured in substitutions per base.

There is no clear pattern of plasmodesmal localisation within the tree, but as β-1,3-glucanases are known to be essential for plasmodesmata function I decided to clone two proteins to explore whether this is true of *P. patens*. I selected BG_PPAP (Q9FHX5, At5g42100) to clone as it has the most proteome support of all the *A. thaliana* proteins (3/3, Table 3-2) and has been characterised in the literature (Levy *et al.*, 2007). A sister clade to BG_PPAP contained three *P. patens* proteins. I chose Pp3c10_5480 (A0A2K1JXU2) to clone, as it was the only member with proteomic support (Figure 3-16).

3.3.6.3.2   OG004 – SKU5
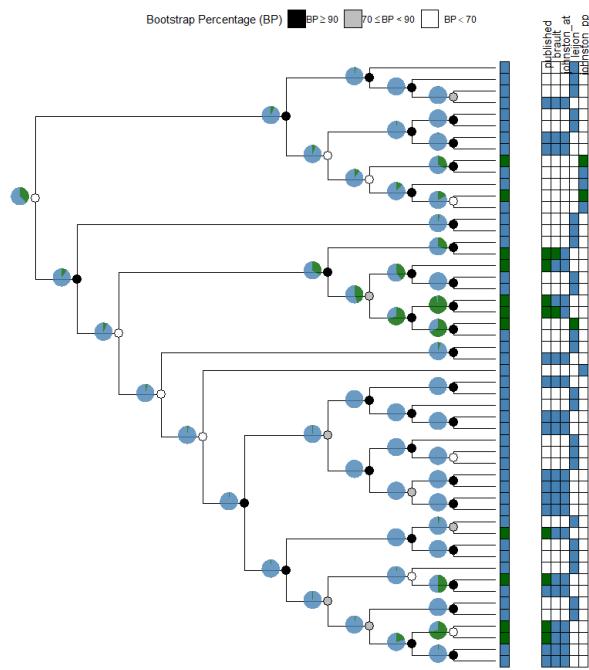
Figure 3-17 Phylogenetic tree of SKU5 proteins.
Left) Unrooted cladogram of homologues of OG004 from *A. thaliana*, *P. trichocarpa*, and *P. patens* as defined by a hmmsearch with a threshold of E < 1 x 10^{-100}. Proteome hits are marked in green. Pie charts estimate the likely ancestral plasmodesmal localisation by phylogenetic backpropagation. Node support is indicated by a greyscale circles. Right) A heatmap of proteome matches for each protein.

SKU5 was identified in an EMS screen of *A. thaliana* plants with a root skewing phenotype (Sedbrook *et al.*, 2002). There appears to be three clades within the SKU5 group in which plasmodesmal proteins reside: a *P. patens* clade in which only *P. patens* proteins are represented, and two clades in which proteins from *A. thaliana* and *P. trichocarpa* are represented (Figure 3-17). These clades may warrant further investigation, though without clear candidates I did not follow this up.

3.3.6.3.3   OG010 – Leucine-rich repeat extensin-like
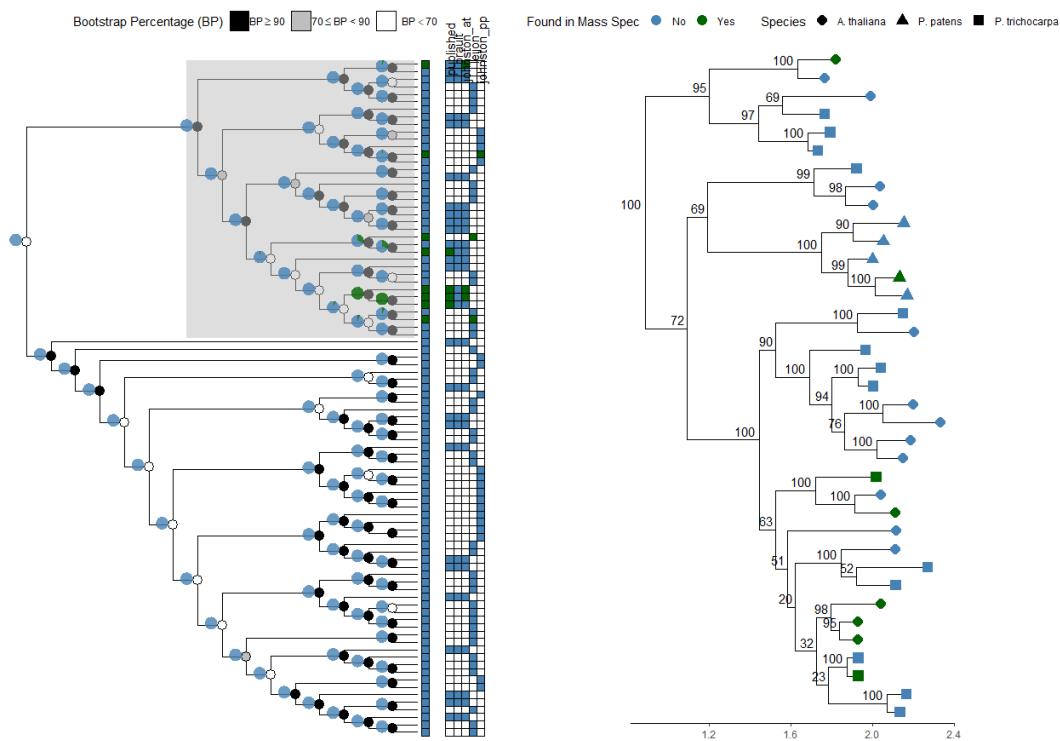
Figure 3-18 Phylogenetic tree of leucine-rich repeat extensin-like proteins.
Left) Unrooted cladogram of homologues of OG010 from *A. thaliana*, *P. trichocarpa*, and *P. patens* as defined by a hmmsearch with a threshold of $E < 1 \times 10^{-100}$. Proteome hits are marked in green. Pie charts estimate the likely ancestral plasmodesmal localisation by phylogenetic backpropagation. Node support is indicated by a greyscale circles. Centre) A heatmap of proteome matches for each protein. Right) Enlarged phylogenetic tree of grey box (left). Bootstrap values from 100 bootstraps are marked. Distance is measured in substitutions per base.

The phylogenetic tree of leucine-rich repeat extensin-like proteins is topologically split into two halves, into which all the plasmodesmal proteins fit into one half (Figure 3-18). Within this half, there is an ancient *P. patens* clade with few proteins from other species, and two exclusive *A. thaliana* and *P. trichocarpa* clades. There were no obvious candidates to clone and follow up.

3.3.6.3.4   OG050 – Calcium-dependent lipid-binding
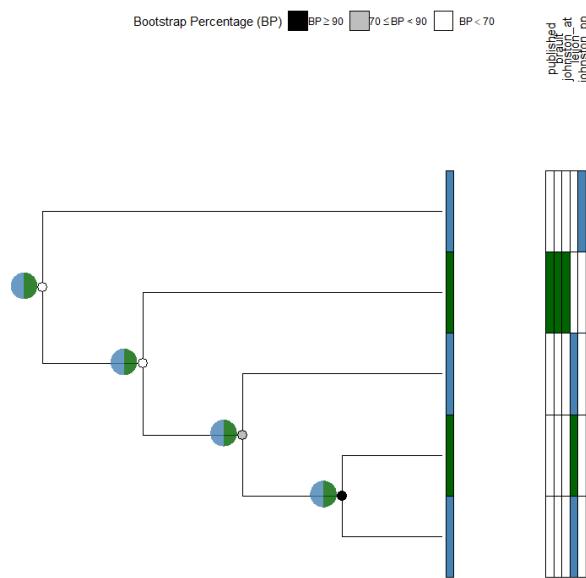
Figure 3-19 Phylogenetic tree of calcium-dependent lipid-binding proteins.
Left) Unrooted cladogram of homologues of OG050 from *A. thaliana*, *P. trichocarpa*, and *P. patens* as defined by a hmmsearch with a threshold of E < 1 x 10^{-100}. Proteome hits are marked in green. Pie charts estimate the likely ancestral plasmodesmal localisation by phylogenetic backpropagation. Node support is indicated by a greyscale circles. Right) A heatmap of proteome matches for each protein.

There were too few proteins retrieved from the three genomes by HMMER to make any inferences on likely plasmodesmal clustering within the calcium-dependent lipid-binding proteins (Figure 3-19). However, At4g34150 (Q945K9) makes a plausible target as it appears in all three *A. thaliana* proteomes (3/3, Table 3-2). Yet, without any phylogenetic signal I chose not to clone any proteins from OG050.

In summary, I chose to examine the localisations of 10 proteins (Table 3-10). Six of which had a clear phylogenetic pattern in which *A. thaliana* and *P. patens* homologues clustered together (OG005, OG016) or where most family members were found at plasmodesmata (OG006). Four of which there was strong biological evidence for plasmodesmal function (OG000, OG003).

Table 3-10 A summary of proteins chosen to be cloned from highly conserved plasmodesmal orthogroups across Viridiplantae

The table is divided into proteins with a clear phylogenetic pattern (top) and proteins with strong evidence of biological function (bottom).

Orthogroup	Species	UniProt	Identifier	Name	Reasoning
OG005 GDSL esterase/ lipase	*A. thaliana*	Q9LY84	At5g14450		Clustered with support
	P. patens	Q4A3V3	Pp3c18_ 1550		Most proteome support (10,3,1)
OG006 Tetraspanin	*A. thaliana*	Q8S8Q6	At2g23810	TET8	Most proteome support (2/3)
	P. patens	A9RCL2	Pp3c7_ 23740		Clustered with support
OG016 RNA-binding glycine-rich	*A. thaliana*	Q03250	At2g21660	GRP7	Most proteome support (2/3)
	P. patens	Q8LPB1	Pp3c11_ 19620	GRP2	Most proteome support (0,2,2)
OG000 β-1,3-glucanase	*A. thaliana*	Q9FHX5	At5g42100	BG_ PPAP	Most proteome support (3/3)
	P. patens	A0A2K1JXU2	Pp3c10_ 5480		Clustered with support
OG003 C2 lipid-binding	*A. thaliana*	Q9C8H3	At1g51570	MCTP4	Most proteome support (3/3)
	P. patens	A0A2K1IA48	Pp3c27_ 520		Best hit on phmmer search

3.3.7 Expression of conserved plasmodesmal proteins gives a range of localisations when transiently, heterologously expressed in *Nicotiana benthamiana*

I wanted to localise the identified list of highly phylogenetically-conserved proteins *in A. thaliana* (Table 3-10). First, this would confirm the plasmodesmal localisation of the *A. thaliana* proteins. Secondly, it would demonstrate the *in vivo* conservation of the *P. patens* homologue. To do so, I synthesised Golden Gate compatible coding sequences of the proteins of interest (see Table S1). The coding sequences were cloned into vectors with a strong constitutive promoter (35S) and a fluorescent tag (eGFP or mCherry). All proteins were tagged on their C-terminus, bar the β-1,3-glucanases due to a requisite C-terminal glycosylphosphatidylinositol (GPI) anchor (Figure 3-21). Thus, the β-1,3-glucanases were tagged on their N-terminus with a fluorescent tag (citrine) after the signal peptide. In addition, C2 lipid-binding proteins were N-tagged, after Brault *et al.* (2019), as were the RNA-binding glycine-rich proteins (Figure 3-22).

I heterologously expressed these proteins in *N. benthamiana*, outside the native system for all the protein families and both homologues. Although, *N. benthamiana* is a closer relative of *A. thaliana* than *P. patens* and so the localisation of *P. patens* proteins may be less representative of their localisation in *P. patens*. Nonetheless, a plasmodesmal localisation in a heterologous system is a strong indication of conservation of function. Moreover, it would suggest these proteins are integral to plasmodesmata across Viridiplantae. However, the localisation of proteins from closely related species, such as *A. thaliana*, can be different when transiently expressed in *N. benthamiana*. For example, TET3 has been shown to localise to plasmodesmata when stably expressed in *A. thaliana* (Fernandez-Calvino *et al.*, 2011)*,* yet localised to the endoplasmic reticulum when expressed transient in *N. benthamiana* (Faulkner, C., *pers. comms.*). It is unknown whether this is due to the transient expression of the protein, or the heterologous expression. Thus, a failure of a protein to localise at plasmodesmata in *N. benthamiana* is not conclusive as it might still be a plasmodesmal protein when expressed endogenously. However, a plasmodesmal localisation is strong evidence of the protein being a *bona fide* plasmodesmal protein.

Proteins of three of the five selected orthogroups have a distinct punctate localisation in *N. benthamiana*, indicative of a plasmodesmata localisation: tetraspanin, C2 lipid-binding, and β-1,3-glucanase. In the latter case, only the *A. thaliana* BG_PPAP localises to plasmodesmata, whereas Pp3c10_5480 does not. This may be because BG_PPAP has a GPI-anchor (99.9%, PredGPI), whereas Pp3c10_5480 does not (9.1%, PredGPI). In this case, the GPI anchor is required to tether the protein to the membrane: explaining the diffuse cell wall localisation

3.3.7 Expression of conserved plasmodesmal proteins gives a range of localisations when transiently, heterologously expressed in Nicotiana benthamiana

of Pp3c10_5480 with no GPI-anchor (Figure 3-22). Moreover, GPI anchors are known to localise proteins to plasmodesmata (Zavaliev *et al.*, 2016).

Both GDSL esterase/lipases are predicted to have a secretory signal peptide (SignalP 5.0) and localise to the endoplasmic reticulum when expressed in *N. benthamiana* (Figure 3-21). The RNA-binding glycine-rich proteins also did not localise to plasmodesmata, rather they had a nuclear/ cytosolic localisation with an enrichment of signal in the nucleolus (Figure 3-21, Figure 3-22).

The punctate localisations of the tetraspanins, C2 lipid-binding proteins, and BG_PPAP (as a positive control, Levy *et al.* (2007)) were confirmed to co-localise with plasmodesmata with aniline blue staining. Aniline blue stains callose, which accumulates at plasmodesmata and so can be used as a marker for plasmodesmata (Radford *et al.*, 1998). In all cases, the punctae co-localised with aniline blue, confirming the plasmodesmata localisation of these proteins when transiently expressed in *N. benthamiana* (Figure 3-23, Figure 3-24, Figure 3-20). As previously demonstrated in the literature, BG_PPAP co-localised with callose stained by aniline blue at plasmodesmata (Figure 3-20). Pp3c10_5480 maintained a cell wall localisation.
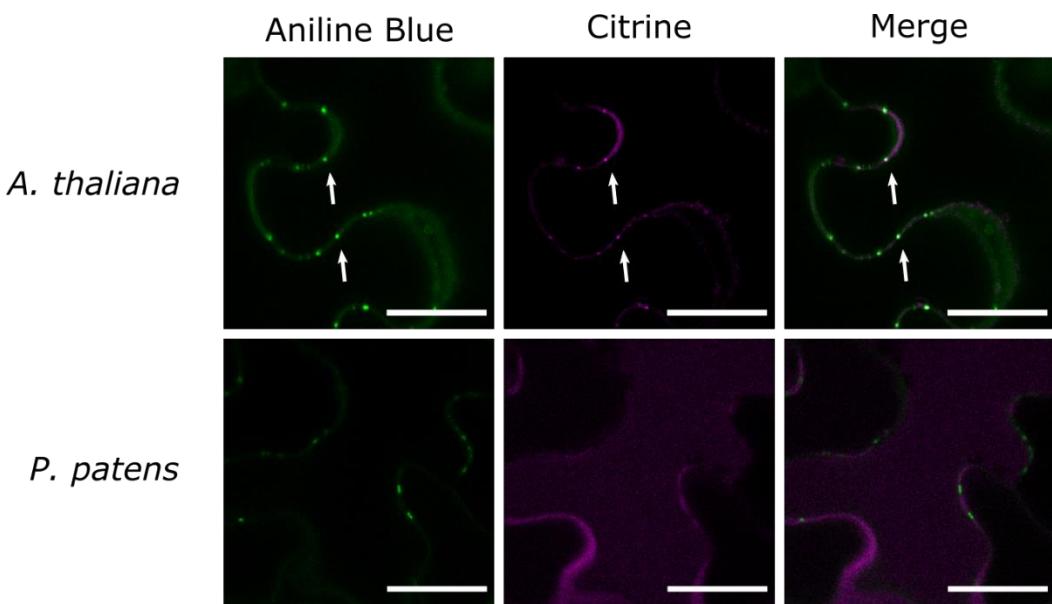
Figure 3-20 BG_PPAP co-localises with callose at plasmodesmata
BG_PPAP (At5g42100) localises to plasmodesmata, as defined by aniline blue staining of callose (examples shown by arrows). Pp3c10_5480 maintains a cell wall localisation. Scale bar = 25 μm.
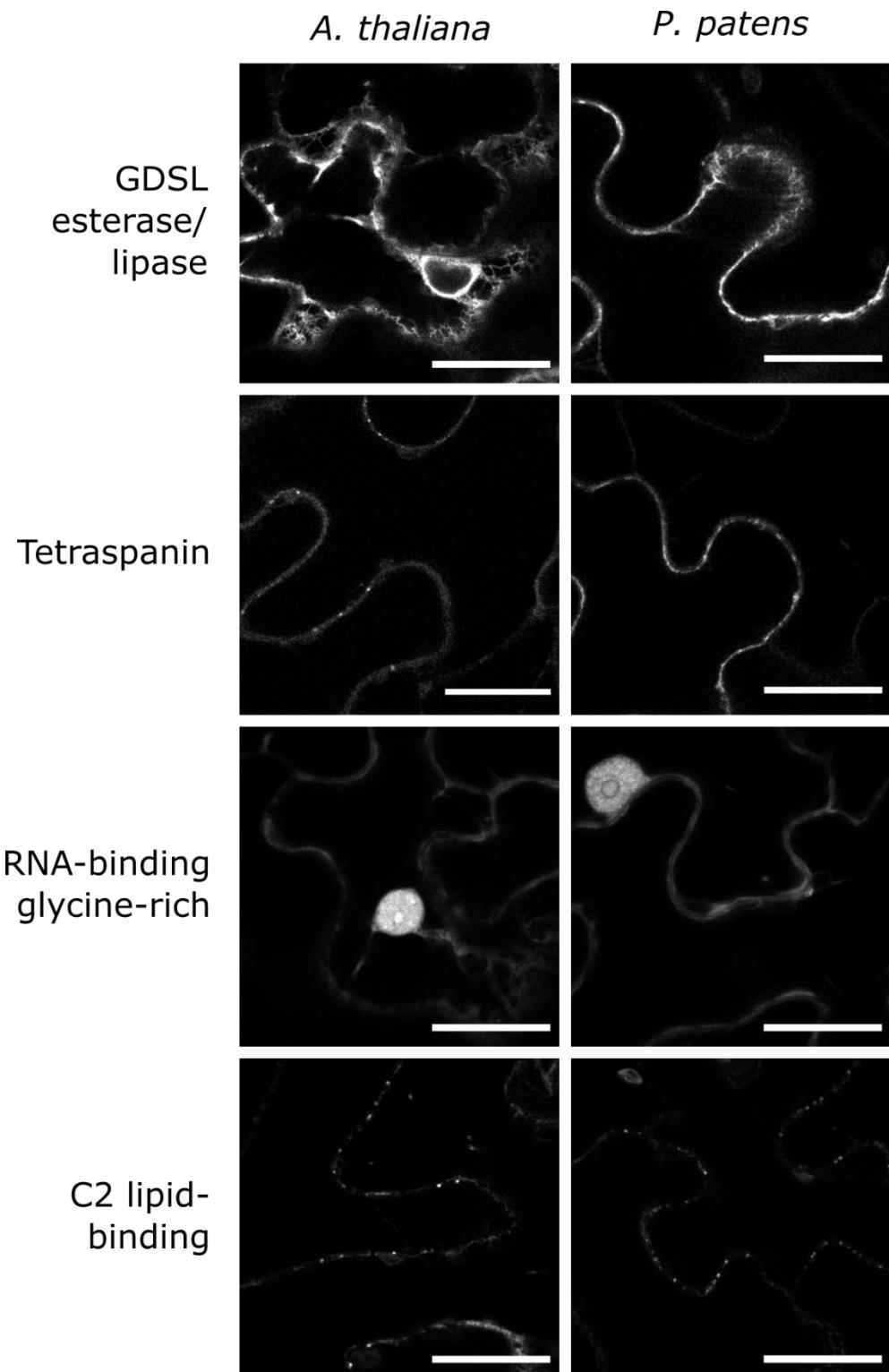
Figure 3-21 Localisation of conserved plasmodesmal proteins when C-tagged with fluorescent proteins and transiently expressed in *N. benthamiana*.

A. thaliana and *P. patens* homologues of phylogenetically conserved plasmodesmal proteins (Table 3-10) were tagged at the C-terminus with eGFP or mCherry and imaged 2 dpi in *N. benthamiana*. Tetraspanins and C2-lipid binding proteins had a punctate localisation. GDSL esterase/lipases were observed in the endoplasmic reticulum. RNA-binding glycine-rich proteins were enriched in the nucleolus and had a nuclear-cytoplasmic localisation. Scale bar = 25 μm.
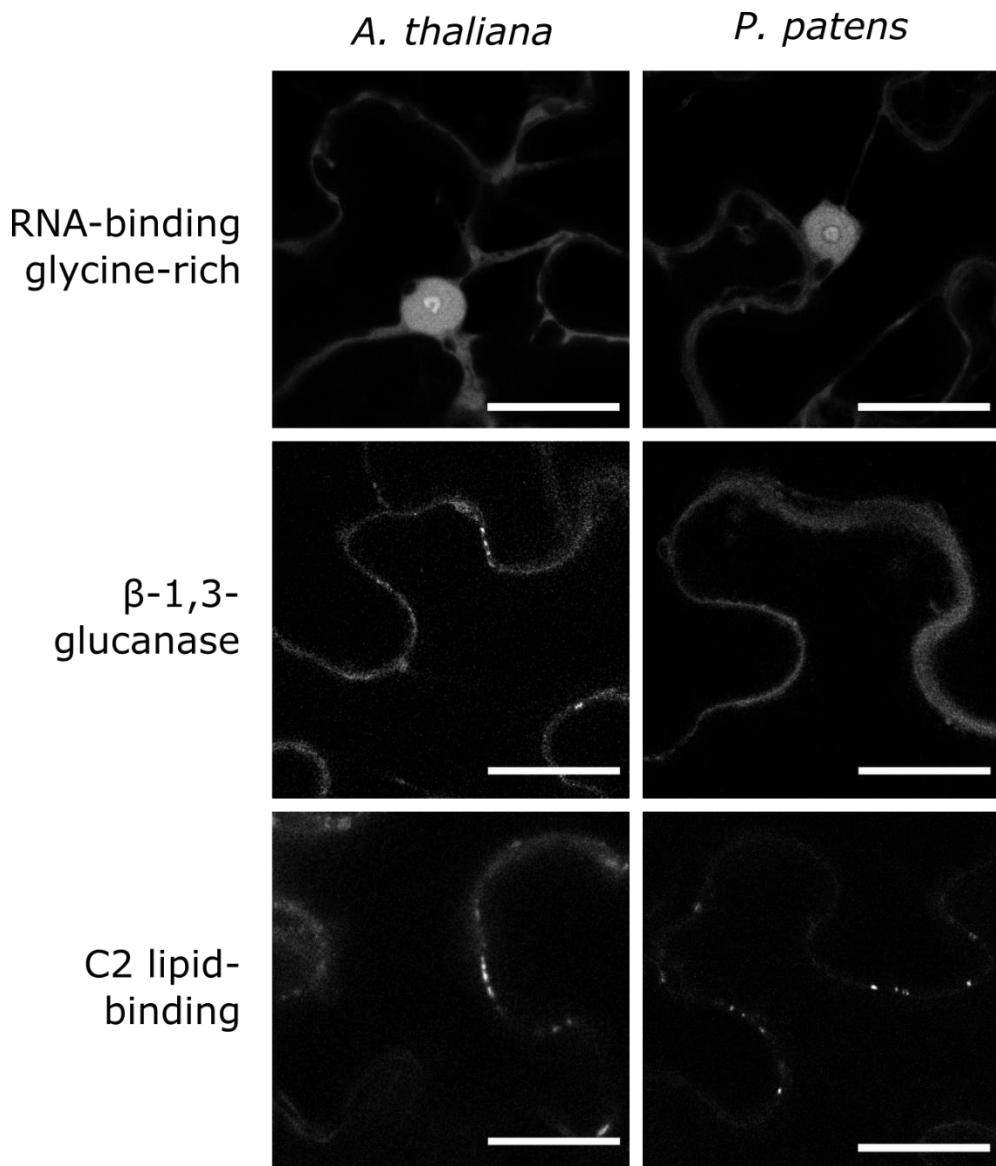
Figure 3-22 Localisation of conserved plasmodesmal proteins when N-tagged with fluorescent proteins and transiently expressed in *N. benthamiana*.

A. thaliana and *P. patens* homologues of phylogenetically conserved plasmodesmal proteins (Table 3-10) were tagged at the N-terminus with eGFP or mCherry and imaged 2 dpi in *N. benthamiana*. C2-lipid binding proteins had a punctate localisation. RNA-binding glycine-rich proteins were enriched in the nucleolus and had a nuclear-cytoplasmic localisation. The β-1,3-glucanase homologues had differing localisations. BG_PPAP (*A. thaliana*) was punctate, whereas Pp3c10_5480 (*P. patens*) was in the cell wall. Scale bar = 25 μm.

3.3.7 Expression of conserved plasmodesmal proteins gives a range of localisations when transiently, heterologously expressed in Nicotiana benthamiana

I observed a stronger and clearer localisation of C2 lipid-binding proteins at plasmodesmata when C-tagged rather than N-tagged (Figure 3-23). However, both constructs were observed at plasmodesmata. This is in agreement with the plasmodesmal localisation of C-tagged MCTP1/ FT INTERACTING PROTEIN (Liu *et al.*, 2012), but contrary to the localisation of MCTP15/ QUIRKY. MCTP15 was only observed at plasmodesmata when N-tagged, and was plasma membrane localised when C-tagged (Trehin *et al.*, 2013; Vaddepalli *et al.*, 2014).
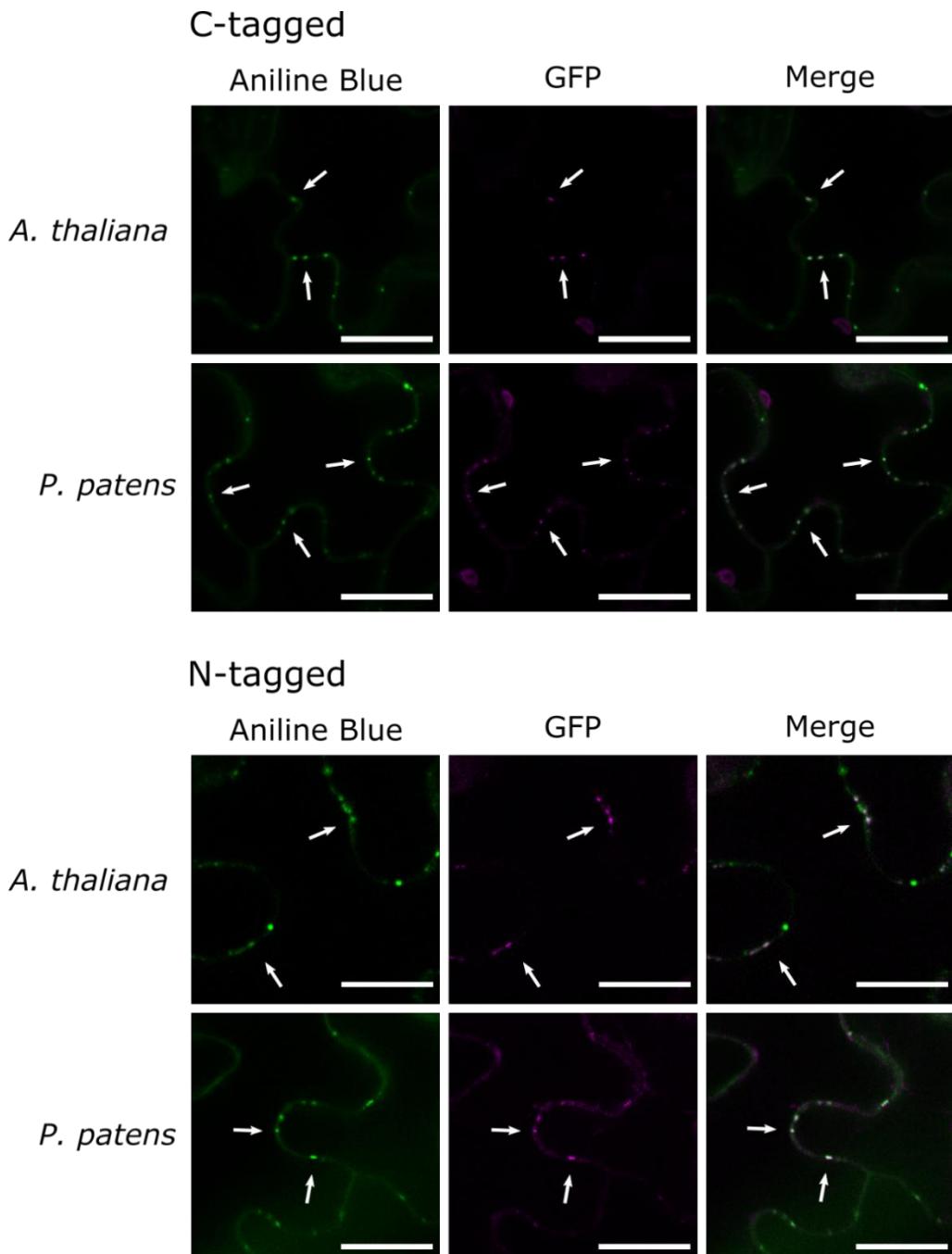
Figure 3-23 C2 lipid-binding proteins co-localise with callose at plasmodesmata
Both At1g51570 and Pp3c27_520 localised to punctae. These punctae co-localised with callose stained by aniline blue (examples shown by arrows), indicating a plasmodesmata localisation. The same pattern was observed whether the protein was tagged at the N- or C-terminus with a fluorescent protein. Scale bar = 25 μm.

3.3.7 Expression of conserved plasmodesmal proteins gives a range of localisations when transiently, heterologously expressed in Nicotiana benthamiana

Both tetraspanin proteins also co-localised with callose stained by aniline blue (Figure 3-24). This localisation was more consistent and apparent for TET8 (*A. thaliana* homologue), standing out more strongly from a secondary plasma membrane localisation. Pp3c7_23740 (*P. patens* homologue) also co-localised with callose, however this was less frequent than a smooth plasma membrane localisation. There was no apparent cause in the change of localisation, with both being observed within in the same leaf. It could be due to the nature of transient expression and varying levels of expression or localised stress among cells altering the cellular state and subsequent protein localisation.
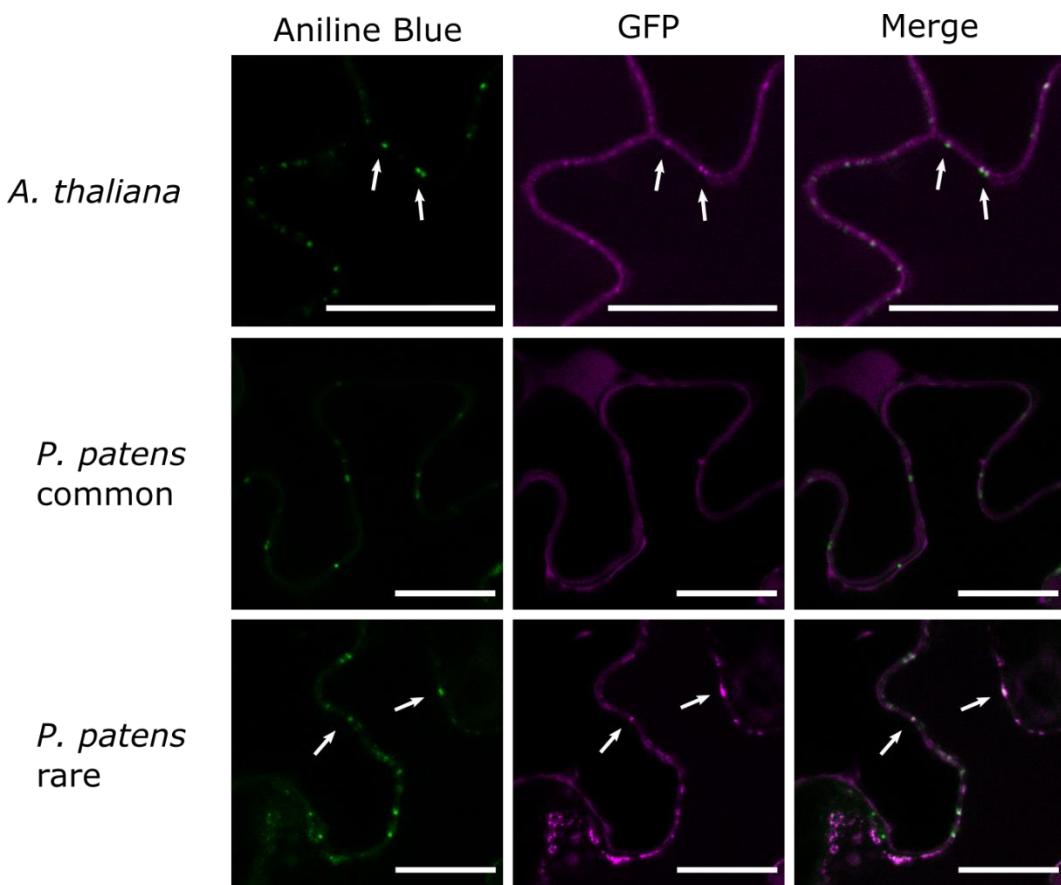
Figure 3-24 Tetraspanin proteins co-localise with callose at plasmodesmata
Both TET8 (At2g23810) and Pp3c7_23740 localise to plasmodesmata, as defined by aniline blue staining of callose (examples shown by arrows). The *P. patens* punctate localisation is more uncommon than a smooth plasma membrane localisation. Scale bar = 25 μm. The *A. thaliana* (TET8) micrograph was taken by S. Samwald.

Overall of the five orthogroups examined, four had homologues with similar localisations when transiently expressed in *N. benthamiana* (Figure 3-21, Figure 3-22). Only the β-1,3-glucanases, had differing localisations with the *A. thaliana* homologue localising to plasmodesmata and the *P. patens* homologue localising to the cell wall. This suggests that both are being passed through the secretory system, but the *A. thaliana* protein is tethered to the plasma membrane by a GPI-anchor, whereas the *P. patens* homologue is not. Of the

remaining four orthogroups, two have a punctate, plasmodesmal localisation: the tetraspanins and C2 lipid-binding proteins. Both classes of proteins have been localised to plasmodesmata before (Fernandez-Calvino *et al.*, 2011; Brault *et al.*, 2019). The remaining two classes of protein, RNA-binding glycine-rich and GDSL esterase/lipases had a nuclear/cytoplasmic localisation and endoplasmic reticulum, respectively. These localisations may be confounded by transient expression, and so were followed up by stable expression in *A. thaliana*.

3.3.8 Localisation of conserved plasmodesmal proteins homologues stably expressed in *Arabidopsis thaliana*

I wanted to confirm the localisations observed transiently in *N. benthamiana* were accurate, and so attempted to make stable lines of the same constructs in *A. thaliana*. However, due to time limitations, I only managed to create a subset of plants with constructs stably expressed in the wild type background (Col-0), namely: *P. patens* C2 lipid-binding protein (Pp3c27_520), *P. patens* tetraspanin (Pp3c7_23740), and both *A. thaliana* and *P. patens* RNA-binding glycine-rich proteins (At2g21660 and Pp3c11_19620, respectively).

The *A. thaliana* homologue of the C2 lipid-binding protein orthogroup, MCTP4, has been previously shown to be localised at plasmodesmata (Brault *et al.*, 2019). I find that the closest *P. patens* homologue also localises to punctae, likely to be plasmodesmata, in *A. thaliana* (Figure 3-25). This is surprising, as no C2 lipid-binding proteins were identified in PpPlant1, despite them being the most abundant plasmodesmal proteins found in AtCells2 (Brault *et al.*, 2019). Nonetheless, it is clear that Pp3c27_520 is recruited to plasmodesmata in *A. thaliana* and *N. benthamiana.*
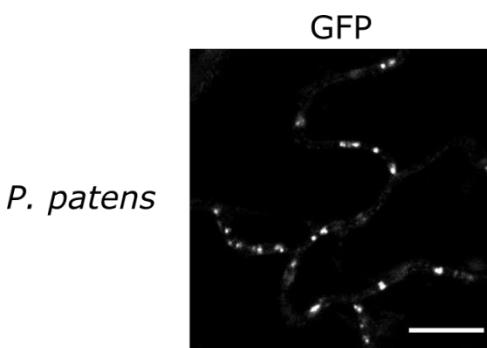
Figure 3-25 *P. patens* C2 lipid-binding protein Pp3c27_520 localises to punctae in *A. thaliana*
C-tagged Pp3c27_520 with GFP localises to distinct punctae visually similar to plasmodesmata, when stably expressed in *A. thaliana*. This image is from a T1 plant (n = 1). Scale bar = 10 μm.

3.3.8 Localisation of conserved plasmodesmal proteins homologues stably expressed in Arabidopsis thaliana

The selected tetraspanin *A. thaliana* homologue TET8 (Table 3-10) was not stably transformed into *A. thaliana*. However, *A. thaliana* TET3 has been localised to plasmodesmata previously (Fernandez-Calvino *et al.*, 2011). Conversely, the tetraspanin family has been largely observed at the plasma membrane, although the micrographs shown were not at a high magnification (Boavida *et al.*, 2013).

I also localised *P. patens* tetraspanin Pp3c7_23740 to the plasma membrane, but with clear distinct punctae as well, likely to be plasmodesmata. In addition, Pp3c7_23740 appeared to localise to the nuclear envelope (Figure 3-26). Unlike when Pp3c7_23740 was transiently expressed in *N. benthamiana* (Figure 3-24), punctae were ubiquitous and common in stable lines.
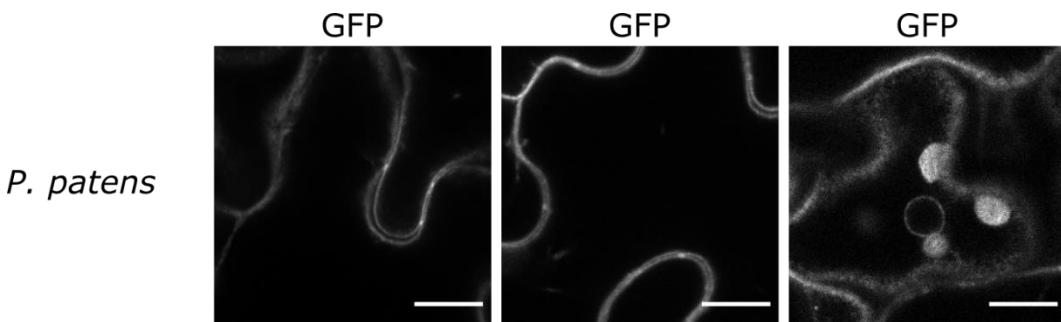
Figure 3-26 *P. patens* tetraspanin Pp3c7_23740 localises to punctae in *A. thaliana*
(Left and centre) C-tagged Pp3c7_23740 with GFP localises to distinct punctae visually similar to plasmodesmata, when stably expressed in *A. thaliana*. In addition, Pp3c7_23740 localises less strongly to the plasma membrane and nuclear envelope (right). These images are representative of T1 plants (*n* = 10). Scale bar = 10 µm.

Both the punctae observed from Pp3c27_520 and Pp3c7_23740 are highly likely to be plasmodesmata, given the punctae overlap with callose staining in *N. benthamiana*. Nonetheless, it would be beneficial to follow this work up with aniline blue staining and plasmodesmata-marker colocalization (e.g. PDLP1) to re-confirm these proteins are localised at plasmodesmata.

GRP7 (At2g21660), an *A. thaliana* RNA-binding glycine-rich protein, has been shown to be essential for a considerable number of physiological processes from drought tolerance to floral transition (Streitner *et al.*, 2008; Yang *et al.*, 2014). Further, GRP7 has been shown to be important for plant defence, including against *tobacco mosaic virus* which is known to use plasmodesmata to move cell to cell (Lee *et al.*, 2012). Multiple effectors target GRP7, with a similar end result: to modulate the RNA-binding capacity of GRP7, and so reduce plant immunity (Fu *et al.*, 2007; Nicaise *et al.*, 2013).

Both GRP7 and GRP2 (Pp3c11_19620) localised to the nucleus and cytoplasm, independent of the terminus of the protein that was tagged. There was a distinct and unambiguous gap

between each cell, indicating the cell wall between the cytoplasm. I did not observe any signal through the cell wall, which would indicate plasmodesmata or cell-to-cell transport of the protein. Thus, the localisation is distinctly not plasmodesmal. This is at odds with the relatively high Bayer Plasmodesmal Enrichment factor of GRP7: 8.4.
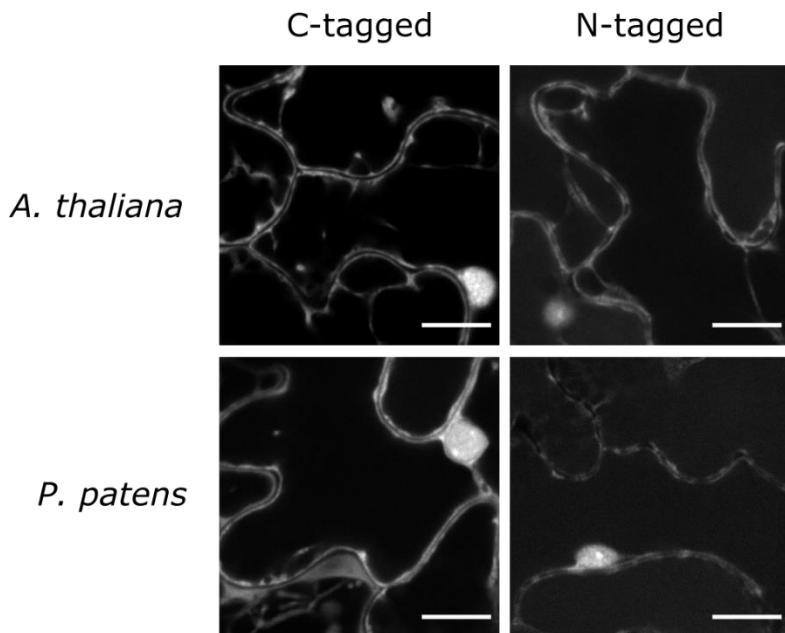
Figure 3-27 RNA-binding glycine-rich homologues do not localise to plasmodesmata in *A. thaliana*
At2g21660 and Pp3c11_19620 localise to the nucleus and cytoplasm when stably expressed in *A. thaliana* under a 35s promoter, irrespective of which terminus of the protein was tagged. These images are from T1 plants (n = 1, each). Scale bar = 10 μm.

Overall, two *P. patens* proteins have been localised to plasmodesmata when expressed transiently in *N. benthamiana* and stably in *A. thaliana*. Both protein families had been observed at plasmodesmata in *A. thaliana* endogenously, yet this work demonstrates that *P. patens* proteins, which are at least 460 million years diverged, have a conserved localisation at plasmodesmata. This underscores their probable importance within plasmodesmal functioning. Additionally, RNA-binding glycine-rich proteins which are common in plasmodesmata proteomes were found not be localised at plasmodesmata: opening the research question as to how this discrepancy could arise.

3.4 Discussion

3.4.1 *Arabidopsis thaliana* plasmodesmata proteomes contain a conserved subset

3.4.1.1 Presence denotes abundance

Mass spectrometry is an inherently noisy technique (Cargile *et al.*, 2004), with noise derived from both sample preparation and the equipment, akin to shotgun DNA sequencing. When purifying a recalcitrant, nanoscopic structure, such as a plasmodesmata, the mantra that "every protocol step loses material" (Faulkner, pers. comms.) rings true. Moreover, variability is introduced at every step: are there more chloroplasts in sample X than Y or was the cellulase digestion less complete in sample Z? This is coupled with the shotgun nature of peptide sequencing, where only a subset of peptides input are measured by the spectrometer. The variability in itself is not catastrophic: it can be overcome by increased sample replicates and lane run time. However, it means that comparing between two identical preparations, even from within the same laboratory, can show reasonable variation. The caveat being that the most abundant proteins in the preparations will be the most consistently identified, and so some qualitative metric of abundance can be assigned to the repeated presence of a protein. This same idea, that repeated presence denotes abundance, holds true between experiments and research groups, as well as between replicates.

3.4.1.2 Variability can be used to estimate the number of plasmodesmal proteins

Due to the inherently variability in mass spectrometry, each defined proteome will be a snapshot of a subset of plasmodesmal proteins and contaminants from sample preparation. This can be viewed akin the capturing of animals from wild populations in ecology, sampling members from a population (3.3.1). In this light, AtCells1 and AtCells2 can be used in a mark-release-recapture experiment, where in the first sample of the plasmodesmal population of proteins (AtCells1) the proteins were marked (by identification). The population was then resampled (AtCells2) (Equation 3-1). However, as the most abundant plasmodesmal proteins should be present in every proteome the proportion of marked proteins found in AtCells2 may be inflated: deflating the total population estimate. On the other hand, some members of AtCells1 and AtCells2 are likely to be contaminants rather than true plasmodesmal proteins, inflating the total population estimate. With these two opposing effects, we can make a rough estimate of the number of plasmodesmal proteins in *A. thaliana* to be about 2,000.

Surprisingly, this estimate is larger than all proteins in the three *A. thaliana* proteomes combined (AtCells1, AtCells2, and AtPlant1), which have 1,576 unique members (Figure 3-6).

It will be interesting to follow this in future work and revise the estimate of plasmodesmal proteins. Currently, it is impossible to tell whether the estimate is over-inflated by contaminants, or there are many more low abundance plasmodesmata proteins still to be found.

3.4.1.3 Variability can be used to define 'proteome-recurrent proteins

In addition to combining proteomes to estimate the number of proteins in a subcellular fraction, the "recaptured" proteins can be informative in themselves. Multiple proteomes can be overlapped to define a proteome-recurrent set of proteins. These proteins will be the most abundant in the sample, as argued in 3.4.1.1. Moreover, this set can filter out contaminants, if the plasmodesmal proteins remain constant and contaminants vary between samples. This can be achieved by having various different methods to extract and identify plasmodesmal proteins, whether that be alternative sample preparations technique, or alternate techniques to generate protein lists (e.g. high-throughput protein localisation, or plasmodesmal protein co-immunoprecipitation (Ham *et al.*, 2012))

Here, I took an alternate approach, whereby I used novel sample preparation to generate a proteome from a different tissue type. I generated a plasmodesmal proteome from *A. thaliana* var Columbia rosette leaf tissue from plants grown on soil, which is substantially different from cell suspension culture tissue. Cell suspension cultures can be made from leaf or root tissue, in this case the proteome was made from *A. thaliana* var *Landsberg erecta* stem explants, from which green photosynthetic cells were made (May & Leaver, 1993). These cells are large, easy to lyse, and contain largely simple plasmodesmata (Bayer *et al.*, 2004) and grow in liquid culture. In this manner, there should be differing contaminants between AtPlant1 and AtCells1/2 proteomes, as they are derived from different protein backgrounds due to the differing tissue types and growth environments, whilst still containing plasmodesmal proteins. Naturally though, the plasmodesmal components between the two proteomes also may vary, especially any proteins exclusive to simple or complex plasmodesmata. Nonetheless, core plasmodesmal proteins required in both simple and complex plasmodesmata will be conserved. This technique was used to produce a list of 64 proteome-recurrent proteins (Table 3-2).

This technique can be combined with the stringent contaminant filtering in AtCells2_filtered to give three overlapping proteins (BG_PPAP, MCTP4 and At4g34150, Figure 3-6). BG_PPAP and MCTP4 have already been shown to be at plasmodesmata (Levy et al., 2007; Brault et al., 2019, Figure 2-20). At4g34150 has yet to be studied, but should be a high priority for

future investigation. At4g34150 contains two annotated protein domains (InterPro): a C2 lipid binding domain and a disordered SoxC DNA-binding domain. The C2 domain has been shown previously to localise MCTP proteins to the plasma membrane (Liu *et al.*, 2018; Brault *et al.*, 2019), while the DNA-binding domain would localise At4g34150 to the nucleus; two seemingly opposing localisations. A recent review put forward a "organelles-nucleus-plasmodesmata signalling (ONPS)" hypothesis (Azim & Burch-Smith, 2020), where organelles signal to the nucleus in a retrograde manner to then unidirectionally control plasmodesmata. As the C2 domain binding of the membrane is reversible, At4g34150 provides a possible retrograde signal from the plasmodesmata to the nucleus: extending the ONPS signalling pathway to include plasmodesmata to nuclear feedback.

3.4.2   Finding novel plasmodesmal proteins with AtPlant1 and PpPlant1

I argued above that AtPlant1 is substantially different from AtCells1 and AtCells2 proteomes, and thereby the contaminants in the two differed. By the same token, it is likely that a different subset of plasmodesmal proteins are captured. This goes beyond the noise of mass spectrometry picking up different protein subsets each experiment, and rather stems from mature tissue containing complex plasmodesmata (Roberts *et al.*, 2001). The fact that TOBACCO MOSAIC VIRUS MOVEMENT PROTEIN (TMV MP) only targets complex plasmodesmata, suggests there is likely a different protein composition in comparison to simple plasmodesmata (Ding *et al.*, 1992a). This was further demonstrated by the requirement for CHER1 for plasmodesmal maturation (Kraner *et al.*, 2017a). Thus, this fraction is expected to contain different proteins to fractions containing mostly simple plasmodesmata, such as cell suspension cultures (Bayer *et al.*, 2004). Moreover, a different subset of proteins may be enriched in a mature plant subcellular environment.

114 proteins within AtPlant1 did not overlap with either AtCells1 or AtCells2 (Figure 3-6). This list will contain both contaminants that were selected against by comparison, in addition to novel plasmodesmata proteins. As proof of concept, the list contains two proteins with a GO plasmodesmata representation ($p > 0.05$ for over representation): GERMIN-LIKE PROTEIN SUBFAMILY 2 MEMBER(GLP)1 and GLP2 (AT1G09560 and AT1G02335). However, the needle is in the haystack for the remaining 107, while there may be additional plasmodesmal proteins there is no way to identify them (see 3.4.3.2 for further discussion) without further experimentation. To look further into this, I would filter the list for known contaminants (e.g. remove the 32 GO 'organelle' proteins, $p < 0.05$), and then further search for proteins with transmembrane domains or that are targeted to the secretary pathway (not dissimilar to Caillaud et al., 2014).

The exact same argument applies to the unexamined proteins within PpPlant1. There may be novel Bryophyte-specific plasmodesmal proteins, which are being left undocumented. However, without additional data or a targeted biological question, there is no specific proteins to target. Instead, here, I focussed on interrogating the proteome to find conserved, evolutionary-ancient, plasmodesmal proteins.

3.4.3   The drawbacks of plasmodesma gene ontology (GO:0009506)

Gene ontology (literally 'the study of being a gene') is a laudable effort to annotate every gene in a standardised and comprehensive way from their molecular function to cellular localisation. Currently [15/10/2020], there are 1,048 *A. thaliana* genes labelled with a plasmodesma localisation (GO:0009506). A vast majority of these come from their identification in AtCells1 (819/1048, 78%) (Fernandez-Calvino *et al.*, 2011). Through phylogenetic backpropagation, 84 proteins are labelled as at plasmodesma in *P. patens* (Gaudet *et al.*, 2011). GO annotations allow for a quick, and in my opinion, dirty check of likely gene properties for a list. I used GO analysis to compare different proteome lists, largely for two reasons: to compare contaminant levels, and to verify plasmodesmal enrichment. For the former, I believe it works very well. However, I think using GO to verify plasmodesmal enrichment may be problematic, as outlined in 3.4.3.1.

3.4.3.1   Confounding localisation with extraction method

The GO:plasmodesma term in *A. thaliana* is largely defined by AtCells1, and the second largest annotation after this is phylogenetic propagation (152/1048, 15%), likely also homologues of AtCells1 proteins. Overall, then, ~93% of the ontology is defined by a single experiment. Thus, the enrichment for plasmodesma genes in AtCells1 is approximated to $p = 0$ (Figure 3-5). AtCells1 was shown to contain a significant proportion of plasmodesmal proteins, with 39 previously validated plasmodesmal proteins identified (Fernandez-Calvino *et al.*, 2011). In addition, the authors experimentally validated the proteome, with a selection of membrane proteins (21% of the proteome) of which 33% (5/15) localised to the plasmodesmata (Fernandez-Calvino *et al.*, 2011). However, AtCells1 is an unfiltered proteome and is likely to contain contaminants (the authors estimate 35%), as well as extraction-specific artefacts (e.g. 12% are putative cell wall proteins). This raises the question whether subsequent extractions, such as AtCells2, which follow an identical protocol are really enriched in plasmodesmal proteins or rather plasmodesmata *extraction* proteins: the logic becomes circular.

3.4.3.2 Inability to make inferences on new genes

If ontology is defined by a single study, it raises a more pressing issue: how to find novel plasmodesmal proteins (3.4.2). Novel proteins would be expected to be present in AtPlant1 but not in AtCells1 (AtPlant1 complement AtCells1). However, by definition, none of these are identified with the plasmodesma GO ontology (3.4.3.1). Therefore, while plasmodesmata proteins may be present in this list, they will not be listed as such by GO. This comes to the heart of how ontologies are defined: by experimental evidence and phylogenetic relationships. As there are no other reliable plasmodesmata predictors, novel proteins simply cannot currently be bioinformatically predicted as to whether they are plasmodesmal. New bioinformatic tools are currently being developed to give *de novo* prediction into plasmodesmal localisation motifs (Li *et al.*, 2020). Otherwise, protein localisation can be tested experimentally, as in this thesis. Two tools are making this more feasible on a larger scale: the rapidly decreasing cost of DNA synthesis, and modular cloning tools (e.g. GoldenGate). In theory, hundreds of constructs could be tested transiently within a few months, especially if aided by robotic cloning (Ortiz *et al.*, 2017). Further, while plasmodesmal proteins cannot be predicted accurately, some traits make more hopeful targets, such as the presence of transmembrane domains or lack of chloroplast transit peptides (Fernandez-Calvino *et al.*, 2011). However, there will always be proteins which 'break the rule'.

3.4.3.3 Poor ontology coverage of proteins from evolutionary distant species

A final consideration when using protein ontology tools to check for plasmodesmal enrichment is that *p* values are inherently influenced by *n* values. In other words, the *p* value is informing the reader to the evidence against the null hypothesis, and with more data (*n*) the stronger that rejection can become. Thus, the relatively large *p* value for PpPlant1 (3.19×10^{-5}, 3.3.4) can be explained by the small size of the *P. patens* plasmodesma ontology. By the same token, it explains why the *p* value drops dramatically when PpPlant1 proteins are translated into their *A. thaliana* homologues (Table 3-4). It is probably impossible to determine an unambiguous phylogenetic relationship (especially if a one-to-one relationship is required) between the majority of the 1,048 *A. thaliana* plasmodesmal proteins and their distantly related *P. patens* homologues, and so the GO term cannot be mapped across. Therefore, either a lower confidence in GO annotation must be accepted (attainable by mapping to *A. thaliana* genes at higher E thresholds) or novel experiments must be done on *P. patens*.

3.4.4   Finding homologues from *Physcomitrella patens*

Evolutionary comparison of plasmodesmal proteomes offers the possibility of defining conserved components of plasmodesmata. In the context of the limitations of GO term analysis outlined above, I have deployed more exhaustive strategies to find the homologues of *P. patens* proteins in *A. thaliana* across some 500 million years of evolution. This has enabled me to better characterise the constituents of the *P. patens* plasmodesmal proteome (3.4.3.3). Further, it allows for direct comparison between the *P. patens* proteome and the *A. thaliana* proteomes: allowing the identification of conserved proteins.

3.4.4.1   Why are homologue databases so poor?

The first approach I tried was to use pre-existing homology databases (InParanoid and MetaPhOrs). Each database individually identified ~25% of the proteome, and only ~10% of the proteome was given the same homologue by both databases. Clearly, this approach results in low coverage and reliability. The large amount of time that has passed since their shared ancestor has led to a significant amount of divergence. Divergence itself is not problematic, advanced probabilistic algorithms can find likely ancestors. However, inparalogues (Figure 3-28) create a philosophical dilemma: which is more ancient? It is non-sensical to ascribe which inparalogue is more ancient! The ancestor is most ancient and there are two subsequent daughters. Thus, if *A. thaliana* β was found to be a plasmodesmal protein, it is impossible to determine which of *P. patens* β1 and β2 should also be noted as plasmodesmal. Instead, therefore, I chose to utilise a orthogroup approach to find candidate *P. patens* plasmodesmal proteins.
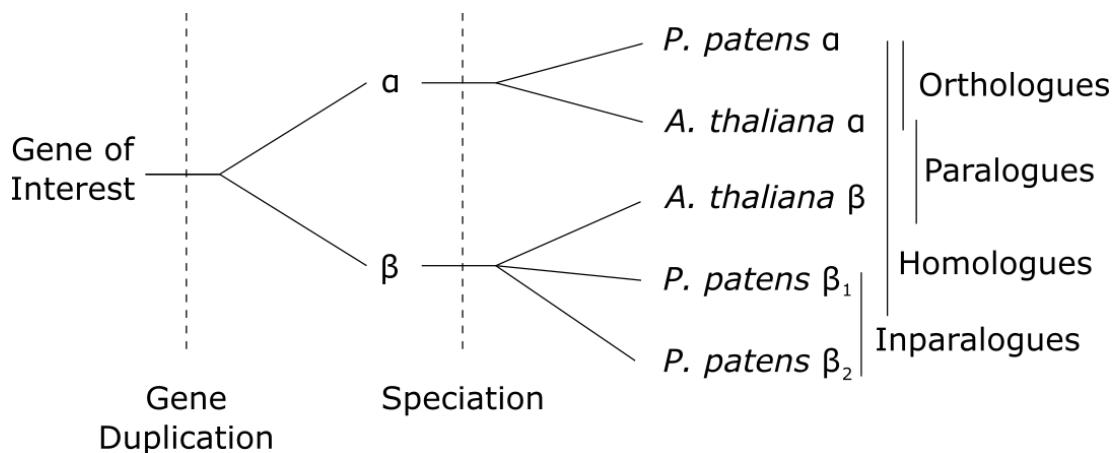
Figure 3-28 Defining homologue terminology
Homologues are all proteins that derive from the same ancestral protein. Paralogues are proteins within a species that have a risen from a gene duplication. An inparalogue is a relative term and is a subtype of paralogue, where the gene duplication event is after a specified speciation event. An orthologues is a protein derived from the same ancestral protein in a different species (Sonnhammer & Koonin, 2002).

3.4.4.2   Orthogroups or Homogroups?

Orthogroups allow the grouping of purported orthologues from across species. Strictly, I do not think Orthofinder achieves this. A strict orthogroup would not contain paralogues by definition. For example, homologues α and β should not be mixed within an orthogroup (Figure 3-28). The package author suggests all same species orthogroup members are paralogues (Emms, D., Orthofinder GitHub Issue 211). This is logically congruent, until there are multiple species with paralogues within a single orthogroup. At this point, orthologues and homologues can no longer be distinguished (e.g. OG000, Table 3-8). Advantageously, this gathers homologous proteins together, irrespective of orthology, as function may be better assessed on this wider level. However, the software is fickle and not all homologues are grouped into an orthogroup, as the programme works on pair-wise BLAST similarity. For example I identified that there are at least four orthogroups (OG52, OG55, OG87, OG95; analysis as in Table 3-8) containing callose synthases. Thus, some homologous conserved groups may be missed because they were not grouped together.

3.4.5   Scarcity in HMMER homologues

3.4.5.1   Similar PpPlant1 recovery by the three A. thaliana *proteomes*

All three unfiltered *A. thaliana* proteomes recovered a similar proportion of the *P. patens* proteome using a HMMER search at (~65%, Table 3-6). On the face of it, this is unsurprising, given they are all *A. thaliana* plasmodesmata proteomes, and may suggest a reliable conservation of plasmodesmal proteins of about 65%. However, the number of proteins compared varies wildly from 238 to 1,000 due to the size of the *A. thaliana* proteomes (AtPlant1 and AtCells1, Table 3-1). Naively, you would expect greater coverage when trialling against more proteins: this is not the case. This could be explained in two ways. First, 238 proteins is sufficient to obtain 65% similarity to PpPlant1. The additional proteins in AtCells1 and AtCells2 may be extraneous. This, however, cannot be true as the conserved intersection between these proteomes only recover 34% of PpPlant1. Alternatively, AtPlant1 may be more similar to PpPlant1, and so fewer proteins are needed for the high coverage. The latter hypothesis is supported by the highest percentage recovery of *A. thaliana* proteins by PpPlant1 of AtPlant1 than the AtCells proteomes (Table 3-5).

3.4.5.2   Few AtCells proteins are found in PpPlant1

In general, the AtCells proteomes poorly translated into *P. patens* homologues (~10%, Table 3-5) suggesting a high dissimilarity between the AtCells proteomes and PpPlant1. This could be derived from many *P. patens* only proteins, or due to the relatively small size of PpPlant1 compared to the *A. thaliana* proteomes. It is hard to distinguish between the two. In fact, it

is likely a combination of small size coupled with a high degree of *P. patens* unique proteins (30%, 51/170 orthogroups).

More interestingly, there was a much greater overlap between AtPlant1 and PpPlant1, than the other proteomes (34%). This could plausibly be because of similar contaminants from mature tissue, the same laboratory conducting the experiment and subsequent mass spectrometry, or the similar preparation method (i.e. mature plasmodesmata).

3.4.6   A Bayesian approach to protein identification

3.4.6.1   *Proteome identification*

The results of a tandem mass spectrometry proteomic experiment are far more detailed than a simple presence/absence of each protein. The data includes the spectra for every identified peptide, with derived probabilities as to the likelihood the peptide was identified properly, and the likelihood of the correct identification of a protein from multiple peptides. The question as to how to assign a peptide to a protein is not trivial and well-debated (Rappsilber & Mann, 2002), and often done incorrectly (Bell *et al.*, 2009). Scaffold is a versatile tool for visualising MS/MS data, which includes easy-to-use false discovery rate (FDR) thresholds for both peptide and protein identification (Searle, 2010). I followed convention in the thesis as used in Leijon *et al.* (2018), with a 95% threshold for both. This is slightly lower than the Scaffold 99% defaults. In addition, to overcome the variability aforementioned (3.4.1), proteins had to be identified in at least two replicates. These thresholds provide suitable confidence in protein identification to create a novel plasmodesmal proteome (Leijon *et al.*, 2018).

3.4.6.2   *Coupling MS/MS with external knowledge*

The probabilistic thresholds in Scaffold are created in isolation from the rest of our knowledge, as is done in frequentist statistics. However, existing knowledge can be used to influence our understanding of probability (Bayes & Price, 1763). For example, we know that PDLP1 is present in plasmodesmata and AHA2 is not from electron gold labelling experiments (Fleurat-Lessard *et al.*, 1995; Grison *et al.*, 2015a). If Scaffold identified both proteins in a sample with 50% probability, we can couple this with our domain knowledge to suggest that the likelihood of AHA2 being identified is actually <50%, while the likelihood of PDLP1 being present >50%. This logic was formalised in Bayes' theorem, which I have not mathematically applied here as the prior likelihoods cannot be calculated. However, the idea can be used. Proteins and peptides with a lower identification probability can be considered more likely to be reliably identified, when they were also identified with high likelihood in other

proteomes. Thus, I generated 'loose' proteomes with a lower threshold of 50% FDR for both peptide and protein identification, and only present in one proteome and compared this list to the existing literature using OrthoFinder. This highlighted protein groups which are likely to be conserved, but are identified with less confidence (Figure 3-10, Table 3-9). The merit of this approach is underscored by the inclusion of callose synthases (OG004) in Table 3-9, but not the more stringent Table 3-8, as they are well-documented to be essential for plasmodesmal function (Guseman *et al.*, 2010; Saatian *et al.*, 2018).

3.4.7   Phylogenetically conserved proteins

3.4.7.1   Phylogenetic-based approach to choosing localisation candidates

I have defined a list of highly phylogenetically-conserved protein families at plasmodesmata, by using orthogroups to combine the results of five proteomes from three species and four laboratories (Table 3-8). I wanted to validate this list and confirm it contained *bona fide* plasmodesmal protein families. Further, I tested the possibility that proteins within this list are so central to plasmodesmata, they would localise to plasmodesmata when heterologously expressed, be that in *P. patens* proteins stably in *A. thaliana* or transiently in *N. benthamiana*.

I chose a subset of these families to take forward for further analysis (43%, 9/19). I chose families which were not known to be essential at other localisations (e.g. ribosomal proteins) and were not part of large multi-gene families where conserved orthology could be due to chance (e.g. ATP-binding cassette). For these nine selected orthogroups, I conducted a wider phylogenetical analysis. I searched for families where there was a clear clade of plasmodesmal localised proteins (44%, 4/9) or where plasmodesmal localisation was ubiquitous (11%, 1/9). I chose similar candidates from *P. patens* and *A. thaliana* from each of these clades, where the species clustered together (60%, 3/5). In addition, I took forward a fourth group, C2 lipid-binding proteins, as they have been recently described as essential for plasmodesmal function (Brault *et al.*, 2019). Lastly, I selected a fifth group to clone (26%, 5/19) as a positive control: β-1,3-glucanases.

3.4.7.2   Endogenous localisations of A. thaliana homologues

Of the five families of proteins I chose to localise, I only managed to observe the endogenous localisation of GRP2 (RNA-binding glycine-rich protein) (Figure 3-27). GRP2 did not localise to plasmodesmata in *A. thaliana*, rather to the nucleolus, nucleus and cytoplasm. Of the remaining proteins MCTP4 (C2 lipid binding) has been localised to plasmodesmata in *A. thaliana* independently by two research groups (Liu *et al.*, 2018; Brault *et al.*, 2019).

AtBG_PPAP has also been localised to plasmodesmata in *A. thaliana* (Papp, D., *pers. comms.*). The final two proteins, TET8 (tetraspanins), GELP91 (GDSL esterase/lipase) have not been accurately subcellularly localised to my knowledge. TET8 has been observed at low magnification at the plasma membrane (Boavida *et al.*, 2013). However, other tetraspanins, TET3 and TET5, have been localised to plasmodesmata (Fernandez-Calvino *et al.*, 2011; Boavida *et al.*, 2013). Many GDSL esterase/lipases have been localised to a variety of subcellular localisations (e.g. cell wall, nucleus, plasma membrane, and endoplasmic reticulum), but not yet to plasmodesmata (Ding *et al.*, 2019).

Overall, therefore, of the five proteins two have not been localised yet and, of the remaining three, two are known to be plasmodesmal proteins (66%, 2/3). Thus, there is strong evidence that the phylogenetically-conserved orthogroups contain plasmodesmal proteins. However, this conclusion comes with the caveat that the orthogroups tested, were ones with unknown or plasmodesmal localisations, and so this figure may be an overestimate. On the other hand, some proteins observed at organelles, e.g. VDAC2 and VDAC3 at mitochondria (Lee *et al.*, 2009), have also been localised to plasmodesmata as well (Bellandi A., *pers. comms.*).

3.4.7.3 Exogenous localisations of phylogenetically conserved proteins

The phylogenetically-conserved nature of the proteins I localised to *A. thaliana* and their *P. patens* homologues indicates an essential and central role in plasmodesmata: whether in formation, maintenance or function. To test whether these proteins were conserved, I expressed them exogenously: in an environment they would not have been found in. I transiently expressed both homologues in *N. benthamiana* and the *P. patens* proteins stably in *A. thaliana* (Table 3-11).

Table 3-11 Summary of phylogenetically-conserved orthogroup protein localisations

Orthogroup	Species	Identifier	Name	Transient *N. benthamiana* localisation	Stable *A. thaliana* localisation
OG005 GDSL esterase/ lipase	*A. thaliana*	At5g14450		Endoplasmic reticulum	NA
	P. patens	Pp3c18_ 1550		Endoplasmic reticulum	NA
OG006 Tetraspanin	*A. thaliana*	At2g23810	TET8	PD	NA
	P. patens	Pp3c7_ 23740		PM/PD	PD
OG016 RNA-binding glycine-rich	*A. thaliana*	At2g21660	GRP7	Nucleus/ cytosol	Nucleus/ cytosol
	P. patens	Pp3c11_ 19620	GRP2	Nucleus/ cytosol	Nucleus/ cytosol
OG000 β-1,3-glucanase	*A. thaliana*	At5g42100	BG_ PPAP	PD	NA
	P. patens	Pp3c10_ 5480		Cell wall	NA
OG003 C2 lipid-binding	*A. thaliana*	At1g51570	MCTP4	PD	NA
	P. patens	Pp3c27_ 520		PD	PD

First, it is striking how most homologues have an identical localisation within the orthogroup (all bar β-1,3-glucanase). The mechanisms sorting protein localisations are clearly well conserved between to *A. thaliana, N. benthamiana* and *P. patens*.

BG_ PPAP had already been localised to plasmodesmata when transiently expressed in *N. benthamiana* (Levy *et al.*, 2007). Surprisingly, as this group was included as a positive control, Pp3c10_5480 did not localise to plasmodesmata. I postulated that this was due to the lack of a GPI anchor (Zavaliev *et al.*, 2016). However, it is clear that *P. patens* must contain a callose degrading protein, as callose is dynamically regulated at protonemal cross walls (Kitagawa *et al.*, 2019). Thus, instead of expressing the phylogenetically closest β-1,3-glucanase, in the future I would elect to localise a β-1,3-glucanase with a GPI-anchor and annotated at

plasmodesmata. For example, using HMMER, only two proteins found in PpPlant1 were conserved across all five proteomes: A0A2K1JXU2 and A0A2K1J8R8 (Figure 3-8). I cloned A0A2K1JXU2, due to its phylogenetic position, which does not have a GPI anchor (Figure 3-16). Alternatively, A0A2K1J8R8 is predicted to have a GPI anchor (99.5%, PredGPI).

Both the tetraspanins and C2 lipid-binding proteins localised to plasmodesmata when exogenously expressed in *N. benthamiana*. The *P. patens* homologues also localised to punctae in *A. thaliana*. This evidence suggests a central role for both proteins within plasmodesmata. It has been previously suggested that MCTPs are essential for desmotubule organisation (Brault *et al.*, 2019). Tetraspanins are also thought to be structural proteins, creating or maintaining nanodomains within membranes: tetraspanins webs (Zuidscherwoude *et al.*, 2015). It is possible that conservation of localisation could be because of endogenous mechanisms specific to Tracheophytes: i.e. C2 lipid-binding proteins are recruited to plasmodesmata by a Tracheophyte-specific protein. In this scenario, only the localisation motif is required to be conserved, even if it is not a localisation motif in *P. patens*. This is seen in C_4 photosynthesis, where ancestral C_3 genes have no spatial selectivity in the ancestor (*A. thaliana*), but are spatially segregated in a C_4 plant (*Cleome gynandra*) (Kajala *et al.*, 2012). In this case, a novel trans factor in *C. gynandra* has evolved to use pre-existing motifs. Thus, to confirm the central role of tetraspanins and C2 lipid-binding proteins throughout Viridiplantae, these homologues need to be expressed in *P. patens* as well to show their ancestral localisation. However, it would be surprising if a motif which had no function had been preserved over 450 million years of evolution.

In a similar vein, it would be a fascinating experiment to express known A. thaliana plasmodesmata proteins in *P. patens*. This could ask the question as to whether PDLP proteins localise to plasmodesmata within *P. patens*: is there an ancient plasmodesmata localisation signal that PDLPs have or are PDLPs targeted to plasmodesmata by a novel Tracheophyte-specific protein? In addition, it would be interesting to express BG_PPAP in *P. patens* as well. GPI anchors are known to have an ancient origin certainly within all of Eukaryota and potentially even within Archaea (Eisenhaber *et al.*, 2001), and so it may be an ancient plasmodesmal localisation motif too.

3.4.8   Conclusions

In this chapter, I have described two novel plasmodesmata proteomes derived from mature plant tissue of *A. thaliana* and *P. patens*. Both of these proteomes are enriched for plasmodesmata-annotated genes, and are similar in purity to AtCells1 and AtCells2. I could use the variability inherent within proteomic experiments to define a proteome-recurrent set of proteins from AtCells1 and AtCells2, which were more enriched at plasmodesmata than the proteomes from which they are derived. These proteomes were compared with AtPlant1, derived from a different tissue type, so that plasmodesmata proteins were retained but contaminants were not, to give a final proteome-recurrent list. I extended this technique to compare proteomes between species by generating *de novo* orthogroups. Again, through conservation, I defined a group of phylogenetically-conserved orthogroups, which are protein families that are postulated to be ancestral plasmodesmata proteins. These proteins were validated to be plasmodesmal, specifically the tetraspanins and C2 lipid-binding proteins, which are likely to have been present in plasmodesmata for over 450 million years. Interestingly, both these proteins are structural proteins and may be central, ancient components for holding plasmodesmata together.

4 A four-pronged approach to find PDLP interactors

4.1 Introduction

PLASMO**D**ESMATA-**L**OCALISED **P**ROTEINs (PDLPs) were some of the first plasmodesmata specific proteins to be found (Thomas *et al.*, 2008). The whole family of eight proteins localise to plasmodesmata in *Arabidopsis thaliana*, and have a canonical tandem DOMAIN OF UNKNOWN FUNCTION26 (DUF26) domain. In this chapter, I explore four approaches to elucidate PDLP partners to aid our understanding of how PDLPs function.

4.1.1 DUF26-containing proteins

A DUF26 domain, otherwise known as Ginkbilobin-2 (Gnk2) or stress-antifungal domain (PF01657), is a roughly 100 amino acid extracellular domain with a core conserved cystine motif (C-8X-C-2X-C), which encompasses three of five largely conserved cystines (Figure 4-1) (Chen, 2001; Schuster-Böckler *et al.*, 2004). DUF26-containing proteins can be broken into three major classes: CYSTEINE-RICH RECEPTOR-LIKE SECRETED PROTEINs (CRRSPs), CYSTEINE-RICH RECEPTOR-LIKE PROTEIN KINASES (CRKs), and PDLPs (Figure 4-2).
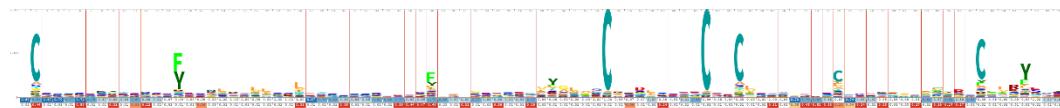
Figure 4-1 HMM logo for stress-antifungal domain PF01657.
THE DUF26 domain (PF01657) is 95 amino acids long and is characterised by the canonical C-8X-C-2X-C motif seen in the middle-right of the image.

4.1.1.1 *CYSTEINE-RICH RECEPTOR-LIKE SECRETED PROTEINs*

The smallest and most ancient group of DUF26 proteins are the CRRSPs, which are present in Bryophytes (Vaattovaara *et al.*, 2019). The best characterised member is Gnk2 from *Gingko biloba*, which has been reported to have anti-fungal activity and act as a mannose-binding lectin (Miyakawa *et al.*, 2014). Other CRRSPs have tandem DUF26 domains, which are the predominant type of CRRSPs in vascular plants (Vaattovaara *et al.*, 2019). Two of these from *Zea mays* have been shown to have the same activities as Gnk2 (Ma *et al.*, 2018). The ancient function of CRRSPs in plant defence appears to have been conserved throughout all DUF26 domain proteins. The short length of CRRSPs indicate that the DUF26 domain itself has an ancestral role in responding to pathogens.

4.1.1.2 *CYSTEINE-RICH RECEPTOR-LIKE SECRETED PROTEINs*

CRKs are thought to have evolved from an ancient CRRSP and a clade III LEUCINE-RICH REPEAT receptor (Zulawski *et al.*, 2014; Vaattovaara *et al.*, 2019), with PDLPs arising from a

subsequent loss of the kinase domain (Figure 4-2). CRKs are an extensive gene family with over 40 members in *A. thaliana* (Chen, 2001).

CRKs have been proposed to be part of the reactive oxygen species (ROS) sensing and signalling pathway (Idänheimo *et al.*, 2014). CRKs are transcriptionally responsive to abiotic stresses, such as ozone, high light and ABA (Wrzaczek *et al.*, 2010a; Tanaka *et al.*, 2012). These conditions are often accompanied by the production of ROS and consequent oxidative stress. CRKs have been shown to be transcriptionally responsive to oxidative stress directly and required for protection from cell death induced by extracellular ROS damage (Idänheimo *et al.*, 2014; Bourdais *et al.*, 2015). This implies that CRKs are required by plant cells to either sense or respond to ROS, and prime downstream responses.

CRKs are also an important part of plant defence, and are required to activate the hypersensitive response leading to rapid cell death. As a part of this, CRKs are also transcriptionally upregulated by salicylic acid (SA) and pathogen infection (Chen *et al.*, 2003, 2004). This may also be linked to ROS, as ROS is produced in response to pathogens by RESPIRATORY BURST OXIDASE HOMOLOG D (RBOHD) (Torres *et al.*, 2002). Exactly how CRKs are involved is poorly understood, but as "tuners" of plant responses to ROS they may be amplifying the plant defence ROS or/and responding to the ROS signal to trigger cell death.

Recent work has shown that CRK2 interacts directly with RESPIRATORY BURST OXIDASE HOMOLOG D (RBOHD), with CRK2-dependent phosphorylation on the C-terminus of RBOHD on S703. Mutating RBOHD S703A reduced ROS production, in a CRK2-dependent fashion (Kimura *et al.*, 2020). This implies that CRK2 lies upstream of ROS production, and so is triggering ROS as part of plant defence. In conjunction with general CRK transcriptional upregulation by ROS, this generates a feed-forward loop, whereby CRK2 activates ROS production and then is upregulated by it. The end point of this loop would be cell death in the hypersensitive response.

CRK2 relocalised to plasmodesmata under salt stress from the plasma membrane (Hunter *et al.*, 2019). This may indicate plasmodesmata-specific ROS production. This phenomena has been hinted at previously in the response of plasmodesmata to chitin (Faulkner *et al.*, 2013; Cheval *et al.*, 2020). RBOHD is required for plasmodesmata to respond to chitin, but abolishing the plasmodesmata-chitin response does not alter the global ROS response. Thus, there must be required plasmodesmata-specific ROS production independent to the global ROS response. The relocalisation of ROS stimulating proteins may provide a mechanism for this.

Further, CRK2 was co-immunoprecipitated with PDLP1 and CALLOSE SYNTHASE1 (CalS1) and was shown to be essential for callose deposition in response to salt stress (Hunter *et al.*, 2019). Note that this runs contrary to prior evidence of CalS1 induced callose deposition in response to pathogen stress and SA, but not hydrogen peroxide treatment (which simulates the effect of ROS production) (Cui & Lee, 2016).
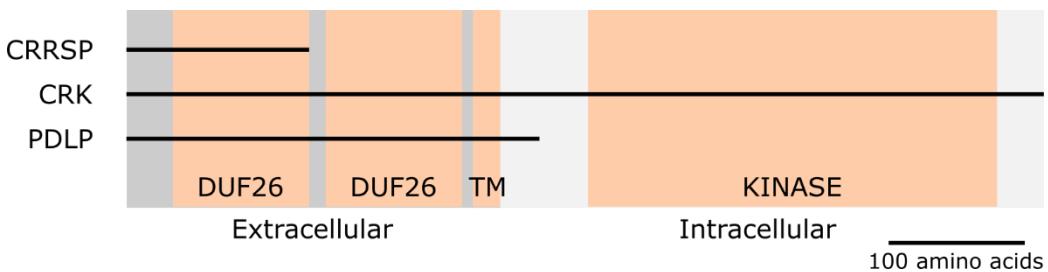
Figure 4-2 Schematic of DUF26-containing proteins to scale.
CYSTEINE-RICH RECEPTOR-LIKE SECRETED PROTEINs (CRRSP) proteins are the smallest of the DUF26-containing proteins exemplified here by Gnk2 from *Gingko biloba*, although some CRRSPs have two DUF26 domains. CYSTEINE-RICH RECEPTOR-LIKE PROTEIN KINASEs (CRKs) contain tandem DUF26 domains, a transmembrane domain (TM) and a kinase domain. *A. thaliana* CRK13 is depicted here. PLASMODESMATA-LOCALISED PROTEINs (PDLPs) are similar to CRKs, although lack the kinase domain. *A. thaliana* PDLP1 is drawn to scale here. Orange boxes indicate conserved domains. Dark grey shows the extracellular region and light grey the intracellular region.

4.1.1.3   PLASMODESMATA-LOCALISED PROTEINs

PDLPs are the most evolutionary recent class of DUF26-containing proteins, arising in the spermatophytes (seed plants) (Vaattovaara *et al.*, 2019). All eight members of this family in *A. thaliana* localise at plasmodesmata (Thomas *et al.*, 2008). PDLPs, akin to CRRSPs and CRKs, have also been shown to have a role in plant defence (Amari *et al.*, 2010; Lee *et al.*, 2011b).

The role of PDLPs in defence are better understood than other DUF26 proteins: PDLPs are required to deposit callose at plasmodesmata (Thomas et al., 2008). This prevents viral spread through plasmodesmata (Amari et al., 2010). Interestingly, PDLPs are also required for immunity against extracellular pathogens: fungi and bacteria (Lee et al., 2011b; Caillaud et al., 2014). This has been linked to PDLPs localised both at plasmodesmata and haustoria, which are fungal penetration structures. This indicates PDLPs are required for callose deposition in general and are not, as their name would suggest, plasmodesmata-specific. PDLPs are required for general defence, including affecting the spread of cell-to-cell signals from calcium to systemic defence hormones (Carella *et al.*, 2015; Lim *et al.*, 2016; Toyota *et al.*, 2018). Furthermore, PDLPs are also required for full resistance from herbivores (Bricchi *et al.*, 2013). Clearly, PDLPs are deeply entwined with plant defence responses, affecting how they communicate and signal plant defence intercellularly with mis-regulation leading to a general increase of susceptibility.

Research has largely focussed on PDLP overexpression transgenics, as some of the PDLP proteins are functionally redundant. While overexpression of PDLP1 significantly reduces cell-to-cell flux of GFP (Thomas *et al.*, 2008), single *pdlp1* mutants have no observed phenotype. This is in contrast to the single *pdlp5* mutant which have been found to have a defence phenotype (Lee *et al.*, 2011b; Wang *et al.*, 2013; Lim *et al.*, 2016). A triple mutant *pdlp1,2,3* has been used in various experiments to overcome the redundancy of some PDLP family members (Bricchi *et al.*, 2013; Caillaud *et al.*, 2014).

The combination of studies of PDLP mutants and overexpression lines have collectively identified that PDLPs positively regulate callose deposition, suggesting they function in plasmodesmal responses to a variety of stimuli. The varying degree of redundancy between PDLPs suggests there has been specific subfunctionalisation within the family. For example, PDLP5 is upregulated by the defence hormone SA, and a single gene knockout, *pdlp5*, has a susceptibility phenotype. Whereas, PDLP1 is not upregulated by SA and a single gene knockout, *pdlp1*, does not affect immunity. However, knocking out several PDLPs, *pdlp1,2,3*, may disrupt plasmodesmata enough that immunity is affected non-specifically, i.e. plasmodesmata have become dysfunctional in general and susceptibility is a consequence of this. If this is the case, the primary function of PDLP1 remains unknown. On the other hand, PDLP1 may work constitutively in plant defence alongside PDLP2 and PDLP3.

PDLP1 and PDLP5 overexpression constitutively closes plasmodesmata, as evidenced by reduced cell-to-cell flux of GFP (Thomas *et al.*, 2008; Lee *et al.*, 2011b). This is thought to contribute to the observed developmental phenotypes of dwarfism, and dwarfism combined with late flowering, for PDLP1 and PDLP5 overexpression, respectively (Thomas *et al.*, 2008; Wang *et al.*, 2013). In both cases, the reduction is cell-to-cell flux has been linked to the overproduction of callose (Thomas *et al.*, 2008; Lee *et al.*, 2011b).

4.1.2 Proteins known to be required for plasmodesmal closure in defence

Plasmodesmata are known to close via callose deposition for a range of physiological responses from dormancy to pathogen stress, and biological functions such as development (Vatén *et al.*, 2011; Faulkner *et al.*, 2013; Tylewicz *et al.*, 2018). In all cases, the end point of the pathway logically must be a CalS. Focussing on plant defence, at the beginning of the pathway there must be a receptor (or receptors) to initiate the signalling cascade.

4.1.2.1 Starting the cascade: PAMP perception

Most defence cascades in plants start with pathogen-associated molecular pattern (PAMP) detection by pattern-recognition receptors (PRRs) (Dodds & Rathjen, 2010). The most well

studied PRR is FLAGELLIN-SENSING2 (FLS2) (Gómez-Gómez & Boller, 2000): the PRR for the highly conserved 22 amino acid peptide flagellin 22 (flg22) (Felix *et al.*, 1999). FLS2 is responsible for all canonical plant defence responses in response to flg22. In addition, FLS2 is also required for flg22-mediated plasmodesmata closure and is slightly enriched at plasmodesmata (Faulkner *et al.*, 2013).

In contrast, CHITIN ELICITOR RECEPTOR KINASE1 (CERK1), the canonical chitin receptor, is not required for chitin-mediated plasmodesmata closure (Miya *et al.*, 2007; Faulkner *et al.*, 2013). Rather a plasmodesmata-enriched PRR, LYSIN MOTIF (LysM) DOMAIN-CONTAINING GLYCOSYLPHOSPHATIDYLINOSITOL-ANCHORED PROTEIN2 (LYM2), is required to initiate callose deposition at plasmodesmata in response to chitin (Faulkner *et al.*, 2013). For LYM2 to correctly transduce the chitin signal, two helper kinases are required (LysM RECEPTOR-LIKE KINASE (LYK) 4 and LYK5), as LYM2 has no intercellular kinase domain (Cheval *et al.*, 2020).

4.1.2.2 *Conferring the defence-induced plasmodesmata-closure signal*

Once the plasmodesmata-closure signal has been initiated by a PRR, the signal must be conveyed to a CalS. This has been shown to be indirect with several intermediate signalling steps between PAMP recognition and callose synthesis using reverse genetic approaches. CALMODULIN-LIKE 41 (CML41) has been shown to be required to specifically transmit the flg22-induced closure signal (Xu *et al.*, 2017). CML41 is a calcium responsive protein that was found to associate with plasmodesmata, where it presumably mediates localised flg22 signalling. Demonstrating that calcium signalling is a key component of plasmodesmal responses, chitin-triggered plasmodesmal closure was found to be dependent on two CALCIUM-DEPENDENT PROTEIN KINASEs (CPKs), CPK6 and CPK11. These CPKs mediate a specific phosphorylation signature of RBOHD, which is also essential for the response (Cheval *et al.*, 2020). It has yet to be determined whether the flg22-signal converges with the chitin-induced signal at RBOHD or further upstream, such as with CPKs.

Pattern triggered immunity (PTI) also induces plasmodesmal closure indirectly. PTI includes the induction of SA, a hormone master regulator of plant defence (Tateda *et al.*, 2014; Palmer *et al.*, 2017). SA also reduces cell-to-cell flux by callose deposition (Wang *et al.*, 2013). PDLP5 is required for SA-induced plasmodesmata closure (Wang *et al.*, 2013). Interestingly though, SA biosynthesis is also required for PDLP5-induced plasmodesmata closure (Lee *et al.*, 2011b; Wang *et al.*, 2013). This creates a model where plasmodesmata require only PDLP5-upregulated SA signalling to signal plasmodesmata closure. However, this raises the question

as to whether PDLP5 directly induces plasmodesmata closure, or simply creates a local maxima of SA. Due to the lack of transcriptional responsiveness of other PDLPs to SA, it is assumed that PDLPs directly signal plasmodesmata closure.

4.1.2.3   Depositing callose

There are 12 known CALLOSE SYNTHASEs (CalSs) in *A. thaliana* (Hong *et al.*, 2001), seven of which have been implicated with plasmodesmal callose deposition (Table 4-1). Five of these, plus an additional two CalS, have been found in plasmodesmata proteomes (Table 4-1). Thus, CalSs show significant diversity at plasmodesmata, with different CalS acting non-redundantly to different stimuli (Cui & Lee, 2016). CalSs have a secondary nomenclature of GLUCAN SYNTHASE-LIKEs (GSLs), which will not be further used here (Richmond & Somerville, 2000).

Callose is deposited at plasmodesmata specifically within 30 minutes of PAMP treatment, which temporally differs from the characteristic macroscopic callose deposition seen at 24 h post treatment (Xu *et al.*, 2017). Of the CalS known to act at plasmodesmata, four are likely candidates to be involved in defence-induced plasmodesmal callose: CalS1, CalS8, CalS10 and CalS12.

CalS1 has been directly implicated in SA plasmodesmal responses, and so is a likely candidate of downstream defence signalling (Dong *et al.*, 2008; Cui & Lee, 2016). CalS8 has been implicated in hydrogen peroxide mediated closure, and so by extension ROS signalling, which also links it to plasmodesmal pathogen responses. Although, SA responses are independent of Cals8, and so Cals8 is a less likely candidate in defence-induced plasmodesmal callose than CalS1 (Cui & Lee, 2016). Furthermore, CalS8 is the only CalS among these four which are not significantly transcriptionally upregulated by SA (Woo *et al.*, 2020).

CalS10 (also known as CHORUS) is known to maintain a basal plasmodesmal callose level at the epidermis (Guseman *et al.*, 2010). Thus, it is possible that CalS10 is involved in defence responses. This is made more likely by the short timescale of the plasmodesmata callose deposition, as it appears more likely an already present protein will be functionally post-translationally upregulated than new proteins transcribed and translated.

CalS12 has been identified as a protein involved with SA-mediated plant defence for a long time. However, this was in the context of depositing callose at papillae (cell wall thickenings) at penetration sites, or as macroscopic callose depositions (Nishimura *et al.*, 2003; Ellinger *et al.*, 2013). Recently, CalS12 has been implicated at plasmodesmata in response to iron stress

(O'Lexy *et al.*, 2018). Thus, it is likely that CalS12 may also be acting to deposit callose at plasmodesmata in a defence context.

Table 4-1 Summary of all known activity of CALLOSE SYNTHASEs (CalSs) at plasmodesmata

Names	Locus	Known plasmodesmal functions	Plasmodesmal reference(s)
CalS1/GSL6	AT1G05570	SA-induced callose	(Fernandez-Calvino *et al.*, 2011; Cui & Lee, 2016; Brault *et al.*, 2019)
CalS2/GSL3	AT2G31960	NA	(Fernandez-Calvino *et al.*, 2011; Brault *et al.*, 2019)
CalS3/GSL12	AT5G13000	Callose in the stele, roots and phloem Ectopic expression closes plasmodesmata (iCals3m)	(Vatén *et al.*, 2011; Fernandez-Calvino *et al.*, 2011; Yadav *et al.*, 2014; Brault *et al.*, 2019)
CalS7/GSL7	AT1G06490	Callose on phloem sieve elements	(Xie *et al.*, 2011)
CalS8/GSL4	AT3G14570	H_2O_2-induced callose Basal callose	(Cui & Lee, 2016)
CalS9/GSL10	AT3G07160	Possibly requires CalS10	(Fernandez-Calvino *et al.*, 2011; Saatian *et al.*, 2018)
CalS10/GSL8/ CHORUS	AT2G36850	Callose during the phototropic response Callose in roots Interacts with PDLP5 Possibly requires CalS9 Epidermal callose	(Guseman *et al.*, 2010; Fernandez-Calvino *et al.*, 2011; Han *et al.*, 2014; Saatian *et al.*, 2018; Brault *et al.*, 2019)
CalS11/GSL1	AT2G31960	NA	(Brault *et al.*, 2018)
CalS12/GSL5/ PMR4	AT4G03550	Callose in response to iron	(Fernandez-Calvino *et al.*, 2011; O'Lexy *et al.*, 2018)

4.1.2.4   Summary

Overall, a complex picture is developing of the control of plasmodesmata during defence. For some pathogen signals plasmodesmal responses are mediated by specific receptors, e.g. LYM2 for chitin perception. In other cases, plasmodesmal signalling specificity is mediated by downstream components such as CML41 for flg22. Once perceived, callose deposition is signalled in a plasmodesmata-specific manner, using proteins shared in other defence responses. For example, ROS is required but not altered on a whole cell level and callose is synthesised specifically at plasmodesmata before other locations in the cell. This suggests there is a signalling nexus that integrates signals from multiple stimuli into a plasmodesmata-response pathway. Given their relevance to plasmodesmal responses to a broad range of stimuli, it is possible that PDLPs might act as such a nexus, or as a component of the pathway that acts downstream of such a nexus, feeding in to activation of callose synthesis by callose synthases.

4.1.3   Co-immunoprecipitation experiments for plasmodesmal interactors

There are many approaches to identifying novel interactor proteins, either through screens of the entire genome, such as in yeast-two-hybrid experiments, or in a more targeted way, such as co-immunoprecipitation coupled to mass spectrometry (Kadota *et al.*, 2016). These approaches could be applied to PDLPs to find interactors. However, plasmodesmata are a very small proportion of the cell's mass and so additional experimental or analytical techniques must be employed to enrich for plasmodesmata interactors. This has been done before for two other plasmodesmata-localised proteins.

To discover interactors of a *Cucurbita maxima* non–cell-autonomous protein, one research group utilised a co-immunoprecipitation technique. Here, they modified the input material to enrich for plasmodesmata proteins, by using a plasmodesmata-enriched cell wall fraction. They found two different proteins that alter growth and development, NtNCAPP1 and NtPDGLP1 (Lee *et al.*, 2003; Ham *et al.*, 2012). In both cases, the initial co-immunoprecipitation was done in *Nicotiana tabacum* 'Bright Yellow 2' (BY-2) cells. Ham *et al.* (2012) extended this technique to *A. thaliana* seedlings. In both BY-2 cells and *A. thaliana*, very few interactors were identified by this method (six and seven, respectively) (Lee *et al.*, 2003; Ham *et al.*, 2012). This is surprisingly low, with a 'normal' co-immunoprecipitation expected to yield tens to hundreds of proteins (for examples see: Krügel et al., 2012; Kadota et al., 2014; Bellati et al., 2016; Junková et al., 2018).

Kriechbaumer *et al.* (2015) took an alternative approach, utilising downstream analysis. They identified 146 putative interactors of RTNLB3 and RTNLB6 by co-immunoprecipitating the proteins from whole tissue extract. These putative interactors were subsequently cross-referenced against a plasmodesmal proteome (AtCells1 (Fernandez-Calvino *et al.*, 2011)) to select likely plasmodesmata interactors. In this way, they identified a subset 12 plasmodesmata proteins for further interaction analysis by FRET-FLIM.

4.1.4   Aims

It has been shown multiple times for fungal, oomycete and bacterial pathogens that disrupting plasmodesmal responses increases susceptibility (Faulkner *et al.*, 2013; Caillaud *et al.*, 2014; Xu *et al.*, 2017). The PDLP proteins have been shown to be important for plasmodesmata-specific callose deposition in plant defence (Lee *et al.*, 2011b). First, this chapter aims to identify novel interactors of PDLPs. A multi-pronged strategy is undertaken to find genetic and physical interactors with PDLP1. The dwarfed phenotype of 35s::PDLP1-GFP plants is used as a readout for both a forward and reverse genetic approach. In addition, I search for physical interactors of PDLP1 and PDLP5 using a split-ubiquitin approach and a plasmodesmata-specific co-immunoprecipitation. Secondly, I aim to characterise a putative interactor with PDLPs, as to whether it interacts downstream of PDLPs in PDLP-mediated plasmodesmata callose deposition and if the candidate protein is involved in PAMP-mediated plasmodesmata callose deposition.

4.2   Methods

4.2.1   Plant growth conditions

Arabidopsis thaliana plants grown on soil were grown in short day conditions (10 h light / 14 h dark) at 22°C. Seed were sown on soil and stratified for two days at 4°C in the dark. Trays were moved to light conditions and covered with a transparent cover for seven days, after which the cover was removed.

Arabidopsis thaliana plants grown on plates were grown in long day conditions (16 h light / 8 h dark) at 20°C. Seeds were surface sterilised by a 5 minute incubation in sterilization solution (100 mL of sterilisation solution contained 95 mL water, 5 mL bleach and 20 µL TWEEN-20 (Sigma-Aldrich)). Seeds were then washed three times with sterile water. Sterilised seeds were sown on 1× MS agar (0.8%) plates supplemented with 1% sucrose. The seeds were stratified in the dark at 4°C for two days, before being moved to light conditions.

4.2.2   Plant lines used in Chapter 4

In this chapter, I used *Arabidopsis thaliana* var Col-0 as wild type (lab stock maintained by the Faulkner lab). I utilised transgenic and mutant plants already described in the literature, as set out in Table 4-2.

Table 4-2 Published *Arabidopsis thaliana* lines used in Chapter 4
NASC: Nottingham Arabidopsis Stock Centre

Name	Line	Source	Reference
PDLP1-GFP	35s::PDLP1-GFP	Carole Thomas	(Thomas *et al.*, 2008)
PDLP5OE	35s::PDLP5	Jung-Youn Lee	(Lee *et al.*, 2011b)
cals1	SALK_142792	Diana Papp	(Dong *et al.*, 2008)
cals7	SALK_048921	Diana Papp	(Barratt *et al.*, 2011)
rbohd	*dSpm* insertion	Cyril Zipfel	(Torres *et al.*, 2002)
nhl3-1	SALK_035427C	NASC	(Humphry *et al.*, 2010)
nhl3-2	SALK_150318C	NASC	(Humphry *et al.*, 2010; Singh *et al.*, 2018)

4.2.3   Crossing of *Arabidopsis thaliana*

In all crosses PDLP1-GFP was used as the male line, due to the dwarfed size of its gynoecium. *Arabidopsis thaliana* plants were grown in long day conditions (16 h light / 8 h dark) at 22°C. Two unopened buds on the first bolt of the female line were selected and the remaining flowers, buds and meristem removed. The sepals, petals and stamens from the selected buds were removed leaving a single immature carpel. The two gynoecia were wrapped in clingfilm and left overnight. One day later, when the stigma had developed to allow pollination, stamens from PDLP1-GFP were brushed against the carpel. Pollinated gynoecia were wrapped in cling film until the first signs of silique development, when the clingfilm was removed. Dissections were carried out with a pair of sharp tweezers.

When the siliques were mature, the F1 seeds were grown on soil (see 4.2.1). F2 seeds were grown on plates (see 4.2.1) and were screened by epifluorescence microscopy for GFP expression and PCR for homozygous mutations. PCR was carried out on gDNA (see 4.2.5.1) using GoTaq® G2 Flexi DNA Polymerase (Promega) as per the manufacturer's instructions and with the primers as listed in Table 4-3. Lines which were had GFP expression and were homozygous for the mutation were transferred to soil and allowed to set seed.

F3 seeds were grown on plates (see 4.2.1) and were screened by epifluorescence microscopy for GFP expression. Lines were expected to segregate 1:2 homozygous:heterozygous for GFP expression. Lines were selected which were homozygous for PDLP1-GFP.

Table 4-3 Primers used to genotype F2 crosses

Name	Wild-type Forward Primer	Wild-type Reverse Primer	Mutant Primer
cals1	AGAAGATCGCAAAGGTCAAACCAAT	AGGAAAGTCAAAGCATTCTGTGTGG	Reverse primer and LBa1
cals7	GCTTCCCTCGCATTTCTTCTACTC	AGGAACATCCACATTCGGTCG	Forward primer and LBa1
rbohd	GTCGCCAAAGGAGGCGCCGA	GGATACTGATCATAGGCGTGGCTCCA	Reverse primer and dSpm1
nhl3-1 and *nhl3-2*	ACGAAGACTGAATGGCGGACTTAAACGGTG	TGGAAGACTACGAACCAAAGTCAACGTCACACTTGG	Forward primer and LBa1
LBa1	NA	NA	TGGTTCACGTAGTGGGCCATCGCTTATTTCAGTAAG
dSpm1	NA	NA	AGTGTGGGGTTTTG

4.2.4   Rosette Area Phenotyping

Seeds were surface sterilized in bleach, and washed with water three times. Seeds were sown onto 1% sucrose 0.8% agar Murashige and Skoog medium (Murashige & Skoog, 1962) square 100 cm^2 plates in a regular 8 × 8 grid. Plates were stratified at 4°C for two days before being moved to a long day growth chamber (16h light / 8h dark). 22 days post germination photos were taken.

RGB images were imported into FIJI (Schindelin *et al.*, 2012; Schneider *et al.*, 2012) and set to scale using the 10 cm square plate as a reference. Images were made binary using a "Moments Colour Threshold" to identify plant tissue. Individual plants were drawn around by hand and the area of plant tissue measured in FIJI. The data were analysed with an ANOVA followed by a post-hoc Tukey test, if required, in R (v4.0.0).

4.2.5   Whole genome resequencing

4.2.5.1   CTAB DNA extractions

Whole plants were pooled in 5 mL Eppendorf tube and flash frozen in liquid nitrogen. Tissue was ground by hand with a micropestle. 500 mg of tissue was transferred into a 2 mL Eppendorf and incubated with CTAB extraction buffer (100 mM Tris-HCl pH 8, 20 mM EDTA, 1.4 M NaCl, 2% (w/v) CTAB, 1% (w/v) PVP 40 kDa) at 65°C for 30 minutes. At room temperature, 500 µL of Chloroform / Isoamylalcohol (24:1) was added, and the mixture centrifuged. The aqueous upper phase was kept and was washed again with 300 µL of Chloroform / Isoamylalcohol (24:1) in the same manner. 500 µL of isopropanol was added and incubated at -20°C for 30 minutes. The solution was centrifuged for 15 minutes and the supernatant discarded. The DNA pellet was washed twice with 300 µL ethanol (70% v/v) and allowed to air dry. The pellet was resuspended in dH$_2$O.

4.2.5.2   Data processing

Genomic DNA pools was sequenced by Novogene with 150 bp paired-end Illumina reads to an approximate depth of 20×. Reads were trimmed with Trimmomatic (v0.33) with a headcrop of 15 bases and sliding window of size 4 bases and minimum average phred quality 20 (Bolger *et al.*, 2014). Quality was checked using FastQC (v 0.11.8) (Simon Andrews, 2020). Reads were aligned to the Araport11 reference genome using the Burrows-Wheeler Aligner (v0.7.5) (Li & Durbin, 2009). The quality of mapping is shown in Table 4-5. Sequence alignment/map (SAM) files were sorted to binary alignment/map files using samtools (v1.4.1) (Li *et al.*, 2009). SNPs were identified using the bcftools (v1.8) 'mpileup' and 'call' commands with a phred quality filter of 20 (Li, 2011). This generated a binary variant call format file

(Danecek *et al.*, 2011). A tab-separated values file of SNPs was generated from the binary variant call file using an in-house Perl script, derived from ("BCF2NGM.pl", (Austin *et al.*, 2011)).

SNPs were filtered for a minimum read depth of 5, and canonical EMS transitions (C -> T and G -> A). SNPs were determined to be fixed if the alternative allele frequency (i.e. T/(C+T) or A/*G+A)) was greater than or equal to 0.8. SNPs that were fixed in Col-0 were disregarded in the downstream analysis. Fixed SNPs in the POEM46 parent were identified. POEM46 SNP ratios were compared in the F2 large and F2 small pools.

4.2.6   Split-ubiquitin screen by DualSystems Biotech

This screen was carried out by DualSystems Biotech on behalf of Christine Faulkner on 7th July 2010 (Möckli *et al.*, 2008). The screen is a split-ubiquitin system for detecting interactions of membrane proteins (Stagljar *et al.*, 1998).

PDLP5 was amplified from cDNA and ligated into pBT3-SUC and sequence verified. pBT3-SUC is used to fuse a cleavable signal sequence derived from SUC2 (invertase) to the N-terminus of PDLP5 and Cub-LexA-VP16 to the C-terminus. The expression of PDLP5-Cub-LexA-VP16 in yeast (strain NMY32) was verified by co-expression with pOst1-NubI prey, which is expected to give a strong interaction. Thus the co-transformed yeast grew on SD-ade-leu-trp medium. 5 mM 3-aminotriazole was added to the screening medium to prevent self-activation of PDLP5-Cub-LexA-VP16. PDLP5-Cub-LexA-VP16 was screened against "*Arabidopsis thaliana* NubG-x, P02210" library with reporter plasmid. 3.6×10^6 transformants were screened for growth on SD-ade-his-leu-trp + 3-aminotriazole medium, of which 123 grew. The interaction identified by autotrophism was combined with a secondary β-galactosidase assay. All 123 clones also activated lacZ. Therefore 123 interacting clones were identified and inserts sequenced. The sequences were searched against the NCBI non-redundant protein sequences database limited to *Arabidopsis thaliana* (id 3702) sequences using the BLASTx algorithm. In total 65 interactors were found from 96 sequences.

4.2.7   Plasmids used in Chapter 4

In this chapter, I constructed plasmids to produce fluorescently-tagged proteins for co-localisation and FRET-FLIM experiments. I constructed level 0 modules of PDLP5 (AT1G70690), NHL3 (AT5G06320) and a peroxidase gene (AT1G71695) with overhangs AATG-TTCG. In addition, I used the level 0 AB15 created my Annalisa Bellandi which is a truncated version of AT5G43020) with overhangs AATG-TTCG. These plasmids and sequences are listed in Table S1.

Level 1 plasmids were made from the level 0 plasmids in a one-pot reaction (see 3.2.8) with the following modules: a forward level 1 position 3 acceptor (pICH47751), a 35s short promoter and 5' UTR (pICH51277), the coding sequence of interest, a 35s and terminator (pICH41414), and one C-terminal fluorescent tag of eGFP (see Table S1) or mCherry (pICSL50004).

I also used a the published plasmid 35s::PDLP1-GFP (PDLP1-GFP) with the backbone pB7FW2.0 (Thomas *et al.*, 2008).

4.2.8   FRET-FLIM

Leaves of 4 − 5-week-old *N. benthamiana* transiently expressing the constructs of interest were used for FRET-FLIM. The abaxial side of the leaf samples were imaged using a 63x/1.2 water immersion objective lens (Leica C-APOCHROMAT 63×/1.2 water). FLIM experiments were performed using a Leica 110TCS SP8X confocal microscope equipped with TCSPC (time correlated single photon 111counting) electronics (PicoHarp 300), photon sensitive detectors (HyD SMD detector), a pulsed laser (white light laser, WLL). The WLL was used to excite GFP at 488nm. GFP emission was collected between 500 nm – 530 nm. The repetition rate was set to 40 Mhz. A 488 nm notch filter was used to reduce interference from reflected light. The FLIM data sets were recorded using the Leica LASX FLIM wizard linked to the PicoQuant SymPhoTime 64 software. The FLIM data sets were acquired by scanning each image until a suitable number of photon counts per pixel (maximum 1,000) was reached. For the acquisition, the image size was set to 125×125 pixels, allowing a pixel dwell time of 19 μs. Data were analysed by obtaining excited-state lifetime values of the all pixels with over 250 photons. Calculations were performed using the PicoQuant SymPhoTime 64 software instructions for FLIM-FRET-Calculation for Multi-Exponential Donors, using a two-exponential decay for GFP. The lifetimes were initially estimated by fitting the data using the Monte Carlo method and then by fitting the data using the Maximum Likely Hood Estimation. The amplitude weighted average donor lifetime with model parameter n = 2 was used to calculate the average FRET-efficiency. FRET efficiency (E) was calculated by comparing the lifetime of the donor in the presence τ_{DA} or absence τ_D of the acceptor according to the following formula: $E = 1 - \frac{\tau_{DA}}{\tau_D}$.

4.2.9   RNA extraction and RT-qPCR

RNA was extracted from rosette leaves from 5-week-old *A. thaliana* plants grown on soil under short day conditions. RNA was extracted using the RNAeasy Qiagen Plant Mini Kit (Qiagen) with buffer RLT as per the manufacturer's instructions. DNA contamination was

removed using TURBO DNAse (ThermoFisher) as per the manufacturer's instructions. cDNA was synthesised from 500 ng of RNA using GoScript Reverse Transcriptase (Promega) with oligio(dT) primers (Applied Bioscience) as per the manufacturer's instructions. cDNA was diluted 1:9 with distilled water before use.

qPCR was performed with a LightCycler 480 (Roche Applied Science) using LightCycler® 480 SYBR Green I Master (Roche Applied Science) following the the manufacturer's protocol with a 60°C annealing temperature for a 10 µL reaction with a melt curve. A CT value was estimated with 100% primer efficiency using a second derivative maximum method. The arithmetic mean of three technical replicates was calculated. The arithmetic mean of three biological replicates was then calculated for each gene. The difference the mean NHL3 CT and the geometric mean of the CT value of UBQ10 and UBC was used to calculate the log fold change in expression by $2^{-\Delta CT}$. The primers for qPCR are outlined in Table 4-4. The location of the *NHL3* qPCR primers relative to *NHL3* and the *nhl3-1* and *nhl3-2* T-DNA insertions is shown in Figure 4-3.

Table 4-4 Primers used for *NHL3* RT-qPCR

	Forward Primer	Reverse Primer
NHL3	CGAATTCACACTCGACCCAAC	TCGCTGATCACCGTAGTATCC
UBQ10	AGTCTACTCTTCACTTGGTCCTGC	GCCCCAAAACACAAACCACCAAAG
UBC	GACGCTTCAGTCTGTGTGTAGAGC	CTTAGAAGATTCCCTGAGTCGCAG
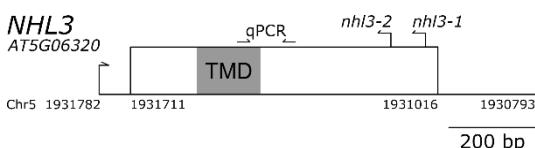
Figure 4-3 The position of RT-qPCR primers and T-DNA insertions for *NHL3*
The transcript for NHL3 (AT5G06320.1) is shown to scale. The arrow denotes the transcription start site. The protein coding sequence is shown by the box, with the location of a predicted transmembrane domain (TMD) in grey. The transcript includes a 5' and 3' untranslated region. The locations of the two T-DNA insertions for *nhl3-1* (SALK_035427C) and *nhl3-2* (SALK_150318C) are noted, as well as the binding sites for the qPCR primers used in this study.

4.2.10 Callose staining

Rosette leaves of 4 – 5 week old *A. thaliana* plants were infiltrated with 0.1% (w/v) aniline blue dissolved in PBS pH 7.4 (Sigma-Aldrich) after no treatment or 30 minutes after adaxial infiltration with 100 nM flg22. Leaves were cut into 1 cm^2 samples and mounted adaxially on a slide. Samples were imaged on a ZEISS LSM800 confocal microscope with a 63×/1.2 water immersion objective lens (C-APOCHROMAT 63×/1.2 water). Aniline blue was excited at 405 nm with a UV laser and collected at 430 – 470 nm. Micrographs were taken as z-stacks with

z-step of 1 μm starting above the mesophyll and ending below the stomata, so as to capture only epidermal plasmodesmata. Images were captured as 64-bit and resolution 1024 by 1024 pixels.

Callose was quantified by measuring aniline blue fluorescence using an in-house FIJI script (Schindelin *et al.*, 2012). The script quantified plasmodesmata callose deposits in each z-frame of an image. First, the image was made binary using a user-defined threshold. The resulting black and white image was "despeckled". The integrated density for each particle with a pixel size between 30 and 250 was then quantified.

The data were analysed in R (v4.0.0) (R Core Team, 2020). A random mixed effects model was fitted using lme4. Genotype, treatment, the interaction of genotype and treatment and experiment ID were treated as fixed effects. Images nested within leaves, plants, treatments, genotypes then experiments were treated as a random variable. The estimated marginal means were compared pairwise within genotypes with a Tukey adjustment for multiple comparisons using emmeans.

4.2.11 Microprojectile bombardment

Cell-to-cell connectivity was assayed by microprojectile bombardment as described by Faulkner *et al.* (2013). *A. thaliana* plants grown in short day conditions were used when 4 – 5 week's old. Expanded leaves were exercised and placed on 0.8% agar MS medium. Leaves were bombarded with 1 nm gold particles (BioRad) coated with pB7WG2.0-GFP, using a Biolostic PDS-1000/He particle delivery system (BioRad) at 1100 PSI (Kikkert, 1993). The leaves were infiltrated adaxially with water or 100 nM flg22 four hours later.

Leaves were imaged on a ZEISS LSM800 confocal microscope with a 20×/0.8 air objective lens (PLAN-APOCHROMAT 20×/0.8 air). GFP was excited at 488 nm with an argon laser and collected at 500 – 545 nm.

4.2.11.1 Analysis

Traditionally, in the literature, bombardment data is analysed with the non-parametric Mann-Whitney-Wilcox (MWW) test (Burch-Smith & Zambryski, 2010). However, as can be seen in Figure 4-23, bombardment data is not exchangeable: i.e. there is heteroskedasticity between treatments. While it is still appropriate to apply the MWW test, the result does not signify if distributions have the same median (Mann & Whitney, 1947; Hart, 2001; Conroy, 2012). Therefore, a bootstrap method was developed and used to analyse these data (Johnston & Faulkner, 2021).

4.2.12 Plasmodesmal purification co-immunoprecipitation

Plasmodesmata were extracted as previously described in section 2.2.5.

HA immunoprecipitation was carried out using Pierce magnetic HA beads. Beads were washed in 1 mL IP buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 1 mM DTT, 1× cOmplete™ ULTRA EDTA-free protease inhibitors (Roche), 1 mM PMSF, 1% v/v IPEGAL® CA-630 (Sigma)) twice. 25 µL of beads were mixed with a PD fraction and rotated for two hours. Beads were collected and washed with 300 µL of IP buffer three times. The beads were washed with 300 µL of IP buffer (0% v/v IPEGAL® CA-630 (Sigma)). For John Innes Samples, 30 µL 6x SDS Laemmli buffer was added to the beads and samples were boiled at 95°C for 5 minutes. For Cambridge Centre for Proteomics samples, proteins were acid eluted at RT in 50 µL pH 2.2 0.1 M glycine for 10 minutes. The beads were separated and the supernatant neutralised with 7 uL 1 M pH 8.5 Tris-HCl. 10 µL 6x SDS Laemmli buffer was added and samples were boiled at 70°C for 10 minutes.

4.2.12.1 Sample preparation and mass spectrometry at the John Innes Centre

Samples were run 7 mm into a 1.5 mm thick 10% TRIS resolving gel (containing 0.1% SDS) without a stacking gel, ran in a glycine 0.1% SDS running buffer. The gel was washed in dH$_2$O and then the band was exercised. Bands were washed at 65°C for 30 min in 30% EtOH; 55°C for 15 min in 50 mM TEAB/50% ACN; 55°C for 30 min in 10 mM DTT/50 mM TEAB; RT in the dark for 30 min in 30 mM Iodoacetamide (IAA)/50 mM TEAB; RT for 15 min in 50 mM TEAB/50% ACN; then RT for 15 min in 50 mM TEAB. Bands were then cut into 1 mm^2 pieces. The pieces were washed at RT for 15 min in 50 mM TEAB/50% ACN; RT for 15 min in 100% ACN; and again at RT for 15 min in 100% ACN. Excess ACN was removed by vacuum concentrators before being given to the proteomics department.

The gel bands were treated and digested with trypsin according to standard procedures. Peptides were extracted from the gel and aliquots analysed by nanoLC-HDMSE on a Synapt G2Si mass spectrometer coupled to a nanoAcquity UPLC system (Waters, Manchester, UK). Peptides were trapped using a pre-column (Symmetry C18, 5µm, 180 µm x 20 mm, Waters Ltd) which was then switched in-line to an analytical column (BEH C18, 1.7 µm, 75 µm x 250 mm, Waters Ltd) for separation. Peptides were eluted with the following gradient of solvents A (water, 0.1% formic acid) and B (acetonitrile, 0.1% formic acid) at a flow rate of 200 nl min-1 and a temperature of 40°C : 0-1 min at 2% B; 1-10 min linear increase B to 11%; 10-74 min increase B to 31%; 74-89 min increase B to 44%; followed to a ramp to 99% B and re-equilibration to 2 % B. The column was connected to a 10 µm SilicaTip™ nanospray emitter

(New Objective, Woburn, MA, USA) for infusion into the mass spectrometer. Glu-Fibrinogen peptide (1 pmole μl-1, Sigma-Aldrich) was infused at 0.5 μl min-1 as a lock mass (m/z 785.842) for recalibration and measured every 60 s. The mass spectrometer was controlled by the Masslynx 4.1 (SCN957) software (Waters) and operated in positive HDMSE and resolution mode with capillary voltage of 3.6 kV, cone voltage of 40 V, cone temperature of 90°C. Scan time was 1 s over the range of 50-2000 m/z. The quadrupole profile was set to 300 (15% dwell), 450 (20% ramp and 30% dwell, 600 (40% ramp). For the low energy scan the trap and transfer collision (CE) was off, and for the high energy scan the CE was applied from look-up table generated from previously acquired HDMSE data. IMS settings were 650 for wave velocity and 40 for wave height.

Data were processed in Protein Lynx Global Server 3.0.2 (Waters) with the following parameters: chromatographic peak width and TOF resolution: automatic; lock mass for charge 2: 785.842 Da; Thresholds 80 counts for low and 20 counts for high energy; intensity threshold: 750 counts. For the database search a custom database was generated from Nicotiana benthamiana protein sequences downloaded from uniprot.org (744 sequences, as of Jan 2018), the MaxQuant contaminants database (246 entries), and the sequences of the constructs used in this experiment (3 sequences). The search ws performed with the following parameters: peptide/fragment tolerance automatic, min fragment ions per peptide/protein =2,4, min peptides per protein =1, enzyme = trypsin, 2 missed cleavages, fixed modification = carbamidomethyl (C), variable modifications = oxidation (M), deamidation (N,Q). Results were exported into Excel spreadsheets.

4.2.12.2 Sample preparation and mass spectrometry at the Cambridge Centre for Proteomics,

Samples were run 5 mm into a 1.5 mm thick 10% TRIS resolving gel (containing 0.1% SDS) without a stacking gel, ran in a glycine 0.1% SDS running buffer. The gel was washed in dH_2O and then the band was exercised. The bands were washed four times in 20% acetonitrile at 40°C for 15 minutes to remove detergents, and then stored at 4°C in a 1.5 mL Eppendorf with 100 μL of dH_2O. Samples were analysed as previously described in section 3.2.2.

4.3   Results

4.3.1   PDLP1 and PDLP5 overexpression caused dwarfing in *Arabidopsis thaliana*

I chose to phenotype PDLP1 and PDLP5 overexpression plants, to find the most suitable phenotype to characterise. It has been previously reported in the literature that the overexpression of PDLP1 (35s::PDLP1-GFP, denoted as PDLP1-GFP (Thomas *et al.*, 2008)) and PDLP5 (35s::PDLP5, denoted as PDLP5OE (Lee *et al.*, 2011b)) causes dwarfing in *A. thaliana*. I tested the same lines as those published to ascertain the degree of dwarfism in my experimental system (Figure 4-4). Both PDLP1-GFP and PDLP5OE plants were significantly dwarfed. PDLP1-GFP overexpressing plants were 36% the size of Col-0 wild type (0.36 cm^2), whereas the PDLP5OE plants were 1.78× bigger than PDLP1 plants and 65% the size of Col-0 (0.65 cm^2). With this result, I was able to follow up on a second-site suppressor screen for dwarfism conducted previously on PDLP1-GFP EMS mutants.
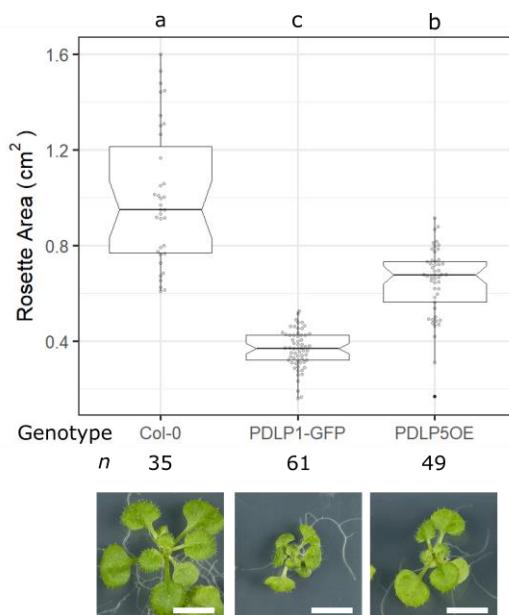
Figure 4-4 Overexpression of PDLP proteins dwarfed *A. thaliana*.
Transgenic *A. thaliana* plants were grown on agar for 22 days. Plants overexpressing PDLP1-GFP (35s::PDLP1-GFP) and PDLP5OE (35s::PDLP5) had a significantly smaller rosette area compared to Col-0 (wild type). Letters indicate significantly different Tukey groups (p < 0.05). Scale bar = 0.5 cm.

4.3.2   PDLP1 Overexpression Mutant (POEM) plants were no longer dwarfed

To find novel proteins interacting with PDLP1 and consequently functioning in plasmodesmata closure, I took advantage of mutants identified in a previous screen for suppression of the PDLP1-GFP dwarfed phenotype. The screen and mutant selection was carried out by Lourdes Fernandez-Calvino in 2009. The screen was of 500 M2 plants from an EMS treated PDLP1-GFP population in the Col-0 background. Initially, 46 lines were found to have reverted to a wild type size. Of these 46 lines, 14 lines were expressing GFP when

observed by confocal microscopy. Lastly, the expression of full length PDLP1-GFP was checked by Western blot, and 11 of the 14 lines contained full length protein. Christine Faulkner carried out further characterisation of the remaining 11 lines and selected the four with the strongest phenotype. I have continued to work on these four lines. We hypothesised that plants with a wild type size would have a mutation in the pathway downstream of PDLP1 action. These plants were termed POEM (**P**DLP1 **O**ver**e**xpression **M**utant).

First, I confirmed the phenotype of the four POEM M4 lines selected by Christine Faulkner in my hands. The POEM plants were no longer dwarfed, yet still had plasmodesmata-localised PDLP1-GFP expression (Figure 4-5, Figure 4-6).
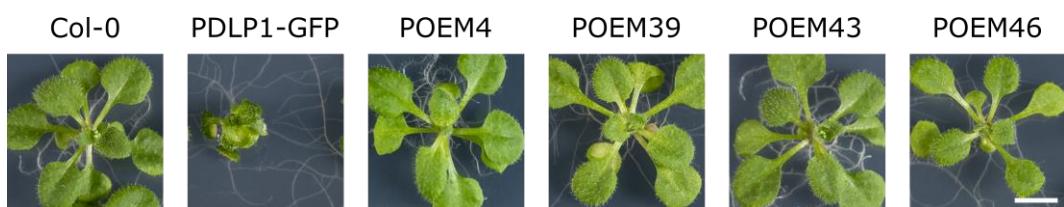
Figure 4-5 POEM plants no longer have a dwarfed size.
PDLP1-GFP plants were EMS mutagenized and revertant mutants were selected, termed POEM (**P**DLP1 **O**ver**e**xpression **M**utant). Scale bar = 0.5 cm.
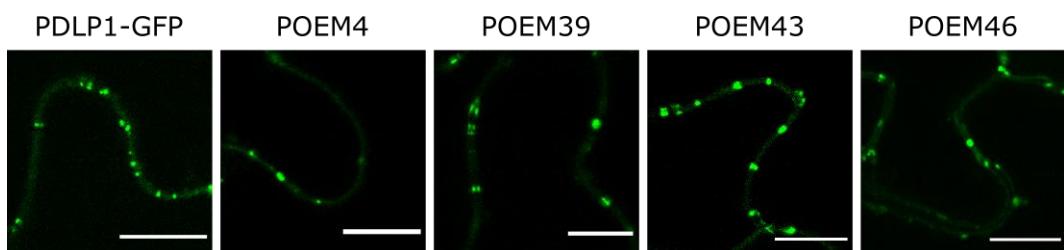
Figure 4-6 POEM plants maintain plasmodesmata-localised PDLP1-GFP expression.
PDLP1-GFP expression and plasmodesmal localisation was reconfirmed after mutagenesis by confocal microscopy of the GFP transgene. Scale bar = 10 µm.

However, as part of my phenotyping I also checked the homozygosity of the PDLP1-GFP transgene through BASTA selection. Unfortunately, the M4 lines of POEM4, POEM39 and POEM43 had plants which were not resistant to BASTA and so did not contain the transgene. These lines could not be used for further investigation until the transgene was homozygous again. I started the work of backcrossing the lines through single seed decent, but did not manage to further characterise them.

Therefore, I chose to focus on POEM46 and further characterised its phenotype, as the PDLP1-GFP was fixed in this line. POEM46 showed an intermediate phenotype with significant reversion from dwarfism, with a mean rosette area 2.5× greater than PDLP1-GFP

(1.0 cm^2 compared to 0.42 cm^2, respectively), yet still not as large as wild type plants (2.2 cm^2, Figure 4-7).
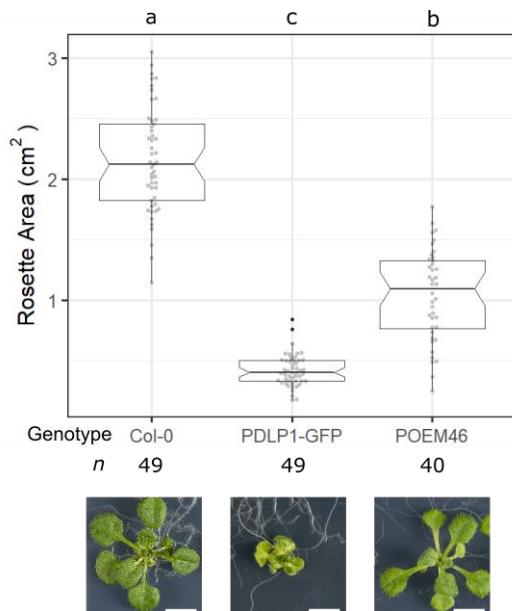
Figure 4-7 POEM46 is significantly larger than PDLP1-GFP.
POEM46 has an intermediate size phenotype, with a significantly larger rosette area than PDLP1-GFP but a smaller rosette area than Col-0 (wild type). Letters indicate significantly different Tukey groups (p < 0.05). Scale bar = 0.5 cm.

4.3.3 Candidate Single Nucleotide Polymorphisms in POEM46

I wanted to locate the causal mutation in the POEM46 line that partially reverted PDLP1-GFP induced dwarfism. To do so, I took a bulk segregant analysis approach, where a pool of F2 backcrossed plants can be used to find non-segregating mutations (Austin *et al.*, 2011). Normally, one would outcross an EMS mutant in the Col-0 background to another ecotype (usually *Landsberg erecta*) and then use the segregation of the ≈500,000 known single-nucleotide polymorphisms (SNPs; 1 every 250 bp) to determine where the mutation is (Schneeberger *et al.*, 2011). However, the PDLP1-GFP transgene must be homozygous to induce strong dwarfing, as the heterozygous state has an intermediate phenotype that may be confused with the POEM mutation (Thomas *et al.*, 2008). Therefore, I backcrossed POEM46 to the PDLP1-GFP parent. In this manner, the transgene remains homozygous and the dwarfing phenotype stable. Consequently, however, this also reduces the number of single nucleotide polymorphisms (SNPs) available to map the mutation with to only EMS mutations – dramatically reducing the mapping power. This removes the possibility for a "sliding window" mobig average analysis of SNP frequencies. However, the sequencing nature of the bulk segregant analysis approach means instead each individual SNP can be interrogated for its alternative allele frequency (EMS mutant reads/Col-0 reads). The crossed F1 plant was selfed to produce an F2 mapping population (Figure 4-9).

The POEM46 mutation appeared semidominant, much like PDLP1-GFP, as intermediate phenotypes are also observed in the F2 mapping population (Figure 4-8). However, this does not appear to be simply semi-dominance in the expected ratio 1:2:1, rather 1.8:1:1.9 (small:intermediate:large, n = 37, χ^2_2 = 12.0, $p < 0.01$).

Figure 4-8 Size variation within the POEM46 F2 mapping population.
Within the POEM46 F2 mapping population there was a variation in plant size of 1.8:1:1.9 small (S) : intermediate (I) : large (L). This is not a simple Mendelian ratio for semi-dominance of 1:2:1. Scale bar = 0.5 cm

To overcome this, when pooling plants I did not select any intermediate-sized plants (Figure 4-9). Thus, the POEM46 mutation should be in 100% of reads in large plants as these plants must be homozygous. There should be 0% POEM46 reads in the small plants pool, if the mutation is semi-dominant, otherwise 33% POEM46 SNP reads are expected in a recessive allele (a 2:1 mix of heterozygous and homozygous wild type plants).

Figure 4-9 POEM46 mapping population.
POEM46 was backcrossed to the parent PDLP1-GFP line to create an F2 mapping population. This F2 mapping population is homozygous for the PDLP1-GFP transgene, but segregating for the POEM46 causal SNP. Only large or small plants were picked for the F2 sequencing pools. Scale bar 1 cm.

To find the causal mutation, pools of DNA for each line were created and sent for whole genome resequencing. The PDLP1-GFP pool (40 plants), show the genetic state before mutagenesis. The POEM46 pool (18 plants) identified all SNPs that were fixed in the parent

line, and were determined to not be background noise by comparison to the results of PDLP1-GFP. The F2 mapping pools (105 large, 95 small plants) were then used to examine the POEM46 mutations, and see how their proportions differed when in large and small plant pools.

The DNA pools were sequenced by Novogene with 150 bp paired-end reads to an approximate depth of 20× (Table 4-5). Reads were trimmed with Trimmomatic (v0.33 (Bolger *et al.*, 2014)) and aligned to the Araport11 reference genome using the Burrows-Wheeler Aligner (v0.7.5 (Li & Durbin, 2009), Table 4-5). In all cases, the libraries mapped well to the *A. thaliana* genome, with almost all reads mapping within their pairs.

Table 4-5 Mapping statistics of POEM46-associated whole genome DNA libraries.
Libraries were created by Novogene and sequenced with 150 bp paired-end reads. Reads were mapped with Burrows-Wheeler Aligner.

Library	Paired Reads	Mapped (%)	Mapped as Pairs (%)	Mean Depth (± SD)	Positions with < 5 reads (%)
PDLP1-GFP	12148252	99.95	99.01	26.2 ± 142.2	0.35
POEM46	9354383	99.82	98.94	20.2 ± 118.1	1.79
F2 Large	8696619	99.70	98.67	18.9 ± 126.1	2.05
F2 Small	13423597	98.40	97.48	28.6 ± 175.1	0.66

SNPs were called from binary alignment files (BAM) using bcftools (v1.8), with low quality SNPs removed (<20 phred Q score). Binary call format (BCF) files were processed by an in-house Perl script, derived from ("BCF2NGM.pl", (Austin *et al.*, 2011)) into .snp files. SNPs were filtered for a minimum read depth of 5, and canonical EMS transitions (C -> T and G -> A). There were 2,825 SNPs in PDLP1-GFP, of which 285 (10%) were fixed (defined as an alternative allele frequency greater than 80%). There were 2,745 SNPs in POEM46, of which 271 were fixed (9.9%). 227 (84%) of these POEM46 SNPs were already fixed in PDLP1-GFP, leaving 44 (16%) candidate POEM46 SNPs.

Of the 44 candidate SNPs, 37 were recovered in F2 Large plants and 41 in F2 Small plants. I filtered the list of candidate SNPs, for where they remained fixed in the F2 Large mapping populations (alternative allele frequency greater than 80%) and unfixed in the F2 Small plants (alternative allele frequency less than or equal to 80%), leaving a final eight candidate SNPs (Figure 4-10).
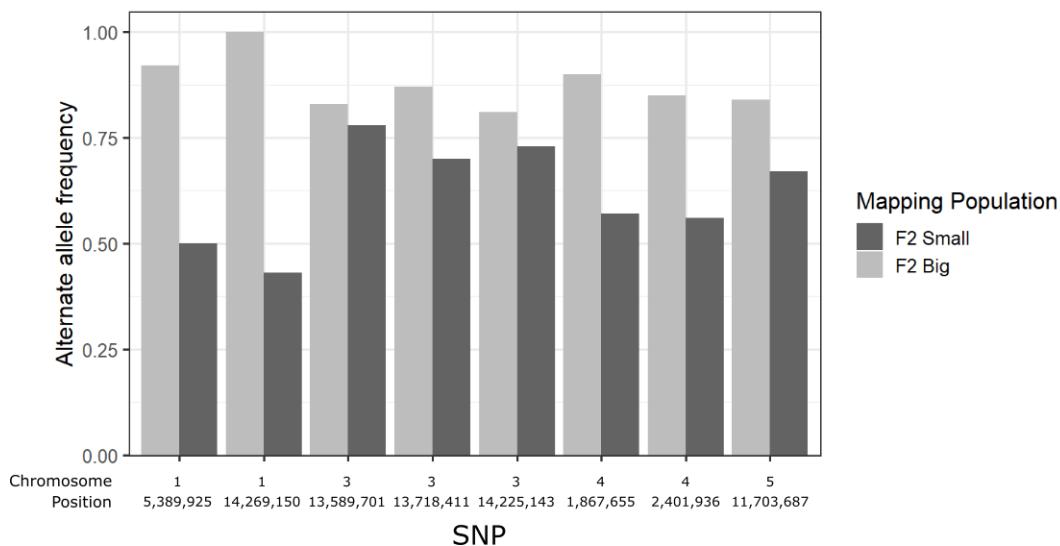
Figure 4-10 POEM46 candidate SNP alternate allele frequencies in the F2 mapping populations.
Genome resequencing of POEM46 identified 44 fixed transition SNP mutations were identified. These SNPs were filtered by alternative allele frequency (EMS SNP reads / (EMS SNP reads + WT reads)) in the F2 mapping pools. SNPs were considered to be candidate POEM46 SNPs if they were fixes (alternative allele frequency > 0.8) in the large pool and not fixed in the small pool.

I then considered the location of each SNP, using Integrative Genomics Viewer Web App, to determine whether the SNP is near or in a gene (Table 4-6). I identified a likely candidate SNP in the promoter of AT1G15670 (KISS ME DEADLY2). All 21 reads at Chr1:5 389 925 were "C" in PDLP1-GFP, whereas all 13 reads were "T" in POEM46. In F2 Big plants, 12/13 (92%) reads were "T", reduced to 12/24 (50%) in the F2 Small pool. The C -> T mutation occurred at position -27 relative to the transcription start site of KMD2.

Table 4-6 POEM46 SNP Candidates compared to Araport 11 gene annotations.
Candidate POEM46 SNP were identified by fixed alleles in the POEM46 parent genome, fixed in revertant F2 mapping plants, and segregating in wild-type F2 mapping plants. Their locations with respect to Araport11 gene annotations was noted.

SNP Location

Chromosome:Position	Araport11 Annotation
1: 5 389 925	Promoter of AT1G15670
1:14 269 150	Transposable element
3:13 589 701	Intergenic
3:13 718 411	Intergenic
3:14 225 143	Intergenic
4: 1 867 655	Intergenic
4: 2 401 936	Ta11-like transposon
5:11 703 687	Intergenic

To definitively conclude that the KISS ME DEADLY2 (KMD2) SNP is the causative SNP of the size reversal there is further work to be done. For instance, the dwarf phenotype of POEM46 should be recovered by transgenic expression of KMD2. In addition, independent confirmation of the phenotype should be demonstrated in a further mutant *kmd2* lines, such as a CRISPR knockout plant or t-DNA insertion mutants in KMD2. Further work should also be carried out on POEM46, namely to show that *KMD2* expression is reduced in POEM46 by qPCR or RNA-seq. Nonetheless, I have identified KMD2 as likely candidate protein to be required for PDLP1-mediated dwarfism.

4.3.4   PDLP1 induced dwarfing was reduced in CALLOSE SYNTHASE (CalS) mutants

Concurrently, I also took a reverse genetics approach to find proteins likely to be acting downstream of PDLP1. It is unknown how PDLP1-GFP causes a dwarf phenotype, but the literature suggested it may be overactive callose deposition (Caillaud *et al.*, 2014). The final component of any callose deposition pathway must be a callose synthase (CalS, see 4.1.2.3), and so I considered whether a *cals* mutant might reverse PDLP1-induced dwarfism.

I identified CalS1/8/10/12 as the most likely CalS to be involved with plasmodesmata callose deposits. I was only able to utilise the *cals1* mutant in this work, with *cals7* (phloem localised CalS) as a control. This was due to a lack of time to obtain mutants in CalS8 and CalS12. CalS10 was not used, as the *cals10* mutant is homozygous lethal preventing it from being used in a reverse genetic screen (Chen *et al.*, 2009; Thiele *et al.*, 2009). Moreover, weak alleles of *cals10* (*chorus* or *essp8*) are so stunted that a dwarf-based screening approach would not succeed (Guseman *et al.*, 2010; Saatian *et al.*, 2018).

Thus, I investigated CalS1 and CalS7. CalS1 is known to act at plasmodesmata, whereas CalS7 was chosen as a control as it acts in the phloem. I crossed mutants into PDLP1-GFP and, in the same assay as above, checked their size for reversion. CalS1 was required for PDLP1-induced dwarfism, and a *cals1* mutant had a full reversion to wild type size (Figure 4-11). Whereas, a *cals7* mutant had partial reversion in size (Figure 4-11). The reversion in *cals7* is unexpected, and may be due to the ectopic expression of PDLP1-GFP in the vasculature suggesting PDLP1 is promiscuous. On the other hand, the total reversion in *cals1* suggests that PDLP1 is not promiscuous and has a total dependence on CALS1. Alternatively, *cals7* is known to create an aberrant developmental phenotype itself, as the mutant blocks sugar transport, leading to increasing leaf starch levels (Figure 4-11) (Barratt *et al.*, 2011; Xie *et al.*, 2011). Thus, the decrease in dwarfing in *cals7*/PDLP1-GFP plants may be due to an increase in sugar availability in the leaves, compared to wild type: i.e. the PDLP1-GFP may be rectifying

the dwarfing caused by *cals7* in the vasculature, but causing dwarfing in the leaves. In either case, CalS1 is required for PDLP1-induced dwarfism.

As noted before, PDLP1 localises to plasmodesmata (Thomas *et al.*, 2008). I wanted to confirm that this localisation was unchanged in *cals* mutants, as this may be the cause of the revertant phenotype. In both mutants, PDLP1-GFP still localised to plasmodesmata thus indicating the absence of a callose synthase does not change localisation of PDLP1 (Figure 4-24).
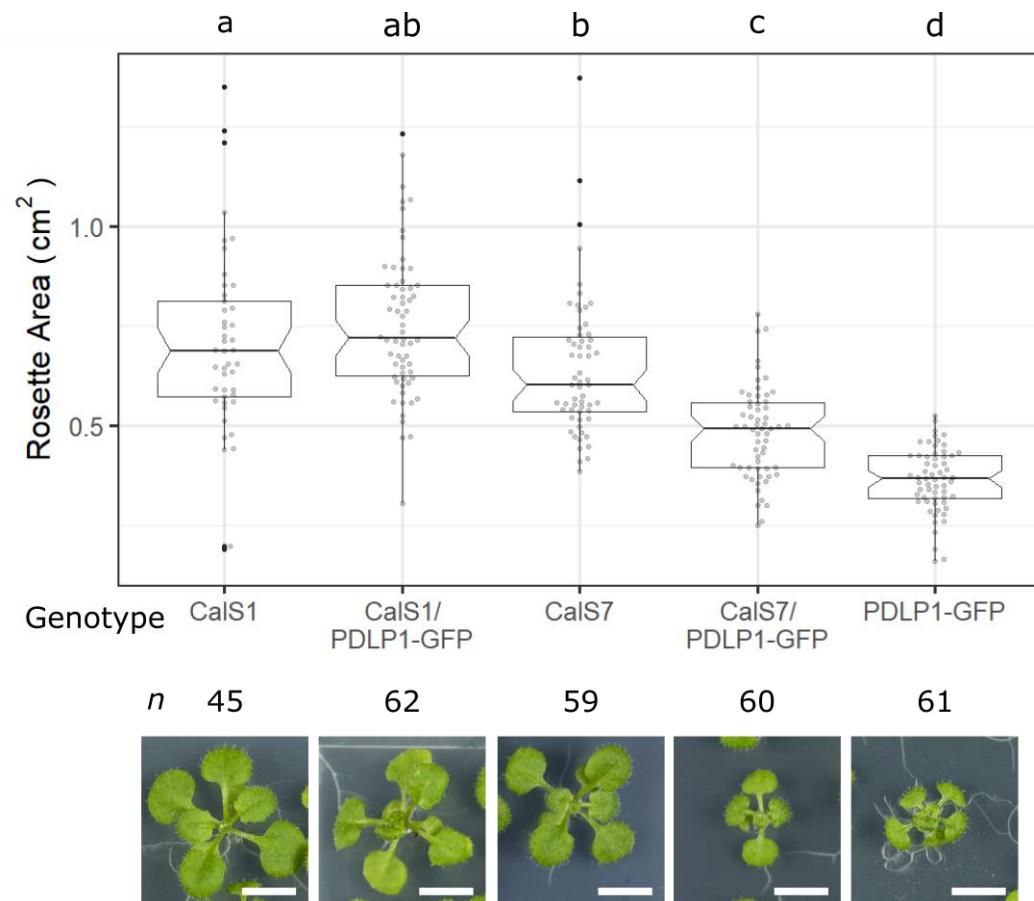
Figure 4-11 CalS1 is required for PDLP1-induced dwarfism.
PDLP1-GFP was crossed into *cals1* and *cals7* mutants. Loss of CalS1 caused full reversion of the PDLP1-induced phenotype, whereas *cals7* caused partial reversion of size. Letters indicate significantly different Tukey groups (p < 0.05). Scale bar = 0.5 cm.

Figure 4-12 PDLP1 maintains a plasmodesmal localisation in *cals* mutants.
PDLP1-GFP expression and plasmodesmal localisation was reconfirmed in the *cals1* and *cals7* backgrounds by confocal microscopy of the GFP transgene. Mutations in CalS1 and CalS7 did not alter the localisation of PDLP1. Scale bar = 10 µm.

4.3.5 PDLP1-induced dwarfism does not require RBOHD

In the same manner as the CalSs, I investigated whether RBOHD was upstream or downstream of PDLP1-GFP by crossing in a *rbohd* mutant into PDLP1-GFP. I did not manage to generate a double homozygous PDLP1-GFP/*rbohd* line. However I obtained multiple F2 homozygous *rbohd* heterozygous PDLP1-GFP lines. If RBOHD was upstream of PDLP1, an expected 3:1 (large:small) ratio would be expected in the F3. If RBOHD is downstream of PDLP1, all plants would be expected to be large.

For three different F3 heterozygous PDLP1-GFP lines, I scored plant size ($n > 50$ for each). In all three cases, there was insufficient evidence to reject the null hypothesis that the plants were segregating in a 3:1 ratio (Table 4-7, Figure 4-13). Therefore, RBOHD acts upstream or independently of PDLP1 as it does not prevent PDLP1-induced dwarfism.

Table 4-7 Chi-squared test for size segregation in rbohd/rbohd PDLP1-GFP/+ plants.
In all cases the degrees of freedom was 1, and the $X^2_{critical} = 3.84$.

Large	Small	X^2 (3:1 large:small ratio)	P
50	11	1.6	>0.20
42	14	0	1
48	16	0	1

Figure 4-13 Representative image of F3 progeny of F2 *rbohd/rbohd* and PDLP1-GFP/+ plants
The F3 generation of selfed F2 homozygous *rbohd* heterozygous PDLP1-GFP plants segregated in a 3:1 (large:small; L:S) ratio (Table 4-7). This demonstrates that the PDLP1-GFP phenotype is independent of RBOHD. Scale bar = 0.5 cm.

4.3.6   Candidate PDLP5 interactors were identified with a split-ubiquitin screen

An alternative approach to finding PDLP interactors was to search for physical interactors, as opposed to a genetic approach. One such method is a split-ubiquitin assay, a form of yeast-2-hybrid for membrane proteins. I analysed the results of a PDLP5 split-ubiquitin screen ran by DualSystems Biotech, Switzerland, in 2010 (Christine Faulkner, *pers. comm.*, now Hybrigenics Services, France).

To find the most likely plasmodesmal interactors, I cross-compared the list of prey hits to the plasmodesmal proteomes (Chapter 3, Table 3-1). There were 65 preys, of which 22 overlapped with the plasmodesmal proteomes (Table 4-8). In addition to these 21, nine further proteins have multiple prey hits and so were considered likely interactors. In total, therefore 31 PDLP5 interactors were identified.

Table 4-8 Cross-comparison of PDLP5 split-ubiquitin yeast-2-hybrid interactors with *A. thaliana* plasmodesmal proteomes.
For AtCells2 ++ indicates the protein was also in AtCells2_filtered. For AtPlant1_loose ++ indicates the protein was also in AtPlant1. Bayer enrichment factor is the plasmodesmata/plasma membrane protein ratio from Brault *et al.* (2019). Proteins are grouped by the number of matching proteomes.

Locus	Split Ubiquitin	Number of Hits	AtCells 1	AtCells 2	Bayer Enrichment Factor	AtPlant1 _loose	Name
AT5G06320	+	1	+	++	48	+	NHL3
AT1G71695	+	1	+	+	16	++	*peroxi-dase*
AT4G23630	+	1	+	+	0	+	RTNLB1
AT1G20330	+	1	+	+	9		SMT2
AT1G44575	+	2		+	8	++	NPQ4
AT1G11260	+	1	+	+	3		STP1
AT3G54140	+	1	+	+	3		NPF8.1
AT4G35100	+	9	+	+	2		PIP3
AT1G53210	+	1	+	+	2		NCL
AT2G45470	+	1		+	1	+	FLA8
AT3G53420	+	1	+	+	1		PIP2A
AT4G30190	+	1	+	+	1		AHA2
AT4G25810	+	1		++	816		XTR6
AT3G55360	+	2		+	21		CER10
AT2G21600	+	1		+	8		RER1B
AT1G02130	+	1		+	5		RA5
AT3G26520	+	8		+	4		TIP2
AT5G66190	+	1		+	4		FNR1
AT1G50430	+	1		+	3		DWF5
AT2G37170	+	3	+				PIP2B
AT1G09560	+	1				++	GLP5
AT2G01470	+	1	+				STL2P
AT4G30950	+	1				+	FAD6
AT1G44920	+	4					*transmembrane*
AT3G12610	+	3					DRT100
AT5G46110	+	3					APE2
AT2G20920	+	2					*chaper-one*
AT3G15850	+	2					FAD5
AT3G16240	+	2					DELTA-TIP
AT3G21690	+	2					MATE efflux
AT5G44920	+	2					TIK

4.3.7 NHL3 was present at plasmodesmata

While the split-ubiquitin screen identified likely PDLP5 interactors, I wanted to find plasmodesmata-specific interactors. Therefore, I cloned the two proteins with present in every proteome: NDR1/HIN1-LIKE3 (NHL3, AT5G06320) and AT1G71695 (peroxidase, cloned by Adrien Gregorio under my supervision) and tagged them with a fluorescent protein to determine their localisation when transiently expressed in *Nicotiana benthamiana*. I did not clone RETICULAN LIKE PROTEIN B1 (RTNLB1, AT4G23630), as its close family members RTNLB3 and RTNLB6 have been localised to the desmotubule (Knox *et al.*, 2015). RTNLB1 showed a similar localisation to RTNLB3/6, when observed at lower magnification and so is unlikely to interact with PDLP5 in the plasma membrane. NHL3 co-localised with PDLP1 at plasmodesmata and the plasma membrane, whereas AT1G71695 (peroxidase) was present in the vacuole and the cell wall (Figure 4-14).
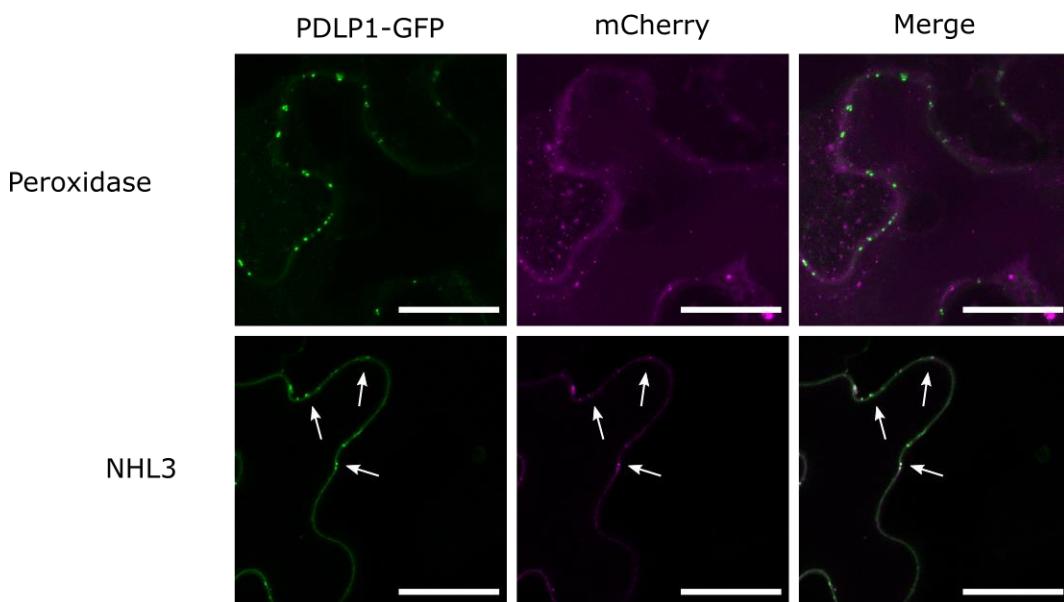
Figure 4-14 NHL3 co-localises with PDLP1-GFP.
PDLP1-GFP was co-expressed in *N. benthamiana* for two days with NHL3-mCherry or AT1G71695-mCherry. AT1G71695 localised to the cell wall and vacuole. NHL3 localised to the plasma membrane and plasmodesmata (arrows). Scale bar = 25 µm.

To confirm the plasmodesmata localisation of NHL3, I quantified the co-occurrence of NHL3-mCh punctae with that of a plasmodesmal marker, PDLP5-GFP (Figure 4-15). NHL3 punctae were plasmodesmal 76.6% (± 4.29% SE, *n* = 8 micrographs) of the time, confirming that NHL3 is a *bona fide* plasmodesmal protein.
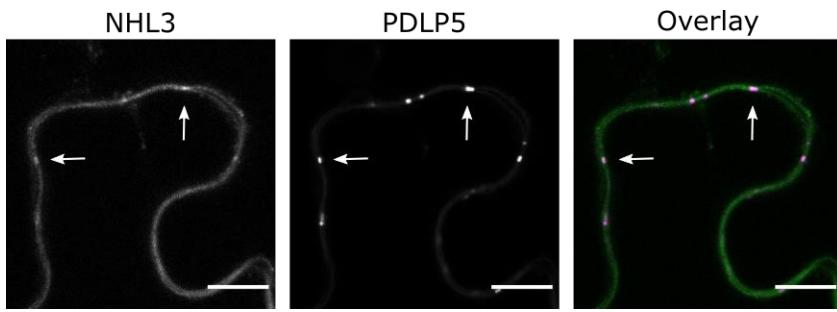
Figure 4-15 NHL3 co-localises with PDLP5.
PDLP5-GFP was co-expressed in *N. benthamiana* for two days with NHL3-mCherry. The overlay has NHL3 in green and PDLP5 in magenta. NHL3 and PDLP5 co-localised at plasmodesmata (arrows). Scale bar = 10 μm.

4.3.8   NHL3 interacted with PDLP1

As NHL3 and PDLPs interacts as suggested by its detection in the split ubiquitin screen described above, I wanted to explore if this interaction occurs *in planta*. Additionally to the DualSystems Biotech data, NHL3 has been shown to interact with PDLP8 in separate high-throughput split-ubiquitin yeast-2-hybird screen (Jones *et al.*, 2014b). Caillaud *et al.* (2014) conducted an IP MS/MS experiment looking for PDLP1-GFP interactors, but did not identify NHL3 as an interactor. However, on reanalysis of their data I found there was a 2-fold enrichment in NHL3 peptides in the PDLP1 sample compared to control. This was below the authors' threshold of 4-fold enrichment, yet can still be considered weak evidence for a PDLP1/NHL3 interaction (Caillaud *et al.*, 2014) (Table S3). Therefore, there is evidence that NHL3 interacts with three of the eight PDLP proteins.

I decided to follow up on the interaction of NHL3 with PDLP proteins using PDLP1, as the PDLP1-GFP overexpression phenotype is more clear cut and is congruent with the other results presented here (Figure 4-4). First, I wanted to confirm NHL3 interacts with PDLP1. I used FRET-FLIM microscopy, using transient expression *N. benthamiana* (Figure 4-16). I used a transient system after attempting to make stable lines with a construct co-expressing PDLP5-GFP and NHL3-mCh, as all transformants had only GFP expression (*n* = 4), despite both fluorescing when transiently expressed.
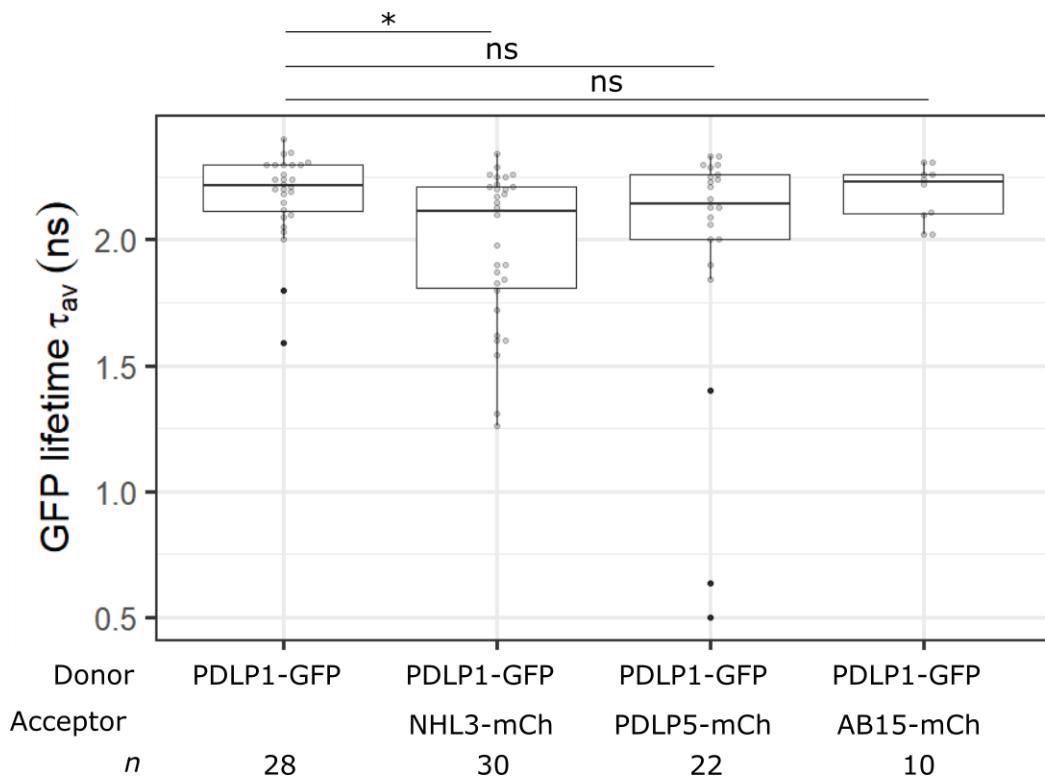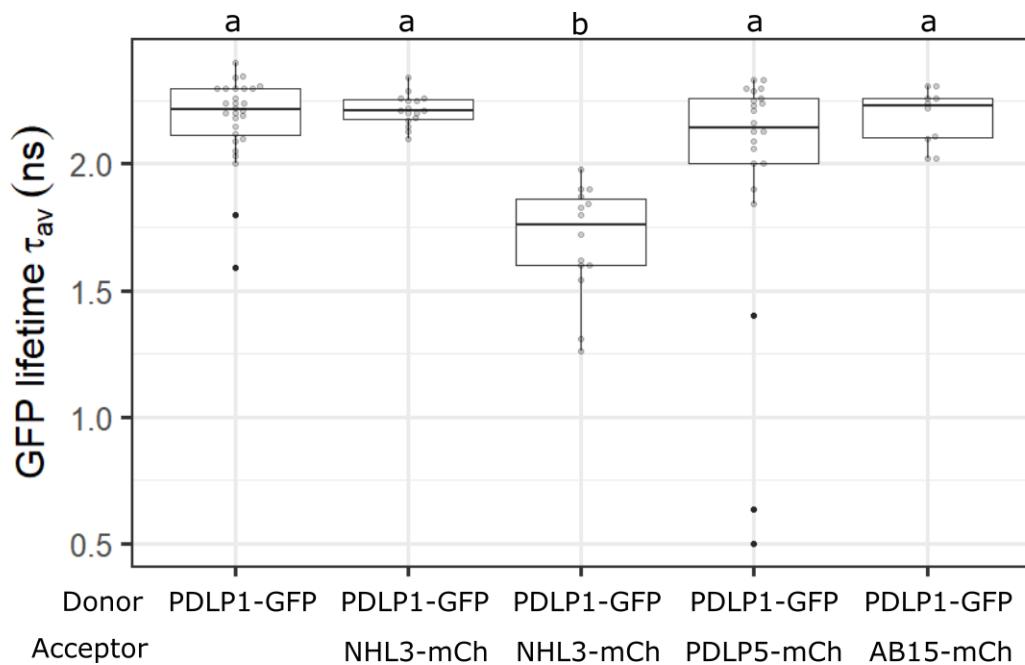
Figure 4-16 NHL3 interacts with PDLP1.
FRET-FLIM analysis of PDLP1-GFP in the presence of acceptors NHL3-mCh, PDLP5-mCh, or AB15-mCh. Fluorescence lifetime was measured in *N. benthamiana* tissue transiently co-expressing the indicated constructs as donors or acceptors. The y-axis represents GFP fluorescence-weighted average lifetime (τ_{Av}, ns). Data were analysed by ANOVA with a multivariate *t* correction for multiple comparisons. * = p < 0.05, ns = p > 0.05.

There was a significant decrease in GFP lifetime from donor alone (PDLP1-GFP), than when co-expressed with NHL3-mCh. However, the data for NHL3-mCh is clearly bimodal, with a cluster of data points where there is no interaction (lifetime is similar to PDLP1-GFP) and a cluster with strong interaction, and so I split the dataset at 5% FRET efficiency (~2 ns) (Figure 4-17).

Figure 4-17 NHL3 has two pools of data: with and without PDLP1 interaction
FRET-FLIM analysis of PDLP1-GFP in the presence of acceptors NHL3-mCh, PDLP5-mCh, or AB15-mCh. Fluorescence lifetime was measured in *N. benthamiana* tissue transiently co-expressing the indicated constructs as donors or acceptors. The y-axis represents GFP fluorescence-weighted average lifetime (τ_{Av}, ns). NHL3 datapoints were split at 5% FRET efficiency (2.07 ns). Data were analysed by ANOVA with a post hoc Tukey test. Letters indicate significantly different Tukey groups (p < 0.05).

In this case, I observed a strong reduction in fluorescence lifetime from 2.18 ns (± 0.032 SEM) in PDLP1-GFP to 1.70 (± 0.059 SEM) in interacting NHL3-mCh, representing a FRET efficiency of 22%. This is a strong FRET efficiency, about half that of the theoretical maximum of 40-60% (Bajar *et al.*, 2016), which indicates the proteins are within 2 nm (1.9 nm) of each other (Equation 4-1).

$$Distance\ (nm) = r - 2.4 \times 2 = \sqrt[6]{5.4^6 \times \left(\frac{1}{E} - 1\right)} - 2.4 \times 2$$

Equation 4-1 Approximation of distance between fluorescent proteins.
Derived from (Lam *et al.*, 2012; Bajar *et al.*, 2016). The distance between the chromophore and the edge of the β-barrel (2.4 nm) is removed from *r* (the distance between donor and acceptor dipoles) to approximate the distance between the two proteins. E = FRET efficiency.

Plasmodesmata are exceedingly small structures with an diameter between 20 - 50 nm (Nicolas *et al.*, 2017). Thus, it may be possible that any plasmodesmal protein would be close enough to PDLP1 for FRET to occur when both are overexpressed. However, neither PDLP5 nor AB15 (a truncated form of AT5G43020) reduced the lifetime of PDLP1-GFP. Thus, it is very likely NHL3 and PDLP1 directly interact.

Interestingly, while GFP lifetime is not effected by concentration or viscosity, the physical environment does impact lifetime, specifically the refractive index (Suhling *et al.*, 2002). The

reduction of lifetime of GFP from 2.7 ns in buffer (Suhling *et al.*, 2002) to 2.18 ns in planta for PDLP1-GFP, suggests a theoretical refractive index of 1.48 (Suhling *et al.*, 2002). This is in agreement with the measured refractive index of cell walls 1.48 (Woolley, 1975) and membranes (1.46 – 1.54) (Meyer, 1979), thus presenting a possible explanation for the low donor only lifetime of GFP in this experiment.

The constructs were transiently expressed in *N. benthamiana*, using a co-infiltration of two clones of *A. tumefaciens*. Thus, there was variability between data points of the ratio of NHL3-mCh to PDLP1-GFP (Figure 4-18). In this case, this did not confound the experiment, as there was no relationship between fluorophore ratio and GFP lifetime.
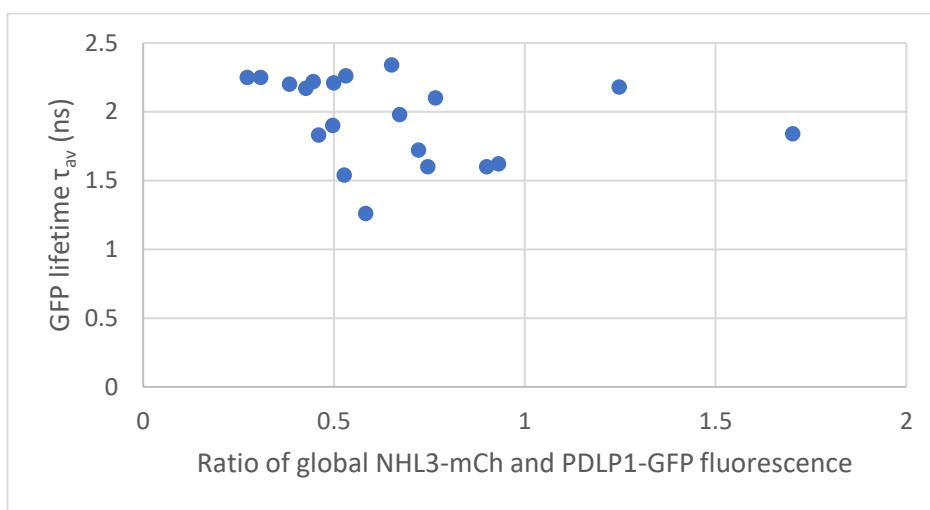
Figure 4-18 GFP lifetime was not affected by the NHL/PDLP1 protein ratio.
The ratio of NHL3-mCh to PDLP1-GFP was generated in Fiji by dividing the total fluorescence in each image. This ratio was plotted against the GFP fluorescence-weighted average lifetime of the GFP in the image to access any relationship between the two variables. There was no correlation between fluorophore ratio and GFP lifetime. Pearson's R^2 = 0.07.

4.3.9 *nhl3* mutants have reduced NHL3 expression

NHL3 has been characterised in the literature as being essential for plant defence (Humphry *et al.*, 2010; Singh *et al.*, 2018), however the mutants used were not fully characterised. The T-DNA insertions were not ideal SALK mutants, as the insertions were towards the 3' end of the gene (i.e. most of the transcript was transcribed). However, in 78% of cases these 3' coding sequence insertions still either prevented or reduced transcription (Wang, 2008). Moreover, these mutants have been shown to have a phenotype indicating the SALK insertion is disrupting the gene.

Nevertheless, I characterised their basal levels of *NHL3* transcript to confirm the effect of the SALK insertion. To do this, I extracted the RNA of 5-week-old soil grown rosette leaves and quantified *NHL3* transcript levels with RT-qPCR (Figure 4-19). I specifically targeted the

primers to be downstream of the T-DNA insertion sites, so as to observe the effect on the transcript stability (Figure 4-3). Both *nhl3-1* (SALK_035427C (Humphry *et al.*, 2010)) and *nhl3-2* (SALK_150318C (Humphry *et al.*, 2010; Singh *et al.*, 2018)) had significantly reduced levels of *NHL3* transcript ($p < 0.001$, Tukey post-hoc test).
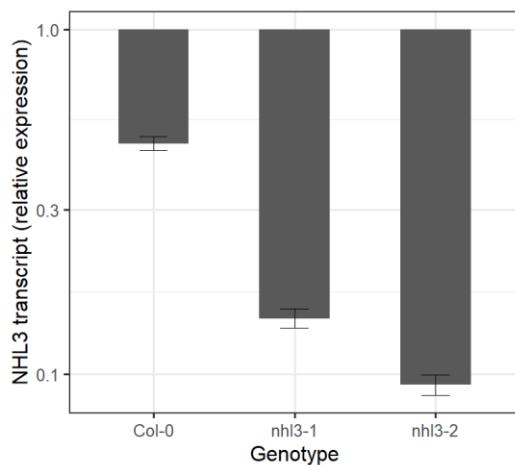
Figure 4-19 NHL3 expression is significantly reduced in *nhl3* SALK lines.
Transcript stability of *NHL3* transcripts upstream of the SALK insertion site (see Figure 4-3) was assessed in control conditions of 5-week-old rosette leaves of soil grown *A. thaliana*. There was significantly less transcript in both *nhl3-1* (SALK_035427C) and *nhl3-2* (SALK_150318C) ($p < 0.001$). Expression is relative to the geometric mean of two housekeeper genes: UBQ10 and UBC. *n* = 3

4.3.10 PDLP1 induced dwarfing was abolished in NHL3 knockdown plants

Taking a reverse genetic approach as before, I aimed to demonstrate a genetic as well as physical interaction for NHL3 and PDLP1. If NHL3 acts downstream of PDLP1, the *nhl3* mutant should revert the PDLP1-GFP phenotype. Thus, I crossed *nhl3-1* with PDLP1-GFP. As before PDLP1-GFP plants were severely dwarfed (55% reduction compared to Col-0) while knocking down *NHL3* expression had no effect on size in *nhl3-1* plants (Figure 4-20). When PDLP1-GFP was introduced, by crossing, into the *nhl3-1* background, there was no reduction in size (Figure 4-20). This epistasis reveals that NHL3 is downstream of the action of PDLP1.

Figure 4-20 NHL3 is required for PDLP1-induced dwarfism.
PDLP1-GFP was crossed into the *nhl3-1* mutant. The *nhl3-1* mutation had no effect on size (rosette area) in the wild type background. *nhl3-1* plants were also wild type size in the PDLP1-GFP background, reverting the dwarfism phenotype of PDLP1-GFP. This demonstrates PDLP1 is upstream of NHL3. Letters indicate significantly different Tukey groups ($p < 0.001$). Scale bar = 0.5 cm

Overexpression of PDLP1 may no longer be producing dwarfism because NHL3 is required for its plasmodesmal localisation. To test this hypothesis, I imaged the leaves of *nhl3-1/*PDLP1-GFP plants (Figure 4-21). PDLP1 localises to plasmodesmata irrespective of the presence of NHL3. Thus, NHL3 is not required for the plasmodesmal localisation of PDLP1.
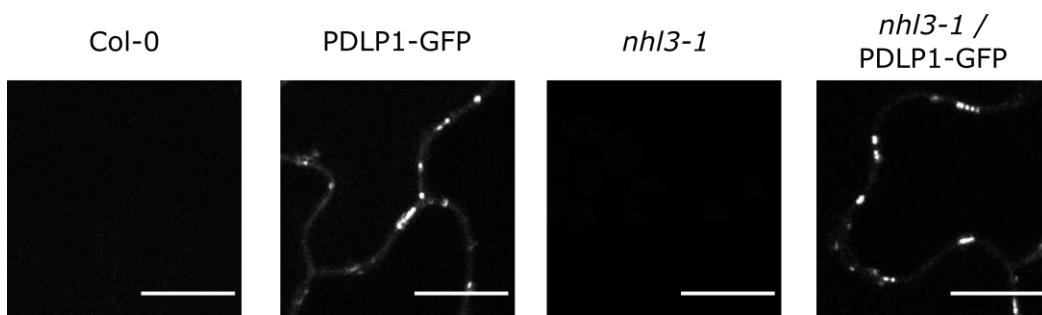
Figure 4-21 NHL3 does not affect PDLP1 localisation.
PDLP1-GFP expression and plasmodesmal localisation was reconfirmed in *nhl3-1* by confocal microscopy of the GFP transgene. Mutation of *nhl3* did not alter the localisation of PDLP1. Scale bar = 10 μm.

4.3.11 NHL3 mutants were defective in callose deposition in response to flg22

PDLP proteins are known to regulate callose deposition at haustoria and plasmodesmata (Thomas *et al.*, 2008; Lee *et al.*, 2011b; Caillaud *et al.*, 2014). Given that NHL3 is a plasmodesmal protein that physically and genetically interacts with PDLP proteins, I wanted to see if plasmodesmal responses were altered in *nhl3* mutant plants.

As flg22 is known to induce callose deposition, and *NHL3* is flg22 responsive (Dörmann *et al.*, 2000; Varet *et al.*, 2003), I considered the effect of *nhl3* on flg22-induced callose deposition. I quantified plasmodesmata-associated callose stained with aniline blue in the rosette leaves of 4 – 5 week old *A. thaliana* plants infiltrated with flg22 (100 nM) (Figure 4-22).
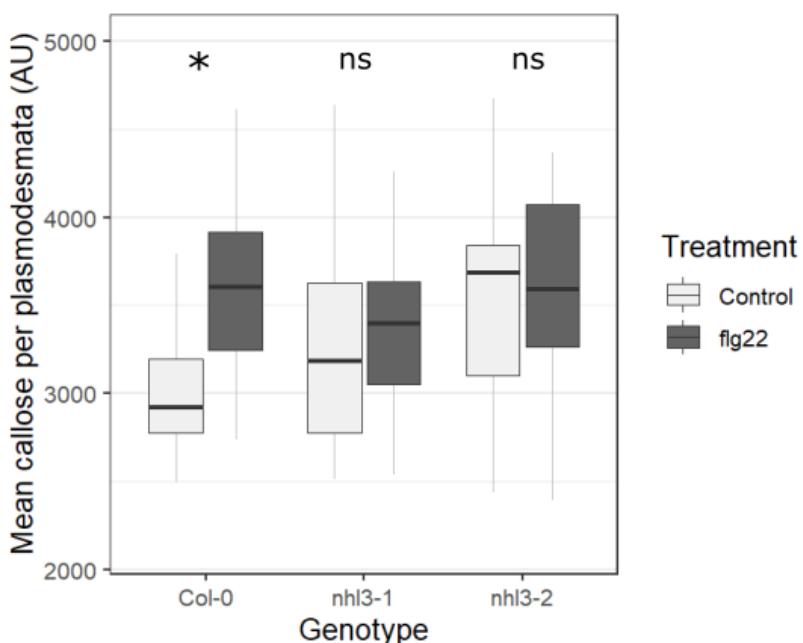
Figure 4-22 *nhl3* mutants cannot properly regulate callose deposition.
Automated image quantification of plasmodesmata-associated callose via confocal microscopy of aniline blue fluorescence. Rosette leaves of 4 – 5 week old *A. thaliana* plants were either untreated (control) or infiltrated with flg22 (100 nM) for 30 minutes. There was a significant increase in mean integrated density of aniline blue fluorescent spots in Col-0, which was not observed in *nhl3-1* or *nhl3-2* due to increased variability. Combined results from 5 control and 3 flg22 experiments, with the mean integrated density of plasmodesmata in each averaged of each leaf from at least 2 images per leaf and three leaves per experiment. Pairwise contrasts of a linear mixed model with a Tukey adjustment for multiple comparisons (ns = $p > 0.05$, * = $p < 0.05$).

As shown before in the literature flg22 treatment increased the integrated density of aniline blue fluorescent spots within one hour, indicating increased quantity of callose at each plasmodesmata (Xu *et al.*, 2017). However, this was not replicated in either *nhl3* mutant, which showed much greater variability in callose levels (Figure 4-22). Thus, callose deposition is mis-regulated in *nhl3* mutants.

4.3.12 NHL3 is required for plasmodesmata closure in response to flg22

Having ascertained that *nhl3* plants have aberrant callose phenotypes in response to flg22, I wanted to test whether this translated into an inability to reduce cell-to-cell flux in response to flg22 as well. Thus, I infiltrated 4-5 week old *A. thaliana* plants with water (Mock) or flg22 (100 nM) after bombarding them with GFP-coated gold particles (Faulkner *et al.*, 2013). As expected, there was a reduction in the mean cell-to-cell spread of GFP in Col-0 (Xu *et al.*, 2017). The two independent T-DNA *nhl3* lines showed differing responses, where *nhl3-2* behaved the same as Col-0. Whereas, *nhl3-1* could not close its plasmodesmata in response to flg22 (Figure 4-23).

This somewhat contradicts the callose results finding in Figure 2-21. The callose data indicates that *nhl3-1* and *nhl3-2* should have similar patterns of cell-to-cell flux, as the deposition of callose is responsible for the restriction of cell-to-cell flux (Faulkner *et al.*, 2013). However, the callose data shows greater variability of callose deposition in *nhl3-1* and *nhl3-2* than Col-0, so had I had more time I would have included more bombardment replicates to overcome this variation. This is borne out in the bombardment data where *nhl3-2* has an increased inter-whisker range of flg22-treated GFP spread than Col-0 (10 cells opposed to 7 cells (Figure 4-23)).
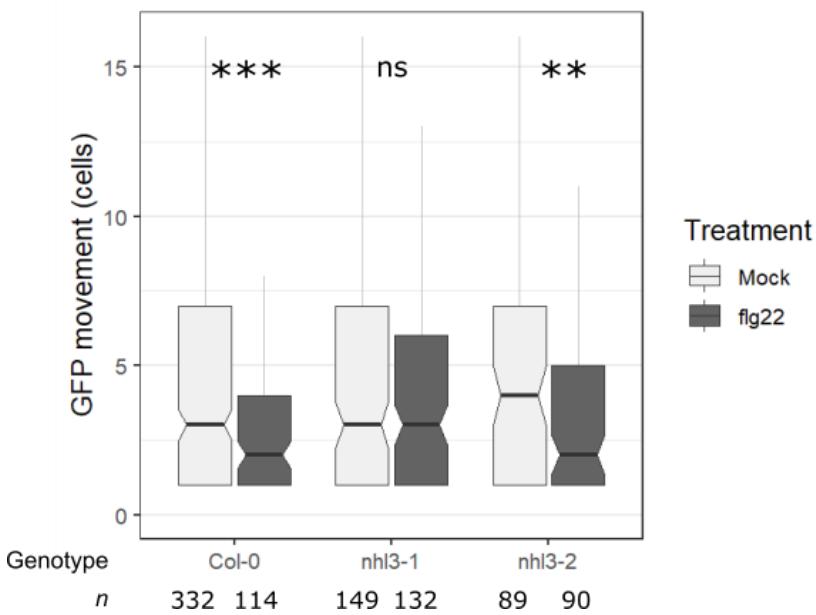
Figure 4-23 *nhl3-1* plants cannot close their plasmodesmata in response to flg22.
Fully-expanded adult *A. thaliana* leaves were bombarded with gold particles coated with a plasmid for cytosolic GFP expression. Four hours after bombardment leaves were infiltrated with water (Mock) or flg22 (100 nM). The number of cells which expressed GFP at each bombard site was counted 16 hours after bombardment. The mean cell-to-cell flux was significantly reduced by flg22 treatment in Col-0 and *nhl3-2*, but not in *nhl3-1*. The data are from three independent experiments. Difference were tested by a mean bootstrap method (ns = p > 0.05, ** = p < 0.01, *** = p < 0.001).

4.3.13 Plasmodesmal co-IP MS to find novel PDLP interactors

As with the split-ubiquitin screen, I tried another biophysical interactor approach to discover what PDLP1 and PDLP5 may be interacting with at plasmodesmata. I also included analysis of the *Tomato yellow leaf curl virus protein* C4 in this approach (Rosas-Diaz *et al.*, 2018). I wanted to find plasmodesmal specific interactors of these proteins, and so sought to extend the plasmodesmal extraction technique developed in Chapter 2 to include co-immunoprecipitation.

I attempted to co-immunoprecipitate proteins from a plasmodesmata extract. I generated plasmodesmata extracts from *N. benthamiana* tissue transformed with P19 alone, and P19 with PDLP1-HA, PDLP5-HA and C4-HA. I chose the HA tag as I found it the most stable through plasmodesmata extractions (see 2.3.3). I then used magnetic α-HA beads to specifically purify the tagged proteins. For all protein samples, bar the P19 negative control, a band was seen in the IP lane of the Western (Figure 4-24). In the case of C4 (Figure 4-24b), IP dramatically increased signal intensity from imperceptible to a clear band. Whereas for PDLP1 and PDLP5, the IP kept similar levels of protein to the input. Thus, proteins can be specifically enriched by immunoprecipitate within plasmodesmata extracts.
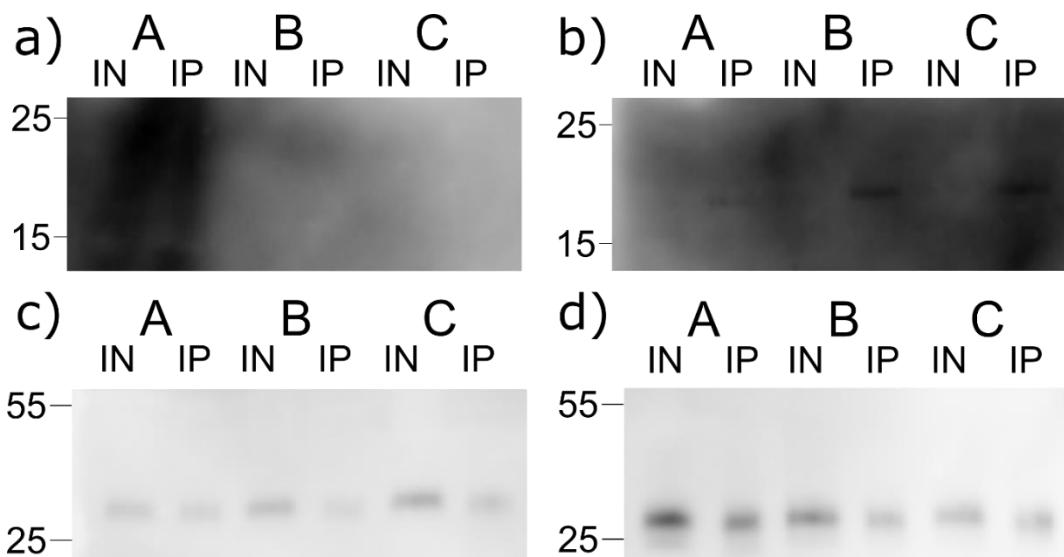
Figure 4-24 Plasmodesmal coimmunoprecipitations in N. benthamiana.
Proteins were transiently expressed for two days in *N. benthamiana*: a) P19, b) C4-HA and P19, c) PDLP1-HA and P19, and d) PDLP5-HA and P19. Plasmodesmata were extracted from four leaves and used as the input (IN) for the immunoprecipitation (IP). Each Western blot shows three independent biological replicates (A, B and C) from one experiment. Blots were probed with an α-HA-HRP antibody HRP antibody (abcam, ab173826). The location of size markers are noted in kDa.

The immunoprecipitate samples were run on a Waters Synapt G2-Si mass spectrometer to identify proteins that co-immunoprecipitated with the tagged proteins. 50 proteins were detected overall across all four samples, 30 of which were in the P19 control. Thus, there

were 20 proteins which may be interacting with C4, PDLP1 and/or PDLP5 (Table 4-9). A majority of the proteins exclusively precipitated with C4 (15/20, 75%). However, these were predominantly chloroplastic proteins, which is likely because C4 has a dual chloroplastic localisation, and so are likely off-target hits (Rosas-Diaz *et al.*, 2018). There were no PDLP5 specific interactors and three PDLP1 specific proteins: a geranylgeranyl reductase and two catalases. All three of these proteins are unlikely candidates for PDLP1 interactors as geranylgeranyl reductases are involved with chlorophyll biosynthesis (Meadows *et al.*, 2018) and both catalases convert hydrogen peroxide into water, as is common in respiring organisms. The most likely hit to be plasmodesmal was a pectinesterase that was found in all three samples (UniProt: Q84V57, closest *A. thaliana* homologue: PECTIN METHYLESTERASE 3; AT3G14310). However, Q84V57 is annotated to be in the cell wall by UniProt. It is also suspicious that the protein was found in all three samples, as there is no reason to expect common interactors between PDLPs and C4. Q84V57 may be a contaminant which was missed in the P19 control.

Table 4-9 *N. benthamiana* **proteins which co-immunoprecipitated with (C4/PDLP1/PDLP5)-HA and not in a control pulldown.**

Protein	C4	PDLP1	PDLP5	Description
A0A0A8K9V3_NICBE		+		Geranylgeranyl reductase
A0A0F7R532_NICBE	+			S-adenosylmethionine synthase
A0A248QJL3_NICBE	+			S-adenosylmethionine synthase
A0A286RNF7_NICBE	+			Carbonic anhydrase
A4D0J9_NICBE	+			Carbonic anhydrase
A4D0K0_NICBE	+			Carbonic anhydrase
A7L4B4_NICBE	+			Histone H3
B8R518_NICBE	+			Ubiquitin
C6FFS2_NICBE		+		Catalase
C9DFB9_NICBE	+			Heat shock protein 70-like protein
I0B7J5_NICBE	+			Chloroplast PsbP1
I0B7J6_NICBE	+			Chloroplast PsbP2
K0IBB4_NICBE		+		Catalase
Q2LAH0_NICBE	+			Chloroplast photosystem II 22 kDa component
Q2LAH1_NICBE	+			Chloroplast photosystem II 22 kDa component
Q58H59_NICBE	+			Chloroplast photosynthetic oxygen-evolving protein 23 kDa subunit
Q5EEQ1_NICBE	+		+	Photosystem I reaction center subunit X
Q5I6U1_NICBE	+			Ubiquitin/s27a 40s ribosomal protein
Q84V57_NICBE	+	+	+	Pectinesterase
R9W4N2_NICBE	+			2-Cys peroxiredoxin

167

Concerningly, neither C4, PDLP1, nor PDLP5 was recovered in the mass spectrometry, despite being observed on the Western blot (Figure 4-24). This could be due to the reduced sensitivity of the Synapt machine. Therefore, I sent new PDLP1 and PDLP5 samples to Cambridge Centre for Proteomics to be run on an Orbitrap spectrometer. Cambridge Centre for Proteomics follow a different downstream protocol to JIC Proteomics, and so the first samples sent could not be analysed due to PEG-like contamination (Deery, M., *pers. comms.*). The protocol does not contain any PEG directly, but many detergents, such as Triton-X 100 (which is used), contain a PEG-repeat. Thus, the protocol was revised with additional washing steps to remove excess detergent.

Novel samples were sent to Cambridge. In all six samples, a clear band can be seen in the IP lane, as before (Figure 4-25). This time, mass spectrometry with the Orbitrap mass spectrometer identified PDLP1-HA and PDLP5-HA in the samples, along with nine other proteins. Unfortunately, all nine proteins were also detected in the prior P19 control (Figure 4-24a). Thus, while this confirms the immunoprecipitate was working, as detected by Western blot and tandem mass spectrometry, no PDLP interactors were found to co-immunoprecipitate.
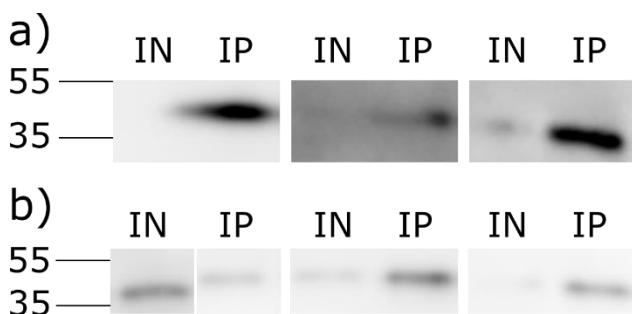
Figure 4-25 Plasmodesmal coimmunoprecipitations samples sent to Cambridge Centre for Proteomics.
Proteins were transiently expressed for two days in *N. benthamiana*: a) PDLP1-HA and P19, and b) PDLP5-HA and P19. Plasmodesmata were extracted from four leaves and used as the input (IN) for the immunoprecipitation (IP). Each gel shows three independent biological replicates from one experiment. Blots were probed with an α-HA-HRP antibody (abcam, ab173826). The location of size markers are noted in kDa.

4.4 Discussion

4.4.1 Adding new components to the PDLP-mediated plasmodesmal closure pathway

Callose has been known to be at plasmodesmata for over 20 years (Turner *et al.*, 1994; Roy *et al.*, 1997), yet the pathway controlling its deposition remains unknown. PDLPs have previously been implicated in this pathway specifically (PDLP5, Wang *et al.*, 2013), as well as localised to other callose deposition areas (PDLP1 at haustoria, Caillaud *et al.*, 2014). Intuitively, the callose deposition must be linked to PDLPs via a CalS.

Here, the PDLP1 overexpression-induced phenotype was leveraged to find novel components in the callose deposition pathway. Mutants that revert the dwarfed phenotype likely effect the action of PDLP1. However, it is possible that the mutants may increase plant size and so counteract the PDLP1 phenotype. In this case, the plants would be expected to be larger than wild type without the PDLP1 transgene. This has not been ascertained for the POEM mutants, as the mutation is background specific. This would first require full characterisation and validation of the causative mutation in either the PDLP1-GFP parent background or complementation of the mutation in the POEM background.

I also exploited the dwarfed phenotype of PDLP1-GFP in reverse genetic approaches. Here, I identified three genes that genetically interact with PDLP1-GFP. Mutations in each of these three genes did not increase plant growth in their own right, confirming that their role lie the PDLP pathway, rather than growth in general (Figure 4-11, Figure 4-20).

There are several plausible ways the action of PDLP1 could be altered: removing a downstream signalling component, removing a required partner for PDLP function, or preventing PDLP localisation to plasmodesmata. For all mutants examined (including the POEM mutations), PDLP1 still localised at plasmodesmata (Figure 4-6, Figure 4-12, Figure 4-21). Thus, the mutants here must act with or downstream of PDLP1 to signal constitutive increased callose.

4.4.2 NDR1/HIN1-like proteins

4.4.2.1 Overview of NHL protein family

NDR1/HIN1-LIKE (NHL) proteins were defined by homology to *Nicotiana tabacum* HARPIN-INUDCED1 (HIN1) and *A. thaliana* NON-RACE-SPECIFIC DISEASE RESISTANCE1 (NDR1) (Dörmann *et al.*, 2000). Both genes have been implicated in plant defence (Gopalan *et al.*, 1996; Century *et al.*, 1997). The NHLs are a large 45 member family in *A. thaliana* (Bao *et al.*,

2016a). NHLs have been implicated in diverse functions from pathogen responses to seed germination (Varet *et al.*, 2003; Zheng *et al.*, 2004; Bao *et al.*, 2016b,a).

NHL proteins are a subset of LATE EMBRYOGENESIS ABUNDANT (LEA) proteins. They were first discovered in cotton (*Gossypium hirsutum*) seeds, and are thought to provide defence against dehydration at the late embryogenesis stage (Galau *et al.*, 1986). NDR1 has homology to LEA14 (At1G01470) and mammalian integrins. Integrins play a dual function of providing and signalling cell-cell adhesion (Hynes, 2002). Overexpression of LEA14 enhances *A. thaliana* salt tolerance (Jia *et al.*, 2015). These domains may be the reason why NDR1 has a role in regulating fluid loss in response to pathogens (Knepper *et al.*, 2011). Both NHL3 and NHL26 have a LEA2 domain (confusingly from AtLEA14) (Jones *et al.*, 2014a). The LEA2 domain also has no known function, though it is noted to be unusually hydrophobic compared to other LEA domains (Hundertmark & Hincha, 2008). LEA domains are often thought to act similarly to chaperonins, and protect other proteins through a disordered domain. However, LEA2 is an unusual LEA domain and is folded, so is less likely to be acting as a chaperonin (Singh *et al.*, 2005).

Several *A. thaliana* NHL proteins have been characterised. For example, *nhl6* mutants are insensitive to ABA (Bao *et al.*, 2016a). Only one NHL, NHL26, has been previously localised to plasmodesmata (Vilaine *et al.*, 2013). NHL26 was localised to the plasmodesmata in the phloem and was found to either have an effect on plasmodesmal permeability or sugar signalling within the vasculature when overexpressed (Vilaine *et al.*, 2013). The *nhl26* mutant line used by Vilaine *et al.* (2013), which produced a truncated form of NHL26, had no observable phenotype. Overexpression of NHL26 gave a similar phenotype to *cals7* mutants: an increase in leaf sugar concentrations (Xie *et al.*, 2011; Vilaine *et al.*, 2013). *cals7* mutants have reduced phloem callose and, consequently, have poorly formed and blocked sieve plates, as callose is required to form the sieve pores (Xie *et al.*, 2011). So, counter-intuitively, NHL26 may be increasing callose deposition to phenocopy the *cals7* mutant, and so block the sieve pores with callose.

In this chapter, I identified NHL3 as a protein of interest. NHL3 has been studied previously, where it was found to be significantly transcriptionally upregulated in response to viruses, fungal and bacterial pathogens (especially in non-compatible interactions) (Dörmann *et al.*, 2000; Zheng *et al.*, 2004; Wang *et al.*, 2018). NHL3 was phenotypically confirmed to be involved in plant defence with knockout mutants being more susceptible to bacteria, and overexpressors more resistant (Varet *et al.*, 2003; Singh *et al.*, 2018). This is the same

phenotype of mutants which cannot close their plasmodesmata, e.g. *cml41* or *pdlp5*. Thus, it adds support to NHL3 acting downstream of PDLP proteins in the defence-induced plasmodesmata callose pathway.

4.4.2.2   NHL3 localises to plasmodesmata and the plasma membrane

Prior work on NHL3 localised the protein solely to the plasma membrane (Varet *et al.*, 2003; Caesar *et al.*, 2011). Varet *et al.* (2003) used a combination of methods to conclude that NHL3 was at the plasma membrane, using a NHL3-HA expressing line coupled with *in situ* hybridisation and immunogold labelling. In addition, they took a biochemical approach and used two-phase separation to separate the plasma membrane from internal membranes. Caesar *et al.* (2011) transiently expressed NHL3-RFP as a plasma membrane marker in *N. benthamiana*. However, the images presented are to show that AHK3 (*ARABIDOPSIS HISTIDINE KINASE3*) is in the endoplasmic reticulum, and so the Z-frame presented does not allow for the observation of plasmodesmata.

I found similar results with NHL3-mCherry localising to the plasma membrane (Figure 4-14, Figure 4-15). However, I also observed enrichment at punctae that co-localised with PDLP proteins suggesting it accumulates at plasmodesmata. This is congruent with biochemical data that shows NHL3 is present in plasmodesmata fractions (Table S3, (Fernandez-Calvino *et al.*, 2011; Brault *et al.*, 2019)). These data coupled with an interaction with PDLP5 identified by split-ubiquitin and PDLP1 by FLIM FRET, indicated the true localisation of NHL3 is plasma membrane and plasmodesmata.

This localisation is similar to that of other known plasmodesmal proteins, for example LYM2 (Faulkner *et al.*, 2013). LYM2 was recently shown to increase its plasmodesmata:plasma membrane ratio in response to its ligand (Cheval *et al.*, 2020). Other plasma membrane proteins have been shown to relocate to plasmodesmata upon osmotic stress (Grison *et al.*, 2019; Hunter *et al.*, 2019). Thus, as NHL3 is required for flg22 plasmodesmal responses (Figure 4-22, Figure 4-23), it would be worthwhile to investigate whether NHL3 also localises more strongly to plasmodesmata upon flg22 treatment.

4.4.2.3   NHL3 dynamically interacts with PDLP1

NHL3 has been shown here to directly interact with PDLPs through two independent techniques: an *in vitro* split-ubiquitin system (Table 4-8) and *in vivo* transient FRET-FLIM method (Figure 4-16). FRET-FLIM allows spatial-temporal imaging of *in vivo* interactions. Intriguingly, NHL3 showed an extremely strong interaction in 14/30 (47%) micrographs when

averaged over an entire image, and there was no interaction in the other 16 (53%). This suggests that NHL3 can be in two states of close association (1.9 nm apart) or dissociated.

These experiments were done after water infiltration of the *N. benthamiana* leaf tissue to keep the refractive index constant with the water lens. NHL3 is known to be transcriptionally responsive to flg22 and required for callose deposition at plasmodesmata in response to flg22 (Figure 4-22). It would be exciting to follow this up with FRET-FLIM measurements of flg22 treated leaves. I would expect an increased proportion of associated PDLP1-NHL3, coupling the flg22 treatment with signalling. Further, the temporal aspect of FRET-FLIM could be used to visualise the interaction forming in real time.

FRET-FLIM measures the decrease in fluorescence lifetime of a donor fluorophore due to Förster resonance energy transfer. This can be used to give a concentration-independent FRET efficiency, unlike a direct FRET measurement. However, a FRET-FLIM measurement is still dependent on the ratio between fluorophores. Thus, the ideal experimental set-up is to use a stable line, so that concentration and ratio are constant. Next best is transiently expressing a dual expression construct, so that the ratio in all cell is the same as each other (but concentrations vary). Lastly, infiltrating a mix of agrobacterium, meaning that both concentration and ratio vary between cells. Thus, as in my case, when using a transient system it is important to demonstrate that the ratio of fluorophores were not contributing to the effect seen (Figure 4-18).

4.4.2.4 The method of NHL3 function remains unknown

Just like PDLP1, NHL3 has a single transmembrane domain and a smaller (predicted) cytoplasmic domain than extracellular domain (Figure 4-3). Also similarly to PDLP1, the method of NHL-signalling is completely unknown. NHL3 does not contain any known "active" domains, such as a kinase domain. The only annotated domain of NHL3 is a LEA2 domain. The enigmatic LEA2 domain also has no known function, though it is noted to be unusually hydrophobic compared to other LEA domains and be folded (Singh *et al.*, 2005; Hundertmark & Hincha, 2008). The founding member of the LEA2 domain family is LATE EMBRYOGENESIS ABUNDANT14 (LEA14, At1G01470). Overexpression of LEA14 enhances *A. thaliana* salt tolerance (Jia *et al.*, 2015). Thus, little can be predicted from its amino acid composition as to how NHL3 acts to translate the PDLP signal, except that it probably does not signal directly to a CalS itself, as it has no known interacting domains.

NHL3 is predicted to have 361 unique interactors on BioGRID, a curated repository for protein interaction data (Stark *et al.*, 2006). Taking a naïve approach and looking at the top functional

GO terms for the interactors, may suggest how NHL3 is functioning. The top three different most significant GO terms are transmembrane transport (FDR = 2.43×10^{-31}), establishment of localization (FDR = 3.30×10^{-27}), and protein phosphorylation (FDR = 7.20×10^{-25}). These categories lend themselves to some easy hypotheses. Perhaps, NHL3 activates an efflux of ions that signals CalS initiation. Alternatively, NHL3 may recruit or degrade relevant proteins to initiate callose synthesis. The possibility of NHL3 affecting PDLP1 localisation was unsupported (Figure 4-21). Finally, NHL3 may recruit or activate a phosphorylation cascade that ultimately modifies CalS activity.

The NHL family was defined by homology to *Nicotiana tabacum* HIN1 and *A. thaliana* NDR1. NDR1, as well as having structural similarity to LEA14, has a high degree of homology with mammalian integrins (Knepper *et al.*, 2011). Integrins are well characterised in cell adhesion and signalling (Harburger & Calderwood, 2009). Integrins are known to have a large extracellular domain, with a small connecting intracellular tail, akin to NHL3 (Harburger & Calderwood, 2009). Plasmodesmata are also known as membrane contact sites (Pankratenko *et al.*, 2020). Thus, joining these ideas there is precedence that NHL3 may act in plasmodesmata to connect the cell to the cell wall, and signal to downstream components.

4.4.3   Callose Synthases

Callose synthesis is the end point for the PDLP1-mediated callose deposition pathway (Caillaud *et al.*, 2014). Here, it is shown that *cals1* mutant plants are not dwarfed at all by the overexpression of PDLP1 (Figure 4-11). Thus, PDLP1-mediated callose deposition occurs via CalS1.

However, *cals7/*PDLP1-GFP plants also show slight reversion from the dwarf phenotype. Thus PDLP1 can also act via CalS7. The expression domain of CalS7 is restricted to the phloem, whereas PDLP1 is not expressed in the vasculature (Xie *et al.*, 2011; Shi *et al.*, 2020). Thus, while PDLP1-GFP may be dwarfing the plant by initiating callose deposition via CalS7, this is due to ectopic expression. This highlights an interesting possibility: that PDLP1 is promiscuous and will activate callose deposition from multiple CalS proteins. Alternatively, *cals7* is known to create an aberrant developmental phenotype itself, as the mutant blocks sugar transport, leading to increasing leaf starch levels (Barratt *et al.*, 2011; Xie *et al.*, 2011). Thus, the decrease in dwarfing in *cals7/*PDLP1-GFP plants may be due to an increase in sugar availability in the leaves, compared to wild type.

It is a surprising, if PDLP1 can act through CalS1 and CalS7, that the *cals1* mutants fully revert the PDLP1-induced phenotype. This would suggest that the main mode of PDLP1-induced

dwarfing that is being measured is phloem-independent. Alternatively, it may suggest that CalS7 requires CalS1 to function, but this would be unlikely given that other CalSs have been shown to be functionally independent (Cui & Lee, 2016).

Cui and Lee (2016) implicated both CalS1 and PDLP5 in the response of plasmodesmata to SA. PDLP5 was previously shown to be transcriptionally regulated by SA. Here, I identify CalS1 as a required downstream component to PDLP1 signalling. This is the same result as found for PDLP5 (Cui & Lee, 2016). In the same way that *cals7* was found to partially revert the PDLP1 phenotype, so was *cals8* found to partially revert the PDLP5 overexpression phenotype. This builds a picture where PDLPs mainly signal through CalS1, but other CalS proteins are required for basal plasmodesmal function, as propose by Cui and Lee (2016).

Despite multiple PDLPs having been shown to integrate into a single CalS, CalS also have non-PDLP dependent functions at plasmodesmata. For example. CalS8 was linked to H_2O_2 mediated plasmodesmal closure, in a PDLP5 independent fashion. Therefore, while PDLPs are important determinants of plasmodesmata function, CalSs still ultimately represent the end of the plasmodesmata regulation pathway.

4.4.4   Connecting the pathway

Two additional proteins have been added to the pathway of PDLP1-triggered plasmodesmal closure: NHL3 and CalS1 (Figure 4-26). A direct interaction between PDLP1 and NHL3 has been shown, however it remains unknown how NHL3 signals to CalS1. Further, the upstream pathway of PDLP1 remains elusive, with no known phenotypes of the *pdlp1* mutant (Amari *et al.*, 2010). Therefore, the role of PDLP1 signalling remains obscure, but the method of downstream signalling has started to be unravelled.
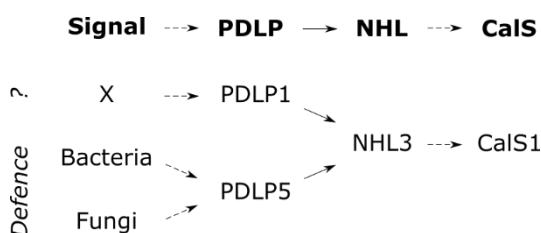
Figure 4-26 NHL3 is a nexus for PDLP signalling
PDLP5 is a common pathway component in plant defence signalling against both fungal and bacterial, whereas the elicitor for PDLP1 function (X) remains unknown. PDLP1 and PDLP5 both directly interact with NHL3. NHL3 is required for PDLP1-induced dwarfism and is assumed to be required for PDLP5 signalling also. CalS1 lies downstream of both PDLP1 signalling and PDLP5 signalling. Dashed arrow indicate an indirect interaction and solid arrow indicate a direct physical interaction.

The split-ubiquitin data indicates that NHL3 also interacts with PDLP5, suggesting that NHL3 is a nexus for PDLP-induced callose signalling (Figure 4-26). This would indicate that while

PDLPs have different upstream signalling triggers, the downstream execution of callose deposition is through a common PDLP – NHL3 – CALS pathway.

These data fit together into the bigger picture of PAMP-triggered plasmodesmal closure (Figure 4-27). PDLP5 is known to be essential for plant defence against bacteria, through a SA-mediated pathway (Lee *et al.*, 2011b; Wang *et al.*, 2013). The importance for PDLP5 to mediate plasmodesmata closure in response to PAMPs directly has also been shown, whereby *pdlp5* mutants did not reduce cell-to-cell flux for chitin nor flg22 (Papp, D., *pers. comms.*; Tee, E., *pers. comms.*). I have proposed that *nhl3* is downstream of PDLP5 responses (Figure 4-26), and this fits with the observed result that *nhl3* plants do not close their plasmodesmata in response to flg22 similarly to *pdlp5* (Figure 4-22, Figure 4-23). It would be exciting to follow up this work with a test of *nhl3* responses to chitin, where the model predicts that there should be no plasmodesmata response. Thus, PDLP5 is a common link in pathogen-defence signalling, acting upstream of NHL3 (Figure 4-26).

RBOHD was not required for PDLP1-induced plasmodesmal closure, either being upstream or unrelated to this pathway (Table 4-7, Figure 4-13). However, it has been established that RBOHD is required for chitin-mediated plasmodesmal closure (Cheval *et al.*, 2020). To test whether RBOHD was common to PAMP-induced plasmodesmata closure, the response the *rbohd* mutant to flg22 was observed. Similarly to chitin, *rbohd* plants did not reduce cell-to-cell flux of GFP in response to flg22 (Tee, E., *pers. comms.;* Figure 4-27). Thus, RBOHD is a common component to pathogen signalling at plasmodesmata. I have placed RBOHD upstream of PDLP5 in my model of plasmodesmata signalling, due to the independence of PDLP1-GFP dwarfing from RBOHD, making it the nexus of PAMP-induced plasmodesmal signalling (Figure 4-27).

Placing RBOHD upstream of PDLPs is in agreement with the fact that CRKs (PDLP homologues) are transcriptionally responsive to H_2O_2 (Czernic *et al.*, 1999; Wrzaczek *et al.*, 2010a). Moreover, it has been suggested in the literature that the conserved thiol groups (Figure 4-1) may act as redox sensors (Bourdais *et al.*, 2015). However, this runs counter to CRK2 phosphorylating RBOHD (Kimura *et al.*, 2020). In addition, recent crystallographic analysis suggest that the disulphide bridges are for structural stability over redox sensing (Vaattovaara *et al.*, 2019). Although, the authors do not argue why a change in stability cannot be used as a redox sensing mechanism. Yet, PDLPs do not contain kinase domains as CRKs do. A simple experiment to test this hypothesis would be to cross PDLP5OE with *rbohd*, and look for the reversion to wild type size as used with PDLP1-GFP.

As of yet, only CalS1 has been implicated in pathogen-mediated plasmodesmal closure (Cui & Lee, 2016). However, CalS12 has been put forward as a likely candidate too (Dong *et al.*, 2008; De Storme & Geelen, 2014). CalS1 was also found to be downstream of PDLP1-induced dwarfism (Figure 4-11). This result implicates CalS1 in a PDLP – NHL3 – CALS1 pathway, where CalS1 is responding to PDLP-mediated signals rather than defence signals specifically (Figure 4-26).

CalS8 was shown to be independent of SA plasmodesmal closure, but required for H_2O_2-induced closure (Cui & Lee, 2016). H_2O_2 treatment produces ROS, akin to RBOHD. This raises the question why CalS8 is not required for RBOHD-mediated signalling. I think this has to do with local specificity of signalling, exemplified in the plant response to chitin. There is a global ROS burst mediated by CERK1 (Miya *et al.*, 2007). The *lym2* mutant has no effect on the signature of the ROS burst (Faulkner *et al.*, 2013). Yet, RBOHD is required for plasmodesmal closure (Cheval *et al.*, 2020). This indicates that plasmodesmata-specific ROS is required for closure, on a scale that is much smaller than the cell as a whole. Therefore, a H_2O_2 treatment may be more akin to a global ROS burst, rather than a localised plasmodesmata-specific ROS burst. It remains to be seen as to why the CERK1-mediated global ROS burst does not induce plasmodesmal closure in the *lym2* mutant, whereas presumably a H_2O_2 treatment does. There must be a fundamental different between endogenous ROS production, and exogenous application of ROS: activating different pathways ending in CalS1 and CalS8, respectively. Indeed, a *pdlp5* mutant still has reduced cell-to-cell flux with a H_2O_2 treatment, suggesting a second ROS sensing mechanism (Cui & Lee, 2016).

The remaining parts of the defence plasmodesmata callose pathway is how the signal is relayed between known components (Figure 4-27). In bacterial signalling, CML41 has been shown to be essential (Xu *et al.*, 2017). It has been hypothesised that CML41 acts upstream of RBOHD, akin to CALCIUM-DEPENDENT PROTEIN KINASEs (CPKs) (Cheval & Faulkner, 2018). Alternatively, CML20 has been shown to directly interact with CalSs suggesting that CML41 may interact with CalS1 (Vu *et al.*, 2019). However, CML41 has been shown to be flg22-specific, where *cml41* knock-down lines respond to chitin but not to flg22. Thus, CML41 could not act in a common PDLP – NHL – CML41 – CALS pathway. Therefore, I agree with Cheval and Faulkner (2018) that CML41 is more likely to be upsteam of RBOHD. However, it affirmed how CML41 signals the plasmodesmata-closure signal. This contrasts flg22-plasmodesmal signalling, where CPK6 and CPK11 have been shown to directly interact with RBOHD (Cheval *et al.*, 2020) (Figure 4-27). Yet, how both classes of these calcium-dependent proteins are activated remains unknown.

Overall, a clearer picture is emerging of how PDLPs are signalling plasmodesmata-specific callose deposition (Figure 4-26, Figure 4-27). I propose a conserved RBOHD – PDLP – NHL – CalS signaling module that activates CalS1 downstream of at least PDLP1 and PDLP5. How NHL3 signals to CalS1 remains an open question (see 4.4.2.4). In a defence context, RBOHD is likely activated by calcium-responsive proteins, though how the calcium influx is signalled is unknown.
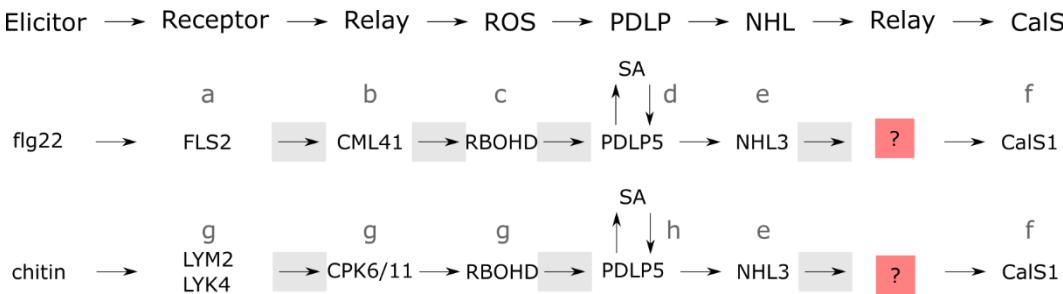
Figure 4-27 Proposed pathway to callose deposition.
Callose deposition at plasmodesmata is elicited by PAMPs, which are perceived by their cognate receptors. This signal is in relayed to RBOHD through calcium-responsive proteins. RBOHD acts upstream of PDLPs and is proposed a nexus of plasmodesmal-PAMP signalling. PDLPs interact directly with NHL3, which serves as a nexus for PDLP-signalling. NHL3 activates callose synthesis by a callose synthase via another relay protein. a) Faulkner et al., 2013; b) Xu et al., 2017; c) Tee, E., pers. comms.; d) Lee et al., 2011b; Wang et al., 2013; e) Table 4-8; f) Cui and Lee, 2016; g) Cheval et al., 2020; h) Papp, D., pers. comms. Shaded boxes indicate assumptions without direct evidence. Red question marks are unknown factors.

4.4.5 POEM46: KISS ME DEADLY2?

4.4.5.1 Cytokinin signalling overview

Cytokinins are potent hormones that regulate growth and development (Mok, 1994). They were originally discovered due to their ability to promote cell division of plant cells in culture (Miller *et al.*, 1956). Subsequent research has found an essential role for cytokinins in regulating development at stem cell niches (Werner *et al.*, 2001; Martínez-Fernández *et al.*, 2020), modulating shoot and root development (Mok, 1994). Additionally, exogenous application of cytokinins and endogenous increases in cytokinin increases the frequency of plasmodesmata in *Sinapis alba* (Ormenese *et al.*, 2006).

Cytokinins are perceived by three receptors ARABIDOPSIS HISTIDINE KINASE 2 (AHK2), AHK3 and AHK4 (Inoue *et al.*, 2001) (Figure 4-28). These kinases autophosphorylate on a conserved histidine (Ueguchi *et al.*, 2001). The reactive phosphoryl group is transferred to type-B *ARABIDOPSIS* RESPONSE REGULATORs (ARRs), via histidine-containing phosphotransfer protein (Kim *et al.*, 2006). These type-B ARRs control downstream growth responses by binding the promoter DNA of homeodomain transcription factors such as *WUSCHEL* (Meng

et al., 2017; Zubo *et al.*, 2017; Xie *et al.*, 2018). This signalling cascade is a two component system with type-A ARRs also being activated by type-B ARRs. Type-A ARRs then repress type-B ARRs (To *et al.*, 2004).

Type-B ARRs are also negatively regulated by an E3 ubiquitin ligase family: KISS ME DEADLY (KMD) (Kim *et al.*, 2013) (Figure 4-28). KMDs are S-PHASE KINASE-ASSOCIATED PROTEIN1(SKP1)-CULLIN-F-BOX PROTEIN (SCF) E3 ligases, the largest of the E3 ubiquitin ligases and responsible for the ubiquitin-mediated turnover of many key regulatory proteins. Strong overexpression of KMD2, using the *Cassava vein mosaic virus* promoter gave a similar phenotype to *arr1/10/12* mutant plants and *ahk2/3/4* mutants, namely severe dwarfism (Riefler *et al.*, 2006; Argyros *et al.*, 2008; Kim *et al.*, 2013). Conversely, ARR10 overexpression gives larger than wild type plants (Zubo *et al.*, 2017). Thus, it would be expected that a *kmd2* knock-out would also be larger than wild-type, but it has not been characterised. Furthermore, a *kmd2* mutant would be hypersensitive to cytokinins, as type-B ARRs are not degraded. Thus, as cytokinins induce plasmodesmata formation, a *kmd2* mutant would be expected to have more plasmodesmata than a wild-type plant. This could be characterised by electron microscopy. However, there may be redundancy in the family, as a triple *kmd1/2/4* mutant is required to delay meristem arrest with *kmd2* having no effect (Martínez-Fernández *et al.*, 2020). KMD2 was also published as KELCH REPEAT F-BOX1 (KFB01) (Zhang *et al.*, 2013).
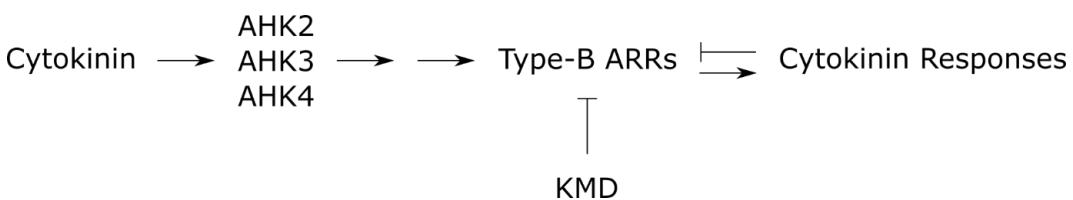
Figure 4-28 Simplified cytokinin signalling pathway
Cytokinins are perceived by ARABIDOPSIS HISTIDINE KINASE (AHK) receptor proteins, which autophosphorylate. The phosphoryl group is indirectly transferred to type-B *ARABIDOPSIS* RESPONSE REGULATORs (ARRs), activating them. Type-B ARRs initiate downstream cytokinin responses, including their own inhibition by type-A ARRs. Type-B ARRs are marked for proteolysis by SCF[KISS ME DEADLY] (KMD) E3 ligase mediated ubiquitination.

4.4.5.2   KISS ME DEADLY2 phenotypes may explain the size reversion phenotype

PDLP1 **o**ver**e**xpression **m**utant (POEM) plants are PDLP1-GFP EMS reversion mutants. The POEM46 mutant was identified as a partial revertant, significantly larger than PDLP1-GFP plants yet smaller than Col-0 (Figure 4-7). The POEM46 SNP was putatively mapped to the promoter of KISS ME DEADLY2 (KMD2, At1G15670). Mutations in the promoter can modulate gene expression, either positively or negatively. Knowing that KMD2 is a negative modulator of cytokinin responses, I hypothesise that *KMD2* expression is reduced thereby,

hypersensitising the POEM46 plants to cytokinins. An increased response to cytokinin could lead to a reversion of the dwarfing phenotype in two ways. On one hand, cytokinin is known to increase the formation (three-fold) of secondary plasmodesmata in *Sinapis alba* (Ormenese *et al.*, 2006). This could counteract the reduction in flux due to callose deposition by PDLP1 (Thomas *et al.*, 2008). On the other hand, cytokinin is a well-characterised growth modulator, and an increase in cytokinin response may lead to a basal increase in plant growth. Naturally, it could be a combination of both effects. To follow up this work, the gene expression levels of *KMD2* in POEM46 plants could be measured to identify an effect of the SNP on expression. Further, *kmd2* and *kmd1/2/4* plants are available from Nottingham Arabidopsis Stock Centre. Thus, the transgene-free size could be determined of mutant plants, to consider whether KMD2 affects plant size globally. KMD2 overexpression has been documented to reduce plant size, so this is plausible (Martínez-Fernández *et al.*, 2020).

The observed size ratio of the POEM46 mutant was not the expected 1:2:1 (l:m:s) or 1:3 (l:s) of semi-dominance or a recessive allele, but closer to 2:1:2 (l:m:s) (Figure 4-8). This is probably due to the fact that there is some variability in the PDLP1-GFP parent, where some plants are as large as medium sized POEM46 plants (Figure 4-5). So, it is possible that while the allele is recessive (1:3 (l:s)), variability in the PDLP1 overexpression phenotype skews the proportions of the sizes. This is borne out in the next-generation mapping, where in both cases we expected no wild-type reads in the large F2 pool (7% was observed). In the small pool, 33% of reads are expected in the recessive case and 0% in the semi-dominant case. By conducting a χ^2 test on the 12 wild-type and 12 mutant reads, there was in sufficient evidence to reject the 33% case (χ^2 = 3, df = 1, p > 0.05). Thus, from the sequence data, I conclude that POEM46 is a recessive mutation.

4.4.6   PDLP Interactors

As well as finding genetic interactors of PDLP1, I worked to find novel physical interactors of the PDLPs. This took a two-pronged approach of analysing a split-ubiquitin screen (carried out DualSystems Biotech under the supervision of Faulkner, C.) and a plasmodesmata-specific co-immunoprecipitation.

4.4.6.1   Putative PDLP5 interactors

I leveraged the growing literature of plasmodesmal proteomes, as well as my own data from Chapter 3 and cross referenced it with the interacting clones found in the split-ubiquitin screen (Table 4-8). A crude score of confidence in the interaction is given by the number of clones identified in the experiment, with multiple hits suggesting a more likely partner.

However, this may be confounded by the make-up of the screening library and would ideally be normalised to the count of each gene within the library. Unfortunately, that information is proprietary, and so a raw count was used. Another metric for confidence was the number of plasmodesmal proteomes the gene overlapped with. Using this, I found three proteins which were in all datasets: NHL3, a peroxidase, and RETICULON LIKE PROTEIN B1 (RTNLB1). Only NHL3 of the two tested genes localised to plasmodesmata (Figure 4-14, discussed above section 4.4.2.2), demonstrating that even the highest confidence candidates may not be plasmodesmal proteins.

I did not follow up on the localisation of RTNLB1, as it is more likely to localise to the desmotubule of the plasmodesmata, rather than the plasma membrane. This is based on evidence that the close family members RTNLB3 and RTNLB6 have been localised to the endoplasmic reticulum and desmotubule (Knox *et al.*, 2015). Knox *et al.* (2015) showed similar localisations of RTNLB1, as RTNLB3/6 in *N. tabacum* 'Bright Yellow 2' BY2 cells, but was not chosen to follow-up in *N. benthamiana* with a plasmodesmata marker (Knox *et al.*, 2015). Additional evidence that RTNLB1 may be located at plasmodesmata is that RTNLB1 and RTNLB2 associate with and are required for the intracellular trafficking of FLS2 (Lee *et al.*, 2011a). In turn, FLS2 has been shown to localise to plasmodesmata (Faulkner *et al.*, 2013). Moreover, RTNLB1 has been shown to interact with the plasmodesmally-located RTNLB3 (Sparkes *et al.*, 2010). Sparkes *et al.* (2010) found a FRET efficiency of 8.3%, equating to a distance of 3.3 nm apart (Equation 4-1). Therefore, it is very likely that RTNLB1 associates with plasmodesmata.

With more time, I would have investigated the possibility of a PDLP5-RTNLB1 interaction, as reticulon proteins are potentially an integral plant of *A. thaliana* plasmodesmata. Reticulon proteins are candidates for constricting the desmotubule into its highly curved state (one of the most constricted membranes in nature (Tilsner *et al.*, 2011)), as they prefer to localise at curved membranes (Sparkes *et al.*, 2010). Moreover, RNLBs can constrict the endoplasmic reticulum to such an extent GFP-HDEL is restricted from the lumen of the endoplasmic reticulum (Tolley *et al.*, 2008, 2010). The same phenomenon is observed within the desmotubule (Crawford & Zambryski, 2000).

Given that PDLP proteins reside in the plasma membrane at plasmodesmata and RTNLBs reside in the endoplasmic reticulum, it raises the question how they would interact *in vivo*? Plasmodesmata have recently be shown to be in two states – Type I and Type II – via electron tomography (Nicolas *et al.*, 2017). Type I plasmodesmata have no cytoplasmic sleeve and so

have the plasma membrane and endoplasmic reticulum in contact: this may be when PDLP5 and RTNLB1 interact. Further work has shown that some mutants can alter the Type I to Type II transition (Yan *et al.*, 2019); a similar study could be carried out on *rtnlb* and *pdlp* mutants to show whether they lower the number of Type I plasmodesmata. Alternatively, RTN proteins may help traffic PDLP proteins to plasmodesmata, and so interact before localising at plasmodesmata. The method of plasmodesmal trafficking remains enigmatic.

Looking at the list of potential PDLP5 interactors as a whole, they are significantly enriched for plasmodesmal genes (FDR = 8.16×10^{-06}, *n* = 10) and plasma membrane genes (FDR = 1.65×10^{-06}, *n* = 16). Neither result is surprising as the list was in-part curated by the plasmodesmal proteome paper (which largely defines the GO term, see 3.4.3 (Fernandez-Calvino *et al.*, 2011)). In addition, split-ubiquitin methods specifically work for membrane bound proteins, and so an increase in membranes genes is expected. Beyond GO terms, there is one other protein in the list directly associated with plasmodesmata apart from NHL3: GERMIN-LIKE PROTEIN 5 (GLP5), also known as PLASMODESMATA GLP1 (PDGLP1) (Ham *et al.*, 2012). PDGLP1 was shown to have a role in root development, whereby reducing growth upon overexpression. This corresponds well with the PDLP5 phenotype (Lee *et al.*, 2011b). Thus, GLP5 may also be a *bona fide* PDLP5 interactor.

4.4.6.2   Plasmodesmata co-immunoprecipitation

Several plasmodesmal proteins have been identified by co-immunoprecipitation of plasmodesmata enriched fractions (Lee *et al.*, 2003; Ham *et al.*, 2012). Here, I attempted to co-immunoprecipitate proteins from plasmodesmal fractions. I successfully immunoprecipitated several plasmodesmal proteins (PDLP1/PDLP5/C4), as indicated by their presence after washing antibody labelled beads (Figure 4-24). However, no interacting proteins were identified alongside this (Table 4-9). This is probably due to the low material volume going into the extraction (four leaves), and purifying for a very small subsection of the mass. This is on the lower end of the mass recommended for a plasma membrane extraction, which is a much larger structure (Kadota *et al.*, 2016). Future work could simply upscale the amount of tissue being put into the protocol, as the immunoprecipitation process itself was successful.

The method required troubleshooting due to PEG-like contaminants. This was solved by repeated washes in 50% acetonitrile. Each 500 µL wash dilutes all soluble contaminants ~100-fold if 5 µL of buffer is left behind. With four washes, the levels of contamination rapidly become undetectable (8 orders of magnitude less). In an elegant experiment, Rosenfeld *et*

al. (1992) demonstrated that repeated washes remove contaminants, but proteins remain within the polyacrylamide gel. This was demonstrated by showing that the levels of radiolabelled protein are not diminished (Rosenfeld *et al.*, 1992). Therefore, I do not think the low levels of protein detected by mass spectrometry were due to the increased washes, rather the amount of input.

4.4.7   Conclusions and Future Directions

This chapter aimed to find novel interactors of PDLP proteins, to further understand how plasmodesmata are closed in defence. Here, a new component has been established: NHL3; as well as providing several putative candidates for further characterisation: RTNLB1 and GLP5. NHL3 physically interacts with PDLP1, and is required for PDLP1-induced dwarfism. Moreover, *nhl3* mutants do not close their plasmodesmata in response to flg22, and are more susceptible to bacteria and fungi (Dörmann *et al.*, 2000; Zheng *et al.*, 2004; Wang *et al.*, 2018). In addition, the order of the PDLP pathway has been further established, with known players being tested in the PDLP1-GFP background. CalS1, already known for SA and PDLP5-mediated plasmodesmata closure, is required for PDLP1-induced closure as well. In addition, RBOHD was shown to be upstream or independent of the PDLPs. Lastly, a putative novel interactor was identified: KMD2. However, KMD2 may be a global plant size regulator, and so further study is needed.

Several questions remained unanswered about the components discovered here. Differing *nhl3* T-DNA lines had opposing phenotypes. This was suggested to be due to an increased variability in callose control. This could be further investigated with NHL3 overexpression transgenic lines, where an expected increase in callose deposition would be expected upon flg22 treatment. More generally, how NHL3 functions upstream to CalS remains unknown, but a GO analysis suggests it may be by protein phosphorylation. The potential for NHL3 to interact with both PDLP1 and PDLP5 raises the question as to whether NHL3 can interact with other PDLP proteins. I have suggested that RBOHD as a potential nexus for PAMP-mediated plasmodesmal closure, therefore *rbohd* mutant plants should also be plasmodesmally unresponsive to salicylic acid. This would test whether the pathway has been properly constructed.

The data presented here provide some areas for future exploration of PDLP-mediated signalling, with additional putative PDLP5 interactors and a potential method for exploring protein interactions at plasmodesmata specifically.

5   Discussion

In this thesis, I explore the proteins found in plasmodesmata and the interactions between them, with specific attention paid to the PLASMODESMATA-LOCALISED PROTEINs (PDLPs). The key results produced by this work are: a novel plasmodesmata extraction method for mature tissue that is functional for higher plants and bryophytes; plasmodesmal proteomes from adult *Arabidopsis thaliana* rosette leaves and from *Physcomitrella patens*; identification of two evolutionary conserved structural plasmodesmata proteins; and the elucidation of some of the callose deposition pathway downstream PDLP1, including a direct physical interactor NHL3 (NDR1/HIN1-LIKE 3). These findings contribute new understanding to plasmodesmal biology, and present new avenues for exploration, such as how does NHL3 function and which other proteins are evolutionarily conserved at plasmodesmata. I will explore some of these questions in detail below.

5.1   What might an ancient plasmodesma look like?

Chapter 3 of this thesis describes my efforts to determine which proteins are conserved in proteomes, both within *A. thaliana* and phylogenetically between *A. thaliana*, *P. patens* and *P. trichocarpa*. Plasmodesmata likely evolved in the common ancestor of Embryophyta and the Charales and so it is also worth considering whether the proteins discussed here were likely to have been in plasmodesmata over 600 million years ago. As with all phylogenetic analysis, we can only compare extant taxa or those in the fossil record. Plasmodesmata have been found in the fossils of land plants and the large pit plugs of red algae, as of yet though no plasmodesmata-sized pores have been found in fossilised streptophytic algae (Vannucci *et al.*, 2000; Edwards, 2003; Raven, 2005). Thus, this discussion will focus on the properties of the last common plasmodesmata (LCPD) between Embryophyta and Charales (Figure 5-1).

a)

b)

Cladogram labels: Plasmodesmata loss · Plasmodesmata gain · LCPD

Species	Clade	Callose synthase	β-1,3-glucanase	Tetraspanin	C2 lipid-binding	RNA-binding glycine-rich	Reticulon	LRR extensin-like	GDSL lipase/esterase	Ca-dependent lipid binding	SKU5	DUF26 protein	NHL protein	Remorin
		Callose		Structural					Other					
G. sulphuraria	Plantae	0	0	0	0	0	1	0	0	0	0	0	0	0
C. reinhardtii	Viridiplantae	7	0	0	0	0	1	5	1	0	0	0	0	0
K. nitens	Streptophyta	2	11	1	1	0	4	29	3	8	0	0	0	0
C. braunii		5	14	1	1	0	2	4	0	1	0	0	0	0
P. margaritaceum		17	18	0	0	0	5	23	5	8	0	0	0	0
P. patens	Embryophyta	43	99	23	16	19	43	183	25	125	3	0	0	0
S. moellendorffii	Tracheophyta	13	69	15	8	3	14	150	26	143	2	0	0	0
A. trichopoda	Angiosperms	9	36	6	8	4	7	97	24	51	1	3	2	1
A. thaliana		30	76	18	24	17	22	187	61	141	1	34	11	6

Figure 5-1 Considering the phylogenetic positioning of the evolution of plasmodesmata proteins
a) Cladogram of species used in this study, and the smallest clade that connects them to the species underneath them. The cladogram is marked with the gain and loss of plasmodesmata, as well as the last common plasmodesmata (LCPD) between Embryophyta and Charales. The predicted proteomes for *Galdieria sulphuraria* (Rhodophyta, red algae), *Chlamydomonas reinhardtii* (Chlorophyta), *Chara braunii* (charale), *Selaginella moellendorffii* (lycophyte), *Amborella trichopoda* (angiosperm)*,* and *Arabidopsis thaliana* (angiosperm) were downloaded from EnsemblPlants. The predicted proteomes from *Klebsormidium nitens* (Klebsormidiales) and *Penium margaritaceum* (Zygnematales) were downloaded from their publications (Hori *et al.*, 2014; Jiao *et al.*, 2020). b) Known plasmodesmal families were selected and plasmodesmal members were used to create an alignment search pattern for HMMER. All the proteomes were searched for the pattern and the number of proteins matched are shown in the table (E < 1 × 10^{-50}). Orange rows are species known to have plasmodesmata.

5.1.1   Did the last common plasmodesmata have a desmotubule?

All embryophytes contain plasmodesmata with a desmotubule. Nonetheless, it has been speculated that plasmodesmata first arose without a desmotubule (Lucas *et al.*, 1993). There is precedent for membrane-lined pores without a desmotubule, for example in the phylogenetically-distant brown algae (Phaeophyceae, Figure 5-2) (La Claire II, 1981). Closer to Streptophyta, *Bulbochaete hiloensis* (Oedogoniales, Chlorophyta, Viridiplantae) also has simple plasmodesmata without a desmotubule (Fraser & Gunning, 1969). However, streptophytic plasmodesmata are thought to be an evolutionary independent from these examples, forming a monophyletic clade as progenitor to embryophytic plasmodesmata. Thus, the monophyletic Streptophyta alone have to considered, when asking questions about the LCPD. However, the precedent set for the possibility of a membrane-lined pore without a desmotubule.

Figure 5-2 *Cutleria hancockii* has membrane-lined pores without a desmotubule
Plasmodesmata between meristematic cells of the brown alga *Cutleria hancockii*. Scale bar = 200 nm. Micrograph by J.W. La Claire II, reproduced with permission from Cook & Graham (1999).

Unlike embryophytes, the presence of desmotubules in basal algae is a contentious subject. Some authors have observed desmotubules in species of Charales and sister Coleochaetales (Pickett-Heaps, 1968; Stewart *et al.*, 1973a; Kwiatkowska & Maszewski, 1976; Cook *et al.*, 1997; Kwiatkowska, 2003; Brecknock *et al.*, 2011). On the other hand, there are members of both orders which have been observed without desmotubules (Spanswick & Costerton, 1967; Marchant & Pickett-Heaps, 1973; Marchant, 1976; Franceschi *et al.*, 1994). Thus, this paints a confusing picture where desmotubules have either been repeatedly gained or repeatedly lost (Figure 5-3).
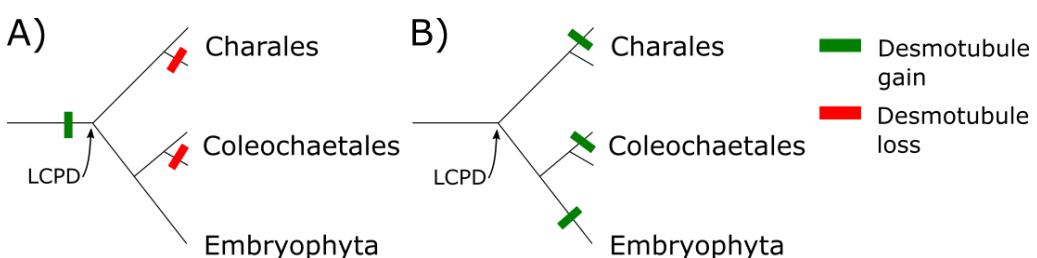
Figure 5-3 Debated origin of desmotubules in Charophytes
The presence and absence of desmotubules have both been observed in the species of Charales and Coleochaetales. This leads to two models of repeated loss of desmotubules in A) and the repeated gain of desmotubules in B). In A) the LCPD has a desmotubule, whereas in B) it does not.

Cook and Graham point out that "chemical fixation of cells in this genus [Coleochaetales] is difficult, however, so further study is required" (Cook & Graham, 1999b). I agree with the sentiment that without further evidence a conclusive picture cannot be drawn. Perhaps the lack of desmotubules observed before Cook *et al.*'s seminal work in 1997 was due to a difference in substrate handling.

In the work of Hepler (1982), the desmotubule was shown to be formed by the entrapment of the endoplasmic reticulum within the cell plate. Thus, primary plasmodesmata must have a desmotubule. In contrast, secondary plasmodesmata could lack a desmotubule, if formed

by the model of tunnelling through the cell wall, rather than endoplasmic driven formation (Faulkner *et al.*, 2008; Ehlers & van Bel, 2010). Thus, an examination of how plasmodesmata form in the Charales and Coleochaetales should inform us on whether a desmotubule is present in these clades. For example, the plasmodesmata of brown algae *Laminaria groenlandic,* lacking a desmotubule, was shown to form plasmodesmata after cytokinesis (Schmitz & Srivastava, 1974; Schmitz & Kühn, 1982). Cook *et al*. (1999) observed primary plasmodesmata formation during cytokinesis in *Chara zeylanica*. However, Franceschi *et al.* (1994) found in *Chara corallina* new cell walls formed without plasmodesmata, and then plasmodesmata were formed subsequnetly without desmotubules. Recently, more modern microscopy techniques (high-resolution scanning electron microscopy) have been used on *C. corallina,* clearly identifying that it possesses desmotubules (Brecknock *et al.*, 2011). This then raises questions over the lack of desmotubules in Coleochaetales, and whether that too was due to limitations of the microscopy or sample preparation techniques. Nonetheless, the new evidence clearly supports a model of LCPD having a desmotubule (Figure 5-3A), with possible loss in the Coleochaetales. For the remainder of the discussion, therefore, I will assert the presence of desmotubule in the LCPD.

It is worth noting that plasmodesmata and phragmoplasts co-occurred (Figure 3-1) in the last common ancestor of Charales and Embryophyta (i.e. after the divergence of Klebsormidiales), long after the evolution of multicellularity. It seems likely then that one of the core advantages of a phragmoplast over centripetal cleavage is the entrapment of endoplasmic reticulum. Thus, plasmodesmata likely co-evolved with the phragmoplast, rather than appearing afterwards.

5.1.2   What proteins may have been in the last common plasmodesmata?

5.1.2.1   Structural

Plasmodesmal membranes are specific microdomains of the plasma membrane, with an altered protein and lipid composition (Fernandez-Calvino *et al.*, 2011; Grison *et al.*, 2015a; Brault *et al.*, 2019; Liu *et al.*, 2020; Cheval *et al.*, 2020). The strongest evidence from the work presented in this thesis points towards the structural protein families of tetraspanins and C2 lipid-binding proteins having been in present in the first plasmodesma. However, as previously discussed (see 3.4.7.3), the plasmodesmata localisation of a *Physcomitrella patens* protein exogenously does not necessarily convey the endogenous localisation, as trans factors may be missing or redirecting the protein.

Caveats aside, the conserved presence of spokes connecting the desmotubule and the plasma membrane in *Chara zeylanica* strongly suggests that this is an ancestral trait (Cook *et al.*, 1997). Brault *et al.* (2019) put forward that these spokes may be C2 lipid-domain proteins. If this were true, we would expect to see C2 lipid-domain proteins in the *Chara braunii* genome as well (Nishiyama *et al.*, 2018). Indeed, a single C2 lipid-domain protein is present in *C. braunii* (Figure 5-1). Moreover, this family of proteins have been lost concurrently with the loss of plasmodesmata in *Penium margaritaceum* (Zygnematales) (Jiao *et al.*, 2020). Interestingly, this family of proteins is also found in *Klebsormidium nitens* (previously known as *Klebsormidium flaccidum* in the order Klebsormidiales) (Hori *et al.*, 2014), implying that C2 lipid-domain proteins have a pre-plasmodesmata role and have been exapted.

Plasmodesmata can be considered as a membrane contact site (MCS) (Tilsner *et al.*, 2011). MCSs are found across all eukaryotes, and most often occur between the endoplasmic reticulum and a second organelle (Prinz, 2014). MCSs require protein tethers to form the membrane-membrane interaction (Eisenberg-Bord *et al.*, 2016). In yeast (*Saccharomyces cerevisiae*), a Δtether mutant was made missing six tether proteins leading to separation of the endoplasmic reticulum from the plasma membrane (Manford *et al.*, 2012). An *A. thaliana* C2 lipid-domain protein (AtMCTP4) could restore MCSs in this yeast mutant pointing to an ancestral function of C2 lipid-domain proteins as tethering proteins (Brault *et al.*, 2019). Interestingly, this may well be a case of convergent evolution, where the *S. cerevisiae* tricalbins also contain an N-terminal endoplasmic reticulum transmembrane domain and multiple C2 lipid-binding domains at the C-terminus (Manford *et al.*, 2012): an almost identical domain composition to AtMCTP4. Both proteins classes are likely to have evolved from the ancient synaptotagmins (which are involved in membrane trafficking, notorious for their role in neurotransmitter release in animals). In plants, synaptotagmins are localised to MCSs but are not thought to localise at plasmodesmata. However, synaptotagmins have been seen to relocalise to plasmodesmata upon viral infection with *Turnip vein clearing virus* (Levy *et al.*, 2015). Potentially the virus movement protein is recreating the same evolutionary event that occurred in C2 lipid-domain proteins to induce their movement to plasmodesmata.

Similarly to the data for C2 lipid-domain proteins, I observed that a *P. patens* tetraspanin protein localised to plasmodesmata in two angiosperms (Figure 3-23, Figure 3-25). Interestingly, the tetraspanins have the same evolutionary pattern as the C2 lipid-domain protein, where they are present in all species with plasmodesmata, but are lost simultaneously with plasmodesmata in in *P. margaritaceum* (Figure 5-1). Tetraspanins have

no confirmed function at plasmodesmata, yet. It appears likely that tetraspanins can control the clustering of signalling proteins by controlling the nanodomain environment, by comparison to the animal literature, (Kummer *et al.*, 2020).

Following this line of thought, ancient plasmodesmata may also have been signalling hubs (Faulkner, 2013; Lee, 2015). In line with this receptor like-kinases (RLKs) are found in the phylogenetically conserved protein families (OG018, INFLORESCENCE MERISTEM RECEPTOR-LIKE KINASE 2 (IMK2), Table 3-8). Moreover, in the less stringent phylogenetically conserved list, receptor-like kinases are the most prominent group with 56 members (OG000, RLK, Table 3-9), twice that of the second most common group. Note, that the stringent IMK2 group is within the less stringent RLK group as RLKs are highly homologous, e.g. the group also includes FLS2 and PHOTOTROPIN 1.

It is probably an impossible task to define what may have induced signalling in the first plasmodesma; nonetheless, it is exciting to speculate. The only evidence currently available on stimuli for plasmodesmal signalling common to both higher and early plants is from ABA. ABA is known to control plasmodesmata in higher plants from avocado to aspen (Moore-Gordon *et al.*, 1998; Botha & Cross, 2000; Tylewicz *et al.*, 2018; Singh *et al.*, 2019). Recently, DENDRA2 (a photoswitchable fluorescent protein) has recently been used to study cell-to-cell movement in *P. patens* (Kitagawa & Fujita, 2013). Kitagawa *et al.* demonstrated that ABA also induces plasmodesmal closure in moss (Kitagawa *et al.*, 2019). Thus, it is probable that the first plasmodesmata, or an early ancestor, also responded to ABA. This is supported by the fact that ABA has been found in *Chara foetida* (Tietz *et al.*, 1989), as well as throughout Plantae (Hartung, 2010). However, the ABA receptor in land plants is cytoplasmic (Ma *et al.*, 2009; Park *et al.*, 2009), and the receptor in *C. braunii* is yet to be found (Nishiyama *et al.*, 2018). Thus, it seems likely that ABA signals indirectly to plasmodesmata rather than having a receptor embedded at the plasmodesma.

Beyond ABA, there are no published studies on the effect of stimuli on plasmodesmata of lower plants and algae to the best of my knowledge. However, given the importance of plasmodesmata in plant defence from pathogens in angiosperms, it would be surprising if there were not receptors for plant pathogens in the first plasmodesmata (Faulkner *et al.*, 2013; Wang *et al.*, 2013).

Similarly to tetraspanins, the remorins are thought to help create the plasmodesma microdomain (Raffaele *et al.*, 2009; Huang *et al.*, 2019). However, as with the PDLPs remorins

are only present in flowering plants (Figure 5-1) (Vaattovaara *et al.*, 2019). Thus, they are unlikely to have been a component of the LCPD.

The final class of structural protein known at plasmodesmata are the reticulons (Knox *et al.*, 2015), which are thought to help induce the tight curvature of the desmotubule (Tolley *et al.*, 2010). Given the presence of a desmotubule in *C. braunii*, it would be plausible for reticulons to be present as well. However, I did not identify any reticulons in *C. braunii* (Figure 5-1). There were proteins in *C. braunii* similar to reticulons but below the cut-off set (E = 1 × 10^{-40}). This may be in part due to the low number of plasmodesmal reticulons identified (n = 2) creating a weaker HMM search pattern. Moreover, while reticulons are present in *P. patens,* none were identified in the proteomics (Table S2). Thus, reticulons may help the tight folding of the desmotubule, but there is insufficient evidence to suggest they were in the LCPD.

5.1.2.2 Callose

Plasmodesmata in higher plants have been shown to respond dynamically to stimuli by depositing callose at the neck of plasmodesmata, reducing cell-to-cell flux (Radford *et al.*, 1998). While this has recently been observed in bryophytes as well (Kitagawa *et al.*, 2019), callose-mediated responses have not yet been shown in the Charales. However, a low level of callose does constitutively associate with plasmodesmata in of *C. corallina* (Blackman *et al.*, 1998; Faulkner *et al.*, 2009). In addition, dynamic plasmodesmata-assocated responses have been observed in in *C. corallina*, such as electrical resistance across plasmodesmata and solute trasnport through plasmodesmata (Côté *et al.*, 1987; Reid & Overall, 1992). Therefore, it is likely that callose is involved with dynamic responses in algae. This could easily and rapidly be tested with aniline blue staining of *C. corallina* with and without a treatment, for example flg22 or ABA addition.

It is irrefutable that callose is present at the plasmodesmata of *C. corallina*, so a CALLOSE SYNTHASE (CalS) must also be present, but there is not yet any evidence for β-1,3-glucanases (i.e. observed callose degradation). As above, it is incredibly likely that callose acts dynamically at plasmodesmata and so a glucanase would be required, non-callose mechanisms of plasmodesmata control have been recently postulated (Yan *et al.*, 2019; Tomoi *et al.*, 2020). In either case, control of plasmodesmal trafficking is an important part of intercellular communication, for which plasmodesmata were evolved. Given the presence of callose at higher plants, moss, and *C. corallina* plasmodesmata and the clear function callose plays in higher plants and moss I postulate that both callose synthases and β-1,3-glucanases were present in the LCPD.

5.1.2.3 Phylogenetically conserved proteins (Table 3-8)

Finally, I considered the remaining phylogenetically conserved proteins which were chosen to focus on and so I drew phylogenetic trees for (Table 3-8). Only the leucine-rich repeat (LRR) extensin-like, GDSL esterase/lipases, and RNA-binding glycine-rich are present in all the species examined which have plasmodesmata (Figure 5-1). LRR extensin-like proteins have been localised to the cell wall, and so may be a common contaminant of plasmodesmal proteomes (Bayer *et al.*, 2006; Herger *et al.*, 2019). Alternatively, they may function in the wall specifically associated with plasmodesmata, as has been suggested for pectin methylesterases (Morvan *et al.*, 1998). Immunogold localisation of LRR extensin-like proteins would help answer this question. I investigated GDSL esterase/lipase proteins in this thesis and found that they localise to the endoplasmic reticulum when transiently expressed. While this does not rule out that they exhibit plasmodesmal localisation when stably expressed, it does make it less likely. However, it may be possible that the GDSL esterase/lipase proteins localise to the desmotubule. Therefore, I would not confidently place LRR extensin-like nor GDSL esterase/lipases at the LCPD.

The RNA-binding glycine-rich proteins were extremely promising candidates, as they are incredibly ancient: in Fig 1-1 they were the only protein to fall in Rhodophyta (*Galdieria sulphuraria*). Moreover, RNA-binding glycine-rich proteins were identified in AtCells2_filtered with a high Bayer Enrichment factor (plasmodesmata quantity / plasma membrane quantity) of 186 (c.f. AtCell2_ filtered median Bayer enrichment factor is 87), and were independently recovered in the AtPlant1 and PpPlant1 proteomes. Furthermore, an RNA-binding glycine-rich protein was identified in be specifically present in plasmodesmata-rich nodal of *C. corallina* (Faulkner *et al.*, 2005)*.* Thus, prior to *in vivo* imaging it was highly likely to be localised to plasmodesmata. However, *in vivo* imaging revealed that the localisation was conclusively in the nucleus (enriched in the nucleolus) and the cytoplasm (Figure 3-27). Despite this, I think it is highly probable that RNA-binding glycine-rich proteins interact with plasmodesmata in some manner, due to the proteomic evidence. Plausibly, under stress, the proteins may relocate to plasmodesmata (Grison *et al.*, 2019; Hunter *et al.*, 2019; Cheval *et al.*, 2020). Additionally, RNA-binding glycine-rich proteins may be involved in the movement of mRNA between cells (Lucas *et al.*, 1995; Thieme *et al.*, 2015). mRNA itself would be too large to move cell-to-cell, so RNA-binding glycine-rich proteins may act as mRNA chaperones. Nonetheless, without evidence of their function at plasmodesmata, I cannot include RNA-binding glycine-rich proteins in a model of the LCPD.

This builds a final model of the LCPD, some 600 million years ago, which in response to hormonal signals can dynamically close by depositing callose and that is held together by tetraspanins with a desmotubule tethered in place by spoke-like elements, possibly C2-lipid proteins (Figure 5-4).
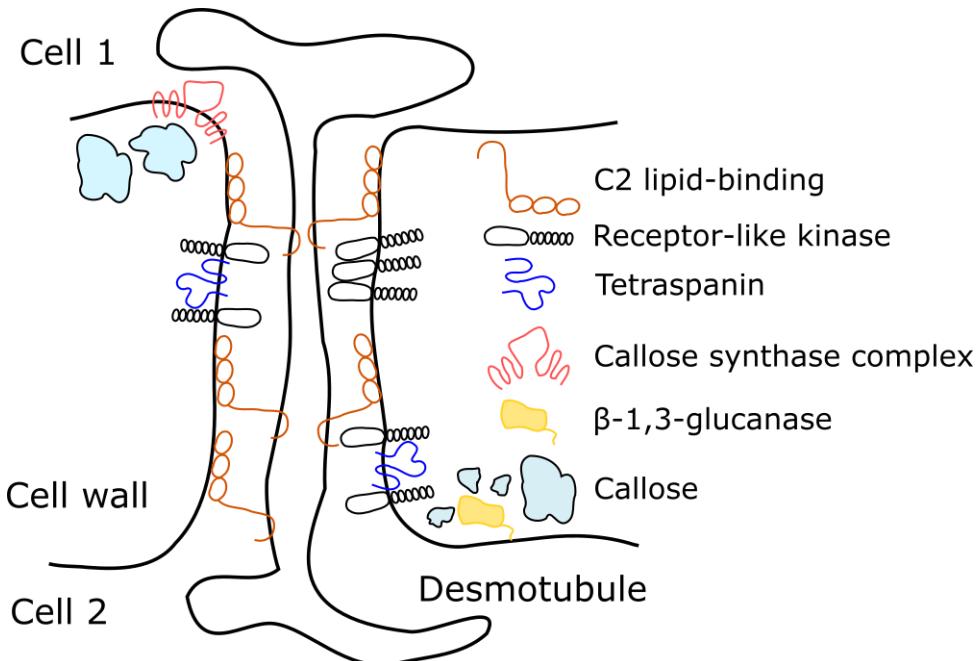
Figure 5-4 Model of the last common plasmodesmata (LCPD)
The LCPD between Charales and Embryophyta proposed based on conservation of proteins in proteomes (C2 lipid-binding, receptor-like kinase, tetraspanins), function (receptor-like kinase), callose (callose synthase, β-1,3-glucanases) and ultrastructure (spokes: C2 lipid-binding proteins).

5.2 What additional proteins may be at plasmodesmata in *Arabidopsis thaliana*?

There have been about 600 million years since the last common plasmodesmata (LCPD, Figure 5-4), giving *A. thaliana* plenty of scope to incorporate novel proteins into the LCPD. The most obvious example are the PLASMODESMATA-LOCALISED PROTEINs (PDLPs), which are only present in angiosperms and have a canonical plasmodesmata localisation (Thomas *et al.*, 2008; Lee, 2014; Vaattovaara *et al.*, 2019). The remorin proteins have also been described at plasmodesmata (Huang *et al.*, 2019). Interestingly, both PDLPs and remorin proteins only arise in the angiosperms (Figure 5-1) and both mediate salicylic acid (SA, a pathogen defence hormone) dependent plasmodesmata closure. This suggests that this pathway has grown in importance from the LCPD and has come under finer control with a more elaborate signalling pathway.

The *A. thaliana* proteome-recurrent proteins (Table 3-2) provide an opportunity to delve into the some of the most abundant *A. thaliana* plasmodesmata proteins (see 3.4.1.1) and

proteins which are conserved between simple and complex plasmodesmata. Three proteins are found shared between the proteomes AtCells1, AtCells2_filtered and AtPlant1 (Table 3-2), namely MCTP4, BG_PPAP and AT4G34150, a calcium-dependent lipid-binding protein. Only the latter protein has not been shown at plasmodesmata, making it an obvious candidate for further work (Levy *et al.*, 2007; Brault *et al.*, 2019). Unfortunately, no work has been carried out on this protein yet, despite its exciting properties of having a lipid-binding domain and DNA-binding domain (see 3.4.1.3).

Several exciting candidates can be identified from the remaining proteins that recurrently appear in the proteomes. One protein has no defined function apart from being labelled as containing both a DUF1191 and a transmembrane domain (AT3G08600). Akin to the PDLPs, AT3G08600 has a large extracellular region and has a short disordered internal tail. Moreover, the PDLoc software predicts a plasmodesmata localisation (z = 1.2) (Li *et al.*, 2020).

Two other classes of protein are thought provoking. First, DYNAMIN RELATED PROTEIN (DRP) 2B (AT1G59610, also known as DYNAMIN-LIKE 3) is part of a 16-member family related to phragmoplastin. DRP1, DRP2 and DRP5a subfamilies are involved in cytokinesis, specifically cell plate formation (Ahn *et al.*, 2017). More to the point, the *trans*-Golgi network (in part orchestrated by DRP2B (Huang *et al.*, 2015)) creates a tubule-vesicular membrane network during cell plate formation, which may be important for plasmodesmata formation (Hepler, 1982; Seguí-Simarro *et al.*, 2004). DRP2B and DRP1A co-localise and function together for both clatherin-mediated endocytosis and cell plate formation (Fujimoto *et al.*, 2008, 2010). Interestingly, DRP1A is crucial for the formation of dynamin-like rings at membrane tubules during cell plate formation (Zhang *et al.*, 2000; Otegui *et al.*, 2001) with the help of other proteins (Ahn *et al.*, 2017). This comes together to paint a picture where DRPs may be essential for plasmodesmata initiation and DRP2B may be part of this process or subsequently maintained at plasmodesmata.

The second class of proteins I think are of interest is more diffuse: pectin-associated cell wall-localised proteins. Generally, cell wall proteins have been disregarded at plasmodesmata, probably as it is harder to draw a causal connection directly to plasmodesmata function. For example, when screening AtCells1 for *bona fide* plasmodesmal proteins only membrane-bound proteins were included (Fernandez-Calvino *et al.*, 2011). However, I found that a number of cell wall proteins were conserved between the three *A. thaliana* proteomes. For example, AT5G25460 (DUF642 L-GALL RESPONSIVE GENE 2 (DGR2)), AT5G06870 (POLYGALACTURONASE INHIBITING PROTEIN 2 (PGIP2)) and AT4G18670 (LEUCINE-RICH

REPEAT EXTENSIN 5). The pectin-rich cell wall environment around plasmodesmata has been postulated to be important for plasmodesmata maturation (Ehlers & Kollmann, 2001; Faulkner *et al.*, 2008). The control of this cell wall environment is, in part, done by leucine-rich repeat extensins (Zhao *et al.*, 2018). PGIP2 functions to prevent pathogenic degradation of pectin (Kirsch *et al.*, 2020), and DGR2 has no ascribed function, but its transcriptional regulation is tightly coupled to pectin methylesterase activity (Cruz-Valderrama *et al.*, 2019). I think these proteins would be a difficult avenue of research due to their cell wall localisation, but an examination of how the wall environment of plasmodesmata is formed and maintained may be very fruitful. For example, localisation of pectin methylesterases to plasmodesmata has been shown (Morvan *et al.*, 1998), and it has been demonstrated that they are required for the cell-to-cell movement of viruses (Chen & Citovsky, 2003). The control of pectin methylation, and subsequent changes in cell wall flexibility, by the proteins outlined above may also be important for the control of cell-to-cell movement of viruses. A potential phenotype to explore would be the frequency of secondary plasmodesmata, as the pectin-rich environment of pit fields has been proposed as essential for their formation (Faulkner *et al.*, 2008).

Overall, there is still a huge potential for plasmodesmata proteins to be characterised. In Chapter 2, I suggested a lower estimate of 2,000 *A. thaliana* plasmodesmal proteins. Currently, 1,048 proteins are annotated as plasmodesmal in the plasmodesmata gene ontology. This leaves plenty of scope for future exploration and characterisation of novel *A. thaliana* plasmodesmata proteins. The proteomes provided in this thesis add to the growing literature of potential plasmodesmal proteins. Moreover, the analysis presented here paves a way to select highly likely candidates from both proteome-recurrent *A. thaliana* proteins and phylogenetically-conserved plasmodesmal proteins.

5.3 How might defence be signalled at plasmodesmata?

Plasmodesmata are part of pattern-triggered immunity (PTI) defence responses in *A. thaliana*. Upon the detection of conserved pathogenic molecules, e.g. chitin for fungi, callose is deposited at plasmodesmata, reducing cell-to-cell flux (Faulkner *et al.*, 2013). It is still unknown the exact benefit this yields to the plant (Cheval & Faulkner, 2018). Nonetheless, it is clear that when this response is prevented plants are more susceptible to a range of pathogens (Lee *et al.*, 2011b; Faulkner *et al.*, 2013; Lim *et al.*, 2016; Xu *et al.*, 2017). Moreover, it has been recently shown that pathogens modulate plasmodesmata to increase cell-to-cell flux with effectors (Aung *et al.*, 2020; Tomczynska *et al.*, 2020; Ohtsu *et al.*, 2021).

5.3.1   PDLP signalling at plasmodesmata

PLASMODESMATA-LOCALISED PROTEINs (PDLPs) have been shown to be a required part of the immune response. The overexpression of PDLP1 and PDLP5 leads to the deposition of callose (Lee *et al.*, 2011b; Caillaud *et al.*, 2014). However, no interaction partners for PDLP proteins (bar themselves (Wang *et al.*, 2020)), or method of action have been identified. I could not replicate the results of Wang *et al.* (2020) and found no interaction between PDLP1 and PDLP5 (Figure 4-16), despite using a similar FRET-based technique. PDLP hetrodimerisation may be induced by stress, and so differences in plant growth or sample processing for microscopy (e.g. I did not pre-infiltrate with water) may have caused the interaction.

Despite the lack of a known signalling pathway, transgenic plants overexpressing PDLP1 or PDLP5 have been described to have a dwarfed pheotype. I utlised this phenotype to find proteins downstream of PDLP1 by crossing in mutantions in candidate genes. In this manner, I idenitfied that both *nhl3* (NDR1/HIN1-LIKE 3) and *cals1* (CALLOSE SYNTHASE 1) both fully reverted the PDLP1 dwarf phenotype. This system could easily be used to consider whether PDLP5 is also essential for PDLP1 function, and whether other known CALSs are required for the downsteam pathway, with CalS8 being an obvious target (see 4.1.2.3) (Cui & Lee, 2016). PDLP1 has also been shown to regulate callose deposition outside of the plasmodesmata-context, as PDLP1-GFP was localised to haustoria prior to callose encasement (Caillaud *et al.*, 2014). Therefore, it would be interesting to see whether NHL3 is also required for PDLP1 roles at haustauria, or whether it has a plasmodesmata-specific function.

I hypothesised in Chapter 3 that NHL3 may act as a nexus for signalling downstream of the all PDLPs. If this were true, one would expect *nhl3* mutant plants not to close their plasmodesmata in reponse to chitin, as well flg22 (Figure 4-23), as *pdlp5* mutants are chitin-insensitive (Papp, D., *pers. comms.*). Further, Cui & Lee (2016) demonstrated that PDLP5 and CALS1 are required for SA-mediated plasmodesmatal closure, but was not required for H_2O_2 induced closure. Thus, if the predicted signalling pathway runs PDLP – NHL3 – CALS, then *nhl3* mutants would remain responsive to H_2O_2. In a similar vein, it would be worthwhile to compare the effects of *nhl3* and *cals1* in PDLP5 overexpression plants which have a similar dwarf phenotype (Figure 4-4). The proposed model (see Figure 4-26) predicts a full reversion to wild-type size in both mutants.

The pathway I proposed was: RBOHD-PDLP-NHL3-CALS. The direct physical interaction of PDLP1 and NHL3 was shown via FRET-FLIM microscopy, with genetic dependence described

above. The genetic independence of *rbohd* of the PDLP1 phenotype demonstrates RESPIRATORY BURST OXIDASE PROTEIN D (RBOHD) may be upstream of PDLP1, but does not indicate interaction. It may be worthwhile testing for a direct interaction to understand if the pathway is missing more interactors. However, as reactive oxygen species are produced apoplastically and could be mobile, there is no reason for a direct RBOHD – PDLP physical interaction. Thus, the activation of PDLPs by RBOHD would have to be tested in a exogenous system, such as ooycetes. To make this experiment work, a readout of PDLP acitivation is required. This may be possible to design by building a downstream pathway to callose deposition.

Looking at the downsteam part of the pathway, *cals1* also reverted the PDLP1 phenotype. Callose synthases, by definition, must be the end of the callose deposition pathway and so NHL3 must sit between PDLP1 and CALS1. Again, it would be interesting to see whether NHL3 directly interacts with CALS1 or whether there are intermediary factors. One such factor may be a CALMODULIN-LIKE (CML) protein. CML41 has been impliacted in the flg22 pathway, and was suggested to activate RBOHD (Xu *et al.*, 2017; Cheval & Faulkner, 2018). However, CML20 has since been shown to interact with CalSs, which indicates that the end of the pathway may be: NHL3-CML-CALS (Vu *et al.*, 2019). If this is the case, then the signalling pathway gets much more complicated as CML41 is specific for bacterial over fungal PAMPs (Xu *et al.*, 2017). In this case, additional information would have to be carried through the central PDLP-NHL pathway. Thus, I concur with Cheval and Faulkner (2018) and hypothesise CML41 is upstream of RBOHD (see Figure 4-27 for model). Regardless of the position of CMLs, NHL3 remains a central, required component for bacterial-induced plasmodesmal closure. The importance of plasmodesmata in defence is underscored by the increased susceptibility phenotype of *nhl3* mutants (Varet *et al.*, 2003; Singh *et al.*, 2018).

5.3.2   Chitin perception at plasmodesmata

LYSIN MOTIF (LysM) DOMAIN-CONTAINING GLYCOSYLPHOSPHATIDYLINOSITOL-ANCHORED PROTEIN 2 (LYM2), present in AtCells1, was shown to be required to mediate chitin-induced plasmodesmata closure (Faulkner *et al.*, 2013). LYM2 is a membrane-anchored extracellular protein, and so requires partner proteins to signal chitin perception through the membrane. Cecilia Cheval found that there was a genetic dependence of plasmodesmata closure in response to chitin on LysM RECEPTOR-LIKE KINASE (LYK) 4 and LYK5, as well as LYM2 (Cheval *et al.*, 2020). In Chapter 2, I demonstrated that only LYK4 resides at plasmodesmata, while LYK5 does not (Figure 2-14). This helped generate a model, where LYK5 is required for post-

translational modification of LYK4 in the plasma membrane, with modified LYK4 interacting with LYM2 at plasmodesmata (Cheval *et al.*, 2020).

LYM2 abundance at plasmodesmata increases upon chitin perception(Cheval *et al.*, 2020). This happens too rapidly to be deposition of new protein, and so LYM2 likely moves from the plasma membrane to plasmodesmata. This movement does not occur when LYK5 is overexpressed. This suggests that the movement of LYM2 to plasmodesmata may be controlled by LYK5 (Cheval *et al.*, 2020). The interaction between LYK4 and LYK5 also weakens upon the addition of chitin. This indicates that upon chitin addition LYK5 may release both LYM2 and LYK4, allowing them to move to plasmodesmata. Thus, it may be that rather than chitin recruiting LYM2 to plasmodesmata, chitin releases LYM2 from the plasma membrane. This hypothesis would indicate that LYM2 would be more enriched at plasmodesmata in *lyk5* knockout plants in standard conditions, and so would not accumulate more upon chitin addition. On the other hand, SA has been shown to dynamically recruit remorins to plasmodesmata, potentially altering the lipid composition of plasmodesmata (Huang *et al.*, 2019). LYM2 may preferentially locate to this SA-induced lipid composition due to its GPI anchor (Zavaliev *et al.*, 2016). In this case, LYM2 movement would be impeded by the overexpression of LYK5 (Cheval *et al.*, 2020). In this case, *lym2* would not be recruited as efficiently to plasmodesmata in a *rem1.2*/*rem1.3* mutant.

5.4 What may be the future direction of plasmodesmata extractions?

The very first extractions of plasmodesmata-enriched tissue were from purified cell wall extract (Monzer & Kloth, 1991; Yahalom *et al.*, 1991). The heart of the process has not changed in the following three decades. The only substantial addition has been the generation of "wall-free plasmodesmata" using cellulases and the modification of buffers (Epel *et al.*, 1995; Faulkner & Bayer, 2017). In this thesis, I presented an extension for this method whereby plasmodesmata from mature tissue could be extracted, using detergents to purify the cell walls from plastids (Table 2-3). This was shown to increase the efficacy of chloroplast removal, as in nuclear preparations (Sikorskaite *et al.*, 2013). However, plasmodesmata lipid preparations from mature tissue have been demonstrated without the addition of detergents (Liu *et al.*, 2020). This confirms that detergents are not required for the extraction, rather to help clean the cell walls.

It would be an interesting exercise to compare the degree of contamination produced in proteomes made by each method. As with all purification protocols, there is a purity-yield trade-off. This is evident in the proteomes generated in this thesis, where the proteomes

contain about 200 proteins (Table 3-1). This is more similar to the filtered proteomes (AtCell2_filtered, PtCells1_filtered), rather than to the whole proteomes which contain closer to 1,000 proteins. Following this line of thought, the preparations of Liu *et al.* (2020) may contain more proteins, but also be more impure.

The use of detergents to remove contaminants also posed the question as to whether the plasmodesmata themselves were being disrupted. This has been observed with electron miscroscopy following stronger detergent treatments (Tilney *et al.*, 1991; Turner *et al.*, 1994). Future work could consider the number of different families of plasmodesmal proteins retrieved in plasmodesmata extraction with and without detergent addition. However, clearly plasmodesmata were not fully removed in the protocol presented here, as plasmodesmata proteins could be observed biochemically (Figure 2-12). When considering the micrographs presented by Tilney *et al.* (1991) and Turner *et al.* (1994), it is reasonable to suggest that they may have been imaging Type I over Type II plasmodesmata, where the cytoplasmic sleeve has collapsed but the desmotubule remains.

Mature plasmodesmata extraction protocols present an opprtunity to explore the effect of mutations on the devlopment of plasmodesmata in greater detail. Previously, there were two methods to do this: cell suspension cultures from mutants and comparative proteomics of cell wall fractions. Theoretically, stable cell suspension cultures could be produced from mutants. However, cell suspension cultures are hard to produce and largely only contain simple plasmodesmata (Bayer *et al.*, 2004), although they contain both Type I and Type II plasmodesmata (Brault *et al.*, 2019). No plasmodesmata mutants have been made into cell suspension cultures to the best of my knowledge. Alternatively, proteomes of cell wall fractions can be compared, such as with *cher1* mutants to find differentially expressed proteins in simple and complex plasmodesmata (Kraner *et al.*, 2017b). A direct plasmodesmata proteome of *cher1* mutants would give greater sensitivity and specificity of plasmodesmal protiens. The benefits of a direct proteome are illustrated in the comparative cell wall proteomes of Col-0 (wild type) and *Atmctp3/Atmctp4* double mutants (Brault *et al.*, 2019). In this experiment, only MCTP4 and not MCTP3 was seen to be reduced in the mutant plasmodesmata, which cannot be possible and must reflect a lack of sensitivity in the methodology.

There are many exciting mutants to profile, such as the *plm* (PHLOEM UNLOADING MODULATOR) mutant. This would directly answer the question of whether C2 lipid-binding proteins are enriched in Type I plasmodesmata, as suggested by an age profile in cell

suspesion cultures (Yan *et al.*, 2019; Brault *et al.*, 2019). Further, an *in vivo* comparison of source and sink plasmodesmata may help elucidate proteins that are unqiue to branched plasmodesmata without the need for a mutant like *cher1*. This experiment may help answer the question as to why GFP traffics more freely through seemingly smaller plasmodesmata (Oparka *et al.*, 1999).

I attempted to extend the plasmodesmata purifications to organelle-specific co-immunoprecipitations (see 4.3.13). While these immunoprecipitations worked well, and an enrichment of bait protein could be observed (Figure 4-24), no prey hits were returned in the proteomic data. Perplexingly, the bait protein was also returned with very few spectra in the proteomic data, despite a strong signal on the Western blot. This occurred with samples sent to two independent proteomic facilities. Thus, either the digested cell wall carbohydrates interfere with mass spectrometry and the downstream sample preparation for plasmodesmata fractions could be improved; or the samples generated are close to the detection limit. In the latter case, the solution may simply be to increase the amount of input material. The low number of proteins in the proteomes suggest this may be the limiting factor. Increasing the quantity of input material is made even simpler by the use of transgenic *Nicotiana benthamiana*, which can be scaled to industrial uses (Huafang Lai & Jake Stahnke, 2013). Therefore, the technique may be easily troubleshooted and would be an exciting line of research to continue.

There is a huge amount of potential in the plasmodesmal co-immunoprecipitations, when scaled to a working level. The utility of the technique has been exemplified on plasmodesmata-enriched tissue, where two proteins in the cell-to-cell movement pathway were identified by co-immunoprecipitation with a *Cucurbita maxima* non–cell-autonomous proteins (Lee *et al.*, 2003; Ham *et al.*, 2012). Their data could have been made more specific for plasmodesmal proteins by further purification. Moreover, plasmodesmal co-immunoprecipitations may allow for the rapid identification for partner proteins of plasmodesmata-interacting effectors (Tomczynska *et al.*, 2020; Ohtsu *et al.*, 2021). This is especially useful when plasmodesmata-interacting proteins are targeted to multiple cell locations (Rosas-Diaz *et al.*, 2018). The difficulty originates from the plasmodesmata fraction of the cell being orders of magnitude smaller compared to the chloroplasts and other components and so plasmodesmata prey proteins are likely to be drowned out without specific purification.

5.5   Final Outlook

Looking back to the aims of this thesis, I hoped to extend plasmodesmata extraction techniques to mature tissue and then use this technique to refine plasmodesmal proteomes. Further, I wanted to investigate the signalling pathways surrounding PDLPs at plasmodesmata.

I am pleased with the development of a new extraction protocol, which can extract plasmodesmata from a variety of species' mature tissue. However, I could not extend the protocol to co-immunoprecipitation experiments. I think as it stands, the ability to biochemically test the localisation of transiently expressed proteins is a very helpful tool. Further, I see the future of plasmodesmata extractions moving into mature tissue to directly assay the difference composition of plasmodesmata between genotypes and treatments, instead of relying upon the comparison of cell wall fractions. With further optimisation, especially an increase in tissue input, the co-immunoprecipitation of interactions specifically at plasmodesmata should be possible.

The proteome presented by Brault *et al.* (2019) (AtCells2_filtered) is regarded as a gold-standard list of 115 proteins that are quantitatively enriched at plasmodesmal fractions. If I were to refine it with the proteome presented here, AtPlant1, the field would be reduced to just three proteins! Thus, I do not think it is helpful to refine an already conservative list of highly likely candidate plasmodesmal proteins. Although, I did produce an alternative refinement from AtCells1 and AtCells2 by finding proteome-recurrent proteins. My work took a phylogenetic angle, in which protein classes conserved in plasmodesmal proteomes across 450 million years of evolution were identified. I think there is tremendous potential in considering the what the constituents of the last common plasmodesmata were. These proteins might point the way towards identifying the core building blocks of plasmodesmata, and characterisation of the elusive spoke-like structures.

Numerous approaches were attempted to elucidate the signalling pathway surrounding the PDLPs. A forward genetic screen for PDLP1 overexpression plants with wild type size produced a potential candidate: KISS ME DEADLY2 (KMD2). Further candidates still remain to be mapped. KMD2 may reverse the dwarf phenotype by being a global regulator of plant size or by altering plasmodesmata density in a cytokinin-mediated pathway, and so warrants further investigation. Forward genetic approaches, again leveraging the dwarfed phenotype of PDLP1, provided candidates that were genetically downstream of PDLP1, reverting the plants to wild-type size. NHL3, a protein previously characterised in plant defence, physically

interacted with PDLP1 and may also interact with PDLP5. Thus, this protein may provide the signalling nexus for PDLP-induced callose deposition. Moreover, as NHL3 and PDLP5 are known to be required for bacterial defence, NHL3 may provide a handle on the signalling cascade downstream of defence-elicited plasmodesmata closure. Ultimately, the uncovering of a PDLP1 interactor provides an opportunity to start probing how PDLP1 interacts with NHL3 to communicate the callose deposition signal.

Plasmodesmata remain enigmatic structures, but the work in this thesis has helped widen our understanding of them. I have identified novel components at plasmodesmata and started to place them in the callose deposition pathway. In addition, I have generated a model of plasmodesmata in the last common ancestor of land plants by comparing the proteomes of divergent species. The results of this thesis provide a foundation for further work investigating the composition of plasmodesmata.

6 References

Abou-Saleh RH, Hernandez-Gomez MC, Amsbury S, Paniagua C, Bourdon M, Miyashima S, Helariutta Y, Fuller M, Budtova T, Connell SD, *et al.* 2018. Interactions between callose and cellulose revealed through the analysis of biopolymer mixtures. *Nature Communications* **9**: 4538.

Ahn G, Kim H, Kim DH, Hanh H, Yoon Y, Singaram I, Wijesinghe KJ, Johnson KA, Zhuang X, Liang Z, *et al.* 2017. SH3 domain-containing protein 2 plays a crucial role at the step of membrane tubulation during cell plate formation. *Plant Cell* **29**: 1388–1405.

Almagro Armenteros JJ, Tsirigos KD, Sønderby CK, Petersen TN, Winther O, Brunak S, von Heijne G, Nielsen H. 2019. SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nature Biotechnology* **37**: 420–423.

Amari K, Boutant E, Hofmann C, Schmitt-Keichinger C, Fernandez-Calvino L, Didier P, Lerich A, Mutterer J, Thomas CL, Heinlein M, *et al.* 2010. A family of plasmodesmal proteins with receptor-like properties for plant viral movement proteins. *PLoS Pathogens* **6**.

Argyros RD, Mathews DE, Chiang YH, Palmer CM, Thibault DM, Etheridge N, Argyros DA, Mason MG, Kieber JJ, Schallera GE. 2008. Type B response regulators of Arabidopsis play key roles in cytokinin signaling and plant development. *Plant Cell* **20**: 2102–2116.

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, *et al.* 2000. Gene ontology: Tool for the unification of biology. *Nature Genetics* **25**: 25–29.

Aung K, Kim P, Li Z, Joe A, Kvitko B, Alfano JR, Hea SY. 2020. Pathogenic bacteria target plant plasmodesmata to colonize and invade surrounding tissues. *Plant Cell* **32**: 595–611.

Austin RS, Vidaurre D, Stamatiou G, Breit R, Provart NJ, Bonetta D, Zhang J, Fung P, Gong Y, Wang PW, *et al.* 2011. Next-generation mapping of Arabidopsis genes. *Plant Journal* **67**: 715–725.

Azim MF, Burch-Smith TM. 2020. Organelles-nucleus-plasmodesmata signaling (ONPS): an update on its roles in plant physiology, metabolism and stress responses. *Current Opinion in Plant Biology* **58**: 48–59.

Badelt K, White RG, Overall RL, Vesk M. 1994. Ultrastructural Specializations of the Cell Wall Sleeve Around Plasmodesmata. *American Journal of Botany* **81**: 1422.

Bajar BT, Wang ES, Zhang S, Lin MZ, Chu J. 2016. A guide to fluorescent protein FRET pairs. *Sensors (Switzerland)* **16**: 1488.

Baluška F, Cvrčková F, Kendrick-Jones J, Volkmann D. 2001. Sink plasmodesmata as gateways for phloem unloading. Myosin VIII and calreticulin as molecular determinants of sink strength? *Plant Physiology* **126**: 39–46.

Bao Y, Song WM, Pan J, Jiang CM, Srivastava R, Li B, Zhu LY, Su HY, Gao XS, Liu H, *et al.* 2016a. Overexpression of the NDR1/HIN1-Like gene NHL6 modifies seed germination in response to abscisic acid and abiotic stresses in arabidopsis. *PLoS ONE* **11**.

Bao Y, Song WM, Zhang HX. 2016b. Role of Arabidopsis NHL family in ABA and stress response. *Plant Signaling and Behavior* **11**.

Bar-On YM, Milo R. 2019. The global mass and average rate of rubisco. *Proceedings of the National Academy of Sciences of the United States of America* **116**: 4738–4743.

Barbez E, Dünser K, Gaidora A, Lendl T, Busch W. **2017**. Auxin steers root cell expansion via apoplastic pH regulation in Arabidopsis thaliana. *Proceedings of the National Academy of Sciences* **114**: E4884–E4893.

Bari R. **2006**. PHO2, MicroRNA399, and PHR1 Define a Phosphate-Signaling Pathway in Plants. *Plant physiology* **141**: 988–999.

Baron-Epel O, Hernandez D, Jiang L-W, Meiners S, Schindler M. **1988**. Dynamic Continuity of Cytoplasmic and Membrane Compartments between Plant Cells. *The Journal of Cell Biology* **106**: 715–721.

Barratt DHP, Kölling K, Graf A, Pike M, Calder G, Findlay K, Zeeman SC, Smith AM. **2011**. Callose synthase GSL7 is necessary for normal phloem transport and inflorescence growth in Arabidopsis. *Plant Physiology* **155**: 328–341.

de Bary A. **1866**. *Handbuch der Physiologischen Botanik, Leipzig, Verlag Wilhelm Engelmann, Morphologie und Physiologie der Pilze, Flechten und Myxomyceten*. Leipzig: Engelmann.

Bauer R, Begerow D, Sampaio JP, Weiβ M, Oberwinkler F. **2006**. The simple-septate basidiomycetes: a synopsis. *Mycological Progress* **5**: 41–66.

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

Bayer E, Thomas CL, Maule AJ. **2004**. Plasmodesmata in Arabidopsis thaliana suspension cells. *Protoplasma* **223**: 93–102.

Bayes T, Price R. **1763**. An essay towards solving a problem in the doctrine of chances. *Philosophical Transactions of the Royal Society of London* **53**: 370–418.

Becker B, Marin B. **2009**. Streptophyte algae and the origin of embryophytes. *Annals of Botany* **103**: 999–1004.

Beckett A. **1981**. The ultrastructure of septal pores and associated structures in the ascogenous hyphae and asci ofSordaria humana. *Protoplasma* **107**: 127–147.

Bell AW, Deutsch EW, Au CE, Kearney RE, Beavis R, Sechi S, Nilsson T, Bergeron JJM, Beardslee TA, Chappell T, *et al.* 2009. A HUPO test sample study reveals common problems in mass spectrometry-based proteomics. *Nature Methods* **6**: 423–430.

Bellati J, Champeyroux C, Hem S, Rofidal V, Krouk G, Maurel C, Santoni V. **2016**. Novel aquaporin regulatory mechanisms revealed by interactomics. *Molecular and Cellular Proteomics* **15**: 3473–3487.

Benitez-Alfonso Y, Faulkner C, Pendle A, Miyashima S, Helariutta Y, Maule A. **2013**. Symplastic intercellular connectivity regulates lateral root patterning. *Developmental cell* **26**: 136–47.

Bertani G. **1951**. Studies on lysogenesis. I. The mode of phage liberation by lysogenic Escherichia coli. *Journal of Bacteriology* **62**: 293–300.

Blackman LM, Gunning BES, Overall RL. **1998**. A 45 kDa protein isolated from the nodal walls of Chara corallina is localised to plasmodesmata. *Plant Journal* **15**: 401–411.

Blackman LM, Harper JDI, Overall RL. **1999**. Localization of a centrin-like protein to higher plant plasmodesmata. *European Journal of Cell Biology* **78**: 297–304.

Blackman LM, Overall RL. **1998**. Immunolocalisation of the cytoskeleton to plasmodesmata

of Chara corallina. *Plant Journal* **14**: 733–741.

Boavida LC, Qin P, Broz M, Becker JD, Mccormick S. **2013**. Arabidopsis tetraspanins are confined to discrete expression domains and cell types in reproductive tissues and form homo and heterodimers when expressed in yeast. *Plant Physiology* **163**: 696–712.

Bolger AM, Lohse M, Usadel B. **2014**. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **30**: 2114–2120.

Botha CEJ, Cross RHM. **2000**. Towards reconciliation of structure with function in plasmodesmata—who is the gatekeeper? *Micron* **31**: 713–721.

Bourdais G, Burdiak P, Gauthier A, Nitsch L, Salojärvi J, Rayapuram C, Idänheimo N, Hunter K, Kimura S, Merilo E, *et al.* 2015. Large-Scale Phenomics Identifies Primary and Fine-Tuning Roles for CRKs in Responses Related to Oxidative Stress. *PLoS Genetics* **11**.

Brault ML, Petit JD, Immel F, Nicolas WJ, Brocard L, Gaston A, Fouché M, Hawkins TJ, Crowet J-M, Grison MS, *et al.* 2018. Multiple C2 domains and Transmembrane region Proteins (MCTPs) tether membranes at plasmodesmata. *bioRxiv*: 423905.

Brault ML, Petit JD, Immel F, Nicolas WJ, Glavier M, Brocard L, Gaston A, Fouché M, Hawkins TJ, Crowet J, *et al.* 2019. Multiple C2 domains and transmembrane region proteins (MCTP s) tether membranes at plasmodesmata. *EMBO reports* **20**: e47182.

Brecknock S, Dibbayawan TP, Vesk M, Vesk PA, Faulkner C, Barton DA, Overall RL. **2011**. High resolution scanning electron microscopy of plasmodesmata. *Planta* **234**: 749–758.

Bremer K, Wanntorp H. **1981**. A cladistic classification of green plants. *Nordic Journal of Botany* **1**: 1–3.

Bricchi I, Occhipinti A, Bertea CM, Zebelo SA, Brillada C, Verrillo F, De Castro C, Molinaro A, Faulkner C, Maule AJ, *et al.* 2013. Separation of early and late responses to herbivory in Arabidopsis by changing plasmodesmal function. *Plant Journal* **73**: 14–25.

Brown NJ, Newell CA, Stanley S, Chen JE, Perrin AJ, Kajala K, Hibberd JM. **2011**. Independent and parallel recruitment of preexisting mechanisms underlying C4 photosynthesis. *Science* **331**: 1436–1439.

Burch-Smith TM, Zambryski PC. **2010**. Loss of increased size exclusion limit (ise)1 or ise2 increases the formation of secondary plasmodesmata. *Current Biology* **20**: 989–993.

Burnell JN. **1988**. An Enzymic Method for Measuring the Molecular Weight Exclusion Limit of Plasmodesmata of Bundle Sheath Cells of C 4 Plants. *Journal of Experimental Botany* **39**: 1575–1580.

Burton RA, Gidley MJ, Fincher GB. **2010**. Heterogeneity in the chemistry, structure and function of plant cell walls. *Nature Chemical Biology* **6**: 724–732.

Caesar K, Thamm AMK, Witthöft J, Elgass K, Huppenberger P, Grefen C, Horak J, Harter K. **2011**. Evidence for the localization of the Arabidopsis cytokinin receptors AHK3 and AHK4 in the endoplasmic reticulum. *Journal of Experimental Botany* **62**: 5571–5580.

Caillaud MC, Wirthmueller L, Sklenar J, Findlay K, Piquerez SJM, Jones AME, Robatzek S, Jones JDG, Faulkner C. **2014**. The plasmodesmal protein PDLP1 localises to haustoria-associated membranes during downy mildew infection and regulates callose deposition. *PLoS pathogens* **10**: e1004496.

Calderwood A, Kopriva S, Morris RJ. **2016**. Transcript abundance explains mRNA mobility

data in Arabidopsis thaliana. *Plant Cell* **28**: 610–615.

Campbell DH. **1924**. A remarkable development of the sporophyte in anthoceros fusiformis, aust. *Annals of Botany* **os**-**38**: 473–483.

Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. **2009**. trimAl: A tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**: 1972–1973.

Carbon S, Douglass E, Dunn N, Good B, Harris NL, Lewis SE, Mungall CJ, Basu S, Chisholm RL, Dodson RJ, *et al.* 2019. The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic Acids Research* **47**: D330–D338.

Carbon S, Mungall C. **2020**. Gene Ontology Data Archive. *Zenodo*.

Carella P, Isaacs M, Cameron RK. **2015**. Plasmodesmata-located protein overexpression negatively impacts the manifestation of systemic acquired resistance and the long-distance movement of Defective in Induced Resistance1 in Arabidopsis. *Plant Biology* **17**: 395–401.

Cargile BJ, Bundy JL, Stephenson JL. **2004**. Potential for false positive identifications from large databases through tandem mass spectrometry. *Journal of Proteome Research* **3**: 1082–1085.

Casero PJ, Knox JP. **1995**. The monoclonal antibody JIM5 indicates patterns of pectin deposition in relation to pit fields at the plasma-membrane-face of tomato pericarp cell walls. *Protoplasma* **188**: 133–137.

Century KS, Shapiro AD, Repetti PP, Dahlbeck D, Holub E, Staskawicz BJ. **1997**. NDR1, a pathogen-induced component required for Arabidopsis disease resistance. *Science (New York, N.Y.)* **278**: 1963–5.

Chang Y, Graham SW. **2011**. Inferring the higher-order phylogeny of mosses (Bryophyta) and relatives using a large, multigene plastid data set. *American Journal of Botany* **98**: 839–849.

Chen Z. **2001**. A superfamily of proteins with novel cysteine-rich repeats. *Plant Physiology* **126**: 473–476.

Chen MH, Citovsky V. **2003**. Systemic movement of a tobamovirus requires host cell pectin methylesterase. *Plant Journal* **35**: 386–392.

Chen K, Du L, Chen Z. **2003**. Sensitization of defense responses and activation of programmed cell death by a pathogen-induced receptor-like protein kinase in Arabidopsis. *Plant Molecular Biology* **53**: 61–74.

Chen K, Fan B, Du L, Chen Z. **2004**. Activation of hypersensitive cell death by pathogen-induced receptor-like protein kinases from Arabidopsis. *Plant Molecular Biology* **56**: 271–283.

Chen XY, Liu L, Lee EK, Han X, Rim Y, Chu H, Kim SW, Sack F, Kim JY. **2009**. The Arabidopsis callose synthase gene gsl8 is required for cytokinesis and cell patterning. *Plant Physiology* **150**: 105–113.

Chen C, Vanneste S, Chen X. **2021**. Membrane tethers control plasmodesmal function and formation. *Plant Science* **304**: 110800.

Cheng CY, Krishnakumar V, Chan AP, Thibaud-Nissen F, Schobel S, Town CD. **2017**. Araport11: a complete reannotation of the Arabidopsis thaliana reference genome. *Plant Journal* **89**: 789–804.

Cheng S, Xian W, Fu Y, Marin B, Keller J, Wu T, Sun W, Li X, Xu Y, Zhang Y, *et al.* 2019.

Genomes of Subaerial Zygnematophyceae Provide Insights into Land Plant Evolution. *Cell* **179**: 1057-1067.e14.

Cheval C, Faulkner C. **2018**. Plasmodesmal regulation during plant-pathogen interactions. *New Phytologist* **217**: 62–67.

Cheval C, Samwald S, Johnston MG, de Keijzer J, Breakspear A, Liu X, Bellandi A, Kadota Y, Zipfel C, Faulkner C. **2020**. Chitin perception in plasmodesmata characterizes submembrane immune-signaling specificity in plants. *Proceedings of the National Academy of Sciences* **117**: 9621–9629.

La Claire II J. **1981**. Occurrence of plasmodesmata during infurrowing in a brown alga. *Biology of the Cell* **40**: 139–142.

Cleland RE, Fujiwara T, Lucas WJ. **1994**. Plasmodesmal-mediated cell-to-cell transport in wheat roots is modulated by anaerobic stress. *Protoplasma* **178**: 81–85.

Clough SJ, Bent AF. **1998**. Floral dip: A simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. *Plant Journal* **16**: 735–743.

Cocquyt E, Verbruggen H, Leliaert F, De Clerck O. **2010**. Evolution and cytological diversification of the green seaweeds (Ulvophyceae). *Molecular Biology and Evolution* **27**: 2052–2061.

Conroy RM. **2012**. What hypotheses do 'nonparametric' two-group tests actually test? *Stata Journal*.

Cook ME, Graham LE. **1999a**. *Evolution of plasmodesmata, in Plasmodesmata: Nanochannels with megatasks*. Springer Verlag, Berlin.

Cook ME, Graham LE. **1999b**. Evolution of Plasmodesmata. In: Plasmodesmata. 101–117.

Cook ME, Graham LE, Botha CEJ, Lavin CA. **1997**. Comparative ultrastructure of plasmodesmata of Chara and selected bryophytes: Toward an elucidation of the evolutionary origin of plant plasmodesmata. *American Journal of Botany* **84**: 1169–1178.

Cooper ED. **2014**. Overly simplistic substitution models obscure green plant phylogeny. *Trends in Plant Science* **19**: 576–582.

Côté R, Thain JF, Fensom DS. **1987**. Increase in electrical resistance of plasmodesmata of Chara induced by an applied pressure gradient across nodes . *Canadian Journal of Botany* **65**: 509–511.

Crawford KM, Zambryski PC. **2000**. Subcellular localization determines the availability of non-targeted proteins to plasmodesmatal transport. *Current Biology* **10**: 1032–1040.

Cruz-Valderrama JE, Gómez-Maqueo X, Salazar-Iribe A, Zúñiga-Sánchez E, Hernández-Barrera A, Quezada-Rodríguez E, Gamboa-Debuen A. **2019**. Overview of the role of cell wall DUF642 proteins in plant development. *International Journal of Molecular Sciences* **20**.

Cui W, Lee JY. **2016**. Arabidopsis callose synthases CalS1/8 regulate plasmodesmal permeability during stress. *Nature Plants* **2**: 16034.

Czernic P, Visser B, Sun W, Savouré A, Deslandes L, Marco Y, Van Montagu M, Verbruggen N. **1999**. Characterization of an Arabidopsis thaliana receptor-like protein kinase gene activated by oxidative stress and pathogen attack. *Plant Journal* **18**: 321–327.

Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, *et al.* 2011. The variant call format and VCFtools. *Bioinformatics* **27**:

2156–2158.

Danila FR, Quick WP, White RG, Furbank RT, von Caemmerer S. **2016**. The metabolite pathway between bundle sheath and mesophyll: Quantification of plasmodesmata in leaves of C3 and C4 monocots. *Plant Cell* **28**: 1461–1471.

Dashevskaya S, Kopito RB, Friedman R, Elbaum M, Epel BL. **2008**. Diffusion of anionic and neutral GFP derivatives through plasmodesmata in epidermal cells of Nicotiana benthamiana. *Protoplasma* **234**: 13–23.

Davis JE, Mcketta JJ. **1960**. Solubility of Ethylene in Water. *Journal of Chemical and Engineering Data* **5**: 374–375.

Deinum EE, Mulder BM, Benitez-Alfonso Y. **2019**. From plasmodesma geometry to effective symplasmic permeability through biophysical modelling. *eLife* **8**.

Delwiche CF, Cooper ED. **2015**. The evolutionary origin of a terrestrial flora. *Current Biology* **25**: R899–R910.

Dettmer J, Ursache R, Campilho A, Miyashima S, Belevich I, O'Regan S, Mullendore DL, Yadav SR, Lanz C, Beverina L, *et al.* 2014. CHOLINE TRANSPORTER-LIKE1 is required for sieve plate development to mediate long-distance cell-to-cell communication. *Nature Communications* **5**: 4276.

Dever L V., Pearson M, Ireland RJ, Leegood RC, Lea PJ. **1998**. The isolation and characterisation of a mutant of the C4 plant Amaranthus edulis deficient in NAD-malic enzyme activity. *Planta* **206**: 649–656.

Devereux R, Loeblich AR, Fox GE. **1990**. Higher plant origins and the phylogeny of green algae. *Journal of Molecular Evolution* **31**: 18–24.

Dick PS, ap Rees T. **1975**. The Pathway of Sugar Transport in Roots of Pisum sativum. *Journal of Experimental Botany* **26**: 305–314.

Ding B, Haudenshield JS, Hull RJ, Wolf S, Beachy RN, Lucas WJ. **1992a**. Secondary plasmodesmata are specific sites of localization of the tobacco mosaic virus movement protein in transgenic tobacco plants. *Plant Cell* **4**: 915–928.

Ding LN, Li M, Wang WJ, Cao J, Wang Z, Zhu KM, Yang YH, Li YL, Tan XL. **2019**. Advances in plant GDSL lipases: from sequences to functional mechanisms. *Acta Physiologiae Plantarum* **41**.

Ding B, Turgeon R, Parthasarathy M V. 1992b. Substructure of freeze-substituted plasmodesmata. *Protoplasma* **169**: 28–41.

Dodds PN, Rathjen JP. **2010**. Plant immunity: towards an integrated view of plant–pathogen interactions. *Nature Reviews Genetics* **11**: 539–548.

Dong X, Hong Z, Chatterjee J, Kim S, Verma DPS. **2008**. Expression of callose synthase genes and its connection with Npr1 signaling pathway during pathogen infection. *Planta* **229**: 87–98.

Dörmann P, Gopalan S, Yang He S, Benning C. **2000**. A gene family in Arabidopsis thaliana with sequence similarity to NDR1 and HIN1. *Plant Physiology and Biochemistry* **38**: 789–796.

Doxey AC, Yaish MWF, Moffatt BA, Griffith M, McConkey BJ. **2007**. Functional Divergence in the Arabidopsis β-1,3-Glucanase Gene Family Inferred by Phylogenetic Reconstruction of Expression States. *Molecular Biology and Evolution* **24**: 1045–1055.

Eddy SR. **1998**. Profile hidden Markov models. *Bioinformatics* **14**: 755–763.

Eddy SR. **2020**. HMMER v3.3.

Edwards D. **2003**. Xylem in early tracheophytes. *Plant, Cell and Environment* **26**: 57–72.

Ehlers K, van Bel AJE. **2010**. Dynamics of plasmodesmal connectivity in successive interfaces of the cambial zone. *Planta* **231**: 371–385.

Ehlers K, Kollmann R. **2001**. Primary and secondary plasmodesmata: structure, origin, and functioning. *Protoplasma* **216**: 1–30.

Eisenberg-Bord M, Shai N, Schuldiner M, Bohnert M. **2016**. A Tether Is a Tether Is a Tether: Tethering at Membrane Contact Sites. *Developmental Cell* **39**: 395–409.

Eisenhaber B, Bork P, Eisenhaber F. **2001**. Post-translational GPI lipid anchor modification of proteins in kingdoms of life: Analysis of protein sequence data from complete genomes. *Protein Engineering* **14**: 17–25.

Ellinger D, Naumann M, Falter C, Zwikowics C, Jamrow T, Manisseri C, Somerville SC, Voigt CA. **2013**. Elevated early callose deposition results in complete penetration resistance to powdery mildew in Arabidopsis. *Plant Physiology* **161**: 1433–1444.

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

Emms DM, Kelly S. **2019**. OrthoFinder: Phylogenetic orthology inference for comparative genomics. *Genome Biology* **20**: 238.

Engler C, Gruetzner R, Kandzia R, Marillonnet S. **2009**. Golden Gate Shuffling: A One-Pot DNA Shuffling Method Based on Type IIs Restriction Enzymes (J Peccoud, Ed.). *PLoS ONE* **4**: e5553.

Engler C, Kandzia R, Marillonnet S. **2008**. A One Pot, One Step, Precision Cloning Method with High Throughput Capability (HA El-Shemy, Ed.). *PLoS ONE* **3**: e3647.

Engler C, Youles M, Gruetzner R, Ehnert T-M, Werner S, Jones JDG, Patron NJ, Marillonnet S. **2014**. A Golden Gate Modular Cloning Toolbox for Plants. *ACS Synthetic Biology* **3**: 839–843.

Epel BL, Erlanger MA. **1991**. Light regulates symplastic communication in etiolated corn seedlings. *Physiologia Plantarum* **83**: 149–153.

Epel BL, Kuchuck B, Kotlizky G, Shurt S, Erlanger M, Yahalom A. **1995**. Isolation and Characterization of Plasmodesmata. *Methods in Cell Biology* **50**: 237–253.

Epel BL, van Lent JWM, Cohen L, Kotlizky G, Katz A, Yahalom A. **1996**. A 41 kDa protein isolated from maize mesocotyl cell walls immunolocalizes to plasmodesmata. *Protoplasma* **191**: 70–78.

Erwee MG, Goodwin PB. **1983**. Characterisation of the Egeria densa Planch. leaf symplast. *Planta* **158**: 320–328.

Faulkner C. **2013**. Receptor-mediated signaling at plasmodesmata. *Frontiers in Plant Science* **4**.

Faulkner C, Akman OE, Bell K, Jeffree C, Oparka K. **2008**. Peeking into pit fields: A multiple twinning model of secondary plasmodesmata formation in tobacco. *Plant Cell* **20**: 1504–1518.

Faulkner C, Bayer EMF. **2017**. Isolation of Plasmodesmata. *Methods in molecular biology (Clifton, N.J.)* **1511**: 187–198.

Faulkner CR, Blackman LM, Collings DA, Cordwell SJ, Overall RL. **2009**. Anti-tropomyosin antibodies co-localise with actin microfilaments and label plasmodesmata. *European Journal of Cell Biology* **88**: 357–369.

Faulkner CR, Blackman LM, Cordwell SJ, Overall RL. **2005**. Proteomic identification of putative plasmodesmatal proteins fromChara corallina. *PROTEOMICS* **5**: 2866–2875.

Faulkner C, Petutschnig E, Benitez-Alfonso Y, Beck M, Robatzek S, Lipka V, Maule AJ. **2013**. LYM2-dependent chitin perception limits molecular flux via plasmodesmata. *Proceedings of the National Academy of Sciences of the United States of America* **110**: 9166–9170.

Felix G, Duran JD, Volko S, Boller T. **1999**. Plants have a sensitive perception system for the most conserved domain of bacterial flagellin. *Plant Journal* **18**: 265–276.

Fernandez-Calvino L, Faulkner C, Walshaw J, Saalbach G, Bayer E, Benitez-Alfonso Y, Maule A. **2011**. Arabidopsis Plasmodesmal Proteome (JD Hoheisel, Ed.). *PLoS ONE* **6**: e18880.

Field KJ, Duckett JG, Cameron DD, Pressel S. **2015**. Stomatal density and aperture in non-vascular land plants are non-responsive to above-ambient atmospheric CO2 concentrations. *Annals of Botany* **115**: 915–922.

Fleurat-Lessard P, Bouché-Pillon S, Leloup C, Lucas WJ, Serrano R, Bonnemain J-L. **1995**. Absence of plasma membrane H+-ATPase in plasmodesmata located in pit-fields of the young reactive pulvinus of Mimosa pudica L. *Protoplasma* **188**: 180–185.

Floyd GL, Stewart KD, Mattox KR. **1971**. Cytokinesis and Plasmodesmata in Ulothrix. *Journal of Phycology* **7**: 306–309.

Folta KM, Kaufman LS. **2006**. Isolation of Arabidopsis nuclei and measurement of gene transcription rates using nuclear run-on assays. *Nature Protocols* **1**: 3094–3100.

Franceschi VR, Ding B, Lucas WJ. **1994**. Mechanism of plasmodesmata formation in characean algae in relation to evolution of intercellular communication in higher plants. *Planta* **192**: 347–358.

Fraser TW, Gunning BES. **1969**. The ultrastructure of plasmodesmata in the filamentous green alga, Bulbochaete hiloensis (Nordst.) tiffany. *Planta* **88**: 244–254.

Fu ZQ, Guo M, Jeong BR, Tian F, Elthon TE, Cerny RL, Staiger D, Alfano JR. **2007**. A type III effector ADP-ribosylates RNA-binding proteins and quells plant immunity. *Nature* **447**: 284–288.

Fu S, Xu Y, Li C, Li Y, Wu J, Zhou X. **2018**. Rice Stripe Virus Interferes with S-acylation of Remorin and Induces Its Autophagic Degradation to Facilitate Virus Infection. *Molecular plant* **11**: 269–287.

Fujimoto M, Arimura SI, Nakazono M, Tsutsumi N. **2008**. Arabidopsis dynamin-related protein DRP2B is co-localized with DRP1A on the leading edge of the forming cell plate. *Plant Cell Reports* **27**: 1581–1586.

Fujimoto M, Arimura SI, Ueda T, Takanashi H, Hayashi Y, Nakano A, Tsutsumi N. **2010**. Arabidopsis dynamin-related proteins DRP2B and DRP1A participate together in clathrin-coated vesicle formation during endocytosis. *Proceedings of the National Academy of Sciences of the United States of America* **107**: 6094–6099.

Galau GA, Hughes DW, Dure L. **1986**. Abscisic acid induction of cloned cotton late embryogenesis-abundant (Lea) mRNAs. *Plant Molecular Biology* **7**: 155–170.

Gaudet P, Livstone MS, Lewis SE, Thomas PD. **2011**. Phylogenetic-based propagation of functional annotations within the Gene Ontology consortium. *Briefings in Bioinformatics* **12**: 449–462.

Gaudioso-Pedraza R, Benitez-Alfonso Y. **2014**. A phylogenetic approach to study the origin and evolution of plasmodesmata-localized glycosyl hydrolases family 17. *Frontiers in Plant Science* **5**: 212.

Van Gestel K, Slegers H, Von Witsch M, Šamaj J, Baluška F, Verbelen JP. **2003**. Immunological evidence for the presence of plant homologues of the actin-related protein Arp3 in tobacco and maize: Subcellular localization to actin-enriched pit fields and emerging root hairs. *Protoplasma* **222**: 45–52.

Giannoutsou E, Sotiriou P, Apostolakos P, Galatis B. **2013**. Early local differentiation of the cell wall matrix defines the contact sites in lobed mesophyll cells of Zea mays. *Annals of Botany* **112**: 1067–1081.

Gómez-Gómez L, Boller T. **2000**. FLS2: An LRR receptor-like kinase involved in the perception of the bacterial elicitor flagellin in Arabidopsis. *Molecular Cell* **5**: 1003–1011.

Goodwin PB, Shepherd V, Erwee MG. **1990**. Compartmentation of fluorescent tracers injected into the epidermal cells of Egeria densa leaves. *Planta* **181**: 129–136.

Gopalan S, Wei W, He SY. **1996**. hrp gene-dependent induction of hin1: a plant gene activated rapidly by both harpins and the avrPto gene-mediated signal. *The Plant Journal* **10**: 591–600.

Grabski S, De Feijter AW, Schindler M. **1993**. Endoplasmic Reticulum Forms a Dynamic Continuum for Lipid Diffusion between Contiguous Soybean Root Cells. *The Plant Cell* **5**: 25–38.

Gray J. **1993**. Major Paleozoic land plant evolutionary bio-events. *Palaeogeography, Palaeoclimatology, Palaeoecology* **104**: 153–169.

Grison MS, Brocard L, Fouillen L, Nicolas W, Wewer V, Dörmann P, Nacir H, Benitez-Alfonso Y, Claverol S, Germain V, *et al.* 2015a. Specific membrane lipid composition is important for plasmodesmata function in Arabidopsis. *The Plant cell* **27**: 1228–50.

Grison MS, Fernandez-Calvino L, Mongrand S, Bayer EMF. **2015b**. Isolation of Plasmodesmata from Arabidopsis Suspension Culture Cells. In: Humana Press, New York, NY, 83–93.

Grison MS, Kirk P, Brault ML, Wu XN, Schulze WX, Benitez-Alfonso Y, Immel F, Bayer EM. **2019**. Plasma membrane-associated receptor-like kinases relocalize to plasmodesmata in response to osmotic stress. *Plant Physiology* **181**: 142–160.

Gronnier J, Crowet J-M, Habenstein B, Nasir MN, Bayle V, Hosy E, Platre MP, Gouguet P, Raffaele S, Martinez D, *et al.* 2017. Structural basis for plant plasma membrane protein dynamics and organization into functional nanodomains. *eLife* **6**: e26404.

Groth-Malonek M, Pruchner D, Grewe F, Knoop V. **2005**. Ancestors of trans-splicing mitochondrial introns support serial sister group relationships of hornworts and mosses with vascular plants. *Molecular Biology and Evolution* **22**: 117–125.

Gui J, Liu C, Shen J, Li L. **2014**. Grain setting defect1, encoding a remorin protein, affects the

209

grain setting in rice through regulating plasmodesmatal conductance. *Plant physiology* **166**: 1463–78.

Gui J, Zheng S, Shen J, Li L. **2015**. Grain setting defect1 (GSD1) function in rice depends on S-acylation and interacts with actin 1 (OsACT1) at its C-terminal. *Frontiers in Plant Science* **6**: 804.

Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. **2010**. New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Systematic Biology* **59**: 307–321.

Guseman JM, Lee JS, Bogenschutz NL, Peterson KM, Virata RE, Xie B, Kanaoka MM, Hong Z, Torii KU. **2010**. Dysregulation of cell-to-cell connectivity and stomatal patterning by loss-of-function mutation in Arabidopsis chorus (GLUCAN SYNTHASE-LIKE 8). *Development* **137**: 1731–1741.

Ham B-K, Li G, Kang B-H, Zeng F, Lucas WJ. **2012**. Overexpression of Arabidopsis Plasmodesmata Germin-Like Proteins Disrupts Root Growth and Development. *The Plant Cell* **24**: 3630–3648.

Han X, Hyun TK, Zhang M, Kumar R, Koh E ji, Kang BH, Lucas WJ, Kim JY. **2014**. Auxin-Callose-Mediated Plasmodesmal Gating Is Essential for Tropic Auxin Gradient Formation and Signaling. *Developmental Cell* **28**: 132–146.

Hanahan D. **1983**. Studies on transformation of Escherichia coli with plasmids. *Journal of Molecular Biology* **166**: 557–580.

Harburger DS, Calderwood DA. **2009**. Integrin signalling at a glance. *Journal of Cell Science* **122**: 159–163.

Harris BJ, Harrison CJ, Hetherington AM, Williams TA. **2020**. Phylogenomic Evidence for the Monophyly of Bryophytes and the Reductive Evolution of Stomata. *Current Biology* **30**: 2001-2012.e2.

Hart A. **2001**. Mann-Whitney test is not just a test of medians: Differences in spread can be important. *British Medical Journal* **323**: 391–393.

Hartung W. **2010**. The evolution of abscisic acid (ABA) and ABA function in lower plants, fungi and lichen. *Functional Plant Biology* **37**: 806–812.

Hepler PK. **1982**. Endoplasmic reticulum in the formation of the cell plate and plasmodesmata. *Protoplasma* **111**: 121–133.

Herger A, Dünser K, Kleine-Vehn J, Ringli C. **2019**. Leucine-Rich Repeat Extensin Proteins and Their Role in Cell Wall Sensing. *Current Biology* **29**: R851–R858.

Hidayat R, Wulandari S, Wiryawan KG, Suryahadi. **2005**. Production and Utilization of Cellulase From Trichoderma Viride. *Biotropia* **0**: 50–59.

Hong Z, Delauney AJ, Verma DPS. **2001**. A cell plate-specific callose synthase and its interaction with phragmoplastin. *Plant Cell* **13**: 755–768.

Hori K, Maruyama F, Fujisawa T, Togashi T, Yamamoto N, Seo M, Sato S, Yamada T, Mori H, Tajima N, *et al.* **2014**. Klebsormidium flaccidum genome reveals primary factors for plant terrestrial adaptation. *Nature Communications* **5**: 3978.

Huafang Lai QC, Jake Stahnke JH. **2013**. Agroinfiltration as an Effective and Scalable Strategy of Gene Delivery for Production of Pharmaceutical Proteins. *Advanced Techniques in Biology*

& Medicine **01**.

Huang J, Fujimoto M, Fujiwara M, Fukao Y, Arimura SI, Tsutsumi N. **2015**. Arabidopsis dynamin-related proteins, DRP2A and DRP2B, function coordinately in post-Golgi trafficking. *Biochemical and Biophysical Research Communications* **456**: 238–244.

Huang D, Sun Y, Ma Z, Ke M, Cui Y, Chen Z, Chen C, Ji C, Tran TM, Yang L, *et al.* 2019. Salicylic acid-mediated plasmodesmal closure via Remorin-dependent lipid organization. *Proceedings of the National Academy of Sciences of the United States of America* **116**: 21274–21284.

Huang S, Yuan S, Dong M, Su J, Yu C, Shen Y, Xie X, Yu Y, Yu X, Chen S, *et al.* 2005. The phylogenetic analysis of tetraspanins projects the evolution of cell-cell interactions from unicellular to multicellular organisms. *Genomics* **86**: 674–684.

Huerta-Cepas J, Serra F, Bork P. **2016**. ETE 3: Reconstruction, Analysis, and Visualization of Phylogenomic Data. *Molecular Biology and Evolution* **33**: 1635–1638.

Humphry M, Bednarek P, Kemmerling B, Koh S, Stein M, Göbel U, Stüber K, Pislewska-Bednarek M, Loraine A, Schulze-Lefert P, *et al.* 2010. A regulon conserved in monocot and dicot plants defines a functional module in antifungal plant immunity. *Proceedings of the National Academy of Sciences of the United States of America* **107**: 21896–901.

Hundertmark M, Hincha DK. **2008**. LEA (Late Embryogenesis Abundant) proteins and their encoding genes in Arabidopsis thaliana. *BMC Genomics* **9**.

Hunter K, Kimura S, Rokka A, Tran HC, Toyota M, Kukkonen JP, Wrzaczek M. **2019**. CRK2 enhances salt tolerance by regulating callose deposition in connection with PLDα1. *Plant Physiology* **180**: 2004–2021.

Hynes RO. **2002**. Integrins: Bidirectional, allosteric signaling machines. *Cell* **110**: 673–687.

Idänheimo N, Gauthier A, Salojärvi J, Siligato R, Brosché M, Kollist H, Mähönen AP, Kangasjärvi J, Wrzaczek M. **2014**. The Arabidopsis thaliana cysteine-rich receptor-like kinases CRK6 and CRK7 protect against apoplastic oxidative stress. *Biochemical and Biophysical Research Communications* **445**: 457–462.

Ignatoski KMW, Verderame MF. **1996**. Lysis buffer composition dramatically affects extraction of phosphotyrosine-containing proteins. *BioTechniques* **20**: 794–796.

Imlau A, Truernit E, Sauer N. **1999**. Cell-to-cell and long-distance trafficking of the green fluorescent protein in the phloem and symplastic unloading of the protein into sink tissues. *Plant Cell* **11**: 309–322.

Inoue T, Higuchi M, Hashimoto Y, Seki M, Kobayashi M, Kato T, Tabata S, Shinozaki K, Kakimoto T. **2001**. Identification of CRE1 as a cytokinin receptor from Arabidopsis. *Nature* **409**: 1060–1063.

Jackson D, Veit B, Hake S. **1994**. Expression of maize KNOTTED1 related homeobox genes in the shoot apical meristem predicts patterns of morphogenesis in the vegetative shoot. *Development* **120**: 405–413.

Jarsch IK, Ott T. **2011**. Perspectives on remorin proteins, membrane rafts, and their role during plant-microbe interactions. *Molecular Plant-Microbe Interactions* **24**: 7–12.

Jarvis P, House CR. **1970**. Evidence for Symplasmic Ion Transport in Maize Roots. *Journal of Experimental Botany* **21**: 83–90.

Jia F, Qi S, Li H, Liu P, Li P, Wu C, Zheng C, Huang J. **2015**. Overexpression of Late

Embryogenesis Abundant 14 enhances Arabidopsis salt stress tolerance. *Biochemical and Biophysical Research Communications* **454**: 505–511.

Jiao C, Sørensen I, Sun X, Sun H, Behar H, Alseekh S, Philippe G, Palacio Lopez K, Sun L, Reed R, *et al.* 2020. The Penium margaritaceum Genome: Hallmarks of the Origins of Land Plants. *Cell* **181**: 1097-1111.e12.

Jo Y, Cho WK, Rim Y, Moon J, Chen X-Y, Chu H, Kim CY, Park Z-Y, Lucas WJ, Kim J-Y. 2011. Plasmodesmal receptor-like kinases identified through analysis of rice cell wall extracted proteins. *Protoplasma* **248**: 191–203.

Johnston MG, Faulkner C. 2021. A bootstrap approach is a superior statistical method for the comparison of non-normal data with differing variances. *New Phytologist* **230**: 23–26.

Jones MGK. 1976. The Origin and Development of Plasmodesmata. In: Intercellular Communication in Plants: Studies on Plasmodesmata. 81–105.

Jones P, Binns D, Chang HY, Fraser M, Li W, McAnulla C, McWilliam H, Maslen J, Mitchell A, Nuka G, *et al.* 2014a. InterProScan 5: Genome-scale protein function classification. *Bioinformatics* **30**: 1236–1240.

Jones MGK, Dropkin VH. 1976. Scanning electron microscopy of nematode induced giant transfer cells. *Cytobios* **15**: 149–161.

Jones AM, Xuan Y, Xu M, Wang RS, Ho CH, Lalonde S, You CH, Sardi MI, Parsa SA, Smith-Valle E, *et al.* 2014b. Border control - A membrane-linked interactome of Arabidopsis. *Science* **344**: 711–716.

Ju C, Van De Poel B, Cooper ED, Thierer JH, Gibbons TR, Delwiche CF, Chang C. 2015. Conservation of ethylene as a plant hormone over 450 million years of evolution. *Nature Plants* **1**: 14004.

Junková P, Daněk M, Kocourková D, Brouzdová J, Kroumanová K, Zelazny E, Janda M, Hynek R, Martinec J, Valentová O. 2018. Mapping of plasma membrane proteins interacting with Arabidopsis thaliana flotillin 2. *Frontiers in Plant Science* **9**.

Kadota Y, Macho AP, Zipfel C. 2016. Immunoprecipitation of plasma membrane receptor-like kinases for identification of phosphorylation sites and associated proteins. In: Methods in Molecular Biology. 133–144.

Kadota Y, Sklenar J, Derbyshire P, Stransfeld L, Asai S, Ntoukakis V, Jones JD, Shirasu K, Menke F, Jones A, *et al.* 2014. Direct Regulation of the NADPH Oxidase RBOHD by the PRR-Associated Kinase BIK1 during Plant Immunity. *Molecular Cell* **54**: 43–55.

Kajala K, Brown NJ, Williams BP, Borrill P, Taylor LE, Hibberd JM. 2012. Multiple Arabidopsis genes primed for recruitment into C 4 photosynthesis. *Plant Journal* **69**: 47–56.

Karol KG, McCourt RM, Cimino MT, Delwiche CF. 2001. The closest living relatives of land plants. *Science* **294**: 2351–2353.

Kawakami S, Watanabe Y, Beachy RN. 2004. Tobacco mosaic virus infection spreads cell to cell as intact replication complexes. *Proceedings of the National Academy of Sciences of the United States of America* **101**: 6291–6296.

Keller A, Nesvizhskii AI, Kolker E, Aebersold R. 2002. Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Analytical Chemistry* **74**: 5383–5392.

Kenrick P, Crane PR. **1991**. Water-conducting cells in early fossil land plants: implications for the early evolution of tracheophytes. *Botanical Gazette* **152**: 335–356.

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

Kenrick P, Wellman CH, Schneider H, Edgecombe GD. **2012**. A timeline for terrestrialization: Consequences for the carbon cycle in the Palaeozoic. *Philosophical Transactions of the Royal Society B: Biological Sciences* **367**: 519–536.

Kikkert JR. **1993**. The Biolistic® PDS-1000/He device. *Plant Cell, Tissue and Organ Culture* **33**: 221–226.

Kim HJ, Chiang YH, Kieber JJ, Schaller GE. **2013**. SCFKMD controls cytokinin signaling by regulating the degradation of type-B response regulators. *Proceedings of the National Academy of Sciences of the United States of America* **110**: 10028–10033.

Kim HJ, Ryu H, Hong SH, Woo HR, Lim PO, Lee IC, Sheen J, Nam HG, Hwang I. **2006**. Cytokinin-mediated control of leaf longevity by AHK3 through phosphorylation of ARR2 in Arabidopsis. *Proceedings of the National Academy of Sciences of the United States of America* **103**: 814–819.

Kim J-Y, Yuan Z, Jackson D. **2003**. Developmental regulation and significance of KNOX protein trafficking in Arabidopsis. *Development (Cambridge, England)* **130**: 4351–62.

Kimura S, Hunter K, Vaahtera L, Tran HC, Citterico M, Vaattovaara A, Rokka A, Stolze SC, Harzen A, Meißner L, *et al.* 2020. CRK2 and C-terminal phosphorylation of NADPH oxidase RBOHD regulate reactive oxygen species production in arabidopsis. *Plant Cell* **32**: 1063–1080.

Kirsch R, Vurmaz E, Schaefer C, Eberl F, Sporer T, Haeger W, Pauchet Y. **2020**. Plants use identical inhibitors to protect their cell wall pectin against microbes and insects. *Ecology and Evolution* **10**: 3814–3824.

Kishi-Kaboshi M, Murata T, Hasebe M, Watanabe Y. **2005**. An extraction method for tobacco mosaic virus movement protein localizing in plasmodesmata. *Protoplasma* **225**: 85–92.

Kitagawa M, Fujita T. **2013**. Quantitative imaging of directional transport through plasmodesmata in moss protonemata via single-cell photoconversion of Dendra2. *Journal of Plant Research* **126**: 577–585.

Kitagawa M, Fujita T. **2015**. A model system for analyzing intercellular communication through plasmodesmata using moss protonemata and leaves. *Journal of Plant Research* **128**: 63–72.

Kitagawa M, Tomoi T, Fukushima T, Sakata Y, Sato M, Toyooka K, Fujita T, Sakakibara H. **2019**. Abscisic Acid Acts as a Regulator of Molecular Trafficking through Plasmodesmata in the Moss Physcomitrella patens. *Plant and Cell Physiology* **60**: 738–751.

Knepper C, Savory EA, Day B. **2011**. Arabidopsis NDR1 is an integrin-like protein with a role in fluid loss and plasma membrane-cell wall adhesion. *Plant Physiology* **156**: 286–300.

Knox JP, Benitez-Alfonso Y. **2014**. Roles and regulation of plant cell walls surrounding plasmodesmata. *Current Opinion in Plant Biology* **22**: 93–100.

Knox K, Wang P, Kriechbaumer V, Tilsner J, Frigerio L, Sparkes I, Hawes C, Oparka K. **2015**. Putting the squeeze on plasmodesmata: A role for reticulons in primary plasmodesmata formation. *Plant Physiology* **168**: 1563–1572.

Koncz C, Martini N, Szabados L, Hrouda M, Bachmair A, Schell J. **1994**. Specialized vectors for gene tagging and expression studies. In: Plant Molecular Biology Manual. 53–74.

Kotlizky G, Shurtz S, Yahalom A, Malik Z, Traub O, Epel BL. **1992**. An improved procedure for the isolation of plasmodesmata embedded in clean maize cell walls. *The Plant Journal* **2**: 623–630.

Kraner ME, Link K, Melzer M, Ekici AB, Uebe S, Tarazona P, Feussner I, Hofmann J, Sonnewald U. **2017a**. Choline transporter-like1 (CHER1) is crucial for plasmodesmata maturation in Arabidopsis thaliana. *Plant Journal* **89**: 394–406.

Kraner ME, Müller C, Sonnewald U. **2017b**. Comparative proteomic profiling of the choline transporter-like1 (CHER1) mutant provides insights into plasmodesmata composition of fully developed *Arabidopsis thaliana* leaves. *The Plant Journal* **92**: 696–709.

Krogh A, Larsson B, Von Heijne G, Sonnhammer ELL. **2001**. Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. *Journal of Molecular Biology* **305**: 567–580.

Krügel U, He HX, Gier K, Reins J, Chincinska I, Grimm B, Schulze WX, Kühn C. **2012**. The potato sucrose transporter StSUT1 interacts with a DRM-associated protein disulfide isomerase. *Molecular Plant* **5**: 43–62.

Kummer D, Steinbacher T, Schwietzer MF, Thölmann S, Ebnet K. **2020**. Tetraspanins: integrating cell surface receptors to functional microdomains in homeostasis and disease. *Medical Microbiology and Immunology* **209**: 397–405.

Kwiatkowska M. **2003**. Plasmodesmal changes are related to different developmental stages of antheridia of Chara species. *Protoplasma* **222**: 1–11.

Kwiatkowska M, Maszewski J. **1976**. Plasmodesmata between synchronously and asynchronously developing cells of the antheridial filaments of Chara vulgaris L. *Protoplasma* **87**: 317–327.

Laemmi UK. **1970**. Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. *Nature* **227**: 680–685.

Lai CP, Huang LM, Chen LFO, Chan MT, Shaw JF. **2017**. Genome-wide analysis of GDSL-type esterases/lipases in Arabidopsis. *Plant Molecular Biology* **95**: 181–197.

Lam AJ, St-Pierre F, Gong Y, Marshall JD, Cranfill PJ, Baird MA, McKeown MR, Wiedenmann J, Davidson MW, Schnitzer MJ, *et al.* **2012**. Improving FRET dynamic range with bright green and red fluorescent proteins. *Nature Methods* **9**: 1005–1012.

Laurin-Lemay S, Brinkmann H, Philippe H. **2012**. Origin of land plants revisited in the light of sequence contamination and missing data. *Current Biology* **22**: R593–R594.

Lee RE. **1971**. The pit connections of some lower red algae: Ultrastructure and phylogenetic significance. *British Phycological Journal* **6**: 29–38.

Lee JY. **2014**. New and old roles of plasmodesmata in immunity and parallels to tunneling nanotubes. *Plant Science* **221**–**222**: 13–20.

Lee JY. **2015**. Plasmodesmata: A signaling hub at the cellular boundary. *Current Opinion in Plant Biology* **27**: 133–140.

Lee HY, Bowen CH, Popescu GV, Kang HG, Kato N, Ma S, Dinesh-Kumar S, Snyder M, Popescu SC. **2011a**. Arabidopsis RTNLB1 and RTNLB2 reticulon-like proteins regulate

intracellular trafficking and activity of the FLS2 immune receptor. *Plant Cell* **23**: 3374–3391.

Lee SM, Hoang MHT, Han HJ, Kim HS, Lee K, Kim KE, Kim DH, Lee SY, Chung WS. **2009**. Pathogen inducible voltage-dependent anion channel (AtVDAC) isoforms are localized to mitochondria membrane in Arabidopsis. *Molecules and Cells* **27**: 321–327.

Lee HJ, Kim JS, Yoo SJ, Kang EY, Han SH, Yang K-Y, Kim YC, McSpadden Gardener B, Kang H. **2012**. Different roles of glycine-rich RNA-binding protein7 in plant defense against Pectobacterium carotovorum, Botrytis cinerea, and tobacco mosaic viruses. *Plant Physiology and Biochemistry* **60**: 46–52.

Lee J-Y, Taoka K, Yoo B-C, Ben-Nissan G, Kim D-J, Lucas WJ. **2005**. Plasmodesmal-Associated Protein Kinase in Tobacco and Arabidopsis Recognizes a Subset of Non-Cell-Autonomous Proteins. *The Plant Cell* **17**: 2817–2831.

Lee JY, Wang X, Cui W, Sager R, Modla S, Czymmek K, Zybaliov B, Van Wijk K, Zhang C, Lu H, *et al.* 2011b. A plasmodesmata-localized protein mediates crosstalk between cell-to-cell communication and innate immunity in arabidopsis. *Plant Cell* **23**: 3353–3373.

Lee J-Y, Yoo B-C, Rojas MR, Gomez-Ospina N, Staehelin LA, Lucas WJ, Hake S. **2003**. Selective trafficking of non-cell-autonomous proteins mediated by NtNCAPP1. *Science (New York, N.Y.)* **299**: 392–6.

Leebens-Mack JH, Barker MS, Carpenter EJ, Deyholos MK, Gitzendanner MA, Graham SW, Grosse I, Li Z, Melkonian M, Mirarab S, *et al.* 2019. One thousand plant transcriptomes and the phylogenomics of green plants. *Nature* **574**: 679–685.

Legrand A. **2020**. Anchoring mechanism of the plant protein remorin to membrane nanodomains.

Leijon F, Melzer M, Zhou Q, Srivastava V, Bulone V. **2018**. Proteomic analysis of plasmodesmata from populus cell suspension cultures in relation with callose biosynthesis. *Frontiers in Plant Science* **871**: 1681.

Lemieux C, Otis C, Turmel M. **2007**. A clade uniting the green algae Mesostigma viride and Chlorokybus atmophyticus represents the deepest branch of the Streptophyta in chloroplast genome-based phylogenies. *BMC Biology* **5**: 2.

Levy A, Erlanger M, Rosenthal M, Epel BL. **2007**. A plasmodesmata-associated β-1,3-glucanase in Arabidopsis. *Plant Journal* **49**: 669–682.

Levy A, Zheng JY, Lazarowitz SG. **2015**. Synaptotagmin SYTA Forms ER-Plasma Membrane Junctions that Are Recruited to Plasmodesmata for Plant Virus Movement. *Current Biology* **25**: 2018–2025.

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

Li H. **2011**. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* **27**: 2987–2993.

Li H, Durbin R. **2009**. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**: 1754–1760.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. **2009**. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**: 2078–2079.

Li J, Lee JY, Liao L. **2020**. Detecting De Novo Plasmodesmata Targeting Signals and Identifying PD Targeting Proteins. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). 1–12.

Liarzi O, Epel BL. **2005**. Development of a quantitative tool for measuring changes in the coefficient of conductivity of plasmodesmata induced by developmental, biotic, and abiotic signals. *Protoplasma* **225**: 67–76.

Ligrone R, Duckett JG. **1994**. Thallus differentiation in the marchantialean liverwort asterella wilmsii (Steph.) with particular reference to longitudinal arrays of endoplasmic microtubules in the inner cells. *Annals of Botany* **73**: 577–586.

Ligrone R, Duckett JG. **1998**. The leafy stems of Sphagnum (Bryophyta) contain highly differentiated polarized cells with axial arrays of endoplasmic microtubules. *New Phytologist* **140**: 567–579.

Ligrone R, Duckett JG, Renzaglia KS. **2012**. Major transitions in the evolution of early land plants: A bryological perspective. *Annals of Botany* **109**: 851–871.

Lim GH, Shine MB, De Lorenzo L, Yu K, Cui W, Navarre D, Hunt AG, Lee JY, Kachroo A, Kachroo P. **2016**. Plasmodesmata Localizing Proteins Regulate Transport and Signaling during Systemic Acquired Immunity in Plants. *Cell Host and Microbe* **19**: 541–549.

Lincoln FC. **1930**. Calculating waterfowl abundance on the basis of banding returns. *U.S. Department of Agriculture* **118**: 1–4.

Liu L, Li C, Liang Z, Yu H. **2018**. Characterization of Multiple C2 Domain and Transmembrane Region Proteins in Arabidopsis. *Plant Physiology* **176**: 2119–2132.

Liu L, Liu C, Hou X, Xi W, Shen L, Tao Z, Wang Y, Yu H. **2012**. FTIP1 is an essential regulator required for florigen transport (O Leyser, Ed.). *PLoS Biology* **10**: e1001313.

Liu N-J, Zhang T, Liu Z-H, Chen X, Guo H-S, Ju B-H, Zhang Y-Y, Li G-Z, Zhou Q-H, Qin Y-M, *et al.* 2020. Phytosphinganine Affects Plasmodesmata Permeability via Facilitating PDLP5-Stimulated Callose Accumulation in Arabidopsis. *Molecular Plant* **13**: 128–143.

Long J, Walker J, She W, Aldridge B, Gao H, Deans S, Feng X. **2021**. Nurse cell-derived small RNAs define paternal epigenetic inheritance in Arabidopsis. *bioRxiv*.

López-Sáez JF, Giménez-Martin G, Risueño MC. **1966**. Fine structure of the plasmodesm. *Protoplasma* **61**: 81–84.

Lu K-J, De Rybel B, van Mourik H, Weijers D. **2018**. Regulation of intercellular TARGET OF MONOPTEROS 7 protein transport in the Arabidopsis root. *Development (Cambridge, England)* **145**: dev152892.

Lucas WJ, Bouché-Pillon S, Jackson DP, Nguyen L, Baker L, Ding B, Hake S. **1995**. Selective trafficking of KNOTTED1 homeodomain protein and its mRNA through plasmodesmata. *Science (New York, N.Y.)* **270**: 1980–3.

Lucas WJ, Ding B, Van Der Schoot C. **1993**. Plasmodesmata and the supracellular nature of plants. *New Phytologist* **125**: 435–476.

Ma Y, Szostkiewicz I, Korte A, Moes D, Yang Y, Christmann A, Grill E. **2009**. Regulators of PP2C phosphatase activity function as abscisic acid sensors. *Science* **324**: 1064–1068.

Ma LS, Wang L, Trippel C, Mendoza-Mendoza A, Ullmann S, Moretti M, Carsten A, Kahnt J, Reissmann S, Zechmann B, *et al.* 2018. The Ustilago maydis repetitive effector Rsp3 blocks

the antifungal activity of mannose-binding maize proteins. *Nature Communications* **9**.

Madeira F, Park YM, Lee J, Buso N, Gur T, Madhusoodanan N, Basutkar P, Tivey ARN, Potter SC, Finn RD, *et al.* **2019**. The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Research* **47**: W636–W641.

Malcevschi A, Marmiroli N. **2012**. Plant Protein Analysis. In: Proteomic Applications in Biology.

Manford AG, Stefan CJ, Yuan HL, MacGurn JA, Emr SD. **2012**. ER-to-Plasma Membrane Tethering Proteins Regulate Cell Signaling and ER Morphology. *Developmental Cell* **23**: 1129–1140.

Manhart JR, Palmer JD. **1990**. The gain of two chloroplast tRNA introns marks the green algal ancestors of land plants. *Nature* **345**: 268–270.

Mann HB, Whitney DR. **1947**. On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other. *The Annals of Mathematical Statistics* **18**: 50–60.

Marchant HJ. **1976**. Plasmodesmata in Algae and Fungi. In: Intercellular Communication in Plants: Studies on Plasmodesmata. 59–80.

Marchant HJ, Pickett-Heaps JD. **1973**. Mitosis and Cytokinesis in Coleochaete scutata. *Journal of Phycology* **9**: 461–471.

Martínez-Fernández I, de Moura SM, Alves-Ferreira M, Ferrándiz C, Balanzà V. **2020**. Identification of players controlling meristem arrest downstream of the FRUITFULL-APETALA2 pathway. *Plant Physiology* **184**: 945–959.

Mattox K, Stewart K. **1984**. Classification of the green algae: a concept based on comparative cytology. In: Systematics of the green algae. 29–72.

May MJ, Leaver CJ. **1993**. Oxidative stimulation of glutathione synthesis in Arabidopsis thaliana suspension cultures. *Plant Physiology* **103**: 621–627.

Meadows CW, Mingardon F, Garabedian BM, Baidoo EEK, Benites VT, Rodrigues A V., Abourjeily R, Chanal A, Lee TS. **2018**. Discovery of novel geranylgeranyl reductases and characterization of their substrate promiscuity 06 Biological Sciences 0605 Microbiology 06 Biological Sciences 0601 Biochemistry and Cell Biology. *Biotechnology for Biofuels* **11**: 340.

Mellor NL, Voß U, Janes G, Bennett MJ, Wells DM, Band LR. **2020**. Auxin fluxes through plasmodesmata modify root-tip auxin distribution. *Development (Cambridge, England)* **147**.

Meng WJ, Cheng ZJ, Sang YL, Zhang MM, Rong XF, Wang ZW, Tang YY, Zhang XS. **2017**. Type-B ARABIDOPSIS RESPONSE REGULATORs specify the shoot stem cell niche by dual regulation of WUSCHEL. *Plant Cell* **29**: 1357–1372.

Meyer RA. **1979**. Light scattering from biological cells: dependence of backscatter radiation on membrane thickness and refractive index. *Applied Optics* **18**: 585.

Mi H, Muruganujan A, Ebert D, Huang X, Thomas PD. **2019**. PANTHER version 14: More genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. *Nucleic Acids Research* **47**: D419–D426.

Miller NG, Behrensmeyer AK, Damuth JD, DiMichele WA, Potts R, Sues H-D, Wing SL. **1996**. *Terrestrial Ecosystems through Time. Evolutionary Paleoecology of Terrestrial Plants and Animals*.

Miller CO, Skoog F, Okumura FS, Von Saltza MH, Strong FM. **1956**. Isolation, Structure and

Synthesis of Kinetin, a Substance Promoting Cell Division. *Journal of the American Chemical Society* **78**: 1375–1380.

Mitra SK, Gantt JA, Ruby JF, Clouse SD, Goshe MB. **2007**. Membrane proteomic analysis of Arabidopsis thaliana using alternative solubilization techniques. *Journal of Proteome Research* **6**: 1933–1950.

Miya A, Albert P, Shinya T, Desaki Y, Ichimura K, Shirasu K, Narusaka Y, Kawakami N, Kaku H, Shibuya N. **2007**. CERK1, a LysM receptor kinase, is essential for chitin elicitor signaling in Arabidopsis. *Proceedings of the National Academy of Sciences of the United States of America* **104**: 19613–19618.

Miyakawa T, Hatano KI, Miyauchi Y, Suwa YI, Sawano Y, Tanokura M. **2014**. A secreted protein with plant-specific cysteine-rich motif functions as a mannose-binding lectin that exhibits antifungal activity. *Plant Physiology* **166**: 766–778.

Möckli N, Deplazes A, Auerbach D. **2008**. Finding new protein interactions using the DUALhunter system. *Nature Methods* **5**: i–ii.

Mok MC. **1994**. Cytokinins and Plant Development — An Overview. In: Cytokinins — Chemistry, activity, and function. Boca Raton, FL: CRC Press, 155–166.

De Moliner F, Knox K, Gordon D, Lee M, Tipping WJ, Geddis A, Reinders A, Ward JM, Oparka K, Vendrell M. **2021**. A Palette of Minimally-Tagged Sucrose Analogues for Real-time Raman Imaging of Intracellular Plant Metabolism. *Angewandte Chemie International Edition* **n/a**: anie.202016802.

Monzer J, Kloth S. **1991**. The Preparation of Plasmodesmata from Plant Tissue Homogenates: Access to the Biochemical Characterization of Plasmodesmata-Related Polypeptides. *Botanica Acta* **104**: 82–84.

Moody LA. **2020**. Three-dimensional growth: a developmental innovation that facilitated plant terrestrialization. *Journal of Plant Research* **133**: 283–290.

Moore-Gordon CS, Cowan AK, Bertling I, Botha CEJ, Cross RHM. **1998**. Symplastic solute transport and avocado fruit development: A decline in cytokinin/ABA ratio is related to appearance of the Hass small fruit variant. *Plant and Cell Physiology* **39**: 1027–1038.

Morris JL, Puttick MN, Clark JW, Edwards D, Kenrick P, Pressel S, Wellman CH, Yang Z, Schneider H, Donoghue PCJ. **2018**. The timescale of early land plant evolution. *Proceedings of the National Academy of Sciences of the United States of America* **115**: E2274–E2283.

Morvan O, Quentin M, Jauneau A, Mareck A, Morvan C. **1998**. Immunogold localization of pectin methylesterases in the cortical tissues of flax hypocotyl. *Protoplasma* **202**: 175–184.

Murashige T, Skoog F. **1962**. A Revised Medium for Rapid Growth and Bio Assays with Tobacco Tissue Cultures. *Physiologia Plantarum* **15**: 473–497.

Nakajima K, Sena G, Nawy T, Benfey PN. **2001**. Intercellular movement of the putative transcription factor SHR in root patterning. *Nature* **413**: 307–311.

Nicaise V, Joe A, Jeong BR, Korneli C, Boutrot F, Westedt I, Staiger D, Alfano JR, Zipfel C. **2013**. Pseudomonas HopU1 modulates plant immune receptor levels by blocking the interaction of their mRNAs with GRP7. *EMBO Journal* **32**: 701–712.

Nicolas WJ, Grison MS, Bayer EM. **2018**. Shaping intercellular channels of plasmodesmata: the structure-to-function missing link. *Journal of Experimental Botany* **69**: 91–103.

Nicolas WJ, Grison MS, Trépout S, Gaston A, Fouché M, Cordelières FP, Oparka K, Tilsner J, Brocard L, Bayer EM. **2017**. Architecture and permeability of post-cytokinesis plasmodesmata lacking cytoplasmic sleeves. *Nature Plants* **3**: 17082.

Nishimura MT, Stein M, Hou BH, Vogel JP, Edwards H, Somerville SC. **2003**. Loss of a callose synthase results in salicylic acid-dependent disease resistance. *Science* **301**: 969–972.

Nishiyama T, Sakayama H, de Vries J, Buschmann H, Saint-Marcoux D, Ullrich KK, Haas FB, Vanderstraeten L, Becker D, Lang D, *et al.* 2018. The Chara Genome: Secondary Complexity and Implications for Plant Terrestrialization. *Cell* **174**: 448-464.e24.

Nishiyama T, Wolf PG, Kugita M, Sinclair RB, Sugita M, Sugiura C, Wakasugi T, Yamada K, Yoshinaga K, Yamaguchi K, *et al.* 2004. Chloroplast phylogeny indicates that bryophytes are monophyletic. *Molecular Biology and Evolution* **21**: 1813–1819.

O'Lexy R, Kasai K, Clark N, Fujiwara T, Sozzani R, Gallagher KL. **2018**. Exposure to heavy metal stress triggers changes in plasmodesmatal permeability via deposition and breakdown of callose. *Journal of Experimental Botany* **69**: 3715–3728.

Ohtsu M, Jennings J, Johnston MG, Liu X, Hughes N, Stark K, Morris RJ, de Keijzer J, Faulkner C. **2021**. Colletotrichum higginsianum effectors exhibit cell-to-cell hypermobility in plant tissues and modulate intercellular connectivity amongst a variety of cellular processes. *bioRxiv*: 2021.01.13.426415.

Oliveros JC. **2007**. Vebby. An interactive tool for comparing lists with Venn's diagrams. *bioinfogp.cnb.csic.es/tools/venny/index.html*.

Oparka KJ, Prior DAM. **1987**. 14C sucrose efflux from the perimedulla of growing potato tubers. *Plant, Cell and Environment* **10**: 667–675.

Oparka KJ, Roberts AG, Boevink P, Cruz SS, Roberts I, Pradel KS, Imlau A, Kotlizky G, Sauer N, Epel B. **1999**. Simple, but not branched, plasmodesmata allow the nonspecific trafficking of proteins in developing tobacco leaves. *Cell* **97**: 743–754.

Orfila C, Knox JP. **2000**. Spatial regulation of pectic polysaccharides in relation to pit fields in cell walls of tomato fruit pericarp. *Plant physiology* **122**: 775–81.

Ormenese S, Bernier G, Périlleux C. **2006**. Cytokinin application to the shoot apical meristem of Sinapis alba enhances secondary plasmodesmata formation. *Planta* **224**: 1481–1484.

Ortiz L, Pavan M, McCarthy L, Timmons J, Densmore DM. **2017**. Automated Robotic Liquid Handling Assembly of Modular DNA Devices. *Journal of Visualized Experiments*.

Otegui MS, Mastronarde DN, Kang BH, Bednarek SY, Staehelin LA. **2001**. Three-dimensional analysis of syncytial-type cell plates during endosperm cellularization visualized by high resolution electron tomography. *Plant Cell* **13**: 2033–2051.

Otulak-Kozieł K, Kozieł E, Bujarski JJ. **2018**. Spatiotemporal changes in xylan-1/xyloglucan and xyloglucan xyloglucosyl transferase (XTH-Xet5) as a step-in of ultrastructural cell wall remodelling in potato–potato virus Y (PVYNTN) hypersensitive and susceptible reaction. *International Journal of Molecular Sciences* **19**.

Overall RL, Blackman LM. **1996**. A model of the macromolecular structure of plasmodesmata. *Trends in Plant Science* **1**: 307–311.

Overall RL, Wolfe J, Gunning BES. **1982**. Intercellular communication in Azolla roots: I. Ultrastructure of plasmodesmata. *Protoplasma* **111**: 134–150.

Palmer IA, Shang Z, Fu ZQ. **2017**. Salicylic acid-mediated plant defense: Recent developments, missing links, and future outlook. *Frontiers in Biology* **12**: 258–270.

Pankratenko A V., Atabekova AK, Morozov SY, Solovyev AG. **2020**. Membrane Contacts in Plasmodesmata: Structural Components and Their Functions. *Biochemistry (Moscow)* **85**: 531–544.

Park SY, Fung P, Nishimura N, Jensen DR, Fujii H, Zhao Y, Lumba S, Santiago J, Rodrigues A, Chow TFF, *et al.* 2009. Abscisic acid inhibits type 2C protein phosphatases via the PYR/PYL family of START proteins. *Science* **324**: 1068–1071.

Van Peer AF, Wang F, Van Driel KGA, De Jong JF, Van Donselaar EG, Müller WH, Boekhout T, Lugones LG, Wösten HAB. **2009**. The septal pore cap is an organelle that functions in vegetative growth and mushroom formation of the wood-rot fungus Schizophyllum commune. *Environmental Microbiology* **12**: 833–844.

Pendle AF, Shaw PJ. **2017**. Isolation of Nuclei and Nucleoli. In: Humana Press, New York, NY, 31–44.

Perbal MC, Haughn G, Saedler H, Schwarz-Sommer Z. **1996**. Non-cell-autonomous function of the Antirrhinum floral homeotic proteins DEFICIENS and GLOBOSA is exerted by their polar cell-to-cell trafficking. *Development (Cambridge, England)* **122**: 3433–41.

Pérez-Sancho J, Tilsner J, Samuels AL, Botella MA, Bayer EM, Rosado A. **2016**. Stitching Organelles: Organization and Function of Specialized Membrane Contact Sites in Plants. *Trends in Cell Biology* **26**: 705–717.

Perraki A, Binaghi M, Mecchia MA, Gronnier J, German-Retana S, Mongrand S, Bayer E, Zelada AM, Germain V. **2014**. StRemorin1.3 hampers Potato virus X TGBp1 ability to increase plasmodesmata permeability, but does not interfere with its silencing suppressor activity. *FEBS Letters* **588**: 1699–1705.

Perraki A, Gronnier J, Gouguet P, Boudsocq M, Deroubaix A-F, Simon V, German-Retana S, Zipfel C, Bayer E, Mongrand S, *et al.* 2017. The plant calcium-dependent protein kinase CPK3 phosphorylates REM1.3 to restrict viral infection. *bioRxiv*: 205765.

Petersen CGJ. **1896**. The yearly immigration of young plaice into the Limfjord from the German Sea. *Report of the Danish Biological Station* **6**: 1–77.

Pickett-Heaps J. **1968**. Ultrastructure and Differentiation in Chara (Fibrosa) IV. Spermatogenesis. *Australian Journal of Biological Sciences* **21**: 655.

Pierleoni A, Martelli P, Casadio R. **2008**. PredGPI: A GPI-anchor predictor. *BMC Bioinformatics* **9**.

Pike LJ. **2006**. Rafts defined: A report on the Keystone symposium on lipid rafts and cell function. *Journal of Lipid Research* **47**: 1597–1598.

Płóciennikowska A, Zdioruk MI, Traczyk G, Światkowska A, Kwiatkowska K. **2015**. LPS-induced clustering of CD14 triggers generation of PI(4,5)P2. *Journal of Cell Science* **128**: 4096–4111.

Porter KR, Palade GE. **1957**. Studies on the endoplasmic reticulum. III. Its form and distribution in striated muscle cells. *The Journal of biophysical and biochemical cytology* **3**: 269–300.

Potter SC, Luciani A, Eddy SR, Park Y, Lopez R, Finn RD. **2018**. HMMER web server: 2018 update. *Nucleic Acids Research* **46**: W200–W204.

Prinz WA. **2014**. Bridging the gap: Membrane contact sites in signaling, metabolism, and organelle dynamics. *Journal of Cell Biology* **205**: 759–769.

Puttick MN, Morris JL, Williams TA, Cox CJ, Edwards D, Kenrick P, Pressel S, Wellman CH, Schneider H, Pisani D, *et al.* 2018. The Interrelationships of Land Plants and the Nature of the Ancestral Embryophyte. *Current Biology* **28**: 733-745.e2.

Qiu YL, Cho Y, Cox JC, Palmer JD. **1998**. The gain of three mitochondrial introns identifies liverworts as the earliest land plants. *Nature* **394**: 671–674.

Qiu YL, Li L, Wang B, Chen Z, Knoop V, Groth-Malonek M, Dombrovska O, Lee J, Kent L, Rest J, *et al.* 2006. The deepest divergences in land plants inferred from phylogenomic evidence. *Proceedings of the National Academy of Sciences of the United States of America* **103**: 15511–15516.

R Core Team. **2020**. R: A Language and Environment for Statistical Computing.

Radford JE, Vesk M, Overall RL. **1998**. Callose deposition at plasmodesmata. *Protoplasma* **201**: 30–37.

Radford JE, White RG. **2001**. Effects of tissue-preparation-induced callose synthesis on est plasmodesma size exclusion limits. *Protoplasma* **216**: 47–55.

Raffaele S, Bayer E, Lafarge D, Cluzet S, German Retana S, Boubekeur T, Leborgne-Castel N, Carde J-P, Lherminier J, Noirot E, *et al.* 2009. Remorin, a solanaceae protein resident in membrane rafts and plasmodesmata, impairs potato virus X movement. *The Plant cell* **21**: 1541–55.

Raffaele S, Mongrand S, Gamas P, Niebel A, Ott T. **2007**. Genome-wide annotation of remorins, a plant-specific protein family: evolutionary and functional perspectives. *Plant physiology* **145**: 593–600.

Rappsilber J, Mann M. **2002**. What does it mean to identify a protein in proteomics? *Trends in Biochemical Sciences* **27**: 74–78.

Raven JA. **1997**. Miniview: Multiple origins of plasmodesmata. *European Journal of Phycology* **32**: 95–101.

Raven JA. **2005**. Evolution of Plasmodesmata. In: Plasmodesmata. Oxford, UK: Blackwell Publishing Ltd, 33–52.

Reid RJ, Overall RL. **1992**. Intercellular communication in Chara: factors affecting transnodal electrical resistance and solute fluxes. *Plant, Cell & Environment* **15**: 507–517.

Richmond TA, Somerville CR. **2000**. The Cellulose Synthase Superfamily. *Plant Physiology* **124**: 495–498.

Riefler M, Novak O, Strnad M, Schmülling T. **2006**. Arabidopsis cytokinin receptors mutants reveal functions in shoot growth, leaf senescence, seed size, germination, root development, and cytokinin metabolism. *Plant Cell* **18**: 40–54.

Rinne PLH, Van der Schoot C. **1998**. Symplasmic fields in the tunica of the shoot apical meristem coordinate morphogenetic events. *Development* **125**: 1477–1485.

Ritzenthaler C, Findlay K, Roberts K, Maule AJ. **2000**. Rapid detection of plasmodesmata in purified cell walls. *Protoplasma* **211**: 165–171.

Roberts IM, Boevink P, Roberts AG, Sauer N, Reichel C, Oparka KJ. **2001**. Dynamic changes in the frequency and architecture of plasmodesmata during the sink-source transition in

tobacco leaves. *Protoplasma* **218**: 31–44.

Robinson-Beers K, Evert RF. **1991**. Fine structure of plasmodesmata in mature leaves of sugarcane. *Planta* **184**: 307–318.

Rosas-Diaz T, Zhang D, Fan P, Wang L, Ding X, Jiang Y, Jimenez-Gongora T, Medina-Puche L, Zhao X, Feng Z, *et al.* **2018**. A virus-targeted plant receptor-like kinase promotes cell-to-cell spread of RNAi. *Proceedings of the National Academy of Sciences of the United States of America* **115**: 1388–1393.

Rosenfeld J, Capdevielle J, Guillemot JC, Ferrara P. **1992**. In-gel digestion of proteins for internal sequence analysis after one- or two-dimensional gel electrophoresis. *Analytical Biochemistry* **203**: 173–179.

Ross-Elliott TJ, Jensen KH, Haaning KS, Wager BM, Knoblauch J, Howell AH, Mullendore DL, Monteith AG, Paultre D, Yan D, *et al.* **2017**. Phloem unloading in Arabidopsis roots is convective and regulated by the phloem-pole pericycle. *eLife* **6**: e24125.

Rowland-Bamford AJ, Baker JT, Allen LH, Bowes G. **1991**. Acclimation of rice to changing atmospheric carbon dioxide concentration. *Plant, Cell & Environment* **14**: 577–583.

Roy S, Watada AE, Wergin WP. **1997**. Characterization of the cell wall microdomain surrounding plasmodesmata in apple fruit. *Plant Physiology* **114**: 539–547.

Rubinstein C V., Gerrienne P, de la Puente GS, Astini RA, Steemans P. **2010**. Early Middle Ordovician evidence for land plants in Argentina (eastern Gondwana). *New Phytologist* **188**: 365–369.

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

Saatian B, Austin RS, Tian G, Chen C, Nguyen V, Kohalmi SE, Geelen D, Cui Y. **2018**. Analysis of a novel mutant allele of GSL8 reveals its key roles in cytokinesis and symplastic trafficking in Arabidopsis. *BMC Plant Biology* **18**.

Sager R, Wang X, Hill K, Yoo BC, Caplan J, Nedo A, Tran T, Bennett MJ, Lee JY. **2020**. Auxin-dependent control of a plasmodesmal regulator creates a negative feedback loop modulating lateral root emergence. *Nature Communications* **11**.

Schatlowski N, Stahl Y, Hohenstatt ML, Goodrich J, Schuberta D. **2010**. The CURLY LEAF interacting protein BLISTER controls expression of polycomb-group target genes and cellular differentiation of Arabidopsis thaliana. *Plant Cell* **22**: 2291–2305.

Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, *et al.* **2012**. Fiji: An open-source platform for biological-image analysis. *Nature Methods* **9**: 676–682.

Schmitz K, Kühn R. **1982**. Fine structure, distribution and frequency of plasmodesmata and pits in the cortex of Laminaria hyperborea and L. saccharina. *Planta* **154**: 385–392.

Schmitz K, Srivastava LM. 1974. Fine structure and development of sieve tubes in Laminaria groenlandica Rosenv. *Cytobiologie* **10**: 66–87.

Schneeberger K, Ossowski S, Ott F, Klein JD, Wang X, Lanz C, Smith LM, Cao J, Fitz J, Warthmann N, *et al.* **2011**. Reference-guided assembly of four diverse Arabidopsis thaliana genomes. *Proceedings of the National Academy of Sciences of the United States of America* **108**: 10249–10254.

Schneider CA, Rasband WS, Eliceiri KW. **2012**. NIH Image to ImageJ: 25 years of image analysis. *Nature Methods* **9**: 671–675.

Schuster-Böckler B, Schultz J, Rahmann S. **2004**. HMM logos for visualization of protein families. *BMC Bioinformatics* **5**.

Searle BC. **2010**. Scaffold: A bioinformatic tool for validating MS/MS-based proteomic studies. *Proteomics* **10**: 1265–1269.

Sedbrook JC, Carroll KL, Hung KF, Masson PH, Somerville CR. **2002**. The Arabidopsis SKU5 gene encodes an extracellular glycosyl phosphatidylinositol-anchored glycoprotein involved in directional root growth. *Plant Cell* **14**: 1635–1648.

Seguí-Simarro JM, Austin JR, White EA, Staehelin LA. **2004**. Electron tomographic analysis of somatic cell plate formation in meristematic cells of arabidopsis preserved by high-pressure freezing. *Plant Cell* **16**: 836–856.

Shepherd VA, Goodwin PB. **1992**. Seasonal patterns of cell-to-cell communication in Chara corallina Klein ex Willd. I. Cell-to-cell communication in vegetative lateral branches during winter and spring. *Plant, Cell & Environment* **15**: 137–150.

Shi D, Jouannet V, Agustí J, Kaul V, Levitsky V, Sanchez P, Mironova V, Greb T. **2020**. Tissue-specific transcriptome profiling of the Arabidopsis thaliana inflorescence stem reveals local cellular signatures. *bioRxiv*: 2020.02.10.941492.

Shijie Y, Mingyi L, Xiaoying Z. **1995**. Changes in plasmodesmatal permeability of the staminal hairs of Setcreasea purpurea during development. *Zhi wu sheng li xue bao = Acta phytophysiologica Sinica* **21**: 355—362.

Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, *et al.* 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology* **7**: 539.

Sikorskaite S, Rajamäki M-L, Baniulis D, Stanys V, Valkonen JP. **2013**. Protocol: Optimised methodology for isolation of nuclei from leaves of species in the Solanaceae and Rosaceae families. *Plant Methods* **9**: 31.

Simon Andrews. **2020**. FastQC: A Quality Control tool for High Throughput Sequence Data. *https://www.bioinformatics.babraham.ac.uk/projects/fastqc/*.

Simons K, Ikonen E. **1997**. Functional rafts in cell membranes. *Nature* **387**: 569–572.

Simpson C, Thomas C, Findlay K, Bayer E, Maule AJ. **2009**. An Arabidopsis GPI-anchor plasmodesmal neck protein with callose binding activity and potential to regulate cell-to-cell trafficking. *The Plant cell* **21**: 581–94.

Singh S, Cornilescu CC, Tyler RC, Cornilescu G, Tonelli M, Lee MS, Markley JL. **2005**. Solution structure of a late embryogenesis abundant protein (LEA14) from Arabidopsis thaliana , a cellular stress-related protein. *Protein Science* **14**: 2601–2609.

Singh RK, Miskolczi P, Maurya JP, Bhalerao RP. **2019**. A Tree Ortholog of SHORT VEGETATIVE PHASE Floral Repressor Mediates Photoperiodic Control of Bud Dormancy. *Current Biology* **29**: 128-133.e2.

Singh N, Swain S, Singh A, Nandi AK. **2018**. AtOZF1 Positively Regulates Defense Against Bacterial Pathogens and NPR1-Independent Salicylic Acid Signaling. *Molecular Plant-Microbe Interactions* **31**: 323–333.

Sinha N, Hake S. **1990**. Mutant characters of Knotted maize leaves are determined in the innermost tissue layers. *Developmental Biology* **141**: 203–210.

Sonnhammer EL, von Heijne G, Krogh A. **1998**. A hidden Markov model for predicting transmembrane helices in protein sequences. *Proceedings / ... International Conference on Intelligent Systems for Molecular Biology ; ISMB. International Conference on Intelligent Systems for Molecular Biology* **6**: 175–182.

Sonnhammer ELL, Koonin E V. 2002. Orthology, paralogy and proposed classification for paralog subtypes. *Trends in Genetics*.

Sørensen TH, Cruys-Bagger N, Windahl MS, Badino SF, Borch K, Westh P. **2015**. Temperature effects on kinetic parameters and substrate affinity of Cel7A cellobiohydrolases. *Journal of Biological Chemistry* **290**: 22193–22202.

Sousa F, Civáň P, Brazão J, Foster PG, Cox CJ. **2020a**. The mitochondrial phylogeny of land plants shows support for Setaphyta under composition-heterogeneous substitution models. *PeerJ* **2020**: e8995.

Sousa F, Civáň P, Foster PG, Cox CJ. **2020b**. The Chloroplast Land Plant Phylogeny: Analyses Employing Better-Fitting Tree- and Site-Heterogeneous Composition Models. *Frontiers in Plant Science* **11**: 1062.

de Sousa F, Foster PG, Donoghue PCJ, Schneider H, Cox CJ. **2019**. Nuclear protein phylogenies support the monophyly of the three bryophyte groups (Bryophyta Schimp.). *New Phytologist* **222**: 565–575.

Spanswick RM, Costerton JW. **1967**. Plasmodesmata in Nitella translucens: structure and electrical resistance. *Journal of Cell Science* **2**: 451–464.

Sparkes I, Tolley N, Aller I, Svozil J, Osterrieder A, Botchway S, Mueller C, Frigerio L, Hawes C. **2010**. Five Arabidopsis reticulon isoforms share endoplasmic reticulum location, topology, and membrane-shaping properties. *Plant Cell* **22**: 1333–1343.

Stagljar I, Korostensky C, Johnsson N, Te Heesen S. **1998**. A genetic system based on split-ubiquitin for the analysis of interactions between membrane proteins in vivo. *Proceedings of the National Academy of Sciences of the United States of America* **95**: 5187–5192.

Stahl Y, Grabowski S, Bleckmann A, Kühnemuth R, Weidtkamp-Peters S, Pinto KG, Kirschner GK, Schmid JB, Wink RH, Hülsewede A, *et al.* 2013. Moderation of arabidopsis root stemness by CLAVATA1 and ARABIDOPSIS CRINKLY4 receptor kinase complexes. *Current Biology* **23**: 362–371.

Stark C, Breitkreutz BJ, Reguly T, Boucher L, Breitkreutz A, Tyers M. **2006**. BioGRID: a general repository for interaction datasets. *Nucleic acids research* **34**.

Steemans P, Hérissé A Le, Melvin J, Miller MA, Paris F, Verniers J, Wellman CH. **2009**. Origin and radiation of the earliest vascular land plants. *Science* **324**: 353.

Stewart KD, Mattox KR, Floyd GL. **1973a**. Mitosis, Cytokinesis, the Distribution of Plasmodesmata, and Other Cytological Characteristics in the Ulotrichales, Ulvales, and Chaetophorales: Phylogenetic and Taxonomic Considerations. *Journal of Phycology* **9**: 128–141.

Stewart KD, Mattox KR, Floyd GL. **1973b**. Mitosis, Cytokinesis, the Distribution of Plasmodesmata, and Other Cytological Characteristics in the Ulotrichales, Ulvales, and Chaetophorales: Phylogenetic and Taxonomic Considerations. *Journal of Phycology* **9**: 128–

141.

Stonebloom S, Burch-Smith T, Kim I, Meinke D, Mindrinos M, Zambryski P. **2009**. Loss of the plant DEAD-box protein ISE1 leads to defective mitochondria and increased cell-to-cell transport via plasmodesmata. *Proceedings of the National Academy of Sciences of the United States of America* **106**: 17229–17234.

De Storme N, Geelen D. **2014**. Callose homeostasis at plasmodesmata: Molecular regulators and developmental relevance. *Frontiers in Plant Science* **5**.

Streitner C, Danisman S, Wehrle F, Schöning JC, Alfano JR, Staiger D. **2008**. The small glycine-rich RNA binding protein AtGRP7 promotes floral transition in Arabidopsis thaliana. *Plant Journal* **56**: 239–250.

Suhling K, Siegel J, Phillips D, French PMW, Lévêque-Fort S, Webb SED, Davis DM. **2002**. Imaging the environment of green fluorescent protein. *Biophysical Journal* **83**: 3589–3595.

Takebe I, Otsuki Y, Aoki S. **1968**. Isolation of tobacco mesophyll cells in intact and active state. *Plant and Cell Physiology* **9**: 115–124.

Tanaka H, Osakabe Y, Katsura S, Mizuno S, Maruyama K, Kusakabe K, Mizoi J, Shinozaki K, Yamaguchi-Shinozaki K. **2012**. Abiotic stress-inducible receptor-like kinases negatively control ABA signaling in Arabidopsis. *Plant Journal* **70**: 599–613.

Tangl E. **1880**. Ueber offene Communicationen zwischen den Zellen des Endosperms einiger Samen. *Jahrbücher für Wissenschaftliche Botanik* **12**: 170–190.

Tateda C, Zhang Z, Shrestha J, Jelenska J, Chinchilla D, Greenberg JT. **2014**. Salicylic acid regulates Arabidopsis microbial pattern receptor kinase levels and signaling. *Plant Cell* **26**: 4171–4187.

Taylor LL, Banwart SA, Valdes PJ, Leake JR, Beerling DJ. **2012**. Evaluating the effects of terrestrial ecosystems, climate and carbon dioxide on weathering over geological time: A global-scale process-based approach. *Philosophical Transactions of the Royal Society B: Biological Sciences* **367**: 565–582.

Terauchi M, Nagasato C, Kajimura N, Mineyuki Y, Okuda K, Katsaros C, Motomura T. **2012**. Ultrastructural study of plasmodesmata in the brown alga Dictyota dichotoma (Dictyotales, Phaeophyceae). *Planta* **236**: 1013–1026.

Terry BR, Robards AW. **1987**. Hydrodynamic radius alone governs the mobility of molecules through plasmodesmata. *Planta* **171**: 145–157.

Thiele K, Wanner G, Kindzierski V, Jürgens G, Mayer U, Pachl F, Assaad FF. **2009**. The timely deposition of callose is essential for cytokinesis in Arabidopsis. *Plant Journal* **58**: 13–26.

Thieme CJ, Rojas-Triana M, Stecyk E, Schudoma C, Zhang W, Yang L, Minãmbres M, Walther D, Schulze WX, Paz-Ares J, *et al.* 2015. Endogenous Arabidopsis messenger RNAs transported to distant tissues. *Nature Plants* **1**.

Thomas CL, Bayer EM, Ritzenthaler C, Fernandez-Calvino L, Maule AJ. **2008**. Specific targeting of a plasmodesmal protein affecting cell-to-cell communication (D Weigel, Ed.). *PLoS Biology* **6**: 0180–0190.

Thomas PD, Campbell MJ, Kejariwal A, Mi H, Karlak B, Daverman R, Diemer K, Muruganujan A, Narechania A. **2003**. PANTHER: A library of protein families and subfamilies indexed by function. *Genome Research* **13**: 2129–2141.

Thomas PD, Kejariwal A, Guo N, Mi H, Campbell MJ, Muruganujan A, Lazareva-Ulitsky B. **2006**. Applications for protein sequence-function evolution data: mRNA/protein expression analysis and coding SNP scoring tools. *Nucleic Acids Research* **34**: W645–W650.

Tietz A, Ruttkowski U, Kohler R, Kasprik W. **1989**. Further Investigations on the Occurrence and the Effects of Abscisic Acid in Algae. *Biochemie und Physiologie der Pflanzen* **184**: 259–266.

Tilney LG, Cooke TJ, Connelly PS, Tilney MS. **1991**. The structure of plasmodesmata as revealed by plasmolysis, detergent extraction, and protease digestion. *Journal of Cell Biology*.

Tilsner J, Amari K, Torrance L. **2011**. Plasmodesmata viewed as specialised membrane adhesion sites. *Protoplasma* **248**: 39–60.

Timme RE, Bachvaroff TR, Delwiche CF. **2012**. Broad phylogenomic sampling and the sister lineage of land plants (S Joly, Ed.). *PLoS ONE* **7**: e29696.

To JPC, Haberer G, Ferreira FJ, Deruère J, Mason MG, Schaller GE, Alonso JM, Ecker JR, Kieber JJ. **2004**. Type-A Arabidopsis response regulators are partially redundant negative regulators of cytokinin signaling. *Plant Cell* **16**: 658–671.

Tolley N, Sparkes I, Craddock CP, Eastmond PJ, Runions J, Hawes C, Frigerio L. **2010**. Transmembrane domain length is responsible for the ability of a plant reticulon to shape endoplasmic reticulum tubules in vivo. *Plant Journal* **64**: 411–418.

Tolley N, Sparkes IA, Hunter PR, Craddock CP, Nuttall J, Roberts LM, Hawes C, Pedrazzini E, Frigerio L. **2008**. Overexpression of a plant reticulon remodels the lumen of the cortical endoplasmic reticulum but does not perturb protein transport. *Traffic* **9**: 94–102.

Tomczynska I, Stumpe M, Doan TG, Mauch F. **2020**. A Phytophthora effector protein promotes symplastic cell-to-cell trafficking by physical interaction with plasmodesmata-localised callose synthases. *New Phytologist* **227**: 1467–1478.

Tomoi T, Kawade K, Kitagawa M, Sakata Y, Tsukaya H, Fujita T. **2020**. Quantitative Imaging Reveals Distinct Contributions of SnRK2 and ABI3 in Plasmodesmatal Permeability in Physcomitrella patens. *Plant and Cell Physiology* **61**: 942–956.

Torres MA, Dangl JL, Jones JDG. **2002**. Arabidopsis gp91phox homologues Atrbohd and Atrbohf are required for accumulation of reactive oxygen intermediates in the plant defense response. *Proceedings of the National Academy of Sciences of the United States of America* **99**: 517–522.

Toyota M, Spencer D, Sawai-Toyota S, Jiaqi W, Zhang T, Koo AJ, Howe GA, Gilroy S. **2018**. Glutamate triggers long-distance, calcium-based plant defense signaling. *Science* **361**: 1112–1115.

Trehin C, Schrempp S, Chauvet A, Berne-Dedieu A, Thierry AM, Faure JE, Negrutiu I, Morel P. **2013**. QUIRKY interacts with STRUBBELIG and PAL OF QUIRKY to regulate cell growth anisotropy during Arabidopsis gynoecium development. *Development (Cambridge)* **140**: 4807–4817.

Tucker EB. **1988**. Inositol bisphosphate and inositol trisphosphate inhibit cell-to-cell passage of carboxyfluorescein in staminal hairs ofSetcreasea purpurea. *Planta* **174**: 358–363.

Turmel M, Otis C, Lemieux C. **2006**. The chloroplast genome sequence of Chara vulgaris sheds new light into the closest green algal relatives of land plants. *Molecular Biology and Evolution* **23**: 1324–1338.

Turner A, Wells B, Roberts K. **1994**. Plasmodesmata of maize root tips: Structure and composition. *Journal of Cell Science* **107**: 3351–3361.

Tylewicz S, Petterle A, Marttila S, Miskolczi P, Azeez A, Singh RK, Immanen J, Mähler N, Hvidsten TR, Eklund DM, *et al.* 2018. Photoperiodic control of seasonal growth is mediated by ABA acting on cell-cell communication. *Science* **360**: 212–215.

Tyree MT. **1970**. The symplast concept a general theory of symplastic transport according to the thermodynamics of irreversible processes. *Journal of Theoretical Biology* **26**: 181–214.

Tyree MT, Tammes PML. **1975**. Translocation of uranin in the symplasm of staminal hairs of Tradescantia. *Canadian Journal of Botany* **53**: 2038–2046.

Ueguchi C, Sato S, Kato T, Tabata S. **2001**. The AHK4 gene involved in the cytokinin-signaling pathway as a direct receptor molecule in Arabidopsis thaliana. *Plant and Cell Physiology* **42**: 751–755.

Umen JG. **2014**. Green Algae and the Origins of Multicellularity in the Plant Kingdom. *Cold Spring Harbor Perspectives in Biology* **6**: a016170–a016170.

Upton C, Buckley JT. **1995**. A new family of lipolytic enzymes? *Trends in Biochemical Sciences* **20**: 178–179.

Vaattovaara A, Brandt B, Rajaraman S, Safronov O, Veidenberg A, Luklová M, Kangasjärvi J, Löytynoja A, Hothorn M, Salojärvi J, *et al.* 2019. Mechanistic insights into the evolution of DUF26-containing proteins in land plants. *Communications Biology*.

Vaddepalli P, Herrmann A, Fulton L, Oelschner M, Hillmer S, Stratil TF, Fastner A, Hammes UZ, Ott T, Robinson DG, *et al.* 2014. The C2-domain protein QUIRKY and the receptor-like kinase STRUBBELIG localize to plasmodesmata and mediate tissue morphogenesis in Arabidopsis thaliana. *Development (Cambridge)*.

Vannucci G, Basso D, Fravega P. **2000**. New observations on the anatomy of the fossil calcareous alga Subterraniphyllum Elliott. *Rivista Italiana di Paleontologia e Stratigrafia* **106**: 237–245.

Varet A, Hause B, Hause G, Scheel D, Lee J. **2003**. The Arabidopsis NHL3 gene encodes a plasma membrane protein and its overexpression correlates with increased resistance to Pseudomonas syringae pv. tomato DC3000. *Plant Physiology* **132**: 2023–2033.

Vatén A, Dettmer J, Wu S, Stierhof YD, Miyashima S, Yadav SR, Roberts CJ, Campilho A, Bulone V, Lichtenberger R, *et al.* 2011. Callose Biosynthesis Regulates Symplastic Trafficking during Root Development. *Developmental Cell* **21**: 1144–1155.

Vilaine F, Kerchev P, Clément G, Batailler B, Cayla T, Bill L, Gissot L, Dinant S. **2013**. Increased expression of a phloem membrane protein encoded by NHL26 alters phloem export and sugar partitioning in Arabidopsis. *Plant Cell* **25**: 1689–1708.

Voesenek LACJ, Benschop JJ, Bou J, Cox MCH, Groeneveld HW, Millenaar FF, Vreeburg RAM, Peeters AJM. **2003**. Interactions between plant hormones regulate submergence-induced shoot elongation in the flooding-tolerant dicot Rumex palustris. *Annals of Botany* **91**: 205–211.

Vogt AB, Spindeldreher S, Kropshofer H. **2002**. Clustering of MHC-peptide complexes prior to their engagement in the immunological synapse: lipid raft and tetraspan microdomains. *Immunological Reviews* **189**: 136–151.

Vu HM, Boedi IAB, Yean KJ, Ritesh K, Jinsu L, Shuwei W. **2019**. CML20: CML20 gene from

Arabidopsis thaliana for controlling plant tropism and root growth and uses thereof.

Vu MH, Iswanto ABB, Lee J, Kim JY. **2020**. The role of plasmodesmata-associated receptor in plant development and environmental response. *Plants* **9**.

Waigmann E, Lucas WJ, Citovsky V, Zambryski P. **1994**. Direct functional assay for tobacco mosaic virus cell-to-cell movement protein and identification of a domain involved in increasing plasmodesmal permeability. *Proceedings of the National Academy of Sciences of the United States of America* **91**: 1433–1437.

Waigmann E, Turner A, Peart J, Roberts K, Zambryski P. **1997**. Ultrastructural analysis of leaf trichome plasmodesmata reveals major differences from mesophyll plasmodesmata. *Planta* **203**: 75–84.

Wang YH. **2008**. How effective is T-DNA insertional mutagenesis in Arabidopsis? *Journal of Biochemical Technology* **1**: 11–20.

Wang FX, Luo YM, Ye ZQ, Cao X, Liang JN, Wang Q, Wu Y, Wu JH, Wang HY, Zhang M, *et al.* 2018. iTRAQ-based proteomics analysis of autophagy-mediated immune responses against the vascular fungal pathogen Verticillium dahliae in Arabidopsis. *Autophagy* **14**: 598–618.

Wang F, Muto A, van de Velde J, Neyt P, Himanen K, Vandepoele K, Van Lijsebettens M. **2015**. Functional analysis of the Arabidopsis tetraspanin gene family in plant growth and development. *Plant Physiology* **169**: 2200–2214.

Wang X, Robles Luna G, Arighi CN, Lee J-Y. **2020**. An evolutionarily conserved motif is required for Plasmodesmata-located protein 5 to regulate cell-to-cell movement. *Communications Biology* **3**: 291.

Wang X, Sager R, Cui W, Zhang C, Lu H, Lee JY. **2013**. Salicylic acid regulates plasmodesmata closure during innate immune responses in Arabidopsis. *Plant Cell* **25**: 2315–2329.

Weber E, Engler C, Gruetzner R, Werner S, Marillonnet S. **2011**. A modular cloning system for standardized assembly of multigene constructs. *PLoS ONE* **6**.

Weiner H, Burnell JN, Woodrow IE, Heldt HW, Hatch MD. **1988**. Metabolite Diffusion into Bundle Sheath Cells from C4 Plants. *Plant Physiol* **88**: 815–822.

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

Wellman CH, Strother PK. **2015**. The terrestrial biota prior to the origin of land plants (embryophytes): A review of the evidence (A Smith, Ed.). *Palaeontology* **58**: 601–627.

Werner T, Motyka V, Strnad M, Schmülling T. **2001**. Regulation of plant growth by cytokinin. *Proceedings of the National Academy of Sciences of the United States of America* **98**: 10487–10492.

White RG, Badelt K, Overall RL, Vesk M. **1994**. Actin associated with plasmodesmata. *Protoplasma* **180**: 169–184.

Wickett NJ, Mirarab S, Nguyen N, Warnow T, Carpenter E, Matasci N, Ayyampalayam S, Barker MS, Burleigh JG, Gitzendanner MA, *et al.* 2014. Phylotranscriptomic analysis of the origin and early diversification of land plants. *Proceedings of the National Academy of Sciences of the United States of America* **111**: E4859–E4868.

Wickham H. **2011**. ggplot2. *Wiley Interdisciplinary Reviews: Computational Statistics* **3**: 180–185.

Wiggins DP. **2020**. xvfb 2:1.20.8-2.

Win J, Kamoun S. **2004**. pCB301-p19: A Binary Plasmid Vector to Enhance TransientExpression of Transgenes by Agroinfiltration.

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

Woo DU, Jeon HH, Park H, Park JH, Lee Y, Kang YJ. **2020**. Database: Web application for visualization of the cumulated RNAseq data against the salicylic acid (SA) and methyl jasmonate (MeJA) treatment of Arabidopsis thaliana. *BMC Plant Biology* **20**: 453.

Woolley JT. **1975**. Refractive Index of Soybean Leaf Cell Walls. *Plant Physiology* **55**: 172–174.

Wróbel-Marek J, Kurczyńska E, Płachno BJ, Kozieradzka-Kiszkurno M. **2017**. Identification of symplasmic domains in the embryo and seed of Sedum acre L. (Crassulaceae). *Planta* **245**: 491–505.

Wrzaczek M, Brosché M, Salojärvi J, Kangasjärvi S, Idänheimo N, Mersmann S, Robatzek S, Karpiński S, Karpińska B, Kangasjärvi J. **2010a**. Transcriptional regulation of the CRK/DUF26 group of Receptor-like protein kinases by ozone and plant hormones in Arabidopsis. *BMC Plant Biology* **10**.

Wrzaczek M, Brosché M, Salojärvi J, Kangasjärvi S, Idänheimo N, Mersmann S, Robatzek S, Karpiński S, Karpińska B, Kangasjärvi J. **2010b**. Transcriptional regulation of the CRK/DUF26 group of Receptor-like protein kinases by ozone and plant hormones in Arabidopsis. *BMC Plant Biology* **10**: 95.

Wu SW, Kumar R, Iswanto ABB, Kim JY. **2018**. Callose balancing at plasmodesmata. *Journal of Experimental Botany* **69**: 5325–5339.

Xie M, Chen H, Huang L, O'Neil RC, Shokhirev MN, Ecker JR. **2018**. A B-ARR-mediated cytokinin transcriptional network directs hormone cross-regulation and shoot development. *Nature Communications* **9**.

Xie B, Wang X, Zhu M, Zhang Z, Hong Z. **2011**. CalS7 encodes a callose synthase responsible for callose deposition in the phloem. *Plant Journal* **65**: 1–14.

Xu B, Cheval C, Laohavisit A, Hocking B, Chiasson D, Olsson TSG, Shirasu K, Faulkner C, Gilliham M. **2017**. A calmodulin-like protein regulates plasmodesmal closure during bacterial immune responses. *New Phytologist* **215**: 77–84.

Xu XM, Wang J, Xuan Z, Goldshmidt A, Borrill PGM, Hariharan N, Kim JY, Jackson D. **2011**. Chaperonins facilitate KNOTTED1 cell-to-cell trafficking and stem cell function. *Science (New York, N.Y.)* **333**: 1141–4.

Yadav SR, Yan D, Sevilem I, Helariutta Y. **2014**. Plasmodesmata-mediated intercellular signaling during plant growth and development. *Frontiers in Plant Science* **5**.

Yahalom A, Warmbrodt RD, Laird DW, Traub O, Revel JP, Willecke K, Epel BL. **1991**. Maize mesocotyl plasmodesmata proteins cross-react with connexin gap junction protein antibodies. *The Plant cell* **3**: 407–17.

Yan D, Yadav SR, Paterlini A, Nicolas WJ, Petit JD, Brocard L, Belevich I, Grison MS, Vaten A, Karami L, *et al.* 2019. Sphingolipid biosynthesis modulates plasmodesmal ultrastructure and phloem unloading. *Nature Plants* **5**: 604–615.

Yang DH, Kwak KJ, Kim MK, Park SJ, Yang K-Y, Kang H. **2014**. Expression of Arabidopsis glycine-rich RNA-binding protein AtGRP2 or AtGRP7 improves grain yield of rice (Oryza sativa) under drought stress conditions. *Plant Science* **214**: 106–112.

Yasumura Y, Pierik R, Fricker MD, Voesenek LACJ, Harberd NP. **2012**. Studies of Physcomitrella patens reveal that ethylenemediated submergence responses arose relatively early in land-plant evolution. *Plant Journal* **72**: 947–959.

Yu G, Smith DK, Zhu H, Guan Y, Lam TTY. **2017**. Ggtree: an R Package for Visualization and Annotation of Phylogenetic Trees With Their Covariates and Other Associated Data. *Methods in Ecology and Evolution* **8**: 28–36.

Zavaliev R, Dong X, Epel BL. **2016**. Glycosylphosphatidylinositol (GPI) modification serves as a primary plasmodesmal sorting signal. *Plant Physiology* **172**: 1061–1073.

Zess EK, Kamoun S. **2019**. Coomassie Brilliant Blue (CBB) staining for Rubisco is an appropriate loading control for western blots from plant material.

Zhang J, Fu XX, Li RQ, Zhao X, Liu Y, Li MH, Zwaenepoel A, Ma H, Goffinet B, Guan YL, *et al.*. **2020**. The hornwort genome and early land plant evolution. *Nature Plants* **6**: 107–118.

Zhang X, Gou M, Liu CJ. **2013**. Arabidopsis kelch repeat F-Box proteins regulate phenylpropanoid biosynthesis via controlling the turnover of phenylalanine ammonia-lyase. *Plant Cell* **25**: 4994–5010.

Zhang Z, Hong Z, Verma DPS. **2000**. Phragmoplastin polymerizes into spiral coiled structures via intermolecular interaction of two self-assembly domains. *Journal of Biological Chemistry* **275**: 8779–8784.

Zhang Z, Ruan Y-L, Zhou N, Wang F, Guan X, Fang L, Shang X, Guo W, Zhu S, Zhang T. **2017**. Suppressing a Putative Sterol Carrier Gene Reduces Plasmodesmal Permeability and Activates Sucrose Transporter Genes during Cotton Fiber Elongation. *The Plant cell* **29**: 2027–2046.

Zhang W, Thieme CJ, Kollwig G, Apelt F, Yang L, Winter N, Andresen N, Walther D, Kragler F. **2016**. tRNA-related sequences trigger systemic mRNA transport in plants. *Plant Cell* **28**: 1237–1249.

Zhao C, Zayed O, Yu Z, Jiang W, Zhu P, Hsu CC, Zhang L, Andy Tao W, Lozano-Durán R, Zhu JK. **2018**. Leucine-rich repeat extensin proteins regulate plant salt tolerance in Arabidopsis. *Proceedings of the National Academy of Sciences of the United States of America* **115**: 13123–13128.

Zheng MS, Takahashi H, Miyazaki A, Hamamoto H, Shah J, Yamaguchi I, Kusano T. **2004**. Up-regulation of Arabidopsis thaliana NHL10 in the hypersensitive response to Cucumber mosaic virus infection and in senescing leaves is controlled by signalling pathways that differ in salicylate involvement. *Planta* **218**: 740–750.

Zubo YO, Blakley IC, Yamburenko M V., Worthen JM, Street IH, Franco-Zorrilla JM, Zhang W, Hill K, Raines T, Solano R, *et al.*. **2017**. Cytokinin induces genome-wide binding of the type-B response regulator ARR10 to regulate growth and development in Arabidopsis. *Proceedings of the National Academy of Sciences of the United States of America* **114**: E5995–E6004.

Zuidscherwoude M, Göttfert F, Dunlock VME, Figdor CG, van den Bogaart G, Spriel AB van. **2015**. The tetraspanin web revisited by super-resolution microscopy. *Scientific Reports* **5**: 12201.

Zulawski M, Schulze G, Braginets R, Hartmann S, Schulze WX. **2014**. The Arabidopsis Kinome: Phylogeny and evolutionary insights into functional diversification. *BMC Genomics* **15**: 548.