Investigation of IRQ domain containing proteins in Arabidopsis thaliana

PRIDGEON, GREG,LIAM

Investigation of IRQ domain containing proteins in *Arabidopsis thaliana*

Greg Liam Krick-Pridgeon

Submitted in accordance with the requirements for the degree of Doctor of Philosophy

Department of Biosciences

Durham University

September 2017

Abstract

The endomembrane system in eukaryotic cells plays a vital role in the movement of membranes and substances around the cell in response to abiotic and biotic stimuli. Recent work on an actin binding protein, NETWORKED4B (NET4B), revealed that this vacuolar localised protein (Deeks *et al.* 2012) contains a domain termed the IRQ domain responsible for interacting with particular regulatory proteins of the endomembrane system. Bioinformatic analysis revealed that this IRQ domain was present in six novel proteins not containing the characteristic NET Actin Binding (NAB) domain. These proteins were termed the IRQ proteins (IRQ1-6). Work outlined in this thesis explores the evolution and localisation of expression of these proteins but in particular looks at IRQ4. Phylogenetic analysis revealed that the IRQ proteins represent a eudicot specific group of proteins and that they evolved from the NET proteins. The IRQ proteins can be subdivided based on sequence similarity into three groups: IRQ1 and IRQ6, IRQ2 and IRQ3, IRQ4 and IRQ5. Using promoter GUS lines for IRQ1 and IRQ6 revealed that these proteins may be involved in the initiation or regulation of lateral root growth. IRQ4 was expressed most strongly in the root. Subcellular localisation analysis using *prom*IRQ4::IRQ4-GFP and live cell imaging showed that IRQ4 localises to the prevacuolar compartment (PVC)/multivesicular body (MVB). Immunogold labelling using an IRQ4 specific antibody revealed an additional localisation to autophagasomes. This project investigates a group of novel eudicot specific proteins and shows that IRQ4 may be involved in key endomembrane pathways in plants.

Declaration

I can confirm that this thesis and the original research described therein are my work. Wherever the contributions of others are involved, appropriate credit has been given, with due reference to the literature and acknowledgement of collaborative research.

No material contained in this thesis has been submitted for the award of a higher degree elsewhere.

Statement of Copyright

Acknowledgements

This thesis could not have happened without the help and support provided by colleagues, friends and family. Firstly, I would like to thank Professor Patrick Hussey who has given me solid guidance; and great advice during my studies. I would also like to thank the BBSRC for their funding of a DTP-Doctoral Studentship to fund this project.

Many thanks to the members of the Hussey lab I have had the pleasure of working alongside while in Durham. First and foremost thank you to Dr Dave Mentlak who, in conjunction with Dr Tim Hawkins, got the IRQ project up and running. I also miss your secret santa Christmas presents. Special thanks must go to Dr Johan Kroon and Dr Tim Hawkins, who have given up so much of their time and expertise to help me over the years. Extra special thanks go to Dr Kroon who introduced me to the wondrous concept of a cell bank to impose order in my work and to a whole host of experimental techniques. I will sorely miss our chats in the office. Additional thanks also to Dr Hawkins who has critiqued my thesis and made it far stronger as a result. I also wish to thank Professor Pengwei Wang who provided helpful advice throughout my project, but especially regarding the autophagy assays of chapter 7. Pat, Simon, Teresa, Jingze thank you for making the lab such a pleasant environment to work in. I wish Pat the best of luck as he picks up this project.

I would also like to thank several a members of the Biosciences department for their involvement in various aspects of this project. Thank you to the staff at the Life Sciences Support Unit for the generation of the polyclonal antibodies. Thanks also to Christine Richardson and Helen Grindley for performing the immunogold labelling experiments. In addition I also wish to thank the wonderful microscopy team of Joanne Robson and Bethany Cole who have always been willing to help me whenever I asked.

Mum, Dad, and Ryan you have provided endless love and support through this process and there is no way to repay that debt. Lastly, but in reality firstly, I wish to thank Katherine whose love has been the rock upon which this whole thing stands and without a shadow of a doubt my biggest source of inspiration. I thank you from the bottom of my heart.

Table of Contents

List of Figures

Chapter 1

Chapter 2

Chapter 4

Chapter 5

Chapter 6

Chapter 7

Chapter 8

List of Tables

List of Supplemental Movies

S1	Treatment of IRQ4::GFP expressing BY2 cells with latrunculin B-Control
S2	Treatment of IRQ4::GFP expressing BY2 cells with latrunculin B-Treated
S3	Treatment of *promIRQ4*::IRQ4-GFP expressing plant root cells with latrunculin B-Control
S4	Treatment of *promIRQ4*::IRQ4-GFP expressing plant root cells with latrunculin B-Treated

List of Abbreviations

ASMH3	ASSOCIATED MOLECULE WITH THE SH3 DOMAIN OF STAM3
AtRGS1	Regulator of G-Protein Signalling 1
ARF	AUXIN RESPONSE FACTOR
AtCASP	*Arabidopsis thaliana* CCAAT-displacement protein alternatively spliced product
ATG	Autophagy related
BAR	Bin-amphiphysin-Rvs
BCECF-AM	2',7'-Bis-(2-carboxyethyl)-5-(6)-carboxyfluorescein, acetoxymethyl ester
BRI1	Brassinosteroid insensitive 1
BY2	Bright yellow 2
BiP	Binding protein
CaMV	Cauliflower mosaic virus
CBM	Carbohydrate binding module
CIP1	COP-1 Interacting Protein 1
CLSV	Cucumber leaf spot virus
COG	Conserved oligomeric Golgi
COP	Constitutive photomorphogenic 1
CORVET	Class C core vacuole/endosome tethering
CRISPR	Clustered regularly interspaced short palindromic repeats
DRP2	Dynamin related protein 2
ECM	Extracellular matrix
EE	Early endosome
EGF	Epidermal growth factor
ER	Endoplasmic reticulum
ERAD	ER associated degradation
ERES	ER exit sites
ESCRT	Endosomal sorting complex required for transport
EXPO	Exocyst positive organelle
FREE1	FYVE domain protein required for endosomal sorting 1
GAP	GTPase activating protein
GARP	Golgi associated retrograde protein
GDF	Guanine dissociation factor

GDI	Guanine dissociation inhibitor
GDP	Guanosine diphosphate
GEF	Guanine nucleotide exchange factor
GlcNAc	N-acetyl glucosamine
GNL1	GNOM-1 like
GT	Glycosyl transferases
GTP	Guanosine-5’-triphosphate
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HOPS	Homotypic fusion and protein sorting
HPFFS	High-pressure freezing and freeze substitution
HSP40	Heat shock protein 40
IL	Inner layer
KEG	KEEP ON GOING
KOR1	KORRIGAN 1
LAF1	Long after far-red light 1
LE	Late endosome
LHW	Lonesome highway
LOB	LATERAL ORGAN BOUNDRY
LR	Lateral root
LRP	Lateral root primordia
LV	Lytic vacuole
MS	Murashige and Skoog
MTOC	Microtubule organising centre
MVB	Multivesicular compartment
NAB	NET actin binding domain
NAP1	Nuclear assembly site 1
NE	Nuclear envelope
NET	NETWORKED
NPC	Nuclear pore complex
NSF	N-ethylmalemide-sensitive factor
OE	Overexpression
OL	Outer layer
OST	Oligosaccharyl transferase

PAS	Phagaphore assembly site
PH1	Pleckstrin homology domain containing protein 1
PHY	Phytochrome
PI3K	Phosphatidylinositol-3-kinase
PM	Plasma membrane
PSV	Protein storage vacuole
PVC	Prevacuolar compartment
REP	Rab escort protein
RSW2	RADIALLY SWOLLEN 2
SCAMP1	Secretory associated membrane carrier protein
SCAR	Suppressor of cAR
SDF2	Stromal cell-derived factor 2 precursor
SERK2	Somatic embryogenesis receptor-like kinase 2
SLR	Solitary root
SNARE	Soluble NSF [N-ethylmalemide-sensitive factor] attachment protein receptor
SRP	Signal recognition particle
SYP	Syntaxin
TGN	Trans Golgi network
TIP	Tonoplast intrinsic protein
TRAF	Tumour Necrosis Factor Associated
TRAPP	Transport particle protein
UDP	Uridine diphosphate glucose
UGGT	UDP-glucose:glycoprotein glucosyltransferase
V-ATPase	Vacuolar H^+-ATPase
VAM	Vacuolar morphology
VAMP	Vesicle associated membrane protein
VCL1	VACUOLELESS 1
VFT	Vps fifty-three
VIT1	Vacuolar iron transport
Vps	Vacuolar protein sorting
VSD	Vacuolar sorting determinants
VSR	Vacuolar sorting receptor
WASP	Wiskott–Aldrich syndrome protein

WAVE	WASP family Verprolin-homologous protein
WT	Wildtype

Chapter 1 Introduction

1.1 The endomembrane system

The endomembrane system was proposed as concept in 1974 to describe the movement of membranes along a continuum from the start point at the nucleus to one of two endpoints: either the plasma membrane (PM) or the vacuole via a number of other organelles, mainly the endoplasmic reticulum (ER), Golgi, trans Golgi network (TGN), vacuole/lysosome and endosomes (Morré & Mollenhauer, 1974). The eukaryotic cell endomembrane system is important because it is intimately involved in the synthesis and correct folding of proteins crucial for eukaryotic life. In addition in plant cells it is how food and fibre products are synthesised, stored, and/or released. The products of the endomembrane system are commercially important and include oils and cell wall components used in biofuel production, and stored proteins, starches, and lipids used in human and other animal nutrition. Understanding how this system is regulated is vital for exploiting the potential of the system for the above applications.

The secretory system consists of two main pathways: the secretory (anterograde) pathway and the endocytic pathway (Bassham *et al.*, 2008). The secretory pathway begins at the ER and proceeds via the Golgi to the PM and potentially on to the extracellular matrix (ECM; Figure 1.1). The endocytic pathway is used for either processing of endocytosed contents from vesicles fusing with the plasma membrane from outside the cell or the recycling of proteins sent to the PM e.g. protein signalling receptors and signalling complexes. The endocytic pathway proceeds from the PM via endosomal compartments to the vacuole. Aside from these two main transport pathways there are a number of retrograde recycling pathways between organelles. Examples include retrograde pathways between the cisternae of the Golgi, transport from the Golgi back to the ER, and from endosomes back the plasma membrane. Substances are moved between organelles via membrane bound vesicles.

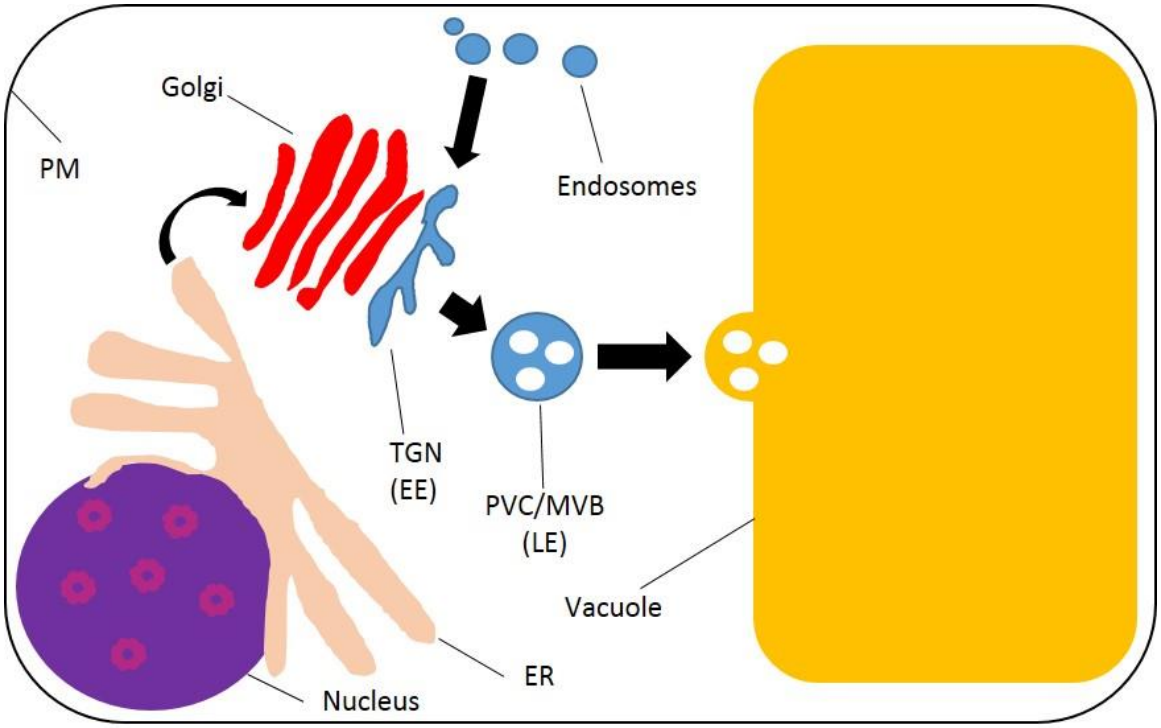


Figure 1.1: The endomembrane system. Illustrated is a typical cell with all the major compartments in the secretory and endocytic pathway labelled. Secretory proteins are translated at the endoplasmic reticulum (ER) membrane by ribosomes (not pictured) before being sent to the Golgi apparatus via membrane bound vesicles. At the Golgi proteins move through the various cisternae of the organelle being modified before they are packaged at the trans Golgi network (TGN) and sent to their destination. This can be the plasma membrane (PM) or to the vacuole for degradation via the prevacuolar compartment (PVC)/multivesicular body (MVB). The PVC/MVB fuses with the vacuole and proteins are degraded. In the endocytic pathway proteins are taken in at the PM in to endosomes and then are transferred to the TGN which functions as an early endosome (EE), packaged and sorted at the TGN, and sent to the vacuole via the PVC/MVB which functions as a late endosome (LE). Arrows indicate direction of traffic. There are various pathways for retrieval of proteins in the pathway (not shown).

Each organelle has a characteristic membrane identity consisting of specific membrane lipids and membrane bound proteins. The contents of each organelle also have their own characteristic pH and protein complement (Schumacher, 2014). To maintain the integrity of each organelle requires tight regulation by a large number of regulatory proteins including coat protein complexes, SNAREs (Soluble NSF [N-ethylmaleimide-sensitive factor] Attachment Protein Receptors), and Rab GTPases (Sanderfoot, 1999; Rutherford & Moore, 2002). These proteins are involved in specific steps of membrane and cargo transport which consist of: cargo selection, formation of membrane bound vesicles for transport at the donor compartment, targeting of the vesicle to the correct organelle, docking and emptying of the contents at the target membrane and recycling of transport machinery back to the donor compartment (Bassham *et al.*, 2008; Rojo & Denecke, 2008; Contento & Bassham, 2012; Morita & Shimada, 2014).

Previous work has identified a novel C-terminal domain in members of the Networked actin binding protein superfamily (Deeks *et al.*, 2012a), NET4A/B, responsible for binding a subset of

Rab GTPases (Mentlak, 2016). Section 1.2.2 discusses functions of Rab GTPases. There are 13 members of the NET superfamily which all contain an N-terminal actin binding domain with members of each subfamily having been shown to bind to a particular membrane (Deeks *et al.*, 2012a). The 13 family members can be split into four phylogenetically distinct subclades (Wang *et al.*, 2014a). Similarity searches using the Rab GTPase binding domain of NET4A/B revealed a set of ten proteins containing the novel C-terminal domain, four of which were NET proteins (NET3A/C, NET4A/B), and were termed 'IRQ' proteins due to the presence of a highly conserved set of three residues 'I', 'R', and 'Q' within the domain (Mentlak, 2016). This thesis focuses on the six unique members containing the IRQ domain, with a primary focus on IRQ4.

1.1.1 The nuclear envelope

The nucleus is where chromatin is stored, and is sequestered from the cytosol by a double membrane structure that allows the separation of transcription and translation activities (Parry, 2015). This double membrane is continuous with the ER, with both organelles having a shared lumen allowing the transfer of substances into the periplasmic space in plants and bacteria (Zhou *et al.*, 2015). The flow of substances such as mRNA and proteins into and/or out of the nucleus is tightly regulated by large multisubunit complexes called nuclear pore complexes (NPCs; Rout *et al.*, 2000; Fiserova, Kiseleva and Goldberg, 2009; Boruc, Zhou and Meier, 2012; Kim *et al.*, 2014). Nuclear pore complexes have an eight-fold rotational symmetry and consist of approximately 30 different proteins (Hoelz *et al.*, 2011). In higher metazoans the nuclear envelope breaks down during mitosis and reforms in a highly controlled manner, called 'open mitosis' (Güttinger *et al.*, 2009). Open mitosis allows for the formation of mitotic spindles composed of microtubules, which are formed from the microtubule organising centres (MTOCs) on opposite sides of the cell. These allow the spindles to access the kinetochores on the condensed chromatids and segregate the genetic material into two, forming daughter cells. This is in contrast to the closed mitosis present in organisms such as yeast where the nuclear envelope remains intact during chromosome segregation (Asakawa *et al.*, 2016).

1.1.2 The endoplasmic reticulum

The ER is a membrane bound organelle that is continuous with the nuclear envelope (NE) and is the major site for the synthesis and correct folding of cytoplasmic and membrane-bound proteins (Vitale & Denecke, 1999). The ER can be classified into two forms: rough and smooth, by the presence (rough) or absence (smooth) of ribosomes (Alberts *et al.*, 2008a). In plant cells, the ER can be more usefully split into cisternal or tubular ER, as though these structures may be different, both forms can support protein synthesis (Hawes *et al.*, 2015). Maintenance of ER

structure is important as altering the structural integrity of the ER leads to severe cell, tissue, and organism wide problems. For example in animal cells mutations in reticulons, genes which are involved in inducing curvature in the ER membrane (Voeltz *et al.*, 2006; Hu *et al.*, 2008; Shibata *et al.*, 2008), are linked to neurological disorders such as Alzheimer's (He *et al.*, 2004; Murayama *et al.*, 2006) and Hereditary Spastic Paraplegia (Hu *et al.*, 2009; Chang *et al.*, 2013; Ulengin *et al.*, 2015). During infection viruses target membranes in order to replicate and infect further cells. Viral targeting can be mediated by viral auxiliary proteins. For example, the auxiliary replicase protein p25 of the Cucumber Leaf Spot Virus (CLSV) has three transmembrane domains at its N-terminus which are sufficient for targeting to the ER. Once there, the virus forms vesicles derived from the ER and also the nuclear ER containing aggregates of viral proteins, leading to a disruption in the ER network. These aggregates and extra vesicles may act as factories for viral replication (Ghoshal *et al.*, 2014).

Approximately 30% of the proteome is destined to be transported to the ER (Akopian *et al.*, 2013). Soluble and membrane bound protein synthesis occurs across the ER membrane in a co-translational process where a signal recognition particle (SRP) binds to the polypeptide as it is being translated on a ribosome and arrests translation. The SRP then directs the SRP-polypeptide-ribosome complex to the translocon (Doudna & Batey, 2004; Akopian *et al.*, 2013). The translocon is a protein complex which creates an aqueous pore allowing the transfer of the nascent polypeptide across the membrane into the ER lumen (Simon & Blobel, 1991; Crowley *et al.*, 1994; Johnson & van Waes, 1999). Similar mechanisms for the transfer of proteins across the mitochondrial and chloroplast membranes also exist (Jackson-Constan & Keegstra, 2001; Lister *et al.*, 2007; Jarvis, 2008).

1.1.2.1 Protein quality control in the ER

Once proteins are synthesised they undergo post-translational modifications and must be folded correctly before they are packaged into vesicles ready for transport. N-linked glycosylation and disulphide bond formation are two post-translational modifications associated with protein quality control (Helenius & Aebi, 2004). N-linked glycosylation is the most common co-translational modification in the ER (Yan & Lennarz, 2005) with common steps conserved across mammals, yeast, and plants (Strasser, 2016). It involves the addition of a 14 sugar molecule ($Glc_3Man_9GlcNAc_2$) to an asparagine residue when it occurs in one of the following sequences: Asn-X-Ser or Asn-X-Thr (where X is any amino acid except proline) (Welply *et al.*, 1983). The reaction is catalysed by oligosaccharyl transferase (OST) which transfers the precursor oligosaccharide to the growing polypeptide chain from a dolichol pyrophosphate donor (Dempski & Imperiali, 2002). These glycoproteins are subsequently further modified in the ER and Golgi

apparatus by trimming, addition and/or removal of sugars and other residues as they proceed along the secretory pathway, something akin to an assembly line of enzymes catalysing reactions (Saint-Jore-Dupas *et al.*, 2006).

In the N-glycan dependent pathway the sugars added by OST described earlier during N-glycosylation consist of three chains: A, B, and C. The A chain is important for the folding machinery, while the B and C chains are used for ERAD (ER associated degradation) or export (Ruddock & Molinari, 2006; Roth & Zuber, 2017). Protein folding is a process that relies on energetics rather than the more directed process of DNA synthesis (Howell, 2013). During folding, the outermost glucose reside on A chain is removed by glucosidase I followed by removal of next outermost glucose residue on the A chain by glucosidase II (Deprez *et al.*, 2005). Removal of these sugar resides causes structural changes important for allowing the binding of calnexin/calreticulin (Helenius & Aebi, 2004; Aebi *et al.*, 2009; Kamiya *et al.*, 2012).

Calnexin and calreticulin are calcium dependent (Michalak *et al.*, 2009) lectins that interact with newly synthesised proteins and promote correct protein folding (Helenius *et al.*, 1997). Calnexin and calreticulin have some overlap in function as removal of calnexin or calreticulin causes improper maturation of some proteins but not others. Removal of both reduces the stringency of ER quality control with misfolded glycoproteins being transported to the plasma membrane, providing evidence that they are essential for correct glycoprotein folding (Molinari *et al.*, 2004). Proteins bind to calnexin/calreticulin as part of normal ER processing after the removal of the two outermost glucose residues from the core $Glc_3Man_9GlcNAc_2$ added by OST. Release from calnexin/calreticulin is followed by cleavage by glucosidase II of the innermost glucose residue. If the protein is folded correctly then it can continue on the secretory pathway. However, if the protein is defective e.g. it has not folded correctly or is a non-native protein it is tagged for re-binding to calnexin/calreticulin by UGGT and retained in the ER (See Figure 1.2; Molinari *et al.*, 2005; Jin *et al.*, 2007). Misfolded glycoproteins are only released from the cycles of folding and refolding in the calnexin/calreticulin system after several failed attempts (Molinari *et al.*, 2005), perhaps due to the energy investment already made in making the protein up to this point. Proteins that are misfolded get degraded by the ERAD pathway or proteasome pathways (Howell, 2013) but there is also a mechanism for degradation involving the vacuole (Foresti *et al.*, 2008). Misfolded proteins destined for the ERAD pathway are recognised by having an exposed α-1,6-linked mannose on the C chain of their oligosaccharide which acts as signal for their export to the cytosol and degradation (Hirsch *et al.*, 2009).

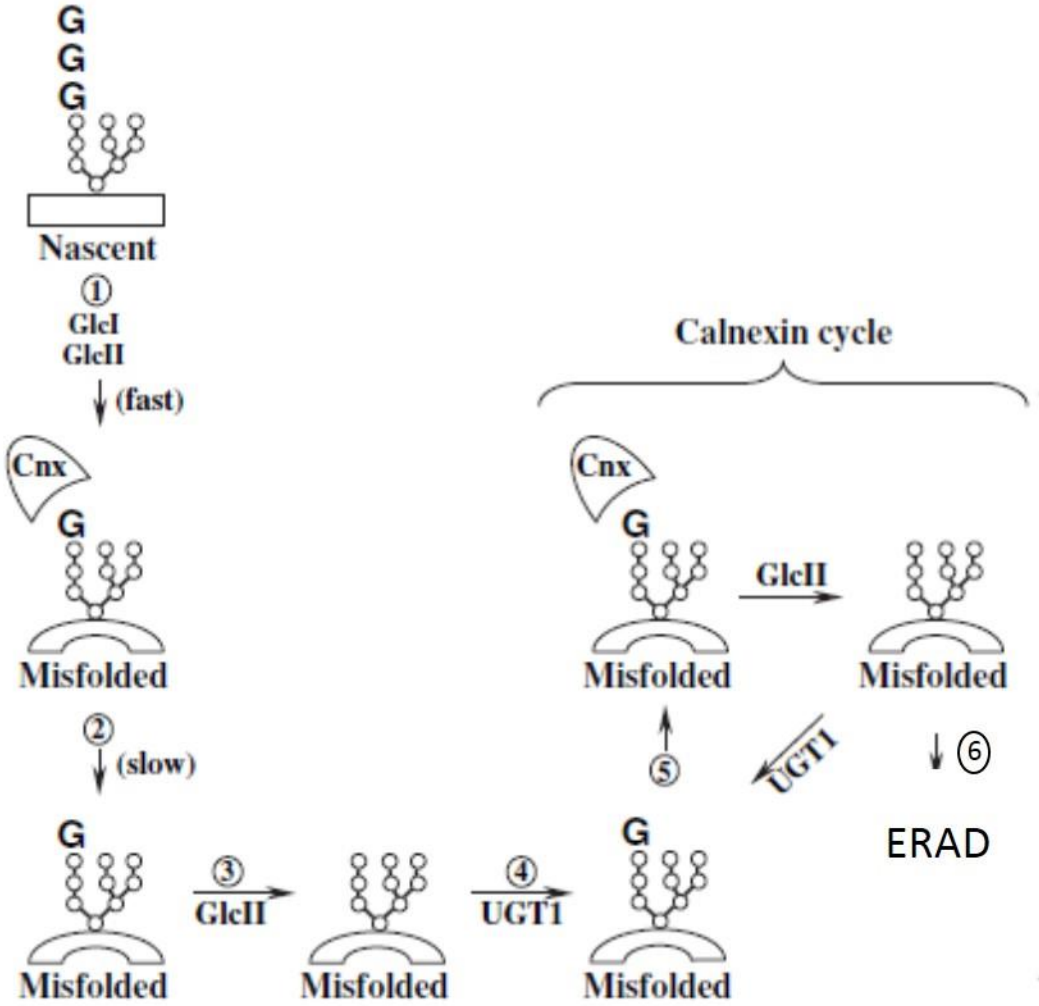


Figure 1.2: The calnexin/calreticulun cycle. Association of the nascent protein with calnexin/calreticulin (CNX) is promoted by the trimming of glucose by ER α-glucosidase I and II (step 1). If a misfolded protein is released from CNX (step 2) this exposes it to action of glucosidase II which cleaves the terminal glucose (step 3). Reglycosylation by UGGT allows the protein to re-enter the CNX cycle again for attempt at refolding the protein (steps 4 and 5). Multiple attempts of folding and refolding can result in protein architecture that is no longer recognised by UGGT. At this point the protein exits the CNX cycle and is degraded via the ERAD pathway (step 6). Modified from Molinari *et al.* (2005).

Protein disulphide isomerases act as catalysts for the formation of disulfide bonds during protein folding (Houston *et al.*, 2005). Disulfide bond formation is important for maintaining tertiary and quaternary structure. Disulfide bond formation can also be important for retention in the ER. For example, prolamins, a class of seed storage protein, are stored as large insoluble polymers in the ER. A fusion of the N-terminal half of the *Zea mays* prolamin γ-zein and the *Phaseolous vulgaris* vacuolar storage protein phaseolin to form so-called Zeolin was found to require disulphide bonds in order to be retained in the ER (Mainieri *et al.*, 2004; Pompa & Vitale, 2006; Vitale & Boston, 2008).

Another quality control mechanism present in the ER is through the Binding Protein (BiP). BiP is involved in the translocation of the nascent polypeptide into the ER lumen. In the BiP binding pathway heatshock 40 protein (HSP40 protein) ERdj3 binds emerging proteins and recruits BiP (Jin *et al.*, 2009; Ohta & Takaiwa, 2014). Plant homologue ERdj3b-recruits SDF2 to form a complex

with BiP (Meunier *et al.*, 2002). Binding of BiP is dependent upon ATP, as hydrolysis of BiP releases the nascent polypeptide (Pedrazzini *et al.*, 1994). BiP binds and stabilises partially folded intermediates to aid in correct folding. When BiP is released from these not correctly folded regions, BiP binds to other sites which have yet to fold into the correct conformation (Knittler & Haas, 1992) BiP is also involved in the folding and assembly of seed storage proteins e.g. the trimeric protein phaseolin. In pools containing both monomeric and trimeric phaseolin, BiP was only seen in association with the monomeric form (Vitale *et al.*, 1995).

1.1.3 The Golgi

The Golgi is a polar organelle consisting of a *cis* face and a *trans* face, and between these faces is a stack of membrane bound compartments called cisternae (Alberts *et al.*, 2008a). The cisternae on the *trans* face of the Golgi contain more polysaccharides than the *cis* face (Dupree & Sherrier, 1998).

The textbook picture of the Golgi in animal cells is of single stack with tubular connections between cisternae and in close association with the nucleus. However, in some animals, e.g. *Drosophila,* the Golgi can have a similar structure to plant Golgi. Plant Golgi consist of smaller, individual stacks (Kondylis & Rabouille, 2009). Plant Golgi are unique in their level of mobility (Hawes *et al.*, 2010b) and number, with hundreds being present in some plant cells (Alberts *et al.*, 2008a). The movement of Golgi is dependent upon the actin cytoskeleton (Boevink *et al.*, 1998; Nebenfuhr *et al.*, 1999; Akkerman *et al.*, 2011) but there are other proteins involved in the motility of Golgi. For example, AtRab1b is required for transport from ER to Golgi. A dominant negative mutation of AtRab1b, AtRab1b(N121I), in which asparagine 121 is substituted for isoleucine in the conserved GTP binding motif, results in the GTPase activity of the Rab being inhibited. AtRab1b(N121I) mutants had reduced or stopped directional movement of Golgi (Batoko *et al.*, 2000).

The Golgi stacks are in close association with the ER (Sparkes *et al.*, 2009, 2011; Sparkes & Brandizzi, 2012) and there is a possibility that a continuum exists between the ER and Golgi stacks involving golgin proteins (Hawes & Satiat-Jeunemaitre, 2005). Golgins are Golgi associated coiled-coil transmembrane proteins involved in capturing vesicles, tethering the Golgi to the ER to maintain Golgi structure and position within cells (Osterrieder, 2012; Gillingham & Munro, 2016; Witkos & Lowe, 2016, 2017). Studies with the golgin AtCASP and optical tweezers (Hawes *et al.*, 2010a) show that in a deletion mutant of AtCASP, AtCASP-ΔCC, the ER-Golgi connection was more easily disrupted and the Golgi bodies themselves were easier to trap (Osterrieder *et al.*, 2017). Optical tweezers are a microscopy tool that uses the radiation pressure generated by focused

infra-red laser beams to trap and move micron sized particles, allowing the micromanipulation of organelles in living cells.

The Golgi is involved in the synthesis of cell wall polysaccharides, membrane lipids, and glycolipids. It also receives proteins from the ER and is involved in packaging and organising those proteins for onward transport (Hawes, 2005). The Golgi contains a large number of glycosyl transferases (GTs): one fifth of proteins found in a proteomics screen of isolated Golgi were putative or confirmed GTs (Nikolovski *et al.*, 2012). Core glycosylation mechanisms in the ER are conserved across eukaryotes (Wang *et al.*, 2017) while the multistep glycosylation modifications that take place in the Golgi vary greatly between plants, animals, and yeasts. For example in N glycosylation in animals and plants the high-mannose N-glycan core is trimmed and acetylglucosamine is attached. These reactions are catalysed by the sequential action of β-1,2 N-acetylglucosaminyltransferase I, Golgi α-mannosidase II and β-1,2 N-acetylglucosaminyltransferase II. Further modifications in plants involve the substitution of the glycan core by a β-1,2-xylose which is catalysed by β-1,2-xylosyltransferase. In addition the proximal GlcNAc is replaced by a α-1,3-fucose. The latter reaction is catalysed by α-1,3-fucosyltransferase (Zhao *et al.*, 2008; Gomord *et al.*, 2010). N-glycosylation in Golgi is required in Golgi for salt tolerance. A mutant defective in N-glycan maturation, *complex glycan 1* (*cgl1*), has been shown to result in a defect in the plasma membrane glycoprotein Korrigan 1/Radially Swollen 2 (KOR1/RSW2). KOR1/RSW2 is needed for correct cellulose biosynthesis. Plants with the *cgl1* mutation are more sensitive to salt stress due to defects in plant cell wall formation (Kang *et al.*, 2008).

Golgi are known to multiply during cell division and at other phases of the cell cycle including G2 (Segui-Simarro & Staehelin, 2006) but also increase in number in response to cell needs (Bourge *et al.*, 2014). The biogenesis of Golgi could occur from ER exit sites (ERES; Hawes, Schoberer, *et al.*, 2010): single Golgi stacks could be generated at the ERES or alternatively the ERES would form two Golgi stacks which fuse together. In either case, both models suggest that Golgi stacks transfer cell contents form once cisternae to another via a cisternal maturation model rather than a static compartment model where materials are shuttled between cisternae via vesicles. In addition, Golgi can also be formed from existing stacks in tobacco Bright Yellow-2 (BY2) cell cultures (Abiodun & Matsuoka, 2013).

The *trans*-Golgi network (TGN) is a compartment associated with the Golgi involved in sorting proteins to the either the PM, secretory vesicles or the vacuole (Griffiths & Simons, 1986; Richter *et al.*, 2009). The TGN is distinct from the Golgi as demonstrated by insensitivity to BFA treatment (Chege & Pfeffer, 1990). The TGN develops from the *trans* cisternae (Staehelin & Kang, 2008; Kang

et al., 2011). Although the TGN is associated with the Golgi, the TGN is able to move independently and the association to the Golgi has been shown to be transient (Viotti *et al.*, 2010). As well as functioning to sort proteins to various compartments, the TGN also acts as an early endosome (Dettmer *et al.*, 2006). In addition, the TGN acts as the source for the pre-vacuolar compartment/multivesicular body (PVC/MVB; Scheuring *et al.*, 2011). The TGN is involved in a variety of processes in plants including cell plate formation (Chow *et al.*, 2008; Toyooka *et al.*, 2009; Park *et al.*, 2013; Teh *et al.*, 2013), cell elongation (Brüx *et al.*, 2008; Gendre *et al.*, 2011, 2013), and abiotic stress responses (Zhu *et al.*, 2002; Kim & Bassham, 2011; Asaoka *et al.*, 2013).

Mutations in the Golgi related proteins have influence beyond the Golgi itself. For example, the *pao* mutant is the result of a mutation in *GNOM-like 1* (*GNL1*). This mutant and shows defects in Golgi morphology and the mutants display redistribution of the Golgi marker sialyltransferase (ST) to circular structures. In addition, ER network and structural defects were also noted (Du *et al.*, 2013). Mutation in Continuous Vascular ring 1 (COV1) causes a reduction in the number of Golgi cisternae and the association between the TGN and Golgi is weakened in these mutants. Abnormal accumulations of seed protein precursors in seeds was noted in the mutant plants indicating a defect in vacuolar protein sorting (Shirakawa *et al.*, 2014). Plants lacking Dynamin Related Protein 2 (DRP2) family members DRP2A and DRP2B were embryonically lethal, with the mutant pollen having branched or irregular cell plates and altered Golgi morphology and ectopic callose deposition (Backues *et al.*, 2010).

1.1.4 The vacuole

The vacuole is a large organelle that can expand to fill 90% of the cell volume (Taiz, 1992). Water flow into the vacuole causes expansion and an increase in turgor pressure which is involved in cell wall expansion and cell growth (Zhang *et al.*, 2014a). The vacuole has many important functions in plant cells including maintaining turgor pressure, storage of metabolic products, and sequestration of compounds generated outside and within the cell e.g. calcium store (Johannes *et al.*, 1992; Pantoja *et al.*, 1992), anthocyanins (Pourcel *et al.*, 2010; Chanoca *et al.*, 2015), and digestion of cytoplasmic constituents (Marty, 1999). Vacuoles have a complement of hydrolytic enzymes equivalent to animal lysozome to aid in the digestion of those cytoplasmic constituents (Boller & Kende, 1979).

The structure of the vacuole is highly dynamic and is controlled largely by the actin cytoskeleton, which associates closely with the tonoplast membrane (Higaki *et al.*, 2006). Few actin binding proteins have been found that associate with the vacuole membrane but this list includes NET4A, NET4B (Deeks *et al.*, 2012b; Mentlak, 2016) and the vacuolar H^+-ATPase (V-ATPase) B subunits

(Wang & Hussey, 2015). NET4A and NET4B have been shown to associate with the tonoplast rather than bind it directly and are hypothesised to potentially have a function in the fusion of the developing vacuole in differentiating root cells (Mentlak, 2016). V-ATPase pumps are multi-subunit proton pumps which modify the pH of the vacuole, the membrane potential created providing the energy for transport of molecules across the tonoplast (Kriegel *et al.*, 2015). The B subunits of this multi-subunit protein complex are able to bind, bundle, and stabilise actin filaments. Certain B subunits are able to cap actin filaments in a non-calcium dependent manner (Ma *et al.*, 2012).

Classification of the vacuole as of a single organelle or of multiple subtypes is controversial. The hypothesis that there should be more than one type of vacuole in a cell makes intuitive sense: vacuoles are multifunctional as they are both a place in which defective proteins and other waste products are sent for disposal, while at the same time the vacuole also acts as a storage place for important metabolic products such as anthocyanins, or for the secondary messenger Ca^{2+} (Frigerio *et al.*, 2008). Vacuoles can be highly convoluted in structure but the lumen of the vacuoles is continuous and there appears to be no specific barrier to movement between different parts of the vacuole itself i.e. there is no evidence of micro-compartmentalisation or domain organisation within the vacuole lumen such that substances can be separated from each other. Therefore, it would appear separate vacuoles would be required to carry out different vacuolar functions.

However, Frigerio *et al.* (2008) reviewed the evidence and suggested that that there is only one type of vacuole in most differentiated plant cell types. The presence of multiple vacuole types only occurs in particular cases of certain species, certain cell types or during specific stages of the plant cell cycle or during cell differentiation (Frigerio *et al.*, 2008). Confusion up to this point mostly occurred due the use of different isoforms of the Tonoplast Intrinsic Protein family (TIP), aquaporins that localise to the vacuole (Hofte *et al.*, 1992; Paris *et al.*, 1996). Evidence from immunolocalisation studies suggested that different vacuole types can exist within the same cell and therefore would be suitable markers for different vacuolar types (Juah *et al.*, 1998, 1999). Use of both constitutively active and endogenously expressed isoforms of TIPs in *Arabidopsis* has shown that TIPS have tissue specific expression patterns, but within the cell will label the same large vacuolar structure (Hunter *et al.*, 2007; Gattolin *et al.*, 2010). Some TIPs will label structures other than the vacuole e.g. TIP3;1 and TIP3;2 are the only TIPs in embryos during seed maturation and the early stages of seed germination and both will label the tonoplast of PSV but also the plasma membrane (Gattolin *et al.*, 2011). Therefore the use of TIPs as markers to label vacuolar types is of limited use.

Despite this confusion, it is generally accepted that there are 2 types of vacuole found in in plant cells: protein storage vacuoles (PSV) and lytic vacuoles (LV; Marty, 1999). This is one area where the localisation of TIP isoforms seems to hold true, with TIP3;1 (formerly δ-TIP; (Johanson *et al.*, 2001) labelling PSV and TIP1;1 (γ-TIP) labelling LV (Juah *et al.*, 1998, 1999).

One recent example of a new type of vacuole is that of vacuolinos in petunia petals described by Faraco *et al.* (2017). Vacuolinos (Italian for "small vacuoles") are reported to be petal epidermis specific and were found to be labelled by PH1, a P_{3B}-ATPase, which interacts with the tonoplast bound PH5, a P_{3A}-ATPase proton pump, in order to acidify the vacuolar lumen. PH1 and PH5 interact with SNAREs PhSYP51 and with PhSYP22 on the membrane of the central vacuole. Vacuolinos are not labelled by TIPs, reiterating the need to use other markers for the vacuole.

The essential nature of a functional vacuole and trafficking pathways to the vacuole is amply demonstrated by mutations in proteins involved in maintaining vacuole morphology. *VACUOLELESS1* (*VCL1*) encodes an ortholog of the yeast protein Vps16p. *VCL1* is predicted to be involved in the regulation of membrane fusion events required for trafficking of proteins to the vacuole and vacuole biogenesis (Rojo *et al.*, 2001). In the *vcl1* mutant lytic vacuole formation is blocked and cell elongation and the orientation of cell division is altered, leading to an embryonically lethal phenotype. An increases in autophagasome number is triggered and vacuole contents are mis-localised to the apoplasm (Rojo *et al.*, 2001). *PROTEIN AFFECT TRAFFICKING 2* (*PAT2*) encodes the β-subunit of adaptor protein complex 3 (AP-3). In the *pat2* mutant, the lytic vacuoles again show altered morphology but this mutation is specific to the lytic vacuole and does not affect the PSV. Mutant seedlings have arrested growth on medium without sucrose, have impaired gravitropic response and produce a reduced number of lateral roots. The intracellular accumulation of proteins such as PIN1 and the soluble protein aluerain is also altered (Feraru *et al.*, 2010; Zwiewka *et al.*, 2011). Associated Molecule with the SH3 domain of STAM 3 (ASMH3) is a deubiquitinating enzyme that hydrolyses K48- and K63-linked ubiquitin chains *in vitro* and accumulates both ubiquitin chain types *in vivo. asmh3* mutants are not able to form lytic vacuole compartments and are embryonically lethal. The mutants also accumulate autophagasomes and mis-sort vacuolar proteins cargo the intercellular space (Isono *et al.*, 2010). Finally, Keep on Going (keg) is a TGN/EE localised protein required for protein transport to the vacuole (Gu & Innes, 2012). *keg* mutant plants have cell expansion defects associated with vacuole morphology defects: in the roots several small vacuoles instead of a large, central lytic vacuole. In epidermal cells mutant have numerous vesicles and compartments, and in cotyledons the vacuole has a more ruffled appearance with bulb structures in the vacuole. In addition, the fusion of MVB with the tonoplast membrane is blocked in the *keg* mutant and apoplastic defence protein secretion is blocked (Gu & Innes, 2012).

1.1.4.5 Vacuole biogenesis

In tobacco root tip cells, the lytic vacuole arises from the fusion of storage vacuoles after the storage compounds in the PSVs have been mobilised (Zheng & Staehelin, 2011). The way this process occurs varies by cell type and following freeze-fracture substitution in tobacco root tips Zheng and Staehelin (2011) showed that in the outer cortex and epidermal cells of the root the PSVs fuse, storage proteins are degraded, and the TIP3;1 marked PSV is gradually replaced by the TIP3;1 marked LV. In contrast, in the inner cortex or cells in the vascular cylinder, the PSV membranes collapse upon themselves as storage proteins are being mobilised. Two domains of membrane are then formed from the collapsed PSV: 1. domains that produce pre-LVs, and 2. multilamellar autophagosomal domains which are later engulfed by the pre-LVs to form the mature LV. The PSV marker TIP1;1 is replaced by the LV TIP3;1 during pre-LV formation (Zheng & Staehelin, 2011).

Arabidopsis root tips lack PSV and studies of vacuole development in the meristematic region of the root have shown that lytic vacuoles can be formed from provacuoles generated from the ER without involvement of ER-Golgi or post-Golgi trafficking (Viotti *et al.*, 2013). Viotti *et al.* (2013) also showed that these provacuoles are autophagasome-like in appearance but do not require any of the core autophagic machinery for their biogenesis.

1.2 Trafficking to the vacuole

Trafficking to the vacuole involves a variety of different pathways but in all cases involves the creation of vesicles carrying the cargo in question, scission of the vesicle from the donor compartment, targeted trafficking to the target compartment, and fusion and release of the contents at the target compartment. These processes will be discussed along with some of the main proteins and complexes involved.

1.2.1 Vesicle generation at donor membrane

Vesicle generation consists of three steps: 1. The cargo is selected and concentrated, 2. A signal is given for the placement of vesicle formation on membrane, 3. Membrane deformation and scission. Coat proteins are key to this process: they select cargo to be transported and recruit effector proteins involved in the targeting and the fusion of vesicle with target membrane (Bassham *et al.*, 2008). Membrane bound proteins are able to directly interact with coat proteins but soluble cargoes must first bind cargo receptors which then interact with coat proteins (Jürgens, 2004).

Each coat protein complex or coatamer is responsible for vesicle formation at a specific organelle. Two of the most well studied coat protein complexes are the COPII and COPI coats. COPII coat complexes are involved in transport from the ER to the Golgi and the COPII coat localises to ERES (ER Exit Sites). Each complex can be seen as three separate protein components: a cargo selective complex, a cage complex, and a coat GTPase. For COPII these are Sec23/24, Sec31/13, and SAR1 respectively. The SAR1 GTPase itself is recruited to the site of vesicle formation on the donor membrane by its guanine nucleotide exchange factor (GEF). The GEF triggers the exchange of GDP for GTP and switches the GTPase into the active conformation. Once active the GTPase can recruit effectors and cargo selective subunits. Sec23/24 directs the formation of the cage subunit complex (Sec31/13) at the correct site on the donor membrane. The cage complex then drives the deformation of the membrane and initiates scission from the donor compartment (Stagg *et al.*, 2007). Once free from the membrane a GTPase activating protein (GAP) triggers the GTPase activity of the coat GTPase and the vesicle is uncoated and free to travel to its target destination. The coat subunits return to the cytosol for further vesicle formation. In the case of the COPII coat complex the GAP, Sec23, is actually part of the coat complex. This may explain why in most eukaryotes the COPII coat complex is relatively short lived (Bassham *et al.*, 2008).

COPI coat GTPases are involved in retrograde trafficking in the Golgi (Waters *et al.*, 1991; Barlowe *et al.*, 1994; Letourneur *et al.*, 1994). Similar to COPII, the COPI coat can be seen as three coat complexes: the B-COP (cage complex), F-COP (cargo selective complex), and the Arf-GTP (ADP-ribosylation factor coat GTPase). B-COP is formed from three proteins: α-COP, β'-COP, and ε-COP. F-COP is formed from four proteins: γ-COP, β-COP, δ-COP, and ζ-COP. The major steps of COPI coat formation are broadly speaking similar to the formation of COPII coats. This includes the ability of the GAP, ARFGAP1, to act as a coating component in a similar way to Sec23 for the COPII coat (Yang *et al.*, 2002; Lee *et al.*, 2005).

Clathrin coats are used to shuttle cargo between the TGN and PM and are also used to sort cargo into endosomes. Clathrin coated vesicles have also been shown to be involved endocytosis at the PM (Dhonukshe *et al.*, 2007). The retromer complex is involved in recycling of VSRs from the PVC back to the Golgi (Oliviusson *et al.*, 2006; Niemes *et al.*, 2010; Yamazaki *et al.*, 2017). ESCRT (Endosomal Sorting Complexes for Transport) is group of complexes (ESCRT0-III) and in plants most of the components localise to PVC/MVB (Richardson *et al.*, 2011), make the internal vesicles of MVB (Scheuring *et al.*, 2011), and are involved in packaging and organising monoubiquitinated cargoes for degradation (Winter & Hauser, 2006). Scission of clathrin coated vesicles is done by GTPases from the dynamin family during endocytosis where the force generated during constriction is used to sever the membrane and release the vesicle (Danino & Hinshaw, 2001; Jin *et al.*, 2001). Coat proteins can also sever the vesicle-membrane connections or scission can

occur as a result of the lipid composition of the neck region (Bassham *et al.*, 2008). Once scission has occurred the GTPase is activated by its respective GAP, the vesicle is uncoated and is translocated to the target membrane.

1.2.2 Targeting and fusion of vesicles-Rabs, tethers, and SNAREs

Ensuring that cargoes reach their intended targets is vital for the proper functioning of the endomembrane system so is tightly controlled. Vesicles move along myosin or actin filaments, with plants mostly using actin for this process. Targeting of vesicles is due to a combination of the sorting signals, Rab GTPases and docking or tethering factors but not SNAREs (Cai *et al.*, 2007).

Sorting signals include the ER retention signals K- or HDEL which recycle cargo back from the Golgi to the ER (Napier *et al.*, 1992; Gomord *et al.*, 1997). Proteins destined for the lytic vacuole also have a set of sequences termed vacuolar sorting determinants (VSD) which are recognised by Vacuolar Sorting Receptors (VSR) (Vitale & Raikhel, 1999; Hadlington & Denecke, 2000). The first VSR to be characterised was BP80 from the pea, *Pisum sativum* (Kirsch *et al.*, 1994; Paris & Neuhaus, 2002; Humair *et al.*, 2007). The N-terminal region of the VSR consisting of the PA domain, Central domain, and EGF-like repeats is responsible for binding the cargo (Cao *et al.*, 2000; Watanabe *et al.*, 2002; Suen *et al.*, 2010). The binding of the cargo to the PA domain induces conformational changes that cause rotation of the C-terminal domain towards the cargo binding site and allowing cargo binding (Luo *et al.*, 2014b).

Once the vesicle reaches its target compartment it needs to be tethered to the membrane to allow fusion. The tethering process happens before interaction between SNAREs. Fusion of vesicles at the target membrane involves either coiled-coil tethers or multiprotein subunit tethering complexes. Examples of a coiled coil tethers include the golgins (discussed in relation to Golgi morphology earlier). There are currently eight conserved multisubunit tethering complexes that have been identified: COG, CORVET, Dsl1, exocyst, GARP/VFT/HOPS/Class C VPS, TRAPP1, and TRAPPII (Cai *et al.*, 2007). Figure 1.2 shows the localisation of these complexes in the endomembrane system. The most pertinent complexes to this thesis are those complexes involved in vacuolar and PVC/MVB protein sorting and trafficking: CORVET (late Golgi-endosome), HOPS (endosome-vacuole), and exocyst (autophagasome-vacuole and autophagasome-PM).

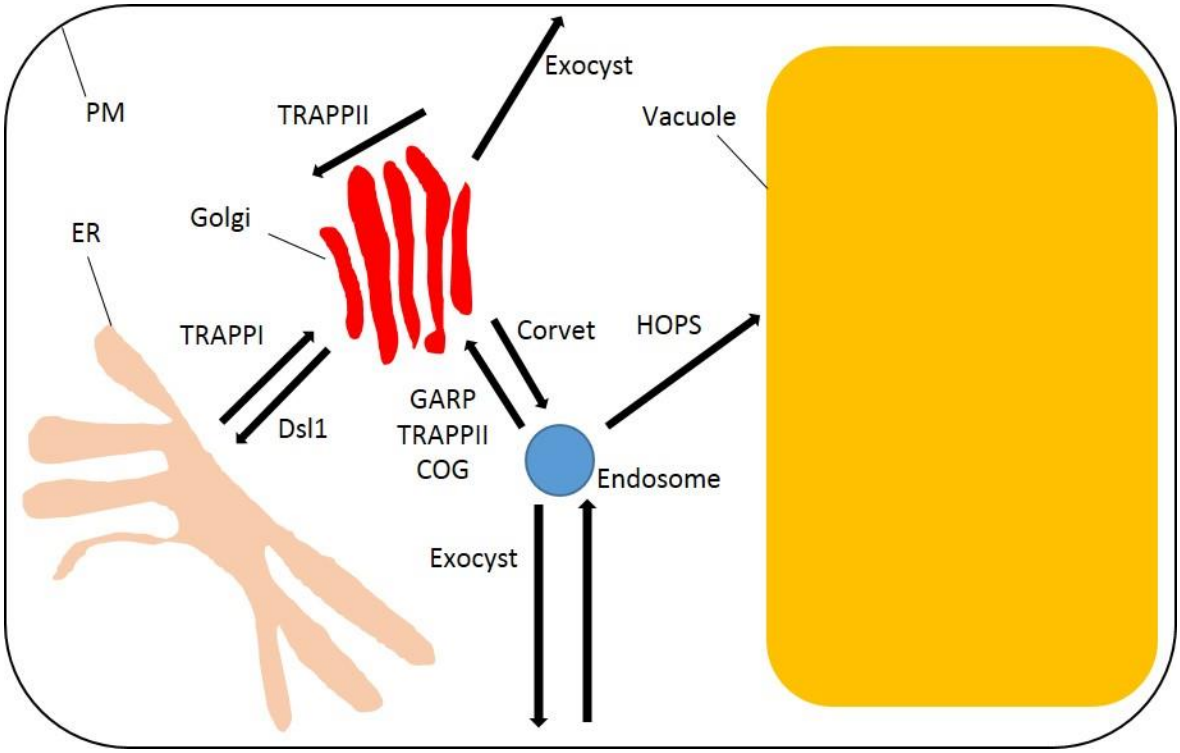


Figure 1.3: Tethering complexes that are involved in the secretory and endocytic pathways. Arrows and labels indicate the location of the tethering complexes at their respective parts of the endomembrane system. TRAPI and Dsl1 mediate traffic between the Golgi and ER. TRAPII is involved in the transport pathways between stacks of the Golgi and also trafficking from the endosome to the Golgi. GARP and COG are also involved in endosome to Golgi transport. COG and CORVET are involved in tethering events between the TGN and endosomes. HOPS is involved in the fusion of the endosome to the vacuole and is involved in homotypic vacuolar fusion. The exocyst complex binds to post-Golgi compartments, endosomes from the PM, and is involved in the creation of internal vesicles in the PVC/MVB (PVC/MVB not shown). Adapted from Cai *et al.* (2007).

The Homotypic Fusion and Protein Sorting (HOPS) complex is required for homotypic fusion of vacuoles. In yeast, HOPS has been shown to promote the formation of a t-SNARE complex on the membrane and protect this complex from disassembly (Hickey & Wickner, 2010; Xu *et al.*, 2010). HOPS binds to the SNARE complex rather than the individual SNAREs and does this through two SNARE binding sites (Kramer & Ungermann, 2011). The subunits present in plants and their yeast and human homologs are shown in Table 1.1. As discussed earlier, the HOPS complex component AtVps16/VCL1 localises to the vacuole and PVC (Rojo *et al.*, 2003) and is essential for vacuolar biogenesis (Rojo *et al.*, 2001). In addition, mutation of AtVps16/VCL1 leads to defects in vacuole biogenesis in the embryo and causes defects in pollen tube growth (Hicks *et al.*, 2004; Tan *et al.*, 2017). AtVps41/VAM2 is a PVC tonoplast localised protein associated with the interaction between the pollen tube and stigma. Loss of function mutants of AtVps41/VAM2 disrupt PVC-vacuole trafficking leading to a failure of the pollen tube to penetrate the transmission tract of the stigma (Hao *et al.*, 2016).

Table 1.1: Subunits in the *Arabidopsis* HOPS and CORVET complexes with their homologs in yeast and humans. Information taken from (Vukasinovic & Žárský, 2016)**.**

Complex	Yeast/Human Subunit	Plant Subunit	Potential *Arabidopsis* homolog
HOPS	Vps11/VPS11	VPS11	At2g05170
	Vps16/VPS18	VPS16/VCL1	At2g38020
	Vps18/ VPS18	VPS18	At1g12470
	Vps33/VPS33	VPS33	At3g54860
	Vps39/VPS39	VPS39/EMB2754	At4g36630
	Vps41/VPS41	VPS41/VAM2	At1g08190
CORVET	Vps3/TGFBRAP1	VPS3	At1g22860
	Vps8/VPS8	VPS8	At4g00800

There is relatively little information available about the plant Class C core Vacuole/Endosome Tethering (CORVET) protein complex. The CORVET complex shares four of its subunits with the HOPS complex (Vps11, Vps16, Vps18, and Vps33; Table 1.1) but also has two unique subunits (Vps3 and Vps8; Table 1.1). All of these subunits have been predicted to be present in the *Arabidopsis* genome (Klinger *et al.*, 2013; Paul *et al.*, 2014). No experimental evidence relating directly to the CORVET complex is available in plants but in yeast the complex acts as an effector of Rab5 and acts sequentially with the HOPS complex during the maturation of endosomes before fusion with the vacuole (Balderhaar & Ungermann, 2013).

The exocyst is an octameric complex (Heider *et al.*, 2016) involved in tethering secretory vesicles to the PM and also regulates cell polarity (Guo *et al.*, 1999; He & Guo, 2009). The plant exocyst has been shown to have homologs of all eight yeast subunits (Exo70, Sec3, Sec5, Sec6, Sec8, Sec10, Sec18, and Sec84) and functions as a complex (Hala *et al.*, 2008; Fendrych *et al.*, 2010). In addition, the plant exocyst is unusual in having 23 EXO70 paralogs compared to the one copy in yeast (Synek *et al.*, 2006). This large expansion is hypothesised to result in there being more than one exocyst complex in plants, with such variation due to the complexity of membrane trafficking in plants (Cvrčková *et al.*, 2012). Mutations in exocyst subunits lead to a wide variety of defects. Mutations in Sec6 cause impaired polar auxin transport and PIN protein recycling in the *Arabidopsis* primary root (Tan *et al.*, 2016). Mutation in EXO70A1 have slower root growth, shorter hypocotyl elongation (when grown on medium supplemented with 4.5% sucrose), impaired flower development, and overall produce stunted plants compared to the WT (Synek *et al.*, 2006). Mutations in SEC3 cause a defect in pollen transmission resulting in impaired pollen germination and pollen tube growth (Bloch *et al.*, 2016).

Recently, research had shown that there was a possibility that the exocyst complex could define a new organelle the 'Exocyst Positive Organelle' or 'EXPO' (Wang *et al.*, 2010). Using EXO70A1 expressed under the 35S promoter Wang *et al.* (2010) showed new compartments which did not

colocalise with standard markers for the Golgi (ManI), PVC/MVB (VSR2), TGN/EE (SYP61), tonoplast (VIT1), or PM/TGN/EE (SCAMP1). Similarly, treatment with inhibitors used to disturb the endomembrane system did not seem to affect the localisation of these EXO70A1 labelled compartments. Wang *et al.* (2010) did not show co-localisation of the 'EXPO' with the autophagsome marker ATG8 indicating that these compartments are also distinct from autophagasomes. Further experiments showed that EXO70E2 was important for the recruitment of other exocyst subunits and the formation of the 'EXPO'. The formation of the 'EXPO' was experimentally tested by transforming a construct overexpressing EXO70E2 in the HEK239A cell line of cultured human embryonic kidney cells (Ding *et al.*, 2013).

This evidence was challenged in work showing that EXO70B1 colocalised with the ATG8 marker and that 'EXPO' formation could be induced when EXO84B was over expressed using the 35S promoter but not when expressed under its own promoter (Kulich *et al.*, 2013). This evidence would indicate that the 'EXPO' was not a true organelle but an artefact created by the overexpression of the exocyst components (Vukasinovic & Žárský, 2016).

However, as argued by Robinson, Ding and Jiang (2015), the 'EXPO' can be labelled using antibodies in immunofluorescence and immunogold labelling experiments in wild type cells (Wang *et al.*, 2010; Ding *et al.*, 2013), indicating that 'EXPO' are not actually artefacts but that their occurrence is dependent upon the level of expression of EXO70E2. Robinson *et al.* (2015) have also argued that EXPO do not simply represent a special form of autophagasome as proposed by Žárský *et al.*, (2013) but that they are distinct, as has been shown in experiments where EXO70E2 was stably expressed in BY2 cell suspension cultures and *Arabidopsis* plants under the 2x35S promoter. EXO70E2 has also been expressed under the ubiquitin10 promoter and co-transformed with the AT8e autophagsome marker (Lin *et al.*, 2015). These experiments showed that although the 'EXPO' and autophagasomes are separate organelles when under non-stress conditions, when autophagy was induced these proteins colocalised indicating some form of relationship between the compartments. Overall, the issue of whether the 'EXPO' is a newly define organelle compartment is still open for debate but does serve to show the complexity of post-Golgi trafficking in plants.

SNAREs (soluble NSF [N-ethylmaleimide sensitive factor] attachment protein receptor) are a group of membrane bound proteins involved in the fusion of the vesicle (v-SNAREs) to the target compartment (t-SNAREs; Sollner *et al.*, 1993; Mcnew *et al.*, 2000). *Arabidopsis thaliana* contains 54 SNARE genes (Sanderfoot *et al.*, 2000; Uemura *et al.*, 2004) and these can be classified into 5 groups based on the similarity of their SNARE domains: Qa, Qb, Qc, R, and Qb/c (Uemura & Ueda, 2014). SNAREs localise to specific intracellular compartments: 6 at the ER, 9 at the Golgi, 4 at the

TGN, 2 at endosomes, 17 at the PM, 7 at both the PVC and vacuole, 2 at the TGN, PVC, and vacuoles, and 1 at the TGN, PVC, and PM (Uemura *et al.*, 2004).

Specific combinations of SNAREs lead to membrane fusion, either combination 1: Qa+Qb+Qc or combination 2: Qa+Qb/c+R. The binding of four SNAREs in a bundle causes fusion by pulling the membranes into contact with one another and through the bending of the α-helices to deform the lipid bilayer and cause the membrane to merge (Sutton *et al.*, 1998). Once the vesicle has fused the four SNARE bundle is left on the target membrane and the SNAPs bind to the SNAREs and, through the binding and ATPase activity of the NSF proteins, disassemble the complex ready for the next vesicle fusion event (Sanderfoot *et al.*, 2000). As an example of a SNARE defect the *ZIG* gene encodes the v-SNARE AtVTI11, a protein involved in vesicle transport. *Zig* mutants have an impaired gravitropic response in the inflorescence stems and hypocotyls. This study links the vacuole to gravitropic responses (Kato *et al.*, 2002).

Rabs are members of eukaryotic group of small G-proteins with conserved members in *Homo sapiens, Drosophila melanogaster, Saccharomyces cerevisiae,* and *Arabidopsis thaliana* (Pereira-Leal & Seabra, 2001; Diekmann *et al.*, 2011). Other small GTPases in *Arabidopsis* include Rho, Arf and Ran GTPases but not Ras GTPases (Vernoud *et al.*, 2003). Rabs couple the hydrolysis of GTP with vesicle formation, targeting and docking processes (Segev, 2001). The C-terminal hypervariable region of 35-40 amino acids is responsible for targeting the Rab to a specific membrane compartment. For example replacing the Rab5 hypervariable region with that of Rab7 causes the Rab to localise to the late endosome which is the location of the Rab7 proteins (Chavrier *et al.*, 1991). Also at the C-terminus of the protein lies the GTPase fold consisting of a six stranded beta sheet surrounded by five alpha helices. The switch I and switch II domains both make contact with the gamma phosphate in the GTP. When GDP bound the switch regions are disordered but upon GTP binding regions become ordered (Pereira-Leal *et al.*, 2003; Lee *et al.*, 2009). The domains of Rab proteins can be structurally different with the greatest variability found in the switch regions and alpha helix 3/beta sheet 5. This variability allows Rabs to recruit a wide variety of effector proteins (Hutagalung & Novick, 2011).

Initially after the Rab is synthesised it is bound to GDP and the Rab Escort Protein (REP). The Rab is then bound by Rab geranylgeranyl transferase to give prenyl tails. The prenyl tails are covalently attached to two cysteine residues with geranylgeranyl moieties located C-terminal to the GTPase fold. The prenyl tails allow the Rab to insert into the target membrane. REP then delivers the Rab to the target membrane but can only insert into the membrane in the GDP bound state. A guanine nucleotide exchange factor (GEF) for that particular Rab exchanges the GTP for the GDP, activating the Rab. The Rab now binds to its specific effectors and performs its particular function.

After it has fulfilled its function the Rab is bound by a GTPase activating protein (GAP) and its GTP is hydrolysed to GDP. The Rab is then removed from the membrane by a guanine nucleotide dissociation inhibitor (GDI) which remains bound to the Rab keeping it in an inactive state. To re-insert the Rab back into membrane to start the cycle again a guanine dissociation factor (GDF) removes the GDI (Hutagalung & Novick, 2011).

In plants Rabs can be organised into 18 functional subclasses based on the sequence similarity between the subclass defining regions of the proteins (Moore *et al.*, 1995; Pereira-Leal & Seabra, 2000, 2001; Rutherford & Moore, 2002). Table 1.2 gives the main subclasses and their localisations in plants. Rab effectors are proteins which bind to a specific Rab in its GTP-bound state and mediate at least one of the downstream processes the Rab is involved in (Grosshans *et al.*, 2006).

NET4A and NET4B were shown to be an effectors for the RabG subclass of GTPases (Mentlak, 2016). The RabG proteins localise to the prevacuolar compartment (PVC)/multivesicular body (MVB). PVC/MVB are intermediate compartments between the TGN and the vacuole (Cui *et al.*, 2016). The existence of MVBs in plants has been known since at least the early 1980's when tracer studies were conducted in order to follow endocytosis (Tanchak *et al.*, 1984, 1988; Tanchak & Fowke, 1987). The earliest PVC type compartments were named the partially coated reticulum and were visualised in soybean protoplasts (Tanchak *et al.*, 1988). PVCs have since been labelled with various antibodies as markers e.g. the cytoplasmic portion of the syntaxin AtPEP12 (da Silva Conceição *et al.*, 1997) and the integral membrane VSR BP-80 (Paris *et al.*, 1997), the latter publication using the term 'prevacuole' to describe these compartments. BP-80 was used again to show that MVBs were PVCs in plant cells (Tse *et al.*, 2004).

Table 1.2: Summary of the identities and localisations of the subclasses of Rab proteins in *Arabidopsis thaliana,* mammals, and *Saccharomyces cerevisiae*.

Rab Subclass			Localisation
A. thaliana	Mammal	Yeast (*S. cerevisiae*)	
RabF	Rab5	Ypt51/52/53	Endosomal compartments
RabG	Rab7	Ypt7	Endosome to vacuole fusion
RabH	Rab6	Ypt6	Endosome to Golgi
RabA	Rab11	Ypt31/31	Endosomes, Golgi, and prevacuolar compartments
RabC	Rab18		Endosomes
RabD	Rab1	Ypt1	ER to Golgi
RabE	Rab8	Sec4	Golgi to PM
RabB	Rab2		Golgi

PVC/MVB are involved in the degradation of integral membrane proteins such as AtLRR84Aand OsSCAMP1 (Cai *et al.*, 2012). In addition, the PVC/MVB is involved in protein degradation during seed germination. Newly synthesised VSR and hydrolytic enzymes such as the cysteine protease alueraine are localised to the MVB/PVC in germinating seeds of mung bean (*Vigna radiata*) acting as a both a storage compartment for proteases and an intermediary for the VSR-mediated delivery of proteases to the PSV to degrade proteins during seed germination (Wang *et al.*, 2007).

Maturation of the PVC/MVB is achieved by a Rab cascade involving Rab5 (RabF) and Rab7 (RabG) proteins. Rab cascades are involved in ensuring the directionality of endomembrane traffic. One set of Rabs and its effectors are replaced by the next set of Rabs and their effectors. The process works in the following way: a Rab, RabA, is recruited to the membrane and activated by its respective GEF (see 1 in Figure 1.3). The activated form of the RabA recruits the GEF for the next Rab in the cascade, RabB (2). RabB when activated and GTP bound recruits the GAP for RabA (3), causing hydrolysis of the GTP in RabA deactivating it, and recruits the GEF for the next Rab in the cascade, RabC (4). Activated RabC recruits the Rab GAP for RabB and the cascade continues (Hutagalung & Novick, 2011).

In human A431 cells conversion between early endosomes (EE) to late endosomes (LE) occurs through a maturation process, where Rab5 is replaced by Rab7 labelled compartments (Rink *et al.*, 2005). Studies in mammalian cells, yeast, and *Caenorhabditis elegans* showed that a complex involving the proteins SAND1 (MON1) and CCZ1 is responsible for the conversion between Rab5 and Rab7 labelled compartments. SAND1 is capable of removing Rab5 to stop the cascade and can also recruit Rab7 (Kinchen & Ravichandran, 2010; Poteryaev *et al.*, 2010). The yeast homolog of SAND1, MON1, acts as the GEF for Rab7 (Nordmann *et al.*, 2010). Rab5 localises to the PVC/MVB while Rab7 localises to the tonoplast (Mazel *et al.*, 2004; Goh *et al.*, 2007; Geldner *et al.*, 2009; Jia *et al.*, 2013). In plants, this conversion takes place between the PVC/MVB to the vacuole (Cui *et al.*, 2014; Ebine *et al.*, 2014; Singh *et al.*, 2014). Rab5 is recruited to the PVC/MVB and recruits the MON1/CCZ1 heterodimer. This heterodimer acts as the GEF for Rab7 leading to the gradual replacement of Rab5 for Rab7.

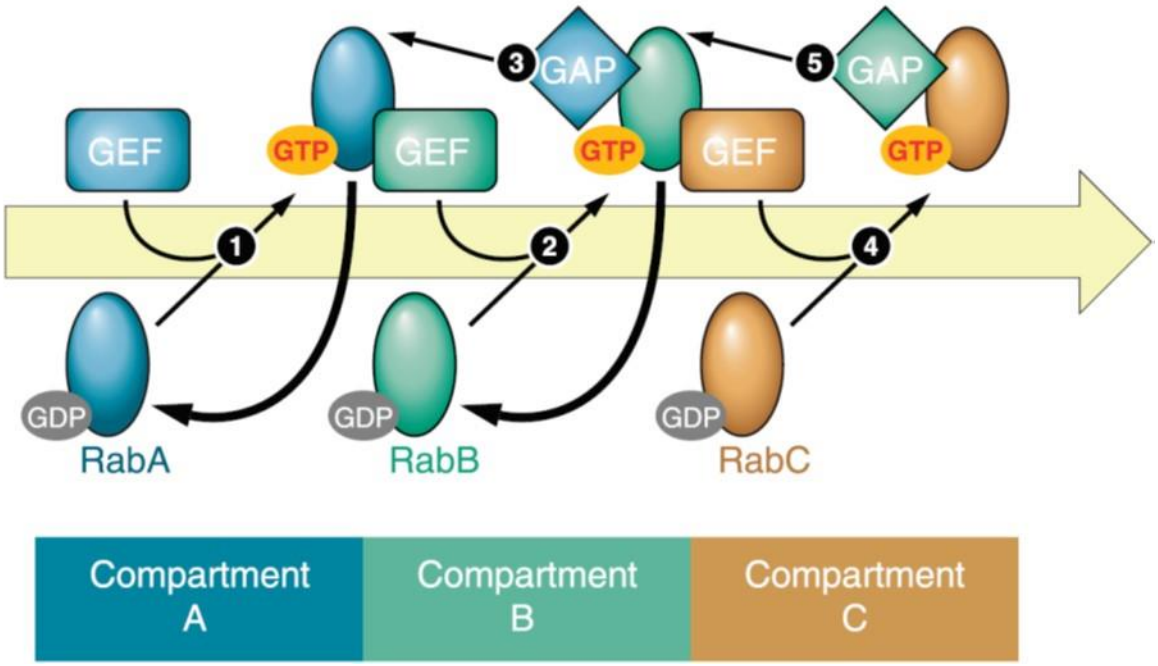


Figure 1.4: Sequential activation and deactivation of Rabs ensures that membrane identity is conserved and organised as compartments mature. Rab A is recruited to the membrane and activated by it respective GEF (1). Activated RabA recruits the GEF for the next Rab in the cascade, RabB (2). Activated Rab2 recruits the GAP for RabA and the GEF for the third Rab in the cascade, RabC (3). Finally, activated RabC recruits the GAP for RabB. Diagram taken from Hutagalung and Novick, (2011).

Examples of defects in plants resulting from mutations in PVC/MVB associated genes include the overexpression (OE) of AtPEP12, which inhibits transport from the PVC/MVB to the vacuole (Foresti *et al.*, 2006). Mutation of the PVC/MVB localised ESCRT-II component Vps36 causes a wide array of defects including delayed embryogenesis, reduced expansion of cotyledons, and impaired root elongation (Wang *et al.*, 2017). The mutation caused the abnormal accumulation of various membrane transporters important for auxin based signalling including PIN proteins, Auxin Resistant 1, and PIP family proteins. This study highlights the importance of Vps36 in the endosomal sorting pathway. In the *vps9a-1* mutant of the Rab5 GEF, embryogenesis is arrested at the torpedo stage while in the *vps9a-2* leaky mutant elongation of the primary root was severely curtailed (Goh *et al.*, 2007).

1.3 Autophagy

Another route to the vacuole is through the process of autophagy. Autophagy or 'self-eating' is the process whereby organisms degrade and recycle cellular components (Ryabovol & Minibayeva, 2016). Autophagy is a universal across eukaryotes (Klionsky, 2007). There are three major types of autophagy: macroautophagy, microautophagy, and chaperone-mediated autophagy (Kim *et al.*, 2012). Microautophagy and macroautophagy are the most well known in plants while macroautophagy is the most well studied (Bassham *et al.*, 2006). Microautophagy involves the cytoplasmic contents being engulfed by the vacuole membrane while macroautophagy involves the formation of a double membrane structure called the

autophagasome which engulfs the cytoplasmic contents to be degraded and subsequently fuses with the tonoplast to release its cargo (Yoshimoto, 2012).

The process of autophagy comprises the following main steps: formation of a pre-autophagosomal structure called the phagaphore at the phagaphore assembly site (PAS), followed by the formation of the double membrane autophagsome. In animals the autophagsome can then fuse with a lysosomal compartment to form an autolysosome. Alternatively, in animals the autophagasome can fuse with am endosome to form an amphisome before fusing with the lysosome to form the autolysosome. In plants, due to the large size of the lytic compartment the autophagsome fuses with and releases a single membrane autophagic compartment into the lumen of the vacuole which is then broken down and its contents released for degradation (Klionsky, 2012).

Formation of the phagaphore is a complex process organised by a group of autophagy-related (ATG) proteins. The membrane origin of the autophagasome is an open question but the ER is a strong candidate. A model involving ATG5 has been proposed whereby ATG5 accumulates at the PAS on the outer membrane of the ER. ATG8 is then recruited and the phagaphore extends out as a cisternal type compartment before forming a cup shaped enclosure within an ER support cradle. ATG5 leaves the phagaphore and the autophagsome is sealed (Le Bars *et al.*, 2014).

Shaping of the phagaphore membrane during this process may be mediated by the actin cytoskeleton. Nuclear Assembly Site 1 (NAP1) is an ER resident protein and part of the Suppressor of cAR (SCAR)/Wisckott-Aldrich syndrome protein (WAVE) complex in plants (Deeks *et al.*, 2004). SCAR/WAVE promotes the formation of branched actin networks by nucleating new filaments at an angle of approximately 70^0 to the mother filament (Deeks & Hussey, 2005). NAP1 localises to the ER and, under autophagy inducing conditions, colocalises with the autophagasome marker ATG8 early during autophagasome formation, and is more susceptible to nitrogen starvation and osmotic shock compared to the WT (Wang *et al.*, 2016a).

ATG9 has recently been shown to regulate autophagy from the ER in plants. ATG9 is conserved in all eukaryotes (Hanaoka *et al.*, 2002a) and is the only ATG protein known which has a transmembrane localisation. The recruitment of ATG9 vesicles to the ER is required for the formation of ER-derived autophagasomes (Zhuang *et al.*, 2017). Mutation of *ATG9* leads to an accumulation of tubular ER structures when cells are subjected to autophagy inducing conditions.

Chaperone mediated autophagy has not been demonstrated in plants but is an important process for the degradation of selected proteins in eukaryotes. The hsc70 chaperone recognises a particular peptide sequence (KFERQ) exposed when proteins are misfolded or when protein

complexes are disassembled. Along with help from co-chaperones the protein is then moved to the lysosome before being translocated across the lysosomal membrane for degradation. The receptor at the membrane is the lysosome-membrane protein type-2A (LAMP-2A). The cytosolic portion of LAMP-2A contacts the targeted protein. Upon binding to the targeted protein LAMP-2A forms a multimeric complex at the membrane required for its translocation. Once translocated across the membrane the LAMP-2A/hsc70 complex dissociates and LAMP-2A monomers are then available for binding to the next protein to be degraded.

1.4 Lateral root formation

During the course of this investigation a potential role of IRQ proteins in lateral root (LR) formation will be speculated. Therefore, the development of lateral roots will be introduced in order to more fully understand later work. LRs are important for acquiring nutrients and water and make a major contribution to the plants' ability to survive stress (Péret *et al.*, 2009). Understanding LR formation and the mechanisms which control it are important for agronomic improvement of crops. LR formation is well studied in *Arabidopsis*, and in general the protein families and the way in which LRs are formed is very similar between monocots such as maize and dicots (Beeckman & Smet, 2014).

LR formation begins at pericycle cells at the xylem pole. The actual location along the length of the root where lateral roots are initiated is mostly due to environmental control (Van Norman *et al.*, 2013; Atkinson *et al.*, 2014). Pericycle cells are a layer of cells which surround the vascular system in the stems and roots of plants and in roots are in turn surrounded by the endodermis (Figure 1.4). In dicot plants, xylem pole pericycle cells are the only cells from which LRs develop. They remain division competent even after exit from the meristem of the root. These division competent cells have been described to be part of an extended meristem of as yet undifferentiated pericycle cells (Casimiro *et al.*, 2003; Smet *et al.*, 2006). Auxin is a key hormone in the formation of LRs (Casimiro *et al.*, 2001; Benková *et al.*, 2003). A mutation in the *solitary-root* (*slr*) gene leads to auxin insensitivity and plants have no LRs. *SLR* encodes IAA14, an Aux/IAA protein family member which acts as a transcriptional repressor of LR formation (Fukaki *et al.*, 2002). During LR initiation auxin accumulation in the pericycle pole cells represses IAA14 and starts a signalling cascade leading to the activation of Auxin Response Factor 7 (ARF7) and ARF19 which in turn activate Lateral Organ Boundry 16 (LOB) and LOB19 and leads to cell cycle activation and the cell divisions involved in the formation of LRs (Péret *et al.*, 2009).

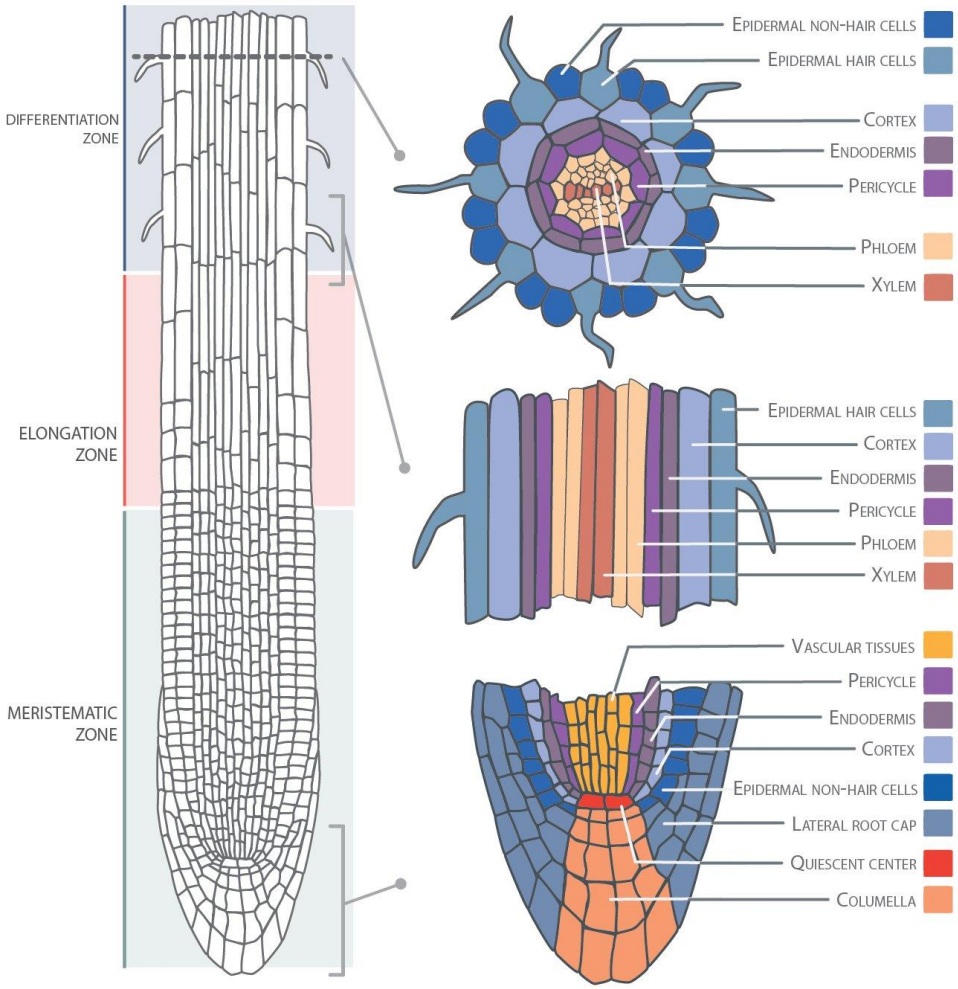


Figure 1.5: Major zones of the *Arabidopsis* root and the cell files within. Pericycle cells are labelled in purple. Image credit: (Bouché, 2017).

LR formation in *Arabidopsis* is well studied and has been mapped extensively leading to a seven stage model for the ordered set of cell divisions leading to lateral root formation (Malamy & Benfey, 1997). These are illustrated in Figure 1.5. Stage I: a particular set of cells in the pericycle layer appear, and are generally shorter with an increased number of anticlinal divisions compared to the cells on the opposite side of the stele. There are 11 of these 'founder' cells in *Arabidopsis* forming the lateral root primordia (LRP). Stage II: A periclinal division creates a division in the cells of the developing LR, with an outer layer (OL) and an inner layer (IL). The most peripheral cells in the LRP do not divide and give the LRP a dome shape. Stage III: The OL now divides producing an additional OL layer. In total there are three layers to the LRP: OL1, OL2, and IL. Stage IV: The IL divides periclinally creating an additional IL layer. The total number of layers in the LRP is now four: OL1, OL2, IL1, and IL2. The LRP also penetrates the endodermis of the root at this stage. Stage V: The central cell in in the OL1 and OL2 layers divides anticlinally and the cells adjacent to these two cells also divide anticlinally creating an OL1 that contains 10-12 cells. The cells in IL2 also divide radially and enlarge. The LRP is now at the midpoint of the parent cortex. Stage VI:

Several events apparently occur simultaneously or at least at approximately at this stage. The cells in the OL2 layer divide periclinally to form OL2a and OL2b. The four central cells of the OL2 divide periclinally. The LRP has now penetrated the epidermis at this stage of development. In the IL2 the cells begin to resemble vascular elements. Stage VI: The LRP now resembles, for the most part, a typical mature root tip with epidermal, cortex, and endodermis layers. The stele and a root cap are also being formed. By stage VI the LR is ready to emerge from the parent primary root. The whole process of lateral root development from the initial divisions of the founder cells to the emergence of the primordium through the epidermal layers (Von Wangenheim *et al.*, 2016) During lateral root emergence the overlaying tissues of the root influence the morphology of the LR more than the patterns of cell division. Inhibition of auxin signalling in the tissues overlaying the LRP causes the formation of abnormal LRP which develop more slowly (Lucas *et al.*, 2013). Recent work using fluorescent reporters and light sheet microscopy, which allows the non-invasive tracking of cell growth over time at a relatively high resolution, has been used to show various properties of LRP biogenesis (Von Wangenheim *et al.*, 2016). Founder cells do not all contribute equally to the development of the LRP, with some more dominant than others and the division pattern of these founders varies between different LRP. Precise control of the orientation of the first cell division is important for the emergence of distinct cell layers but the subsequent divisions appear to proceed in a non-deterministic manner. A relatively simple rule governing the switching of the division plane orientation based on the geometry of the cell is vital for the appearance of the pattern of cell layers seen in divisions after the first division (Von Wangenheim *et al.*, 2016).

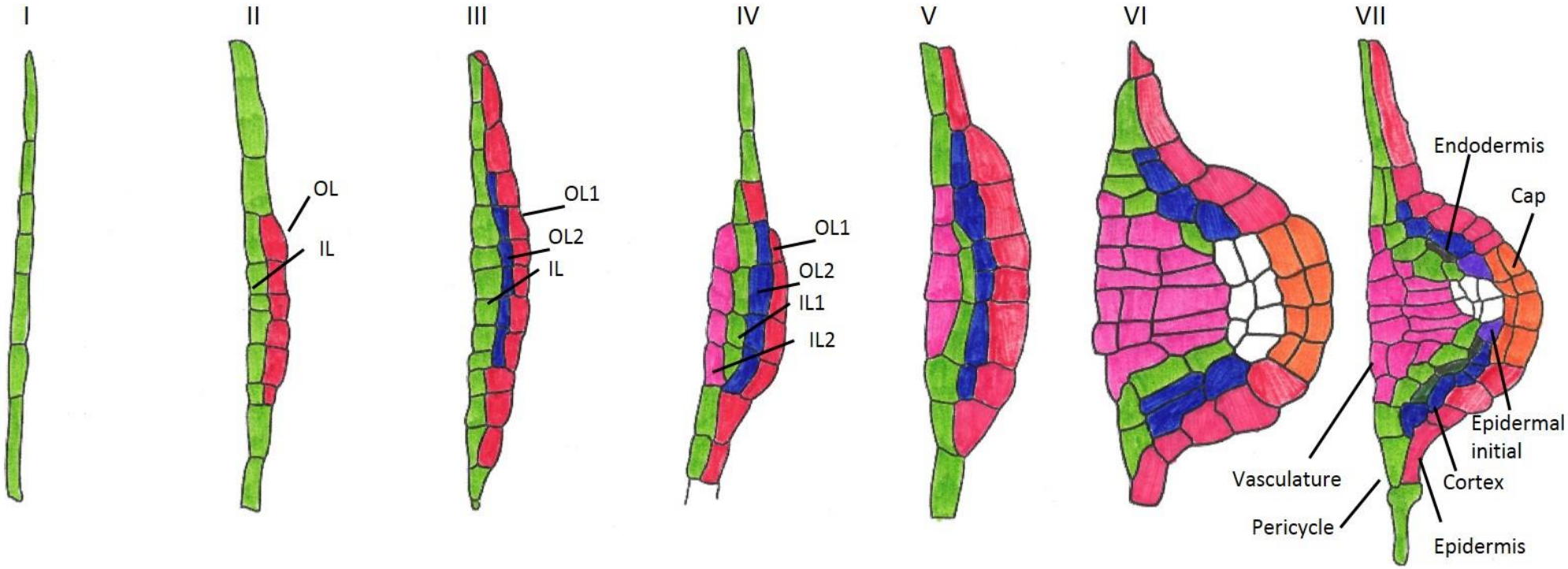


Figure 1.6: **Lateral root development.** Stages of lateral root development described in the text are displayed above each picture. Trace of cell outlines taken from (Maizel, 2017) and modified.

1.5 Introduction to and strategic approach to the analysis of the IRQ group proteins

Work on the IRQ group of proteins developed from previous work on the Networked (NET) family of actin binding proteins (ABP). The NETs were identified by screening a cDNA GFP fusion library transiently expressed in *Nicotiana benthamiana* using *Tobacco mosaic virus* (Deeks *et al.*, 2012b). A 288 residue fragment cloned from the N-terminal end decorated a filamentous network which was confirmed to be actin. The N-terminal fragment was from the founding member of the family, NET1A. The NETs are a superfamily of 13 ABPs split into 4 families (NET1-4) and are conserved across monocots and eudicots, partly due to gene duplication events (Hawkins, 2014). The divergence of the NETs is thought to mirror the increase in land plant complexity through evolutionary time. Each of the families of NETs localises to specific membrane compartments of the cell.

The NETs are an important group of proteins as membranes act as hubs for cell signalling but research into actin-membrane interactions is more advanced in animals and other metazoans compared to plants. The first ABP identified at the membrane was spectrin (Marchesi & Steers, 1968). Spectrin has two actin binding sites 200 nm apart and is linked to peripheral membrane proteins that bind integral membrane proteins beneath the red blood cell membrane (Alberts *et al.*, 2008b). ABPs such as vinculins, talins, paxillins, and integrins are important in the formation of focal adhesions which allow animal cells to move. Plants are sessile organisms and a search of the *Arabidopsis thaliana* genome database revealed no plant homologues of these four types of proteins (Hussey *et al.*, 2002). Plant specific processes such as the formation of the cell plate and the phragmoplast during cytokinesis require specialist cytoskeletal-membrane contacts (Mineyuki & Palevitz, 1990; Eleftheriou & Palevitz, 1992; Twell *et al.*, 2002; Smertenko *et al.*, 2011; McMichael & Bednarek, 2013). Although a greater number of ABPs at the membrane are being discovered (Thomas, 2012) there is as yet no clear idea of how these proteins are co-ordinately regulated to achieve actin-based processes in plants.

The functions of some of the NET proteins is just now being discovered. For example, NET1A was shown to localise to plasmodesmata and may be involved in cell to cell transport (Deeks *et al.*, 2012b). NET3B and NET3C localise to the endoplasmic reticulum (Wang *et al.*, 2014b; Wang & Hussey, 2017). Overexpression of NET3B was found to increase the association between actin and the ER and also cause a reduction in ER membrane diffusion (Wang & Hussey, 2017). NET3C was found to localise to punctae structures at contact sites between the ER and the plasma membrane along with VAP27, a plant homolog of the yeast Scs2 protein (Wang *et al.*, 2014b, 2016b).

Previous work in the lab on two NET proteins, NET4A and NET4B identified the IRQ group of proteins. NET4A and NET4B localise to the vacuole in plant roots and were shown to be effectors for the RabG subclass of Rab GTPases (Mentlak, 2016). The C-terminal domain responsible for this interaction, the IRQ domain, was determined and sequence comparisons revealed that NET3A and NET3C also possessed this domain at their C-terminal ends. In addition, six other proteins were identified which possessed this domain but not did not have true NET Actin Binding (NAB) domains and were termed the IRQ proteins.

Different approaches were used to analyse the IRQ group of proteins. Chapter 3 builds on earlier bioinformatics analyses of the IRQ proteins to examine the primary and secondary structure and localisation of the proteins using *in silico* prediction algorithms and software as well as experimental microarray data. The evolutionary relationship between the IRQ proteins and the NET proteins was investigated using phylogenetic analysis software. Using the IRQ domains of the IRQ proteins sequence databases containing data on different species representing key points in land plant evolution were interrogated.

The transcriptional profile of four of the IRQ proteins (IRQ1, IRQ6, IRQ4, and IRQ5) was studied by expressing the GUS reporter gene under the respective endogenous promoter in stably transformed *Arabidopsis thaliana* plants. GUS assays were then completed at various stages of plant development. The results of this analysis are described in chapter 4.

The *in situ* localisation of IRQ4 was studied in chapter 5 by using a variety of approaches and experimental systems. The full length IRQ4 gene was cloned and fused to GFP under the control of the constitutive 2x35S CaMV promoter and expressed both transiently in the leaves of *Nicotiana tobacum* and stably in *Nicotiana tobacum* BY2 cell culture. In addition, the *in situ* localisation of IRQ4 was studied by fusing the gene to GFP and placing it under the control of the endogenous promoter and stably expressing the resulting construct in *Arabidopsis*. The nature of the endomembrane compartments was investigated by using various endomembrane and cytoskeletal disrupting drugs, as well as using dyes and endomembrane markers to assess colocalisation between the marker and the IRQ4 labelled compartments. Polyclonal antibodies were also raised to IRQ4 and their localisation studied in *Arabidopsis* root tips using immunogold labelling and Transmission Electron Microscopy.

In order to ascertain the function of IRQ4 reverse genetics approaches are detailed in chapter 6. Single T-DNA insertion mutant lines for the *IRQ4* and *IRQ5* genes were obtained and the expression of *IRQ4* and *IRQ5* analysed. In addition, CRISPR mutant plants containing a knockout of the *IRQ5* gene and a truncation of the *IRQ4* gene product was generated. All the mutant plants

were phenotypically tested for a phenotype based on the localisation data disclosed in chapters 3, 4, and 5.

As another approach to elucidate a phenotype for IRQ4, in chapter 7 the yeast 2 hybrid (Y2H) system was used in two ways. The first was to initially screen a suite of representative Rab GTPases for a positive interaction with IRQ4. The second was to conduct a full Y2H library screen using IRQ4 as bait to find potential interacting partners.

Finally, chapter 8 discusses the results obtained in chapters 3 to 7 in relation to the wider research context and gives some scenarios for the potential functions of some of the IRQ proteins studied.

Chapter 2 Materials and Methods

2.1 Materials

2.1.1 Plant Material

Nicotiana benthamiana plants were used for the transient expression of GFP fusion proteins in leaf epidermal cells, including the IRQ4-GFP protein. Tobacco Bright Yellow-2 cells derived from *Nicotiana tobacum* were used for the stable transformation and expression of fluorescent fusion constructs including the IRQ4-GFP encoding construct. *Arabidopsis thaliana* Columbia (Col-0) ecotype plants were used for the stable transformation of fusion proteins including *promIRQ1*::IRQ1-GUS, *promIRQ6*::IRQ6-GUS, *promIRQ4*::IRQ4-GUS, *promIRQ5*::IRQ5-GUS, *promIRQ4*::IRQ4-GFP. Col-0 seeds were obtained from Lehle seeds.

The stable *promIRQ4*::IRQ4-GUS, *promIRQ5*::IRQ5-GUS, and *promIRQ4*::IRQ4-GFP lines were partly created by Dr. D. Mentlak using the *Agrobacterium tumefasciens* floral dipping method (section 2.2.13). The stable *promIRQ1*::IRQ1-GUS lines were created by Dr. J. Kroon using the *Agrobacterium tumefasciens* floral dipping method (section 2.2.15).

The SALK and GABI-KAT insertion lines were obtained from the National Arabidopsis Stock Centre (NASC) and were used for the creation of the *irq4.1* and *irq5.1* homozygous knockout lines.

2.1.2 Bacterial Strains

The *Escherichia coli* strain DH5α was used for cloning. DB3.1 was used for the amplification of empty Gateway vectors and the Rosetta 2 strain was used for recombinant protein expression and production.

2.1.3 Yeast Strains

Two strains of *Saccharomyces cerevisiae* were used: Y187 and AH109. These strains were used in the yeast-2-hybrid, autoactivation, and one-on-one tests.

2.1.4 Vectors and Constructs

All the vectors used in this study (summarised in Table 2.1) are commercially available apart from pGAT4, which was developed by Dr. T. Ketelaar (Wageningen University), and pBI101G, which was developed by Dr. M. Kieffer (University of Leeds) from the commercially available pBI101 vector (Clontech).

Table 2.1: List of vectors used for cloning

Vector Name	Vector Function	Vector Size (bp)	Resistance	Supplier
pDONR207	Gateway entry vector	5585	Gentamycin (bacteria)	Invitrogen
pBI101G	Gateway β-glucoronidase expression vector under the endogenous promoter	12200	Kanamycin (bacteria and plants)	University of Leeds
pBGWFS7	Gateway β-glucoronidase fusion to GFP expression vector under the endogenous promoter	12451	Spectinomycin/Streptomycin (bacteria) BASTA (plants)	University of Zurich via ABRC
pGAT4	Gateway protein expression vector with N-terminal 6xHis-tag	5256	Ampicillin (bacteria)	University of Durham
pMDC43	Gateway expression vector for the expression of N-terminal GFP fusion under 2x35S promoter	12460	Kanamycin (bacteria) hygromycin (plants)	University of Zurich via ABRC
pMDC83	Gateway expression vector for the expression of C-terminal GFP fusion under 2x35S promoter	12513	Kanamycin (bacteria) hygromycin (plants)	University of Zurich via ABRC
pMDC107	Gateway expression vector for the expression of C-terminal GFP fusion under the endogenous promoter	11738	Kanamycin (bacteria) hygromycin (plants)	University of Zurich via ABRC
pGBKT7	Gateway bait vector for Y2H studies. N-terminal GAL4 DNA binding domain fusion	9715	Kanamycin (bacteria) Tryptophan (yeast)	University of Durham
pGADT7	Gateway prey vector for Y2H studies. N-terminal GAL4 DNA activation domain fusion	10399	Kanamycin (bacteria) Leucine (yeast)	University of Durham

Dr. D. Mentlak and Dr T.J. Hawkins originally amplified and cloned the IRQ4 full length DNA coding sequences, the IRQ1, IRQ4 and IRQ6 coding sequences minus their respective promoters, and the promoter coding sequences only for IRQ4 and IRQ5 from *Arabidopsis thaliana* 7-day old whole seedling tissue using first strand cDNA. These constructs were cloned into pDON207 and then into their respective destination vectors. Dr. J. Kroon cloned the promoter sequence for IRQ1 into pDONR207 and then into the pBGWFS7 destination vector.

2.2 Molecular Biology Methods

2.2.1 Amplification of Fragments of Interest Using the Polymerase Chain Reaction

The polymerase chain reaction (PCR) was used to amplify fragments of interest from genomic DNA (gDNA), cDNA, or purified plasmid DNA. MyTaq (Bioline) or BIOTAQ (Bioline) were used for routine PCR applications and higher throughput analyses including colony PCR and genotyping of plant lines. Q5 polymerase (NEB) was used in cloning procedures where higher fidelity was required. Standard sets of cycling conditions are shown for MyTaq (Table 2.2), BIOTAQ (Table 2.3) and Q5 (Table 2.4). Standard reaction mixtures are shown for MyTaq (Table 2.5), BIOTAQ (Table 2.6) and Q5 (Table 2.7).

Table 2.2: Standard PCR Conditions Using MyTaq Polymerase

Stage	Temperature (°C)	Time	Cycles
Initial Denaturing	96	3 min	1
Cycle 1: Denaturing	95	20 secs	10
Cycle 1: Annealing	60	30 secs	
Cycle 1: Extension	72	10 secs/kb	
Cycle 2: Denaturing	95	20 secs	30
Cycle 2: Annealing	56	30 secs	
Cycle 2: Extension	72	10 secs/kb	
Final Extension	72	10 mins	1

Table 2.3: Standard PCR Conditions Using BIOTAQ Polymerase

Stage	Temperature (°C)	Time	Cycles
Initial Denaturing	96	1 min	1
Cycle 1: Denaturing	95	30 secs	10
Cycle 1: Annealing	60-65	30 secs	
Cycle 1: Extension	72	30 secs/kb	
Cycle 2: Denaturing	95	30 secs	30
Cycle 2: Annealing	55-65	30 secs	
Cycle 2: Extension	72	30 secs/kb	
Final Extension	72	10 mins	1

Table 2.4: Standard PCR Conditions Using Q5 Polymerase

Stage	Temperature (°C)	Time	Cycles
Initial Denaturing	98	30 secs	1
Denaturing	98	10 secs	35
Annealing	55	30	
Extension	72	20-30 secs/kb	
Final Extension	72	2 mins	1

Table 2.5: Standard Mixture for MyTaq Reactions

Component	Concentration	Volume (µl)
DNA	-	1
Primer 1	10 µM	0.5
Primer 2	10 µM	0.5
MyTaq Mix*	-	6.25
dH_2O	-	4.25
	Total Volume per reaction	**12.5**

*MyTaq contains a proprietary mixture of components including dNTPs, $MgCl_2$, NH_4, and MyTaq polymerase enzyme.

Table 2.6: Standard Mixture for BIOTAQ Reactions

Component	Concentration	Volume (µl)
DNA	-	1
Primer 1	10 µM	1
Primer 2	10 µM	1
$MgCl_2$	50 mM	1
NH_4	10X Buffer	2.5
dNTPs*	10 mM	0.5
BIOTAQ DNA Polymerase	-	0.15
dH_2O	-	17.85
	Total Volume per reaction	**25**

* Made from 100 mM stocks of dATP, dGTP, dCTP, dTTP.

Table 2.7: Standard Mixture for Q5 Reactions

Component	Concentration	Volume (µl)
cDNA	<1000 ng	1
Primer 1	10 µM	1.25
Primer 2	10 µM	1.25
Q5 Buffer*	-	5
dNTPs#	10 mM	0.5
Q5 DNA Polymerase	-	0.25
dH_2O	-	15.75
	Total Volume per reaction	**25**

* Made from 100 mM stocks of dATP, dGTP, dCTP, dTTP. # Q5 buffer contains a proprietary mixture of components including $MgCl_2$ and NH_4.

2.2.2 Agarose Gel Electrophoresis

Agarose gels for the visualisation of DNA fragments were prepared by dissolving 1.5 % to 3 % (w/v) agarose (Melford) in 1XTAE buffer (40 mM Tris acetate ((Sigma Aldrich), 1 mM EDTA ((Ethylenediaminetetraacetic acid) Melford)) by heating. Once cooled, 8 µl ethidium bromide (BDH) was added per 200 ml of 1XTAE. Gels were then poured and left to set for 30-45 mins. Gels were run in BioRad gel tanks containing 1XTAE running buffer. Samples were loaded with 6X Orange G loading buffer and Bioline Hyperladder I (from a 5x stock) was used a DNA marker. Gels were run for 40 mins at 100 V (1.5 % gels) or 50-90 mins at 70 V (3 % gels) and imaged using a Bio-Rad Gel-Doc 1000 system.

2.2.3 Cloning using the Gateway Cloning System

The Gateway cloning system (Invitrogen) was used for all cloning reactions.

2.2.3.1 BP Reactions

BP reactions were used to insert the fragment of interest into the donor vector pDONR207. The reaction was set up as follows: 4 µl 5x BP Reaction Buffer, 1-8 µl PCR product, 2 µl pDONR207 vector and then made up to 16 µl with 1XTE buffer (10 mM Tris.HCl pH 8.0, 1 mM EDTA). 4 µl of BP Clonase was added and the reaction incubated at 25 ^{o}C for 1 hour. To terminate the reaction 2 µl of Proteinase K solution was added and the sample incubated at 37 ^{o}C for 10 mins. DH5α cells were transformed with the recombinant vector as described in section 2.2.6. Transformed cells were spread on LB plates containing 25 µg/ml gentamicin. Recombinant pDONR207 was purified as described in section 2.2.8, colony PCR carried out as described in section 2.2.1 and positive clones sequenced as described in section 2.2.9 to confirm presence of the insert.

2.2.3.2 LR Reactions

LR reactions were used to transfer the fragment of interest from the entry clone into one of the destination vectors listed in Table 2.1. The reaction was set up as follows: 4 µl 5X LR Reaction Buffer, 2 µl entry clone and 2 µl of destination vector was added and then made up to 16 µl with 1XTE buffer. Four microlitres of LR Clonase was added and the reaction incubated at 25 ^{o}C for 1 hour. 2 µl of Proteinase K solution was added and the sample incubated at 37 ^{o}C for 10 mins. DH5α cells were transformed with the recombinant vector as described in section 2.2.6. Transformed cells were spread on LB plates containing the appropriate antibiotic selection as shown in Table 2.1. The recombinant destination vector was purified as described in section 2.2.8,

colony PCR carried out as described in section 2.2.1 and positive clones sequenced as described in section 2.2.9 to confirm presence of the insert.

2.2.4 Preparation of Chemically Competent Cells

2.2.4.1 *Escherichia coli* strains

Escherichia coli strain DH5α were streaked from glycerol stocks onto solid Luria-Bertani (LB) medium containing 5 g/l yeast extract (Merck), 10 g/l peptone (Merck), 10 g/l NaCl (BDH), 20 g/l micro agar (Sigma-Aldrich), pH 7.0 (Sigma Aldrich) and grown for 2 days at 30 ^{o}C. A single colony was used to inoculate a liquid LB culture and grown at 37 ^{o}C 14-16 hours. One millilitre of this culture was used to inoculate 100 ml of LB and grown at 37 ^{o}C for 2 hours until reaching an OD_{600} of 0.3-0.6. Cells were centrifuged at 3000xg for 5 mins to pellet the cells. The pellet was resuspended in 50 ml ice-cold autoclaved dH_2O and kept on ice for 10 mins. Resuspended cells were centrifuged at 3000xg for 5 mins and then resuspended in 25 ml ice-cold autoclaved dH_2O and kept on ice for 30 mins. Resuspended cells were then centrifuged at 3000xg for 5 mins, resuspended in 1 ml 1.7:0.3 (v/v) 0.1 M $CaCl_2$/glycerol (Melford) solution and then snap frozen in liquid nitrogen and stored at -80 ^{o}C.

2.2.4.2 *Agrobacterium tumefasciens* GV3101 strains

Glycerol stocks of GV3101 cells were streaked onto solid Luria-Bertani (LB) medium containing 5 g/l yeast extract (Merck), 10 g/l peptone (Merck), 10 g/l NaCl (BDH), 20 g/l micro agar (Sigma-Aldrich), pH 7.0 with 40 µg/ml rifampicin (Sigma Aldrich) and grown for 2 days at 30 ^{o}C. A single colony was used to inoculate a 5 ml liquid LB culture containing 40 µg/ml rifampicin and 10 µg/ml gentamicin and grown at 30 ^{o}C for 14-16 hours. One millilitre of this culture was used to inoculate 100 ml of LB containing 40 µg/ml rifampicin and 10 µg/ml gentamicin and grown at 30 ^{o}C for 2 hours until reaching an OD_{600} of 0.3-0.6. Cells were centrifuged at 3000xg for 5 mins to pellet the cells. The pellet was resuspended in 50 ml ice-cold 0.1 M $MgCl_2$ (BDH) and kept on ice for 10 mins. Resuspended cells were centrifuged at 3000xg for 5 mins and then resuspended in 25 ml ice-cold 0.1 M $CaCl_2$ and kept on ice for 30 mins. Resuspended cells were then centrifuged at 3000xg for 5 mins, resuspended in 1 ml 1.7:0.3 0.1 M $CaCl_2$/glycerol (Melford) solution and then snap frozen in liquid nitrogen and stored at -80 ^{o}C.

2.2.5 Preparation of Electro-competent Cells

Electro-competent cells were created for their increased transformation efficiency.

2.2.5.1 *Escherichia coli* strains

A 5 ml overnight culture of *E. coli* cells was grown in LB media containing no antibiotics. These cells were used to inoculate 250 ml fresh LB media containing no antibiotics at a dilution of 1:100 (v/v). Growth of the culture was followed until an OD_{600} of 0.5-1.2 was reached. Cells were then centrifuged at 4000xg for 20 mins at 4 °C and the supernatant discarded. All subsequent steps took place on ice or 4 °C. The pellet was resuspended in 1/3 the initial volume of ice-cold autoclaved dH_2O. Resuspended cells were then centrifuged at 4000xg for 20 mins. Cells were pelleted and resuspended at least twice more decreasing the amount of resuspension solution each time. The final resuspension used 1/75th the initial volume of filter sterilised ice cold 20 % (v/v) glycerol. Cells were then aliquoted in 100 µl amounts and snap frozen in liquid nitrogen.

2.2.5.2 *Agrobacterium tumefasciens* GV3101

Overnight cultures of *A. tumefasciens* GV3101 cells (5 ml) were grown in LB media containing 40 µg/ml rifampicin and 10 µg/ml gentamicin. These cells were used to inoculate fresh LB media containing 40 µg/ml rifampicin and 10 µg/ml gentamicin at a dilution of 1:100 (v/v). Growth of the culture was followed until an OD_{600} of 0.5-1.2 was reached. Cells were then centrifuged at 4000xg for 20 mins at 4 °C and the supernatant discarded. All subsequent steps took place on ice or 4 °C. The pellet was resuspended in 1/3 the initial volume of ice-cold 1 mM HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid)/KOH, pH 7.0 filter sterilised from a 1 M stock. Resuspended cells were then centrifuged at 4000xg for 20 mins. Cells were pelleted and resuspended at least twice more decreasing the amount of resuspension solution each time. The final resuspension used 1/75th the initial volume of filter sterilised ice cold 20 % (v/v) glycerol. Cells were then aliquoted in 100 µl amounts and snap frozen in liquid nitrogen.

2.2.6 Transformation of Chemically Competent Cells

2.2.6.1 *Escherichia coli* strains

Approximately 100 ng of plasmid DNA was added to 50-100 µl of competent *E. coli* cells on ice for 30 mins. Cells were then transferred to a heat block at 42 °C for 45 secs and then put back on ice for 2 mins. After this 400 µl LB was added and then cells were incubated on a shaker at 37 °C for 1 hour at 180 rpm. Cells were then spread on solid LB medium containing the appropriate antibiotic and grown at 37 °C for 14-16 hours.

2.2.6.2 *Agrobacterium tumefasciens* GV3101

Cells were taken from -80 °C storage and defrosted on ice. To the cells 1 µg DNA was added and then gently mixed before being snap frozen in liquid nitrogen. Cells were the heat shocked at 37 °C in a heat block for 5 mins. For recovery 400 µl LB media was added and cells were shaken at 180 rpm at 30 °C for 4-6 hours. Cells were then spread on solid LB agar plates supplemented with 40 µg/ml rifampicin, 10 µg/ml gentamicin and appropriate antibiotic for plasmid selection. Plates were grown at 30 °C for 2-3 days.

2.2.7 Transformation of Electro-competent cells

2.2.7.1 *Escherichia coli* strains

For electroporation, gateway reactions must be desalted to prevent arcing and failure of the transformation. A piece of membrane filter (0.025 µm pore size MCE MF (Merck)) was floated on top of sterile water in a petri dish. The reaction mixture was placed on top of the membrane and left for 30 mins. The reaction mixture was then collected and returned to a clean centrifuge tube ready for electroporation.

For electroporation, electro-competent cells were taken from -80 °C and thawed on ice. Fifty nanograms of DNA was added to the cells and mixed. The whole mixture was then transferred to a chilled electrocuvette (Fisher Scientific) and electroporated at 1350V in an Eppendorf Electroporator 2510 (Fisher Scientific). Cells were immediately resuspended in 400 µl LB medium without antibiotics and left to recover shaking at 1000 rpm for 1 hour. Recovered cells were then spread on solid LB plates containing the appropriate antibiotic for selection and left to grow for 14-16 hours.

2.2.7.2 *Agrobacterium tumefasciens* GV3101

If appropriate, Gateway reactions were desalted as described in section 2.2.7.1 above. Approximately 50 ng of DNA was added to 40 µl electro-competent Agrobacterium GV3101 strain cells. Cells were transferred to an electrocuvette (Fisher Scientific) and electroporated at 1440 V in an Eppendorf Electroporator 2510 (Fisher Scientific). Cells were transferred to a microcentrifuge tube and 400 µl LB was added. Cells were then incubated on a shaker at 30 °C on a shaker for 4-6 hours to allow cells to recover. Cells were then spread on solid LB selective plates (rifampicin 40 µg/ml, gentamicin 10 µg/ml and the appropriate antibiotic for plasmid selection) and grown at 30 °C for 2 days.

2.2.8 Plasmid DNA Purification

Plasmid DNA was purified using the Promega Wizard Plus SV Miniprep DNA purification System according to the manufacturer's protocol. *Escherichia coli* DH5α cells from an overnight culture transformed with the relevant plasmid were centrifuged at 3000xg for 5 mins, resuspended then lysed. Alkaline protease was added to deactivate proteins which degrade DNA. Neutralisation solution was added to stop lysis. The lysate was then centrifuged and decanted into a spin column with a DNA binding nitrocellulose membrane. The column was washed twice and the resulting DNA eluted in nuclease free water and stored at -20 ^{o}C.

2.2.9 DNA Sequencing

DNA sequencing was carried out by Durham University DNA Sequencing service using an Applied Biosystems 3730 DNA Analyser.

2.2.10 RNA Purification

Total RNA was purified from frozen *A. thaliana* tissue using the Spectrum Total Plant RNA extraction Kit (Sigma) according to the manufacturer's protocol. Approximately 100 mg of tissue was cooled using liquid nitrogen and ground into a powder using a pestle and mortar. Five hundred microlitres of previously prepared lysis solution containing 5 µl of β-mercaptoethanol (Sigma) was added to the sample and immediately vortexed without allowing the sample to defrost. Samples were centrifuged at 16,300xg for 3 mins before being put on ice. The supernatant was removed from the sample and added to a filtration column from the kit. The column was centrifuged for 1 min at 16,300xg to clarify the lysate. The filtration column was removed, discarded, and 500 µl of binding solution was added to the clarified lysate before being mixed and added to a RNA binding column from the kit. The sample was centrifuged at 16,300xg for 1 min. The supernatant was discarded and on-column DNase digestion was carried out to remove any excess DNA contamination before cDNA synthesis using the On-Column DNase I Digest Set (Sigma). The membrane was washed using column wash solution, the flow through was discarded and DNase I enzyme mixed with DNase reaction buffer was added to the column and left for 15 mins at room temperature. The column was washed with column wash solution I and the flow through discarded. The column was then washed and centrifuged at 16,300xg twice with 500 µl column wash solution II with the flow through discarded each time. The column was then dried by centrifugation at 16,300xg for 30 secs then RNA was eluted in 50 µl elution solution and quantified using a Nanodrop spectrophotometer (Thermo Fisher) and then either used for cDNA synthesis or stored at -20 ^{o}C.

2.2.11 cDNA Synthesis

cDNA was synthesised from purified RNA produced as described in section 2.2.10 using Superscript III (Invitrogen) and gene specific primers (Eurofins). A G-Storm GS1 (GRI) PCR machine was used to synthesise the cDNA. To a PCR tube 0.2 µl (2 pmoles) of gene specific primer (from 10 pmole/µl stock) and 1 µl 10mM dNTPs (from 100 mM stock of dATP, dTTP, dCTP and dGTP) were added to 8 µl RNA template and the reaction volume made up to 13 µl with RNase-free water. Tubes were heated to 65 ^{o}C for 5 mins and then put on ice for 5 mins. Following this, 4 µl Superscript III 5x First Strand buffer, 1 µl 100 mM DDT, 1 µl RNaseOUT (Invitrogen) and 1 µl Superscript III reverse transcriptase were added to tubes before heating at 55 ^{o}C for 50 mins. The reaction was inactivated by heating to 70 ^{o}C for 15 mins. Complementary RNA was removed by adding 1 µl RNase H (Invitrogen) and incubating the tubes at 37 ^{o}C for 20 mins.

2.2.12 Genomic DNA Extraction

Genomic DNA extraction for genotyping of mutant or transformed *A. thaliana* lines was done using this protocol. All cloning procedures used cDNA (section 2.2.11) made from RNA (section 2.2.10). Genomic DNA extraction was completed using a modified version of a protocol developed by Edwards, Johnstone and Thompson (1991). A piece of rosette leaf (3-5 mg) was placed into an 1.5 ml centrifuge tube and 200 µl of a 1X Edward solution diluted using TE buffer (10 mM Tris-HCL (pH 8.0), 1 mM EDTA) from a 10X stock (200 mM Tris-HCl (pH 7.5) 250 mM NaCl, 25 mM EDTA , 0.5 % SDS (Sodium Dodecyl Sulfate (Sigma)). Two to three metal balls were added to the leaf/Edwards solution mixture. Cells were the broken apart using a TissueLyser (Qiagen) set for 2 mins running at 25 vibrations per sec. The mixture was then centrifuged at 14,000 rpm for 5 mins to pellet cellular debris and 50 µl was transferred into a new 1.5 ml microcentrifuge tube. To stop the SDS from the buffer interfering in the PCR reaction 50 µl of autoclaved MilliQ water was mixed with the sample. The DNA was then either stored at -20 ^{o}C or used directly for PCR as described in section 2.2.1.

2.2.13 Agrobacterium-Mediated Transient Transformation of *Nicotiana benthamiana* Leaf Epidermal Cells

Transient expression of constructs was based upon the method outlined in Brandizzi et al. (2002). Competent GV3101 strain cells were transformed with the relevant plasmid containing the fragment of interest as described in section 2.2.7.2. Single colonies were used to inoculate liquid culture of LB medium containing 40 µg/ml rifampicin, 10 µg/ml gentamicin and appropriate antibiotic for plasmid selection. Cultures were grown on a shaker at 30 ^{o}C overnight. Liquid cultures of GV3101 strain cells containing the p19 protein encoded by the tomato bushy stunt

virus to suppress post-transcriptional gene silencing were also prepared. Five millilitres of each liquid culture was centrifuged at 3000xg and then the pellet was washed and resuspended twice in infiltration solution (10mM MES (2-(N-Morpholino)-ethanesulfonic acid, Melford) pH 6.5, 10 mM $MgCl_2.6H_2O$, 200 µM acetosyringone (3,5-dimethoxy-4-hydroxy-cetophone, Fluka), made from 200 mM stock in DMSO (dimethyl sulfoxide, Sigma-Aldrich). A 10x dilution of the resuspended solution was made and the optical density (OD) measured at 600 nm using an Eppendorf BioPhotometer spectrophotometer (Eppendorf). Before infiltration constructs were mixed 1:1 with p19 solutions.

Nicotiana benthamiana leaves were infiltrated by making 2 punctures using a needle and injecting the solution into the abaxial side of the leaf using a syringe without a needle. Infiltration continued until leaves had a water-soaked appearance. At least 1 leaf on two different plants was infiltrated for each solution. Plants used were 5-6 weeks old and leaves infiltrated were approximately 2.5-3.0 cm in size.

2.2.14 Agrobacterium-Mediated Stable Transformation of Tobacco Bright Yellow-2 Cell Cultures

Transformation of BY2 cell cultures was adapted from (Evans & Graumann, 2014) Agrobacterium strain GV3101 containing the chosen vector was grown overnight in a 5 ml culture of LB containing 40 µg/ml rifampicin, 10 µg/ml gentamicin and the appropriate antibiotic for plasmid selection. Agrobacteria were prepared for transformation by centrifuging the cells at 5000xg for 5 mins. The supernatant was discarded and cells were resuspended in 1 ml of BY2 growth media (see section 2.6.2) supplemented with 200 µM acetosyringone. Centrifugation was repeated and cells were washed 2 more times. After the final wash step 100 µl of agrobacteria were added to 7 ml of a three day old culture of BY2 cells. Cells were gently mixed and spread on a plate containing solid BY2 media with no antibiotics, sealed, and left for three days at 25 ^{o}C in the dark in an incubator (LMS).

After incubation cells were washed off the plate into a sterile tube using BY2 media containing 150 µg/ml carbenicillin, 100 µg/ml timentin and 250 µg/ml cefotaxime to kill remaining agrobacterium. Cells were centrifuged at 1000xg for 5mins, the supernatant was removed and the pellet resuspended in 10 ml BY2 media containing the antibiotics above. This was step was repeated three times in total. After the last wash cells were finally resuspended in 10 ml of BY2 media and spread onto solid BY2 plates containing 150 µg/ml carbenicillin, 100 µg/ml timentin, 250 µg/ml cefotaxime, and the appropriate antibiotic for selection for the plasmid of interest. Plates were left for 4-6 weeks for colony formation. Colonies were checked for fluorescence using

a Leica Stereofluor M165FC microscope with the appropriate filter setting. Independent fluorescing calli were transferred to gridded plates containing solid BY2 medium containing 150 µg/ml carbenicillin, 100 µg/ml timentin, 250 µg/ml cefotaxime, and the appropriate antibiotic for selection for the plasmid of interest. After 2-3 passages the calli transferred to solid BY2 media containing only the appropriate antibiotic for selection of the plasmid and the carbenicillin, timentin and cefotaxime were omitted. After at least two passages on plasmid selecting BY2 media only, calli were transferred into liquid culture. To transfer calli to liquid media for the first time sterile loops were used to remove a small piece of calli and transfer it into liquid media supplemented with the appropriate antibiotic for plasmid selection. The BY2 were grown in conditions as described in section 2.6.2. For calli just transferred cultures were left to grow for two weeks and sub cultured according to section 2.6.2 at least twice before being used for analysis.

2.2.15 Agrobacterium-Mediated Stable Transformation of *Arabidopsis thaliana*

Stable transformation of *A. thaliana* plants was done according to a method based on Clough and Bent (1998). Three to four week old flowering *A. thaliana* plants were grown and all siliques and opened flowers were removed 24 hours prior to dipping. A single colony of *A. tumefasciens* GV3101 harbouring the plasmid of interest was grown in a 5 ml overnight LB culture containing 40 µg/ml rifampicin and 10 µg/ml gentamicin and the appropriate antibiotic for plasmid selection at 30 ^{o}C shaking at 180 rpm. A 1:100 dilution of this overnight culture was added to 400 ml of LB supplemented with 40 µg/ml rifampicin and 10 µg/ml gentamicin and the appropriate antibiotic for plasmid selection and grown at 30 ^{o}C shaking at 180 rpm for 16-24 hours. Prior to dipping the cultures were centrifuged at 6000xg for 20 mins in a J-Lite rota at 4 ^{o}C (BeckmanCoulter). Pellets were resuspended in 800 ml of a 5 % sucrose (w/v) solution in autoclaved dH_2O supplemented with 0.02 % (v/v) Silwet L-77 (Fisher Scientific). Aerial parts of plants were dipped and plants were bagged to maintain humidity and kept in the dark for 24 hours. Bags were then opened and plants moved to growth chambers in conditions described in section 2.6.3. Dipping of the same construct was repeated seven days later but any siliques and flowers were left on the plants.

2.2.16 Genotyping the T-DNA Insertion Lines Using PCR and DNA Sequencing

Genomic DNA was extracted as described in section 2.2.12. An approximate location of the T-DNA insert was obtained using the software available at T-DNA Express. For the wild type (WT) reaction primers were designed approximately 500 bp upstream and downstream of the T-DNA insert. These primers were used to detect the presence of the wild type allele. For Salk T-DNA lines the Salk LBA_1 primer and whichever of the forward or reverse gene specific primers faced it

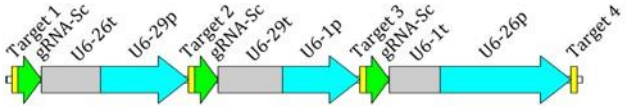
were used to confirm presence of the insert. For Gabi_Kat T-DNA insert lines the left border primer CAP2 08409 and whichever of the forward or reverse gene specific primers faced it was used to confirm the presence of the insert. Azygous plants would produce a band of the correct size in the WT reaction and no band in the insert PCR. Heterozygous plants would produce a band in both the WT PCR and the insert PCR reactions. Homozygous knockout mutants would produce no band in the WT PCR but a band in the insert PCR. As a positive control to confirm that the PCR reactions had worked, mago forward and reverse primers were used to amplify a 1868 bp fragment from the MAGO NASHI gene (At1g02140; see appendix 1 for sequences).

2.2.17 Reverse Transcription-PCR to Check Homozygote Status of Knockout Plants

RNA was purified from 100 mg of frozen seedling tissue and DNase-treated as described in 2.2.10. Using this RNA as a template, cDNA was synthesized using the Superscript III reverse transcriptase and the oligo(dT)20 primer as described in 2.2.11. To determine whether the transcript of the gene of interest was present in the single knockout line plants, gene-specific primers were designed to amplify regions upstream and downstream of the T-DNA insertion point (see Appendix 1 for primer sequences). As a control, gene-specific primers that amplified a 588bp fragment of *NET1A* (At3g22790) were used (see Appendix 1 for primer sequences). DNA amplification was undertaken as described in 2.2.10, using the MyTaq DNA polymerase as described in section 2.2.1. The PCR products were run on a 1 % (w/v) or 3 % (w/v) agarose gel prepared as described in section 2.2.2.

2.2.18 Generation of Double Knockout Lines Using CRISPR/Cas9 Technology

Generation of the double knockout CRISPR lines was performed by Dr J. Kroon. The CRISPR/Cas9 strategy was designed based on a toolkit using a binary vector set which has the pCAMBIA or pGreen backbone and an sgRNA module vector set (Xing *et al.*, 2014). Briefly, two target sites per gene for targeting the Cas protein were chosen within the *IRQ4* (T1 and T2) and *IRQ5* (T3 and T4) sequences using a website of genome-wide prediction of plant CRISPR/Cas9 target sites (http://www.genome.arizona.edu/crispr/CRISPRsearch.html). Target specificities were evaluated using the Cas-OFFinder (Bae *et al.*, 2014). To generate the four gRNAs (guide RNAs), PCR primers containing a *Bsa*I sequence were designed and primary PCR templates containing relevant PolIII regulatory sequences (Xing *et al.*, 2014), were obtained from Addgene (pCBC-DT1T2, pCBC-DT2T3, pCBC-DT3T4).



(Target-1)-(gRNA-Sc)-(U6-26t)-(U6-29p)-(Target-2)-(gRNA-Sc)-(U6-29t)- (U6-1p)-(Target-3)-(gRNA-Sc)-(U6-1t)-(U6-26p)-(Target-4)

Figure 2.1*:* Premade gRNA modules used for the assembly of the four guide RNAs. U6-29p, U6-26p, and U6-1p are three Arabidopsis U6 gene promoters; U6-29t, U6-26t, and U6-1t, corresponding Arabidopsis U6 gene terminators with downstream sequences, gRNA-Sc= gRNA scaffold.

A Gibson assembly method (Gibson *et al.*, 2009) fused the four gRNAs (Figure 2.1) and the assembled product was used in a Golden Gate cloning (Engler *et al.*, 2008) into the vector pHEE401E (obtained from Qi-Jun Chen (Wang *et al.*, 2015)). The vector allows egg cell-specific EC1.2 gene promoter-controlled (Steffen *et al.*, 2007; Sprunck *et al.*, 2012) expression of 3×FLAG-NLS-zCas9-NLS, contains gRNA scaffold for the insertion of a target sequence and encodes Hyg resistance.

The resulting CRISPR/Cas9 construct was transformed into *Agrobacterium tumefasciens* strain GV3101/pMP90 (according to section 2.2.7.2) and used in floral dip plant transformation of *Arabidopsis thaliana* Columbia-0 (according to section 2.2.15).

T1 seed was obtained followed by hygromycin screening for positive independent transformants and T1 CRISPR/Cas9 generated genomic editing events (according to section 2.6.4). Also T2 seed was obtained and T2 plants analysed for the presence and genetic status of CRISPR/cas9 generated gene editing events.

Selection for mutants was via done a T7 Endonuclease I based mutant screening (based on mismatch recognition in a hybrid between wild type DNA sequence and potential mutated DNA sequence) in combination with direct DNA sequencing of purified PCR fragment of the genomic region containing the CRISPR/Cas9 target sequences. The primers used in these procedures are listed in appendix 1.

2.3 Yeast-2-Hybrid

2.3.1 Yeast Transformation

Competent yeast were prepared and transformed using the EZ Yeast Transformation Kit (Zymogen). Glycerol stocks of the AH109 and Y187 yeast strains were streaked onto solid YPDA plates (20 g/l peptone, 10 g/l yeast extract, 0.003 % (w/v) adenine hemisulfate (Sigma), 2 % (w/v) D-glucose, 20 g/l agar, pH 6.5) and grown for 2 days at 30 °C. Single colonies were used to

inoculate a 10 ml YPDA culture that was grown for 12-14 hours at 30 °C or until mid-log phase (~5 x 10^6 - 2 x 10^7 cells/ml or OD_{600} of 0.8-1.0). Cells were centrifuged at 500xg for 4 mins and the supernatant discarded. The pellet was then washed using 10 ml of EZ 1 solution. Cells were then centrifuged at 500xg for 4 mins and the supernatant discarded. The pellet was resuspended in 1 ml of EZ 2 solution. Cells were then aliquoted out into 50-100 µl amounts and either used directly for transformation or frozen slowly and stored at -80 ^oC.

For transformation 50 µl of yeast cells (AH109 or Y197 strain) were mixed with 0.2-1.0 µg DNA and 500 µl EZ 3 solution. Cells were then incubated at 30 ^oC for 45 mins to 1 hour with no shaking but occasionally were gently mixed 2-3 times by inverting the tube. Cells were then spread on appropriate SD plates containing 50 µg/ml kanamycin and incubated at 30 ^oC for 3-4 days.

2.3.2 Autoactivation Tests

Autoactivation tests were used to ensure growth did not occur on selective media in the absence of mating. Autoactivation tests were performed for IRQ4 (pGBKT7/pGADT7), which were both transformed into the AH109 and Y187 strains (according to 2.3.1). Three colonies of each strain were resuspended in 100 µl sdH2O and 10 µl of each colony was dropped onto selective SD dropout plates (containing 50 µg/ml kanamycin). For the autoactivation tests for IRQ4 in pGBKT7 and pGADT7, colonies were dropped onto selective media lacking either leucine or tryptophan with the addition of 2.5 mM 3AT (3-Amino-1,2,4-triazole, Sigma). Colonies were also dropped onto -L and -W plates as a control for colony growth. For NET4B and IRQ + CT fragments in pGBKT7, seven plates were used (-WH, -WH + 2.5 mM 3AT, -WH + 5 mM 3AT, -WH + 10 mM 3AT, -WA, -WHA, -W + 100 µM X-α-Gal (5-Bromo-4-chloro-3-indolyl-α-D-galactopyranoside; Sigma). Colonies were allowed to dry and the plates were incubated for 5-7 days at 30 °C. Plates were imaged using an Epson Perfection 4490 Photo flatbed scanner.

2.3.3 Yeast-2-Hybrid Library Screen

Two separate yeast-2-hybrid screens were performed using IRQ4 as bait. Firstly, an 80 ml overnight SD (-W) culture of the bait construct in Y187 was grown for 24 hours until they reached an OD_{600} of 2.0 and the cells were counted using a haemocytometer. One millilitre of random primed cDNA-prey library cells and 1 ml of oligodT cDNA-prey library cells (both from whole seedling), were mixed with the relevant volume of bait culture containing 1.0 x 10^9 bait cells (40-60 ml). The mixed culture was centrifuged at 1000 g for 5 minutes and resuspended in 10 ml YPDA which was poured onto solid YPDA plates and incubated overnight at 30 °C. The mated cells were washed off the plates in 10 ml YNB-AA, centrifuged as above and resuspended in 10 ml fresh YNB-AA. A hundred microlitres of these cells were used for the mating efficiency tests as described in

2.3.4. The remainder of the cells were plated onto 15 cm plates, containing SD media with -WLHA + 10 mM 3AT dropout (with the addition of 50 µg/ml kanamycin). Five hundred microlitres of cells were used per plate and the plates were incubated at 30 °C for 10 days. Colonies that started to appear after 5 days incubation were streaked onto identical SD dropout plates (50 µg/ml kanamycin), and grown for a further 5 days at 30 °C. The colonies that could survive on the second selection plate were used to inoculate liquid cultures and the cDNA-prey plasmid clone was rescued (according to 2.3.5).

2.3.4 Mating Efficiency Tests

Before spreading mated cells on SD selection plates (section 2.3.3), 100 µl of the cells was reserved for mating efficiency tests. Serial dilutions of the cells were made using sdH2O. Dilutions were made at 1:1000, 1:10,000, 1:100,000, and 100 µl of the dilutions were plated onto SD -L (to select for prey library) and -LW (to select for diploids) plates (containing 50 µg/ml kanamycin) and cells were counted after 3 days growth at 30 °C. The number of diploid colonies screened was calculated by multiplying the number of colonies on the 1:100,000 dilution by the plating volume (100 µl) and total resuspension volume for the mated cells (10 ml). Mating efficiency (% of diploids) was calculated as: (Number of diploid colonies / Number of library colonies) x 100.

2.3.5 Plasmid Rescue

Colonies that successfully grew after being restreaked onto fresh SD dropout media plates (-WLHA + 10 mM 3AT) were taken forward for plasmid rescue. Colonies were inoculated into 5 ml liquid cultures of SD -WHLA +10 mM 3AT (containing 50 µg/ml kanamycin) and grown for 48 hours at 30 °C. Cultures were centrifuged at 5,000xg for 5 minutes and resuspended in 50 µl yeast lysis buffer (YLB; 50 mM Tris.HCl pH 7.5, 0.85 M Sorbitol (BDH), 10 mM EDTA, containing a few grains of lyticase (Sigma) per ml YLB). Cells were incubated at 37 °C for 1 hour before being flash frozen in liquid nitrogen and thawed at 37 °C. The Promega Wizard Plus SV Miniprep DNA Purification System was subsequently used to purify plasmid DNA (as described in 2.2.8) starting with the addition of the resuspension buffer. Purified plasmid DNA was then sequenced, (according to 2.2.9) using the T7 promoter primer (see Appendix 1 for sequence). Prey plasmids that were used for further one-on-one testing were transformed into E. coli DH5α (according to 2.2.7.1) to amplify the DNA before being transformed into the yeast Y187 strain (as described in 2.3.1)

2.3.6 Yeast One-On-One Mating Tests to Verify Interactions

One-on-one mating tests were used to confirm protein-protein interactions. Bait and prey vectors (pGBKT7 and pGADT7) were transformed into the yeast strains AH109 and Y187 respectively. Single colonies of the bait vector (AH109) were first resuspended in 100 µl sdH2O and 10 µl of this suspension was dropped onto solid YPDA plates. Once the drop had dried, single colonies of the prey vector (Y187) were similarly resuspended and 10 µl was placed on top of the dried drop containing the bait vector. The drops were allowed to dry and were then incubated at 30 ^{o}C for 24-48 hours to allow the yeast to mate. As a negative control for autoactivation, the bait and prey vectors containing the plasmids of interest were also mated with the respective empty vectors. Typically, each one-on-one test mated three individual colonies for each bait-prey vector combination.

Once the yeast had grown, the mated cells were resuspended in 100 µl sdH2O and 10 µl was dropped onto solid SD -WL dropout media (containing 50 µg/ml kanamycin) to select for diploid mated cells. Plates were incubated at 30 °C for 3-4 days until colonies started to appear.

Diploid colonies were resuspended in 100 µl sdH2O and 10 µl of this suspension was dropped onto selective SD dropout media (containing 50 µg/ml kanamycin) to test for potential interactions. Different dropout selections were used during this project including:-WL (as a positive control for yeast growth) -WLA, -WLH + 2.5 mM 3AT, -WLH + 10 mM 3AT, -WLHA + 2.5 mM 3AT, -WLHA + 10 mM 3AT. Once the drops had dried, the plates were sealed with cling film and incubated at 30 ^{o}C for 5-7 days. Yeast growth was imaged using an Epson Perfection 4490 Photo flatbed scanner.

2.4 *In vitro* Protein Methods

2.4.1 Expression and Purification of *His*-Tagged Proteins for Antibody Production

Competent Rosetta 2 cells were transformed as described in section 2.2.6.1 with the pGAT4 vector containing the fragment of interest cloned using Gateway Cloning as described in section 2.2.3. Autoinduction of expression was performed by transferring a colony of transformed bacterial cells into 10 ml LB containing no antibiotics and incubating for 6-7 hrs at 37 ^{o}C. One millilitre of this culture was taken and put on ice. This sample was the T0 sample. Three hundred microlitres of the overnight incubation was transferred to 600 ml of 2YT autoinduction medium (AIM; Formedium) and incubated overnight shaking at 300 rpm incubated at 37 ^{o}C. Once the culture was incubated 2 ml of the culture was harvested and stored on ice for processing (AIM

samples; see below for processing). The remaining culture was harvested by centrifugation in at 5000xg for 20 mins at 0 °C using a Beckman Coulter Avanti J-20 XP centrifuge fitted with a JLA 10.500 rotor. The pellet was then resuspended in ice cold tris/HCl (pH 8.0) and transferred to a 50 ml tube. The resuspended pellet was centrifuged at 3000xg for 20 mins at 4 °C. All the supernatant was then removed and the protein was extracted from the pellet.

To process samples (T0 and AIM) the sample was spun for 30 secs at maximum speed in benchtop centrifuge (Eppendorf) cooled to 4 °C. All the supernatant was removed and pellets were resuspended in 1X SDS loading buffer (312 mM Tris-HCl, pH 6.8, 10 (w/v) SDS, 50 % (v/v) glycerol, 5 % ß-mercaptoethanol, 0.05 % bromophenol blue) and heated for 5 minutes at 95 °C before running on an SDS polyacrylamide gel.

Protein was extracted from the overnight AIM cultures by resuspending the pellet in protein extraction buffer (1 M HEPES (N-(2-hydroxyethyl)piperazine-N'-(2- ethanesulfonic acid, Sigma) pH 7.0, 5 M NaCl, 5 mM ß-mercaptoethanol, 8 M Urea (BDH)), containing proteinase inhibitors (10 µg/ml leupeptin (Sigma), 1 mM PMSF (Sigma), 10 µg/ml TAME (Nα-p-Tosyl-L-arginine methyl ester hydrochloride, Sigma) 1 µg/ml pepstatin A (Sigma)). The bacterial cells were lysed by flash freezing in liquid nitrogen and thawing in a 37 °C water bath. The lysed cells were ultracentrifuged at 350,000 g for 15 minutes at 4 °C using a Beckman Optima TLX ultracentrifuge fitted with a TLA 100.4 rotor.

The resultant supernatant containing the protein extract was removed and filtered through a 0.45 µm filter. The filtered extract was added to a column containing nickel nitrilotriacetic acid (Ni-NTA) agarose beads (Qiagen), which had been previously equilibrated with protein extraction buffer. The protein extract and Ni-NTA beads were incubated for 15 minutes on a rotary shaker at room temperature. The extraction buffer was pulled through the column, using a syringe attached to the base, and the beads were subsequently washed using 20 mM, 40 mM and 60 mM imidazole wash buffer (1 M HEPES pH 7.0, 5 M NaCl; 20 mM, 40 mM and 60 mM respectively imidazole (Sigma), and 6 M, 4 M and 3 M respectively urea). Finally, the His-tagged protein was eluted from the column using elution buffer (50 mM HEPES pH 7.0, 300 mM NaCl, 250 mM imidazole, 2 M Urea). Protein concentration was analysed with the Bradford assay (Bio- Rad Protein Assay), using a spectrophotometer set at a wavelength of 595 nm.

2.4.2 Dialysis of Purified Protein

After purification of the 6xHis-tagged IRQ4 antigen fragment ($IRQ4^{125-248}$), the urea was removed by buffer exchange. The eluted protein was dialysed overnight at 4 °C in 1x PBS buffer (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4 (di-sodium hydrogen orthophosphate dihydrate (BDH)), 1.8

mM KH_2PO_4 (potassium phosphate monobasic (Sigma), pH 7.4) using dialysis tubing. Purified protein was analysed using SDS-PAGE gel electrophoresis (according to 2.4.3) and was flash frozen in liquid nitrogen and stored at -80 °C.

2.4.3 SDS Polyacrylamide Gel Electrophoresis

To separate protein fragments based on their size, SDS-PAGE was used. Resolving gels were prepared by adding appropriate amounts of acrylamide (Protogel, 30 % (w/v) acrylamide, 0.8 % (w/v) bisacrylamide solution (37.5:1 ratio), National Diagnostics) to the resolving buffer (0.1 % (w/v) SDS, 375 mM Tris). Polymerization was activated by the addition of 0.1 % (v/v) ammonium persulfate solution (APS, Sigma) and 1.4 µl/ml TEMED (NNN'N'-Tetramethylethylenediamine, BDH). The stacking gel was prepared by adding the appropriate amount of acrylamide to the stacking buffer (0.1 % (w/v) SDS, 125 mM Tris). Polymerization was activated by the addition of 0.1 % (v/v) APS and 4 µl/ml TEMED. Prior to loading, protein was diluted 4:1 using 5x Laemmli sample buffer (312 mM Tris-HCl, pH 6.8, 10 % (w/v) SDS, 50 % (v/v) glycerol, 25 % ß-mercaptoethanol, 0.05 % bromophenol blue) and heated for 5 minutes at 95 °C. ° Electrophoresis was performed at 100-150V in an Bio-Rad Mini Protean II tank filled with electrode buffer (25 mM Tris, 190 mM glycine (Melford), 0.1% (v/v) glycerol, diluted 1:10 with dH2O before use). Precision Plus Dual Colour Protein Ladder (Bio-Rad) was used as a size marker.

2.4.4 Production of Polyclonal Antibodies

Polyclonal antibodies were raised in mice against a fragment of IRQ4 ($IRQ4^{125-248}$), purified from *E. coli* and dialysed into PBS buffer (as described in sections 2.4.1 and 2.4.2). Approximately 90 µg of the purified IRQ4 antigen fragment was used per injection, and was mixed 1:1 with Freund's complete adjuvant (FCA) for the first inoculation and Freund's incomplete adjuvant (FIA) for all subsequent injections. Test-bleeds were taken prior to the first injection. Three mice were inoculated a total of 6 times at day 0, 14, 28, 42 and 56. At day 63, the terminal bleed was collected and the serum was separated. Test-bleeds and terminal bleeds were used for Western blot analysis and immunogold labelling. Antibody production (including animal care, injections, test-bleeds and serum collection) was performed in-house by the Durham University Life Science Support Unit (LSSU).

2.4.5 Total Protein Extraction for Initial Antibody Tests

Protein was extracted from 15-day old vertically grown Arabidopsis Col-0 plants for initial assessment of the mouse terminal bleeds used in the production of the IRQ4 polyclonal antibodies. Plant tissue was ground to a fine powder in liquid nitrogen using a pestle and mortar.

Homogenised tissue was then added to 2x protein extraction buffer (250 mM Tris-HCl pH 6.8, 4 % (w/v) SDS, 20 % (v/v) glycerol, 100 mM ß- mercaptoethanol, 0.02% Bromophenol blue (Sigma)) that was preheated to 95 °C. Samples were heated at 95 °C for 5 minutes and centrifuged at 16,000 g for 2 minutes and supernatants were stored at -80 °C until required

2.4.6 Total Protein Extraction for Quantitative Analysis

Protein quantification was carried out on total protein extracts prior to use in Western blot analysis following an established protocol (Martinez-Garcia *et al.*, 1999). An extraction buffer that is compatible with commercial protein quantification assays was used, as it contains a low concentration of SDS and uses sodium metabisulfate ($Na_2S_2O_5$) as a reducing agent instead of dithriothreitol (DTT) or ß-mercaptoethanol.

Protein was extracted from various sources including Arabidopsis Col-0 tissues (seedling root, seedling leaves, inflorescence stems, flowers, and siliques), 15-day old vertically grown *Arabidopsis irq4* and *irq5* T-DNA and *irq4/5* double CRISPR mutant plants. Plant tissue was collected in microcentrifuge tubes, flash frozen in liquid nitrogen and subsequently homogenised using a liquid nitrogen cooled micro pestle. Protein extraction buffer (125 mM Tris-HCl pH 8.8, 1 % (w/v) SDS, 10 % (v/v) glycerol, 50 mM $Na_2S_2O_5$ (Sigma)) was added to the homogenised tissue in the microcentrifuge tubes and then was warmed to room temperature to resolubilise the SDS. Samples were centrifuged at 16,000 g for 10 minutes and the supernatant decanted into a fresh tube. A small aliquot (approximately 10-20 µl) of the supernatant was used for calculating protein concentration using a spectrophotometer set at 750 nm and the Bio-Rad Detergent-compatible (Dc) Protein Assay (Bio-Rad). The remaining supernatant was diluted 4:1 using 5x SDS loading buffer and samples were heated at 95 °C for 5 minutes. Samples were then centrifuged at 16,000 g for 2 minutes and supernatants were stored at -80 °C until required. Quantity of plant material varied from 100-300 mg and typically 2 µl protein extraction buffer was used per mg of plant material.

2.4.7 Coomassie Brilliant Blue Staining

For visualisation of protein in SDS-PAGE gels, gels were stained with Coomassie Brilliant Blue stain (CBB) (7 % (v/v) glacial acetic acid, 35 % ethanol, 0.125 % (w/v) Coomassie Brilliant Blue R-250 (BDH)) for 20 minutes to overnight and destained using dH_2O for two hours to overnight.

2.4.8 Western Blot Analysis

Western blotting was used to transfer proteins from SDS-PAGE gels onto nitrocellulose membrane (Whatman) for immunological detection using the Bio-Rad Mini Protean II system. Proteins were transferred onto the membrane overnight at 10V at 4 ^{o}C in a tank containing transfer buffer (48 mM Tris, 39 mM glycine, 0.375 % (w/v) SDS, 20 % (v/v) methanol). Following overnight transfer, the membrane was washed in dH2O and protein bands were visualised using Ponceau staining solution (0.1 % (w/v) in 5 % acetic acid (Sigma-Aldrich)). Membranes were stained for 5 minutes to check for transfer of protein and imaged before being washed in dH_2O to remove the stain.

The probing conditions used here were similar in each western blot performed. Nitrocellulose membranes were blocked for 30 minutes in 2x TBST (20 mM Tris pH 7.4, 300 mM NaCl, 0.1 % (v/v) Tween 20 (Sigma)) containing 5 % (w/v) dried skimmed milk powder (Tesco). Membranes were then incubated with the primary antibody diluted in 2x TBST containing 5 % milk, for one hour at room temperature. Primary antibody dilutions varied, but typically 1:500 was used for the anti-IRQ4 polyclonal antibodies. For detection of GFP an anti-GFP primary antibody (Chromtek) was used at a dilution of 1:200.

The membrane was then washed three times for 5 minutes using 2x TBST, before being probed for 30 minutes with the secondary antibody diluted 1:3000 in 2x TBST containing 5 % milk. Secondary antibodies used were conjugated to horseradish peroxidase (HRP). Anti-mouse IgG/HRP antibodies (Dako, raised in pigs) were used for detection of anti-IRQ4 antibody. Anti-Rat IgG/HRP antibodies raised in rabbit were used to detect the anti-GFP antibody.

Following secondary antibody incubation, the membranes were washed as before and antibody binding was visualised with the application of ECL Western Blotting Detection Reagent (Amersham Biosciences). ECL was incubated for 1 minute before being visualised using a Fujifilm Intelligent Dark Box II and Fujifilm film image reader system, or being exposed to X-ray film (SLS) which was developed and fixed in a dark room. X-ray films were scanned using an Epson Perfection 4490 Photo flatbed scanner.

2.5 Cell Biology

2.5.1 High Pressure Freezing and Freeze Substitution of *Arabidopsis thaliana* Root Tips

Preparation of samples for immunogold labelling including fixation, embedding, sectioning and immunogold labelling were performed by the departmental Electron Microscopy Unit.

Root tips were prepared using high-pressure freezing and freeze substitution (HPFFS). *Arabidopsis thaliana* seedling were grown vertically on 1xMS plates for 7 days as described in section 2.6.3. The distal 1-2 mm root tips were removed using a razor blade and immersed in 20 % BSA (w/v). Samples were transferred onto membrane carriers (Leica Microsystems GmbH) and high-pressure freezing was performed using a Leica EMPACT high-pressure freezer.

Samples were then fixed by freeze substitution using a Leica EM automated freeze substitution (AFS) device (Leica Microsystems GmbH). Samples were freeze substituted using anhydrous acetone containing 0.25 % (v/v) glutaraldehyde and 0.1 % (w/v) uranyl acetate. Samples were incubated in this solution for 48 hours at -80 ^{o}C. Thereafter the temperature was raised by hourly 1 ^{o}C increments until it reached 50 ^{o}C. The sample was then washed several times in anhydrous acetone and removed from the membrane carrier using a needle. Infiltration continued using Monostep Lowicryl HM20 (Agar Scientific) at -50 ^{o}C by increasing the ratio of resin to acetone. Three changes of resin/acetone were done at 12 hour intervals using increased concentrations of resin (22, 23, and 66 % resin). Samples were then incubated in 100 % resin for 96 hours. The resin was changed every 24 hours over the 96 hour period. Resin embedding and UV polymerisation were performed at -50 ^{o}C for 48 hours. After which the temperature was slowly raised to 20 ^{o}C and maintained at 20 ^{o}C for 24 hours. Once polymerisation was complete, ultrathin sections were taken using a diamond knife and placed on coated grids.

2.5.2 Immunogold Labelling and Transmission Electron Microscopy

Ultrathin sections were rinsed with 0.1 % (w/v) glycine in PBS and subsequently blocked for 30 min with 1 % (w/v) BSA in PBS. Samples were then incubated three times with PBS containing 0.1 % (w/v) BSA-c (Aurion), pH 7.4. BSA-c is a modified BSA which prevents charged based background noise in samples by acetylating amines on basic amino acid groups. This modified BSA is more negatively charged overall and competes with the gold and antibody particles for the non-specific binding of components which would otherwise cause increased background.

After blocking sections were incubated for 30 mins with the primary antibody (anti-IRQ4 (M3T)) diluted 1:100 in 0.1 % BSA-c. The grids were then washed with 0.1 % BSA-c in three 5 min incubations. Sections were then incubated with colloidal gold-conjugated goat anti-mouse secondary antibody (British Biocell International) diluted 1:20 in 0.1 % BSA-c. Grids were then rinsed with three 5 second washes of PBS, followed by 2 four minute washes with PBS. Following this, the antigen-antibody-gold complex was stabilised by incubating for 5 mins in 1 % (v/v) glutaraldehyde in PBS.

For imaging by transmission electron tomography (TEM), samples were stained with 2 % (w/v) aqueous uranyl acetate, followed by incubation in Reynold lead citrate for 5 mins. Images were taken using a Hitachi H-7600 TEM operating at 100 kV fitted with an AMT Orca-ER digital camera (Advanced Microscopy Techniques).

2.5.3 Confocal Laser Scanning Microscopy

To visualize fluorescent proteins in *N. benthamiana* leaf epidermal cells, the fusion proteins were transiently expressed in *N. benthamiana* leaf epidermal cells as described in section 2.2.13. Imaging was done 2-4 days post-infiltration. To image the epidermal cells, leaf sections were mounted in dH2O under a cover slip and the abaxial surface of the leaf was viewed.

To visualise fluorescent proteins in BY2 cells, cells were sub-cultured and grown for 2-3 days as described in section 2.6.2. Approximately 100 µl of cells were then transferred to a glass slide and a cover slip put gently over the top for imaging. Alternatively, approximately 1 ml of cells was transferred to glass bottomed observation dishes and imaged without disturbance to the cells.

To visualise fluorescent proteins in *A. thaliana* root tips, 5-7 day old seedlings were grown as described in section 2.6.3. All cytoskeletal and endomembrane disrupting drug treatments to root tips were applied to 5-7 day old plants (section 2.5.4). Seedlings were mounted in 1XMS, 1 % sucrose, pH 5.7 (non-drug treatments) or in the media as described in section 2.5.4 for drug treatment assays.

CSLM was performed on a Zeiss 880 confocal laser scanning microscope, using 10X, 20X air and 40X or 63X oil immersion lens as appropriate. GFP was imaged using the 488 nm 20 mW argon laser and emission was detected at 500-560 nm. To image GFP the 488 nm laser settings typically ranged between 5-14 % total laser power, and a smart gain of 500-800 mV was used for PMT detectors, depending on signal strength. To image mRFP, the fluorophore was excited using a HeNe 543 nm laser and emission was detected at 600-680 nm. Laser settings for excitation of mRFP and neutral red dye were typically between 20-30 %, with similar detector settings as used for GFP. Sequential scans were used for the colocalisation experiments using the 488 nm and 543 nm laser, and images were overlaid using the Zen 2 Lite (Zeiss) software. For time series 200 frames were recorded per sample.

2.5.4 Cytoskeletal and Endomembrane Disrupting Drug Treatments in BY2 Cell Cultures and *Arabidopsis* roots

For disruption of the microtubule cytoskeleton in BY2 cells, cells were treated with 37.5 µM oryzalin for 1.5 hours (Sigma, prepared from a 1 mM stock in DMSO). As a negative control, cells were treated with an equivalent amount of DMSO dissolved in BY2 cell culture medium. For *Arabidopsis*, roots of seedlings were incubated in 1XMS containing 37.5 µM oryzalin for 1 hour. As a negative control roots were incubated in 1XMS containing the equivalent amount of DMSO.

For disruption of the actin cytoskeleton in BY2 cells, cells were incubated with 10 µM latrunculin B (Sigma, prepared from a 10 mM stock in DMSO) for 1 hour. As a negative control, cells were treated with an equivalent amount of DMSO dissolved in BY2 cell culture medium. For *Arabidopsis*, roots of seedlings were incubated in 1XMS containing 4 µM latrunculin B for 1 hour. As a negative control roots were incubated in 1XMS containing the equivalent amount of DMSO.

Wortmannin is an inhibitor of phosphatidylinositol 3-kinases (PI3K) and causes disruption to PVC localisation. BY2 cells were treated with 30 µM wortmannin for 3 hours (ThermoFisher, prepared from a 3 mM stock dissolved in DMSO). As a negative control, cells were treated with an equivalent amount of DMSO dissolved in BY2 cell culture medium. For *Arabidopsis*, roots of seedlings were incubated in 1XMS containing 30 µM wortmannin for 3 hours. As a negative control, roots were incubated in 1XMS containing the equivalent amount of DMSO.

Brefeldin A (BFA) inhibits the GTPase activity of ADP-d factor 1 (Arf1p) causing disruption to Golgi localisation. BY2 cells were treated with 10 µM BFA for 1 hour (Sigma, prepared from a 10 mM stock dissolved in DMSO). As a negative control, cells were treated with an equivalent amount of DMSO dissolved in BY2 cell culture medium. For *Arabidopsis*, roots of seedlings were incubated in 1XMS containing 10 µM wortmannin for 1 hour. As a negative control, roots were incubated in 1XMS containing the equivalent amount of DMSO.

2.5.6 GUS Histochemical Staining

Plants were grown to various growth stages on vertical plates (just germinated seedlings, 5-7 day old plants, 2 week old plants) or soil (5-6 week old plants) as described in section 2.6.3. Tissue was incubated in GUS buffer (100 mM phosphate buffer (from 0.2 M stocks of Na_2HPO_4 and NaH_2PO_4), 10 mM EDTA, 0.1% Triton X-100 (Sigma)) supplemented with 20 mM X-Gluc (5-Bromo-4-chloro-3-indonyl-glucuronide (Melford)). All tissues except just germinated seedlings were vacuum infiltrated using an Eppendorf concentrator 5301 (Eppendorf) for 5 mins prior to incubation at 37

°C. Incubation times varied depending upon the tissue and promoter of interest. Table 2.8 gives indicative times for each promoter and tissue type.

For direct imaging, whole seedlings were cleared in an ethanol series at 10 %, 30 %, 50 %, and 70 % to 95 % for 30 mins each at room temperature to remove chlorophyll and clear the tissues. Tissues were first resuspended in 20 % glycerol (v/v) before being placed on coverslips for imaging using a Panasonic 3CCD HD Camera head powered to a Panasonic GP-US932A power unit attached to a Leica Stereofluor M165FC microscope.

2.5.7 Histochemical staining with BCECF-AM

The lytic vacuole in the roots of the *irq4, irq5,* and double *irq4/5* mutant lines was visualised by using BCECF-AM (2',7'-Bis-(2-carboxyethyl)-5-(6)-carboxyfluorescein, acetoxymethyl ester), a pH sensitive dye which labels the lytic vacuole by accumulating in acidic compartments (Swanson & Jones, 1996) . The protocol used here is a modified version of previously described method (Feraru *et al.*, 2010; Kolb *et al.*, 2015).

The F3 generation of *irq4.1, irq5.1, irq4/5CR.1* and *irq4/5CR.2*, azygous, and WT Col-0 plants were grown vertically for 5 days (according to 2.6.3). Seedlings were transferred from the 1XMS plates into liquid 1XMS media containing 10 µM BCECF-AM (1.24 mM stock in DMSO (Sigma)), and incubated for one hour in the dark on an orbital shaker. Seedlings were then mounted in liquid 1XMS media before imaging using CLSM (according to in 2.5.3). BCECF-AM was excited at 488 nm and emission was detected between 500-570 nm. A minimum of 5 seedlings were imaged for each genotype.

2.6 Plant Growth Conditions and Phenotypic Analysis

2.6.1 *Nicotiana benthamiana* Growth Conditions

Nicotiana benthamiana plants were grown centrally by Durham University SBBS. Seeds were sown directly onto compost (seed and modular compost plus sand; F2 + S (Levington)), to establish a seedling nursery. The seeds were covered with a plastic propagator lid to maintain humidity. The seedlings were grown in a growth chamber under a 16/8 hour photoperiod, with temperatures of 24 °C and 21 °C during the day and night respectively. Once the seedlings had grown approximately 1-2 cm tall they were transplanted into individual 10 cm pots and until they were 5-6 weeks old.

2.6.2 Tobacco Bright Yellow-2 Cell Growth Conditions

Tobacco BY2 (Bright Yellow-2) suspension cultures were grown in 1xMS salts, pH 5.8, containing 30 g/L sucrose (Melford), 0.2 mg/L 2,4-D (2,4-dichlorophenoxyacetic acid; Sigma), 200 mg/l KH_2P0_4 (Sigma), and 1 mg/L Vitamin B1. Cells were maintained in 50 ml of this medium and routinely subcultured every 7-days by taking 1 ml of culture and transferring into 50 ml fresh media. Cells were maintained at 25 ^oC shaking at 125 rpm in the dark.

2.6.3 *Arabidopsis thaliana* Growth Conditions

Seeds were surface sterilised by washing in 70 % ethanol for 30 secs, removing and replacing the ethanol with 6% sodium hypochlorite (BDH) for 15 minutes, and finally washing the seeds five times in sterile distilled (sd) H_2O to remove the sodium hypochlorite. Sterilised seeds were sown in 9 cm Petri dishes onto 1XMS (pH 5.7) supplemented with 0.8 % (w/v) plant agar (Duchefa). Root growth assays used 12cm plates to make measurement of root lengths easier and to ensure seedlings grew uniformly. Plates were sealed with micropore tape (3M) and stratified for 3 days at 4 ^oC in darkness. Plants were grown vertically in a temperature controlled Sanyo growth cabinet with a 16/8 hour photoperiod, and temperatures of 22 ^oC and 16 ^oC during the day and night respectively.

If seedlings were transferred to soil, they were sown onto autoclaved compost at 5-7 days old, covered with a plastic propagator lid to maintain humidity, and transferred to a growth room with the same settings as the growth cabinet. The propagator lid was removed after 3-4 days.

2.6.4 Selection for Transgenic *Arabidopsis thaliana* plants

Selection of plants transformed with hygromycin or kanamycin resistance based vectors was based on a previously published method (Harrison 2006). T1 seeds were harvested after dipping (according to section 2.6.5). Seeds were surface sterilised according to section 2.6.3. After the final wash step, the water was removed and seeds were resuspended in an autoclaved 0.1-0.4 % agarose solution made without MS or sucrose but supplemented with either 15 µg/ml hygromycin B (Melford) or 50 µg/ml kanamycin (Melford) for the respective resistance of the vector. The agarose was poured into 15 cm circular petri dishes on top of a basal agar medium consisting of 1XMS (pH 5.3) containing 1 % (w/v) plant agar and supplemented with the appropriate amount of the antibiotic. Excess liquid was drained away before the plates were sealed with micropore tape. Seeds were stratified in the dark at 4 ^oC for two days following which they were transferred to a growth chamber and incubated for 4-6 hours in continuous white light (120 µmol m^{-2} s^{-1}) in order to stimulate germination. Plates were then completely covered in foil and left in the growth

chamber to incubate for two days. After two days the foil was carefully removed from the plates and seedlings were incubated for a further 24-48 hours in continuous white light as before. Seedlings which had green cotyledons were deemed to be positively selected and were carefully transferred to soil and covered with a propagator lid. The lid was removed after 1-2 weeks once seedlings had become established. The selection was verified using PCR as described in section 2.2.1.

Plants which had been transformed with a BASTA resistance based vector were sown onto autoclaved compost in 9 cm pots using sand as a carrier and vernalised as described in section 2.6.3. Plants were then transferred to the growth room. At approximately the two leaf stage (1-2 weeks after germination), plants were evenly sprayed with a 120 mg/ml BASTA solution using a handheld sprayer. Seedlings were placed back in the growth chamber and were sprayed again after a further seven days. Seedlings which retained green cotyledons and leaves were deemed to be positively selected and transferred out of the nursery and into individual pots. The selection was verified using PCR as described in section 2.2.1.

2.6.5 *Arabidopsis thaliana* Seed Collection

The ARACON container system (Betatech) was used to collect *Arabidopsis* seeds. ARACON bases were placed over rosettes just after bolting had occurred. Plastic sleeves were used to segregate plants and stop cross contamination of seed batches. Once plants had reached senescence they were transferred to a seed drying room and left for approximately one month before being collected. Seeds were stored either in microcentrifuge stubs or 50 ml tubes and kept at room temperature in the dark.

2.6.6 Root Growth Assay in Nitrate Starvation Conditions

The nitrogen starvation assays used here were based on Wang *et al.* (2016). The F3 generation of *irq4.1, irq5.1, irq4/5CR.1* and *irq4/5CR.2*, azygous, and WT Col-0 plants were surface sterilised according to section 2.6.3 and sown onto half strength MS ($^1/_2$MS) at pH 5.6 supplemented with 1 % sucrose and 0.8 % plant agar. Plates were vernalised for 2-3 days and then transferred to a Sanyo growth cabinet using the conditions described in section 2.6.3. Plants were grown vertically for 6 days after which images were taken and the primary root length measurements calculated using ImageJ software (Rasband). Mean root lengths were analysed according to section 2.7.3.

2.6.7 Root growth for autophagasome counts on Nitrate Starvation Conditions

Seeds of the *prom*IRQ4::IRQ4-GFP plants were surface sterilised according to section 2.6.3 and sown on to onto half strength MS ($^1/_2$MS) at pH 5.6 supplemented with 1 % sucrose and 0.8 % plant agar. Plates were vernalised for 2-3 days and then transferred to a Sanyo growth cabinet using the conditions described in section 2.6.3. After 5-6 days of vertical growth seedlings were transferred to either a fresh plate containing the same media containing nitrogen (+N) or placed on media where the 1XMS has been replaced with 1XMS N-free media (-N). All other components of the media remained the same. Plants were then placed back in the growth chamber for a further 24 hours before being imaged by CSLM according to section 2.5.3 and autophagasome number calculated according to section 2.7.4.

2.6.7 Salt Stress Conditions Assay

The nitrogen starvation assays used here were based on Wang *et al.* (2016). The F3 generation of *irq4.1, irq5.1, irq4/5CR.1* and *irq4/5CR.2*, azygous, and WT Col-0 plants were surfaces sterilised according to section 2.6.3 and sown onto half strength MS ($^1/_2$MS) at pH 5.6 supplemented with 1 % sucrose and 0.8 % plant agar. Plates were vernalised for 2-3 days and then transferred to a Sanyo growth cabinet using the conditions described in section 2.6.3. Plants were grown vertically for 5 days before being transferred to either the same media or media supplemented with 160 mM NaCl (Melford). Plants were transferred back into the growth chamber and their growth monitored for signs of chlorosis in the leaves. Between 33 and 40 plants were monitored per line.

2.7 Statistical and Image Analysis

2.7.1 Statistical Analysis Software

Statistical analysis was carried out using either SPSS (version 17.0; SPSS) for the 2-way ANOVA and one-way ANOVA or using Microsoft Excel (v13.0; Microsoft) for the student's t-test as described below for different assays. Graphs were generated using Excel.

2.7.2 Quantitative analysis of punctae speed movement in BY2 cell culture and *Arabidopsis* roots

To calculate the speed of IRQ4 GFP-labelled punctae in BY2 and Arabidopsis plants when treated with oryzalin and latrunculin B (according to section 2.5.4), image stacks consisting of 200 frames taken by CSLM were analysed according to section 2.5.3. Stacks were imported into ICY image analysis software (http://icy.bioimageanalysis.org). The spot detector plug-in was used to detect the fluorescent IRQ4 punctae while the Spot Tracking plugin was used to calculate the average

speed of the dots over the time frame of the movie. The averages were imported into Excel and a student's t-test was used to calculate whether significant difference occurred between the average speeds on the different treatments. Punctae speed was analysed for 10 cells on each media composition.

2.7.3 Quantitative Analysis of Root Lengths under Nitrate Starvation Conditions

Mean root lengths for *irq4.1, irq5.1, irq4/5CR.1* and *irq4/5CR.2*, azygous, and WT Col-0 mutant plants grown in N deprivation conditions were measured using line tool in ImageJ. Mean root lengths for each genotype and condition were calculated in excel and imported into SPSS for statistical analysis. To understand whether the genotype and media type (+/-N) factors interacted a 2-way ANOVA was calculated. To compare individual genotype and media types using a post-hoc Tukey test a one-way ANOVA was also calculated. The number of root lengths per genotype and media condition varied between 20-110.

2.7.4 Quantitative Analysis of Autophagasome Number in Roots Grown Under Nitrate Starvation Conditions

Significant difference in autophagasome number count between *prom*IRQ4::IRQ4-GFP plants transferred to –N or +N media, as described in section 2.6.7, were calculated by taking the mean number of GFP labelled punctae and conducting a student's t-test in Excel. Ten images were used per media type.

2.8 Bioinformatics databases searched for phylogenetic analysis

Arabidopsis thaliana IRQ sequences were found by screening TAIR and Genbank sequence databases using the BLASTP and TBLASTN algorithms (Altschul et al., 1997). To identify NET homologs in other species, BLASTP and TBLASTN algorithms were used to screen: non-redundant and EST databases of Genbank; Norway spruce genome (http://congenie. org/); *Selaginella* genome (http://genome.jgi-psf.org/) *Amborella* genome http://www.amborella.org/; Magnolid EST databases, ancestral Angiosperm genome project (http://ancangio.uga. Edu). Further BLAST analysis was conducted using the resources available at phytozome (http://www.phytozome.net/). Sequences were validated by reciprocal TBLASTN screening of the TAIR database.

Chapter 3 Bioinformatics of the IRQ proteins

3.1 Introduction

As discussed in section 1.5 work on the IRQ proteins was a direct result of work on the NET superfamily of actin binding proteins. Evolutionary analysis revealed that the NETs first appeared in the spikemoss *Selaginella moellendorffii* and subsequently duplicated and diverged with increases in plant complexity during evolution (Hawkins *et al.*, 2014). From this point forward collective names for the NET and IRQ proteins will be as follows: proteins containing only the NAB domain will be termed NET proteins (NET1A-D, NET2A-D, NET3B), all proteins containing both a NAB domain and IRQ domain will be named NAB/IRQs (NET3A and NET3C, and NET4A-B), and those proteins containing only an IRQ domain but no NAB domain will be named IRQ proteins (IRQ1-6).

As the first NET protein thought to exist was a member of the NET4 subfamily, and as NET4 proteins contain the IRQ domain, did the NAB and IRQ domains in the ancestral NAB/IRQ protein appear simultaneously? Previous phylogenetic analysis focussed on the NAB domain alone and not the IRQ domain (Hawkins *et al.*, 2014). A protein containing an IRQ domain alone may have appeared with the NAB domain added later to create the NET4 protein. The NAB domain could have then been subsequently lost to form the IRQ proteins. If there are no proteins before the NET4 containing an IRQ domain alone, where in the evolution of the NAB/IRQ proteins did the proteins containing only an IRQ domain separate from the NAB/IRQ proteins?

To answer these questions a range of bioinformatics tools were used. Direct comparisons between IRQ proteins in *A. thaliana* were used to determine residues of importance within the IRQ domain. To provide background information relevant to this chapter and subsequent chapters, several online bioinformatics tools were used to explore possible localisation of the transcription and translation of the IRQs, and to explore their domain architecture. Sequence databases were interrogated to identify the presence of IRQ domain containing proteins in the genomes of different species, with the species chosen representing key changes in plants through evolutionary time. Finally, to better understand the expansion of the IRQ group of proteins the distribution of the IRQ protein encoding genes within the *Arabidopsis* genome was analysed.

3.2 Previous information on the IRQ proteins

All the IRQ proteins apart from IRQ5 are uncharacterised. IRQ5 (At5g41790) is characterised as COP-1 Interacting Protein 1 (CIP1; Matsui and Deng, 1995). Throughout this thesis CIP1 will be

referred to as IRQ5 to aid clarity when discussing the relationship between the IRQ proteins. COP1 (Constitutive Photomorphogenic 1; At2g32950) represses photomorphogenic growth and induces skotomorphogenesis (rapid and exaggerated elongation of the hypocotyl at the expense of root and cotyledon development) in the dark. It is localised to the nucleus in the dark and the cytoplasm in the light (Yi & Deng, 2005). Mutations in COP1 lead to dark grown plants having the same morphology as wild type plants grown in the light (Deng & Quail, 1991). COP1 is an E3 ubiquitin ligase and when in the dark targets light signal transduction regulators such as HY5, LAF1, and phyA and phyB for degradation. In the light, COP1 is exported to the cytosol allowing photomorphogenic responses to occur (Lau & Deng, 2012). COP1 has also recently been found to regulate hypocotyl thermomorphogenesis by integrating temperature and circadian rhythm signals to degrade the repressor protein HY5 allowing hypocotyl elongation at the most optimum time (Park *et al.*, 2017).

IRQ5 was identified from a screen of *Arabidopsis* cDNAs encoding proteins which interact with COP1 (Matsui & Deng, 1995). Immunofluorescence labelling of *Arabidopsis* protoplasts showed that IRQ5 associates with a cytoskeletal structure in the aerial tissues of the plant (hypocotyl and cotyledon cells) but not root cells. The interaction of IRQ5 with COP1 is thought to be mediated by a coiled-coil domain of approximately 100 amino acids in COP1. The cytoskeletal association of IRQ5 was thought to enable sequestration of COP1 under light conditions, preventing it from inhibiting the photomorphogenic response (Matsui & Deng, 1995). Recently it has been shown that IRQ5 acts as a regulator of the ABA response and is sensitive to osmotic stress (Ren *et al.*, 2016).

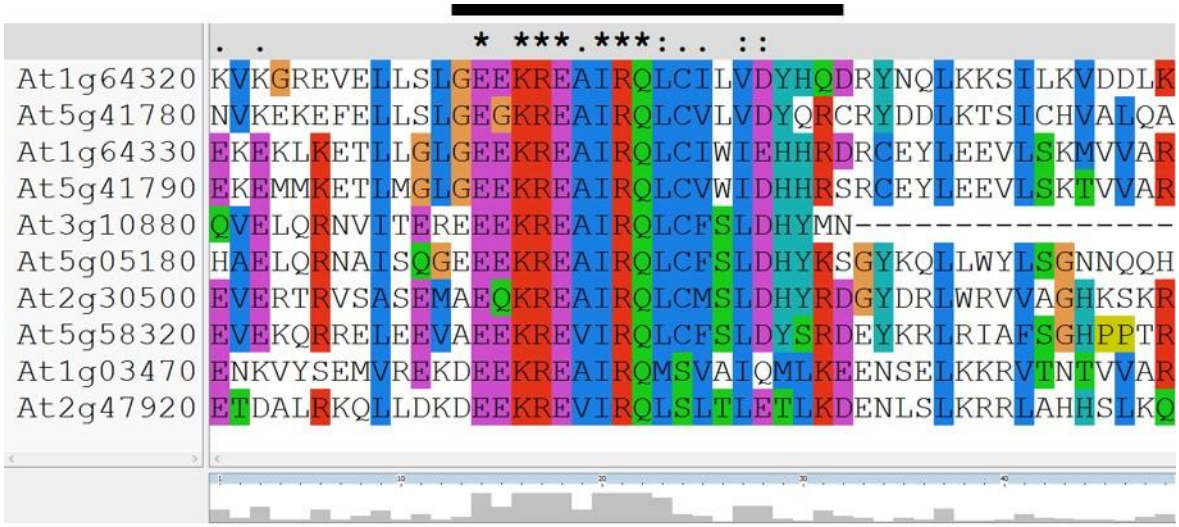
Much less is known about the remaining IRQ proteins as no studies have focussed exclusively on any of the remaining members. Transcriptome analysis has shown that the expression of IRQ3 (At5g41780) is downregulated more than 1.6 times during pollen tube growth (Wang *et al.*, 2008), indicating that IRQ3 may have a negative role in pollen tube growth. IRQ2 (At1g64320), IRQ3, IRQ4 (At1g64330), and IRQ6 (At5g05180) are annotated on the TAIR database as being myosin heavy chain related. Although IRQ2 and IRQ4 are annotated as myosin heavy chain related a study which aimed to find coiled-coil cytoskeletal homologues of metazoan proteins in *Arabidopsis* showed that IRQ2 and IRQ4 are not true motor domain homologues (Gardiner *et al.*, 2011).

The mRNA of IRQ4 and IRQ5 has been shown to be cell-cell mobile (Thieme *et al.*, 2015). IRQ4 and IRQ5 have been linked to the vacuole in a study which characterised the vegetative vacuolar proteome in *Arabidopsis* (Carter *et al.*, 2004).

No information is available about IRQ1 (At3g10880) while Day *et al.*, (2002) showed that IRQ6 (At5g05180) contains a single calcium binding EF-hand domain (Kretsinger & Nockolds, 1973). Ca^{2+} is an important secondary messenger in plant cells involved in responses to various abiotic and biotic stimuli (Bush, 1995; Zhu, 2002; Wahid *et al.*, 2007; Howe & Jander, 2008) so IRQ6 may have a role in calcium signalling. However, the EF hand annotation for IRQ6 was identified by only one database in InterProScan, so caution must be used before following up on this result. IRQ6 was found to show a 2.22 fold change in transcript levels in plants expressing a *35S:miR396* construct (Rodriguez *et al.*, 2015). miR396 is a microRNA that interacts with the Growth Regulating Factor (GRF) class of transcription factors (Rodriguez *et al.*, 2010; Debernardi *et al.*, 2012, 2014). The miR396/GRF pathway was shown to regulate the transition of stem cells to transit amplifying cells in the root meristem (Rodriguez *et al.*, 2015). Transit amplifying cells are progeny of stem cells that undergo rapid cell divisions in meristematic regions allowing organ growth before cell differentiation (Galinha *et al.*, 2007; Matsuzaki *et al.*, 2010). IRQ6 may therefore have a role in cell division or differentiation.

3.3 Comparison of IRQ domains of IRQ domain containing proteins in *Arabidopsis thaliana*

Initial bioinformatics analysis (Mentlak, 2016) forms the basis for the more extensive work in the following sections and showed that the IRQ domain consists of a central core region flanked by areas of lower similarity. Figure 3.1 was produced by alignment of the IRQ domains of the NAB/IRQ and IRQ proteins in ClustalX 2.1 (Larkin *et al.*, 2007). A number of residues are very highly conserved across all the IRQ proteins in *A. thaliana*: (from left to right in the highly conserved region of the IRQ proteins in Figure 3.1) an ‘E’ at position 2, ‘KRE’ at positions 4, 5, and 6, and ‘IRQ’ at positions 8, 9, and 10. This indicates that these amino acids may be important for the biological function of the domain within these proteins.



Colour	Residue at position	(Threshold, residue group)
BLUE	A, I, L, M, F, W, V	(60% WLVIMAFCHP)
	C	(60% WLVIMAFCHP)
RED	R, K	(60% KR) (85% K, R, Q)
GREEN	N	(50% N) (85% N, Y)
	Q	(60% KR) (50% QE) (85% Q, K, E, R)
	S, T	(60% WLVIMAFCHP) (50% TS) (85% S, T)
PINK	C	(100% C)
MAGENTA	E	(60% KR) (50% QE) (85% E, Q, D)
	D	(60% KR), (85% K,R,Q) (50% ED)
ORANGE	G	(0% G)
CYAN	H, Y	(60% WLVIMAFCHP) (85% W,Y,A,C,P,Q,F,H,I,L,M,V)
YELLOW	P	(0% P)

Figure 3.1: Alignment of the IRQ proteins form *Arabidopsis thaliana*. The IRQ domains of the C-terminal regions of the IRQ proteins were aligned in ClustalX 2.1 using the default colour scheme (shown below). The black line above the sequences represents the highly conserved region of the IRQ domain. Stars represent very highly conserved residues, two dots on top of each other represent highly conserved residues and a single dot represents conserved residues. An amino acid is assigned a colour if it meets some minimum criterion for that residue type. The minimum percentage is shown with the amino acids that must meet or exceed this threshold value. If the residues are grouped together such as 'KR' then the colour is applied if any combination of these residues meets or exceeds the threshold percentage. If the residues are separated by commas then the colour is applied when one of these residues individually meets or exceeds the threshold value.

3.4 Cladistic analysis to show the relationship between the NAB/IRQ and IRQ proteins in *Arabidopsis*

As a starting point in understanding the relationship between the NAB/IRQ and IRQ proteins, cladistic analysis was carried out using the IRQ domain alone and the full length proteins. This aimed to test whether the IRQ domains alone separate the NAB/IRQs and IRQ proteins into distinct groups like the NAB domains in the NET proteins (Calcutt, 2009; Deeks *et al.*, 2012b). The full length proteins were analysed to see if any sequence outside the IRQ domain has a greater

influence on the groupings of the proteins within the tree. All the phylogenetic analysis that follows used the same approach of first comparing the IRQ domains alone then comparing the full length proteins.

The IRQ domains of the NAB/IRQ and IRQ proteins shown in Figure 3.1 were aligned in ClustalX 2.1 (Larkin *et al.*, 2007) and a cladogram (Figure 3.2) was generated using a bootstrap neighbour joining algorithm using 1000 iterations (Saitou & Nei, 1987). The full length NAB/IRQ and IRQ proteins are shown in Figure 3.3. Using the IRQ domains alone produced a dendrogram with an altered topology than that made using the full length proteins. In the IRQ domain alone dendrogram there are four groupings: NET3A/C, IRQ4/5, IRQ2/3, and NET4A/B split by a subgroup consisting of IRQ1/6. In contrast, using the full length proteins shows the same groupings for NET3A/C, IRQ4/5, and IRQ2/3 but this time NET4A/B are no longer split by IRQ1/6 and form their own group, making a total of 5 groupings. Overall, this indicates there is a higher similarity between the IRQ domains of IRQ1/6 to NET4A/B than to the NET3A/C or IRQ2/3/4/5 proteins. When the full length proteins are used this similarity is weakened, probably due to the highly similar NAB domains found in the NET3A/C and NET4A/B pairings.

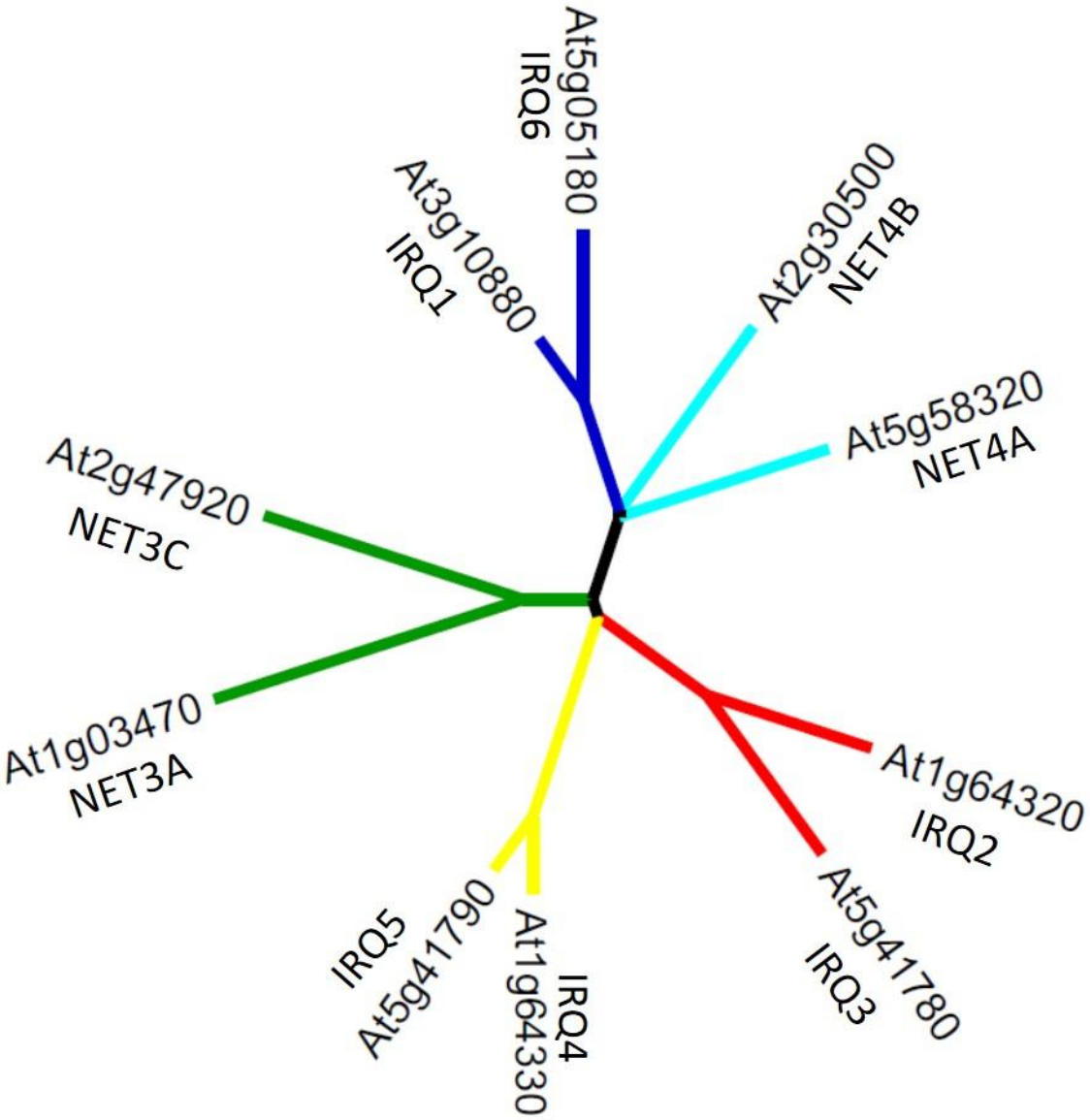


Figure 3.2: Cladogram of the NAB/IRQ and IRQ proteins using the IRQ domain alone. The IRQ domains of the NAB/IRQ and IRQ proteins were aligned using ClustalX 2.1 and a bootstrap neighbour-joining algorithm was used to generate the tree. The tree was drawn using the Hypertree programme. There is a clear division between the NET3 subgroup, the IRQ5 and IRQ4 subgroup, and the IRQ2 and IRQ3 subgroup. The NET4A and NET4B subgroup is split by the IRQ1 and IRQ6 subgroup.

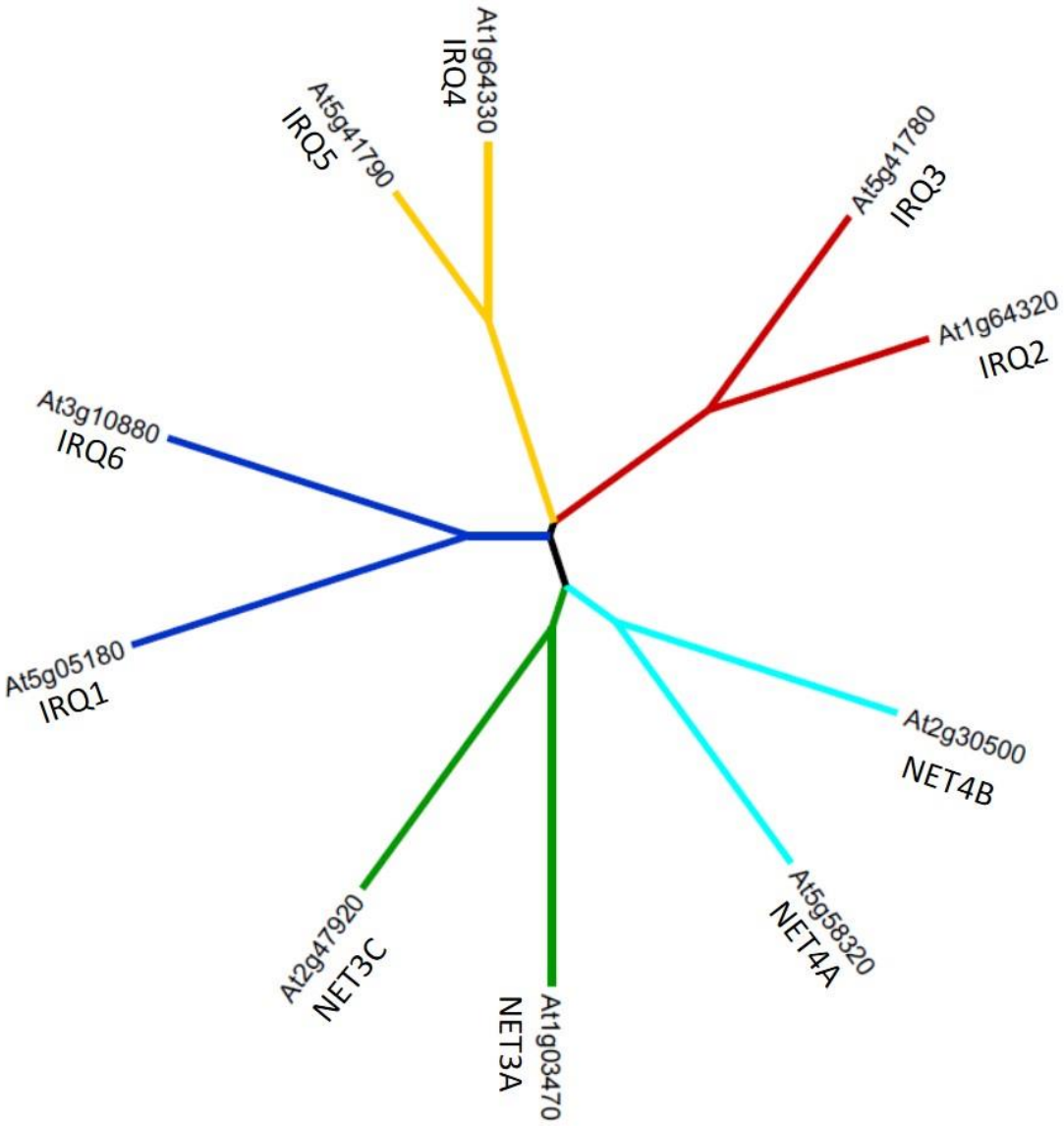


Figure 3.3: Cladogram of the NAB/IRQ and IRQ proteins using the full length sequences. The full length sequences of the NAB/IRQ and IRQ proteins were aligned using ClustalX 2.1 and a bootstrap neighbour-joining algorithm was used to generate the tree. The tree was drawn using the Hypertree programme. The subgroups are arranged slightly differently with the NET4A and NET4B subgroup no longer split by IRQ1 and IRQ6.

3.5 Predicted secondary structures of the IRQ domain and full length IRQ and NAB/IRQ proteins

Like the NET proteins the IRQ proteins consist of numerous coiled-coil domains but unlike these NET proteins do not contain the N-terminal NAB domain (Figure 3.4). IRQ4 and IRQ5, however, both contain an N-terminal divergent region with limited similarity to the NAB domain. This domain has been termed a quasi-NAB domain (qNAB). The IRQ domains for nearly all the NAB/IRQ and IRQ proteins lie at the C-terminus of the protein except for NET3A and NET3C,

where the IRQ domain is found a greater distance away from the C-terminus, positioning which could affect possible interactions involving this domain.

There is a marked difference in size between the smallest IRQ protein (IRQ1) at 37 kDa and the largest IRQ protein (IRQ5) at 182 kDa (Table 3.1). The other IRQ proteins fall in the range of 50-65 kDa. As well as being the shortest IRQ, IRQ1 also has a truncated IRQ domain, terminating at the end of the highly conserved core region of the IRQ domain. It has previously been demonstrated that in NET4B the central region of the IRQ domain alone was insufficient to allow Rab binding (Mentlak, 2016) suggesting that IRQ1 may have a reduced potential to interact with Rab proteins when compared to other family members.

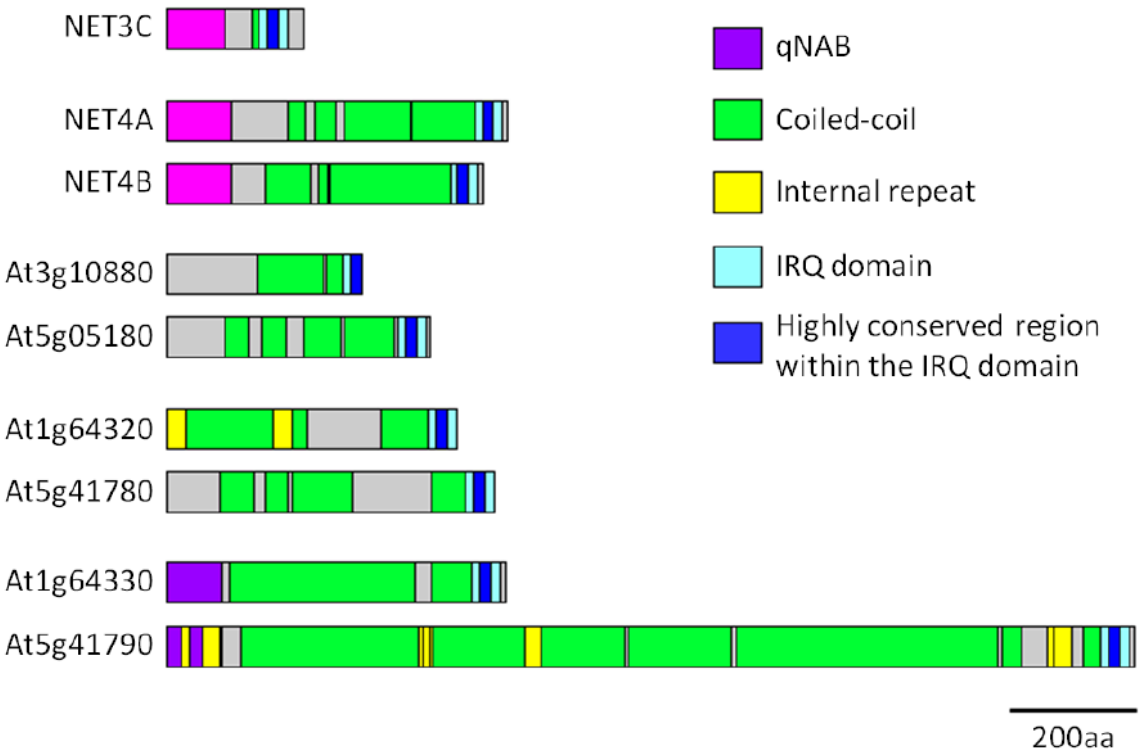


Figure 3.4: Predicted secondary structure of the NAB/IRQ and IRQ proteins. Schematic showing the secondary structural features of the NAB/IRQ and IRQ proteins identified using the SMART and Paircoil2 programmes and protein sequence alignments. Diagram taken from Mentlak, (2015).

Table 3.1: Amino acid lengths and predicted molecular weights of the NAB/IRQ and IRQ proteins

Protein	Amino Acid Length	MW (kDa)
NET3A	269	31
NET3C	225	26
NET4A	558	64
NET4B	517	60
At3g10880 (IRQ1)	319	37
At5g05180 (IRQ6)	432	50
At1g64320 (IRQ2)	476	55
At5g41780 (IRQ3)	537	63
At1g64330 (IRQ4)	555	65
At5g41790 (IRQ5)	1586	182

For secondary structure prediction of the IRQ domains, the IRQ domains of all the NAB/IRQ and IRQ proteins were aligned in ClustalX 2.1 (Larkin *et al.*, 2007) and this alignment was exported to Jalview (Waterhouse *et al.*, (2009); Figure 3.5). The JPred 4 software within Jalview uses alignments and algorithms (in this case JNet4) along with position specific scoring matrices (PSSM) and Hidden Markov Models (HMM) to predict secondary structures such as coiled-coils, α-helices, and β-sheets (Cole *et al.*, 2008; Drozdetskiy *et al.*, 2015). The software predicts a single alpha helix of 28 amino acids for the IRQ domain.

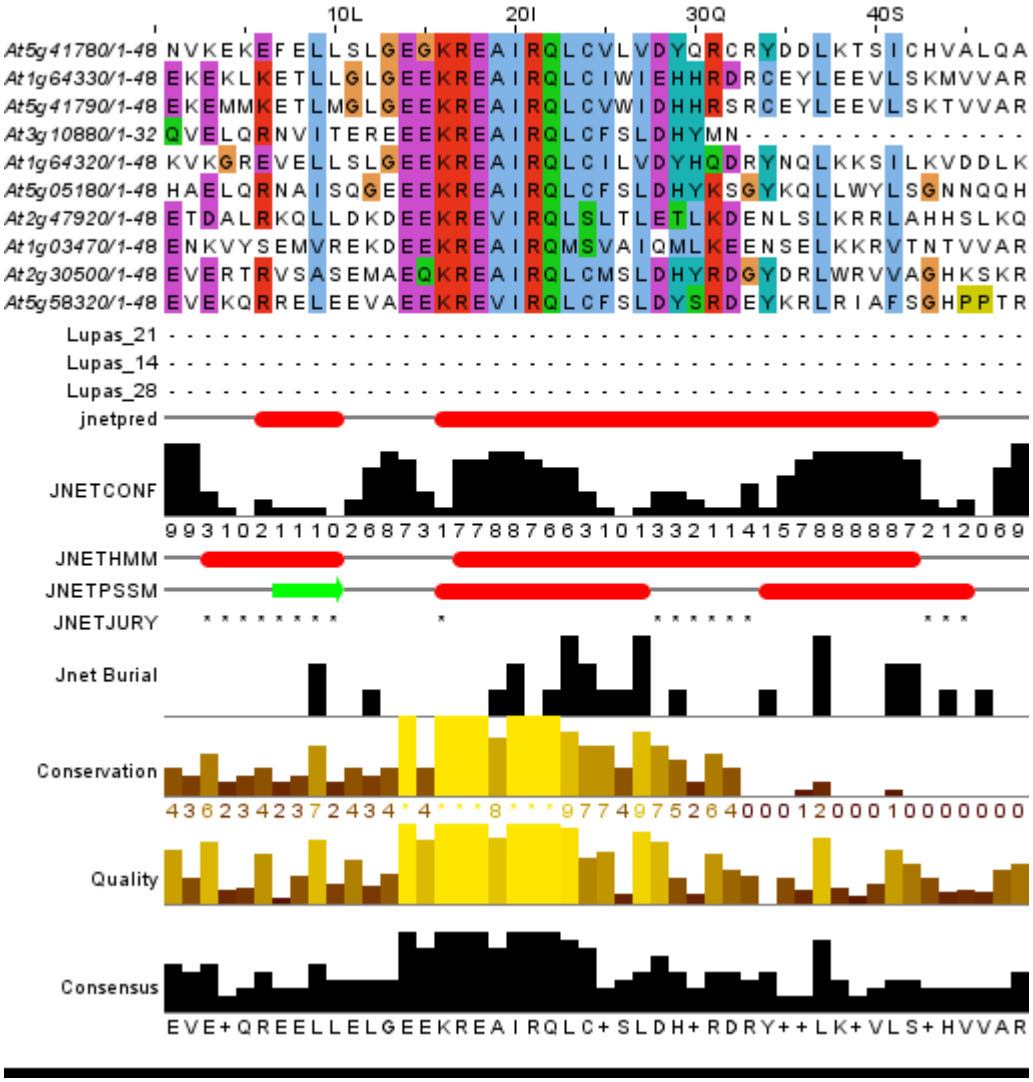


Figure 3.5: Predicted secondary structure of the IRQ domain. The JPred programme on Jalview was used to generate a prediction for the secondary structure of the IRQ domain. The Jpred programme uses a variety of secondary structure prediction algorithms and multiple sequence alignments. The 'jnetpred' line represents the final prediction showing a single alpha helix. 'JNETCONF' represents the confidence the programme assigns to the prediction, with 10 being the highest and 0 being the lowest. 'JNETHMM' and 'JNETPSSM' represent the Hidden Markov Models and position specific scoring matrices algorithms' respectively. 'JNETJURY' shows differences between the two algorithms regarding their predictions for that position.

3.6 Gene expression patterns and subcellular localisation prediction of the IRQ proteins

Understanding the expression pattern of genes of interest and their potential subcellular localisation are important steps in generating hypotheses regarding gene and protein regulation and function. An analysis of the expression patterns of the IRQ proteins was carried out using publically available information found on the *Arabidopsis* electronic Fluorescent Pictograph (eFP) browser. The eFP browser allows the visualisation of gene expression data from the Affymetrix

ATH1 GeneChip containing 22,000 *Arabidopsis* genes (Winter *et al.*, 2007). Using the absolute expression values setting the maps give a gradient of colours ranging from white (no expression) to red (highly expressed). The developmental and subcellular localisation maps of *Arabidopsis* were primarily used for the IRQ proteins with additional information from tissue specific maps of interest. The spatial map of gene expression was generated using data from the AtGenExpress project (Schmid *et al.*, 2005).

For the subcellular analysis, the SUBA3 database was used. SUBA3 combines evidence from fluorescent protein visualisation and interaction data, and manually created subcellular proteomics data sets (Tanz *et al.*, 2013). SUBA3 combines subcellular prediction algorithms from 11 different programmes: SubLoc (Hua & Sun, 2001), Wolf PSORT (Horton *et al.*, 2007), MultiLoc2 (Blum *et al.*, 2009), AdaBoost (Niu *et al.*, 2008), BaCelLo (Pierleoni *et al.*, 2006), Nucleo (Hawkins *et al.*, 2007), Plant-mPloc (Chou & Shen, 2010), PProwler (Hawkins & Boden, 2006), SLPFA (Tamura & Akutsu, 2007), SLP-Local (Matsuda *et al.*, 2005) and YLoc (Briesemeister & Kohlbacher, 2010). These predictions are combined with experimental data to give localisation to 12 different subcellular compartments: the endoplasmic reticulum, mitochondrion, plasma membrane, cytosol, Golgi, peroxisome, vacuole, plastid, extracellular space, cytoskeleton, cell plate, and nucleus. This information can be visualised with cell eFP browser. A confidence gradient for the cell eFP browser is also provided and uses the same colour scheme as the tissue and developmental maps described above (Winter *et al.*, 2007). A summary of the subcellular prediction data is provided in Table 3.2.

With the above programs and algorithms, it should be noted that they are speculative and can only predict subcellular localisation. Results from such analyses must be verified using methods such as transcriptional and translational fusions to proteins of interest e.g. promoter GUS fusions (chapter 4) or fusions to GFP driven by endogenous promoters (chapter 5).

IRQ1 shows the highest level of expression in the shoot apical meristem, with especially high expression during the transition from a vegetative state to flowering (Figure 3.6a). Expression levels are slightly elevated in early stage flowers in the bud stage. Increased expression is also seen in the hypocotyls of young seedlings. Expression is relatively low in other tissues and although there is little elevation of expression seen in any particular subcellular compartment, the SUBA3 database suggests that IRQ1 localises to the cytosol, mitochondrion and nucleus (Table 3.2).

Similar to IRQ1, IRQ6 shows its highest level of expression in the shoot apical meristem during the transition from vegetative to flowering states, the hypocotyl of young seedlings, and shows a slightly raised level of expression in young flower buds (Figure 3.6b). IRQ6 also shows increased

expression in the roots of young seedlings and in the early stage siliques and in the embryo corresponding to the heart and torpedo stages. The localisation of expression at the meristematic regions agrees with the published data described earlier wherein IRQ6 was shown to have a 2.22 fold increase in expression in plants expressing the *35:miR396* construct associated with cell division and differentiation in the meristematic region of the root (Rodriguez *et al.*, 2015). There is also an increase in expression in the stamens of mature flowers and the petiole of fully developed leaves. Expression is relatively low in other tissues and although there is little elevation of expression seen in any particular subcellular compartment the SUBA3 database suggests that IRQ6 localises to the cytosol, mitochondrion, plastids, and nucleus (Table 3.2).

For IRQ2 and IRQ3, expression is low in all tissues except pollen indicating these proteins may be pollen specific (Figure 3.6c and d). In both IRQ2 and IRQ3 expression is strongest in the mature pollen grain but is also elevated in the tricellular pollen grain (Figure 3.6ciii and diii). Both genes show elevated expression during pollen germination, with IRQ3 showing a higher level of absolute expression compared to IRQ2 (compare Figure 3.6civ with Figure 3.6div). Neither gene is strongly expressed in the stigma or ovaries. There is little elevation of expression seen in any particular subcellular compartment for IRQ2 or IRQ3 (Figure 3.6cii and dii) but the SUBA3 database suggests that IRQ2 and IRQ3 localise to the cytosol, mitochondrion, nucleus and the cytoskeleton (Table 3.2).

IRQ4 is expressed in most tissues, showing the highest level of expression in late stage embryos corresponding to the development of cotyledons and the development of the dried seed (Figure 3.6ei). However, the level of expression appears to drop dramatically in the cotyledons of young seedlings and in the imbibed seed. Expression is also elevated in the petals and stamen of open flowers. Increased expression is seen in the hypocotyls and roots of young seedlings and in plants with a vegetative rosette. Looking more closely at the root shows expression is slightly stronger in the late meristematic and early elongation zones (Figure 3.6eiii). Subcellular predictions suggest that IRQ4 localises to the cytosol, mitochondrion, nucleus, and cytoskeleton (Table 3.2).

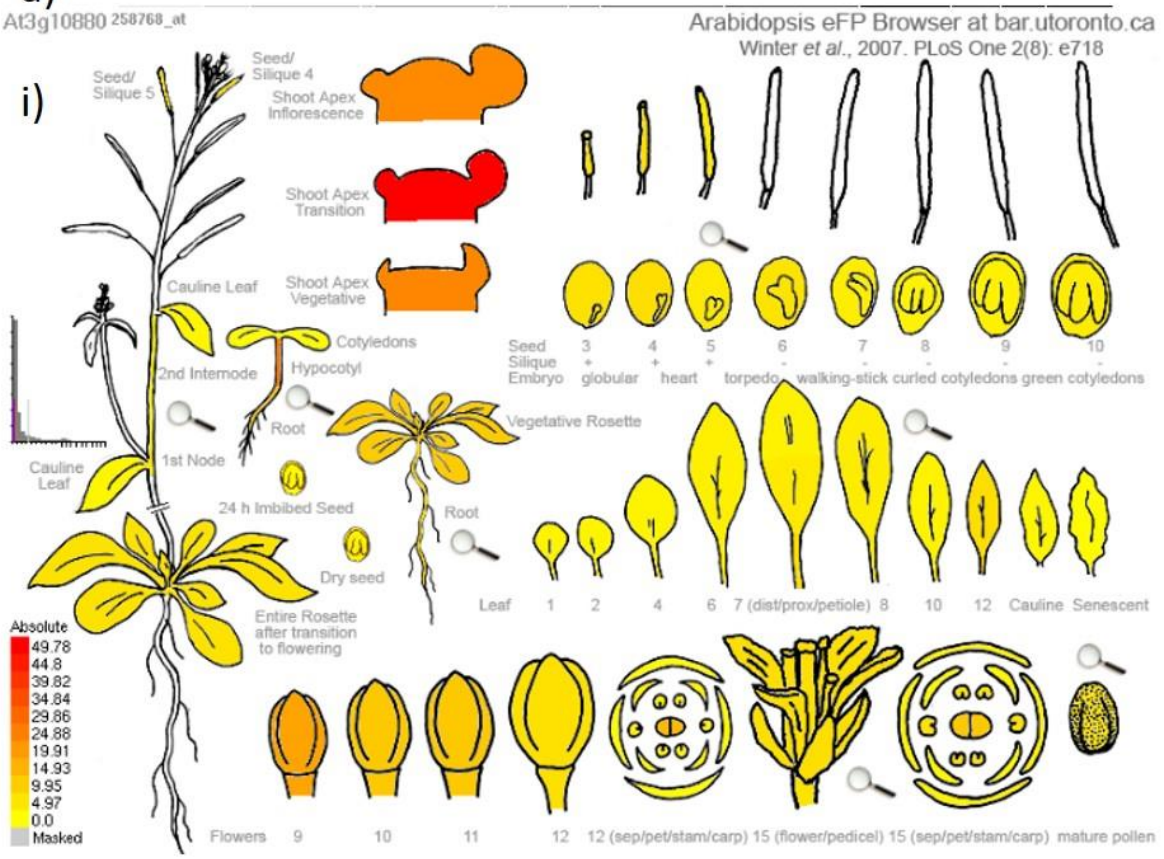
Similar to IRQ4, IRQ5 is expressed relatively ubiquitously with the strongest expression seen in leaf tissue, especially cauline and senescent leaves (Figure 3.6fi). It is also expressed in the sepals of open flowers and the hypocotyls of young seedlings. In contrast to IRQ4, IRQ5 is only weakly expressed in the embryos but more strongly in the cotyledons of young seedlings. IRQ5 shows expression in a similar area of the root to IRQ4, the late meristematic and early elongation zones, but is more restricted to particular cell files (Figure 3.6fiii).

Subcellular localisation prediction suggest that IRQ5 localises to the cytosol, mitochondrion, and nucleus (Table 3.2). In addition experimental evidence from proteomic screens suggests that IRQ5

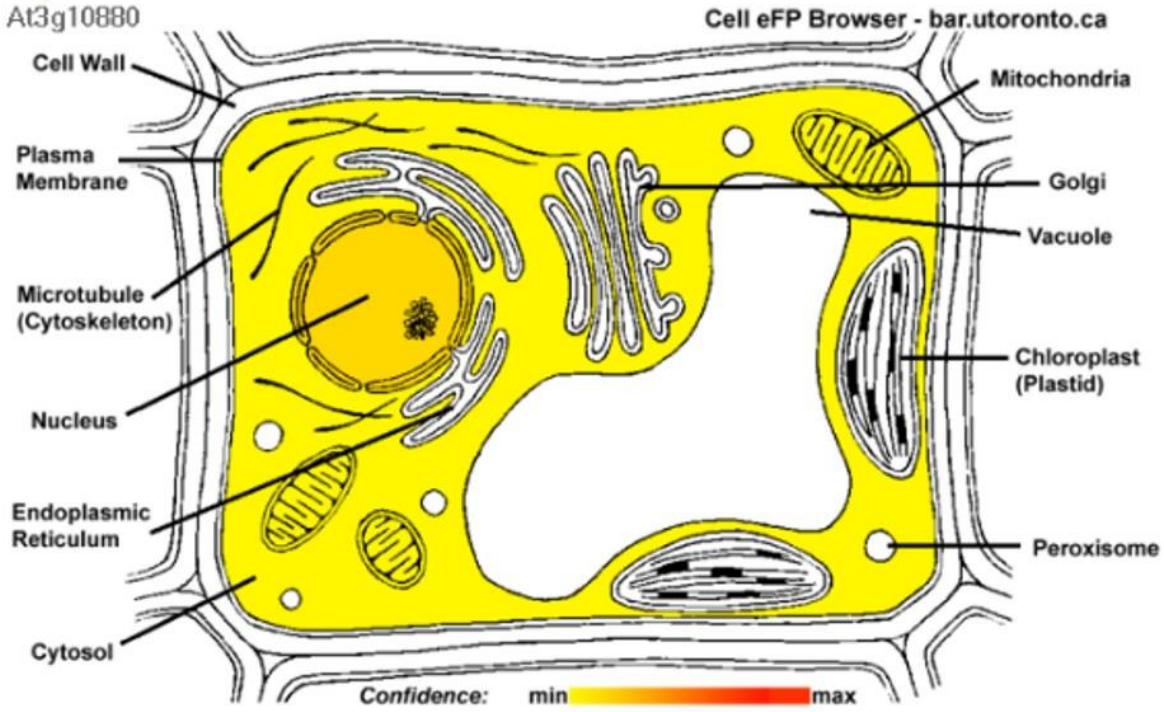
and IRQ4 are localised to the vacuolar and plasma membranes (Carter *et al.*, 2004; Jaquinod *et al.*, 2007; Yoshida *et al.*, 2013). IRQ4 was also detected at the vacuole in a further screen (Nikolovski *et al.*, 2012). The similarity between IRQ4 and IRQ5 may have led to them being indistinguishable in two for the screens (Carter *et al.*, 2004; Jaquinod *et al.*, 2007) as only a few peptides were recovered for each protein or the same score was given to each. But as both have been recovered in multiple screens it suggests that this does represent a true localisation for these two proteins. A further screen suggests that IRQ4 and IRQ5 can be localised to the plasma membrane (Elmore *et al.*, 2012) and that IRQ5 is localised to the chloroplast via a screen looking for chloroplast associated proteins (Kleffmann *et al.*, 2004; Kong *et al.*, 2011). As only one peptide was recorded once from two repeats in the Kong *et al.* study this decreases confidence in this localisation.

Fluorescent protein screens using translational fusions to establish localisation for the IRQs have not been carried out for IRQ1-4, and 6 up to this point but immunofluorescence was used to show that IRQ5 localises to a cytoskeletal, filamentous network in hypocotyl and cotyledon protoplasts (Matsui & Deng, 1995). A transcriptional fusion using promoter-GUS showed that IRQ5 is mainly expressed in the photosynthetic tissues of the plant with only low transcription evident in the root (Ren *et al.*, 2016). In the same study, Ren *et al.* (2016) used a translational fusion to YFP to show that IRQ5 localises to the plasma membrane in hypocotyl cells, root epidermal cells, and root hairs. Immunoblots also showed IRQ5 protein is found within the roots (Ren *et al.*, 2016). The subcellular localisation of the IRQ proteins using transcriptional and translational fusions is discussed in chapters four and five of this thesis.

a)

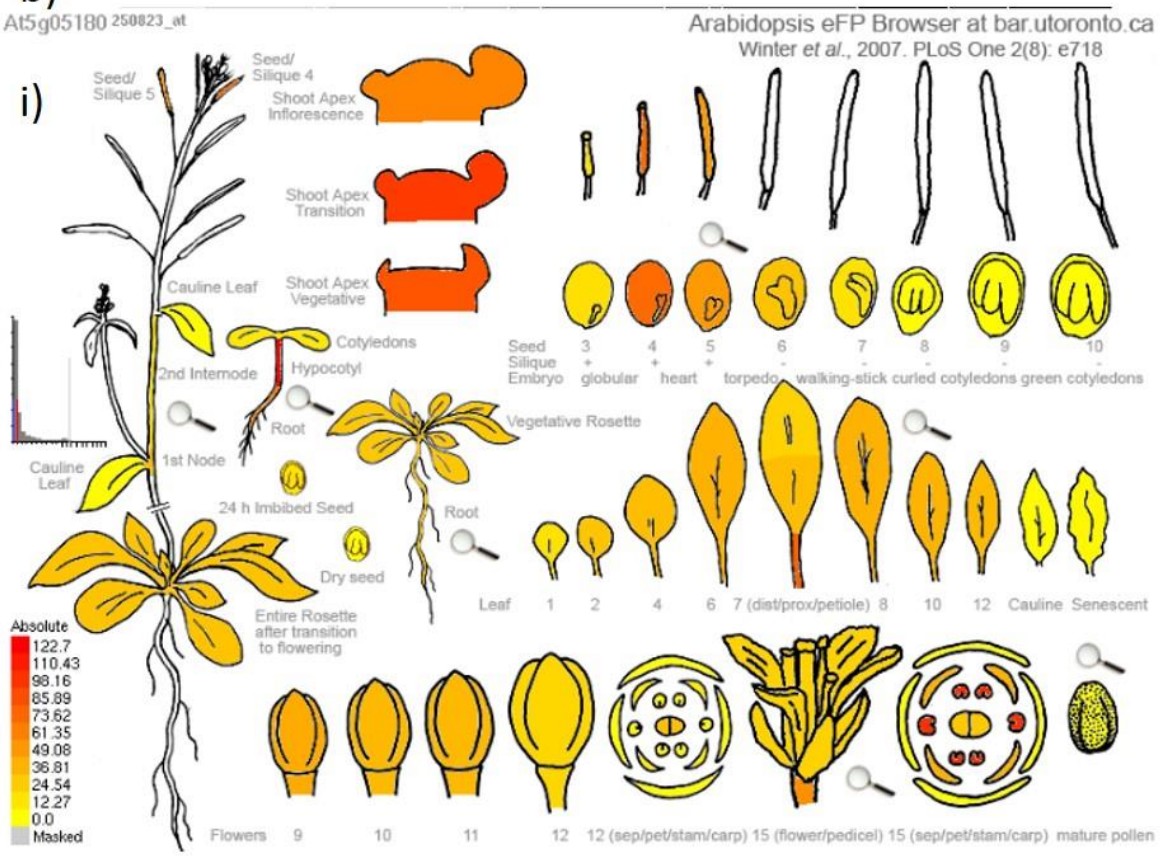

At3g10880 258768_at
Arabidopsis eFP Browser at bar.utoronto.ca
Winter et al., 2007. PLoS One 2(8): e718
i)
Seed/ Silique 5
Seed/ Silique 4
Shoot Apex Inflorescence
Shoot Apex Transition
Shoot Apex Vegetative
Cauline Leaf
Cotyledons
2nd Internode
Hypocotyl
Root
1st Node
Cauline Leaf
24 h Imbibed Seed
Dry seed
Entire Rosette after transition to flowering
Vegetative Rosette
Root
Absolute
49.78
44.8
39.82
34.84
29.86
24.88
19.91
14.93
9.95
4.97
0.0
Masked
Seed 3 4 5 6 7 8 9 10
Silique + + + - - - - -
Embryo globular heart torpedo walking-stick curled cotyledons green cotyledons
Leaf 1 2 4 6 7 (dist/prox/petiole) 8 10 12 Cauline Senescent
Flowers 9 10 11 12 12 (sep/pet/stam/carp) 15 (flower/pedicel) 15 (sep/pet/stam/carp) mature pollen

ii)

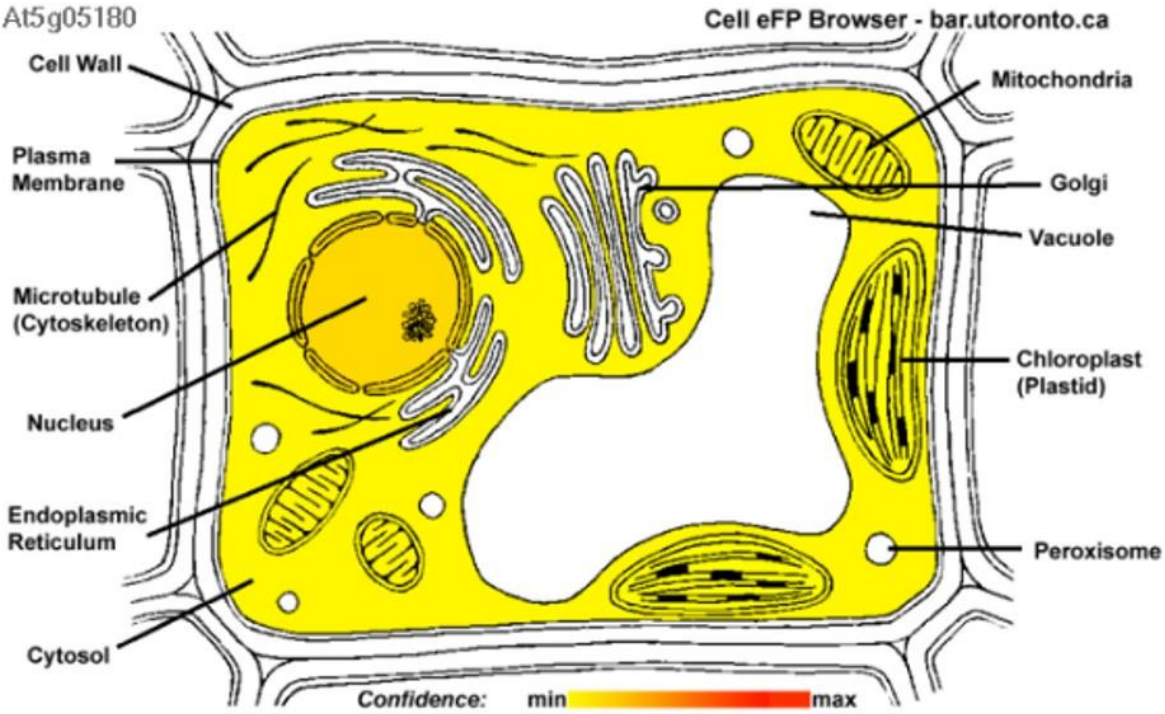

At3g10880
Cell eFP Browser - bar.utoronto.ca
Cell Wall
Plasma Membrane
Microtubule (Cytoskeleton)
Nucleus
Endoplasmic Reticulum
Cytosol
Mitochondria
Golgi
Vacuole
Chloroplast (Plastid)
Peroxisome
Confidence: min max

b)

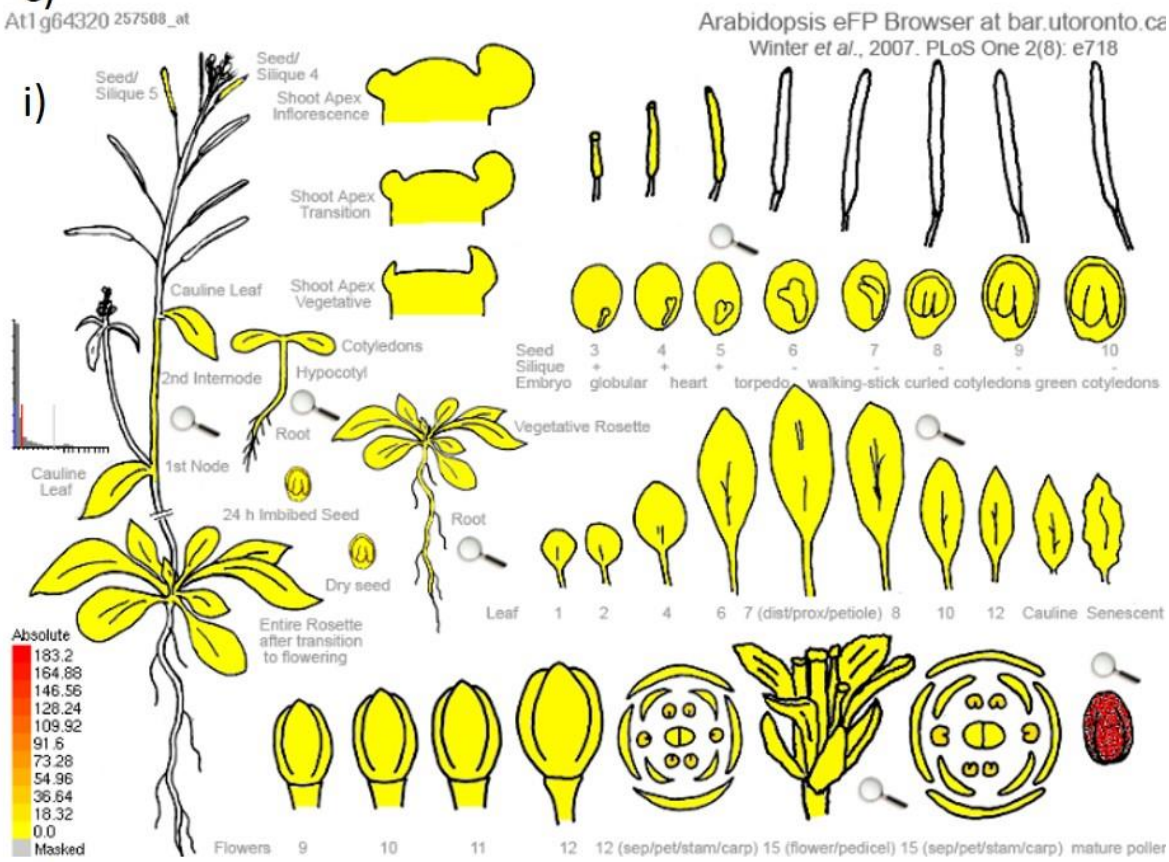
i)


At5g05180 250823_at
Arabidopsis eFP Browser at bar.utoronto.ca
Winter et al., 2007. PLoS One 2(8): e718
Seed/ Silique 5
Seed/ Silique 4
Shoot Apex Inflorescence
Shoot Apex Transition
Shoot Apex Vegetative
Cauline Leaf
Cotyledons
Hypocotyl
2nd Internode
Root
1st Node
Cauline Leaf
24 h Imbibed Seed
Dry seed
Entire Rosette after transition to flowering
Vegetative Rosette
Root
Absolute
122.7
110.43
98.16
85.89
73.62
61.35
49.08
36.81
24.54
12.27
0.0
Masked
Seed 3 4 5 6 7 8 9 10
Silique + + + - - - - -
Embryo globular heart torpedo walking-stick curled cotyledons green cotyledons
Leaf 1 2 4 6 7 (dist/prox/petiole) 8 10 12 Cauline Senescent
Flowers 9 10 11 12 12 (sep/pet/stam/carp) 15 (flower/pedicel) 15 (sep/pet/stam/carp) mature pollen

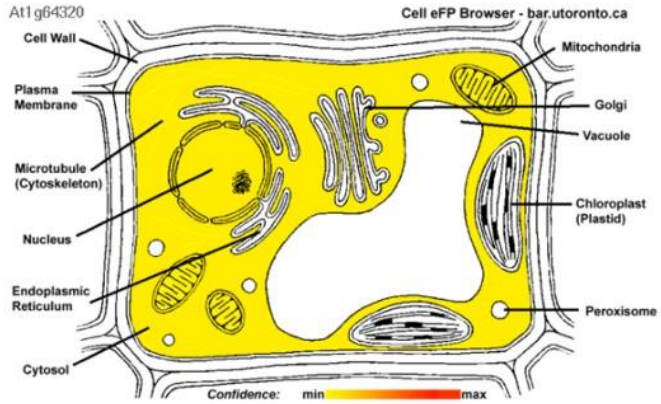
ii)


At5g05180
Cell eFP Browser - bar.utoronto.ca
Cell Wall
Plasma Membrane
Microtubule (Cytoskeleton)
Nucleus
Endoplasmic Reticulum
Cytosol
Mitochondria
Golgi
Vacuole
Chloroplast (Plastid)
Peroxisome
Confidence: min max

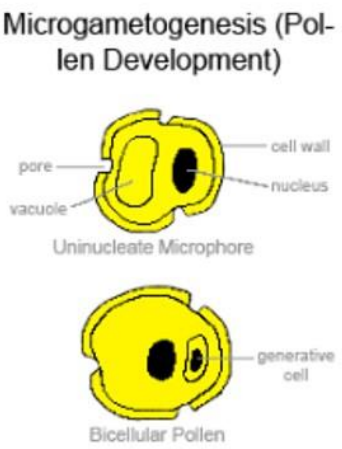
c)


At1g64320 257508_at
Arabidopsis eFP Browser at bar.utoronto.ca
Winter et al., 2007. PLoS One 2(8): e718
i)
Seed/ Silique 5
Seed/ Silique 4
Shoot Apex Inflorescence
Shoot Apex Transition
Shoot Apex Vegetative
Cauline Leaf
Cotyledons
Hypocotyl
2nd Internode
Root
1st Node
Cauline Leaf
24 h Imbibed Seed
Dry seed
Entire Rosette after transition to flowering
Vegetative Rosette
Root
Absolute
183.2
164.88
146.56
128.24
109.92
91.6
73.28
54.96
36.64
18.32
0.0
Masked
Seed 3 4 5 6 7 8 9 10
Silique + + + - - - - -
Embryo globular heart torpedo walking-stick curled cotyledons green cotyledons
Leaf 1 2 4 6 7 (dist/prox/petiole) 8 10 12 Cauline Senescent
Flowers 9 10 11 12 12 (sep/pet/stam/carp) 15 (flower/pedicel) 15 (sep/pet/stam/carp) mature pollen

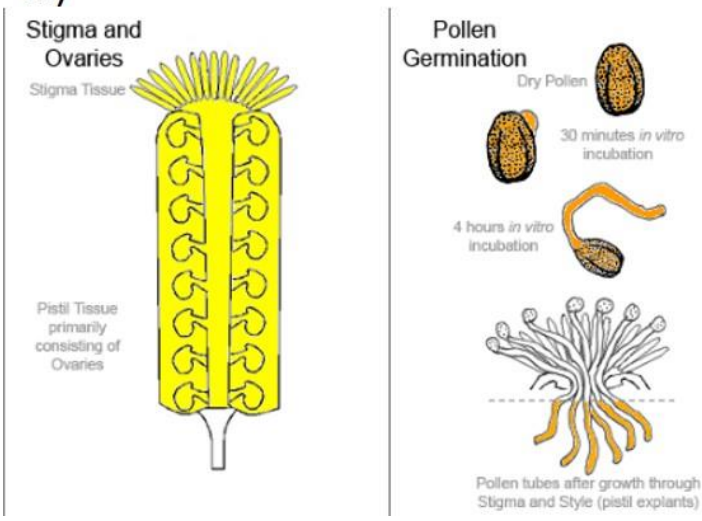
ii)


At1g64320
Cell eFP Browser - bar.utoronto.ca
Cell Wall
Plasma Membrane
Microtubule (Cytoskeleton)
Nucleus
Endoplasmic Reticulum
Cytosol
Mitochondria
Golgi
Vacuole
Chloroplast (Plastid)
Peroxisome
Confidence: min max

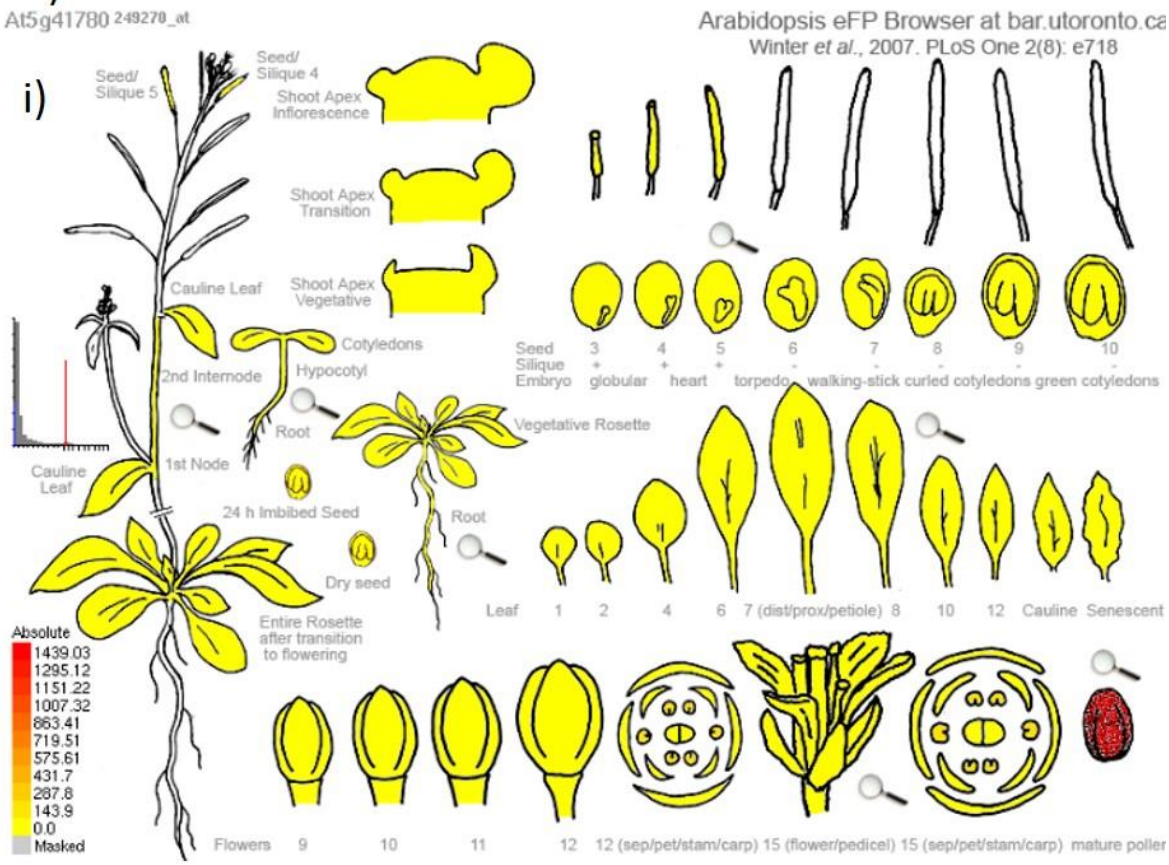
iii)


Microgametogenesis (Pollen Development)
pore
cell wall
nucleus
vacuole
Uninucleate Microphore
generative cell
Bicellular Pollen
Tricellular Pollen
Mature Pollen Grain

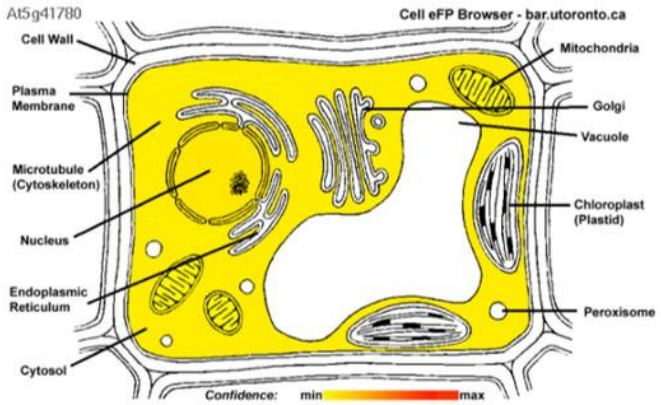
iv)


Stigma and Ovaries
Stigma Tissue
Pistil Tissue primarily consisting of Ovaries
Pollen Germination
Dry Pollen
30 minutes in vitro incubation
4 hours in vitro incubation
Pollen tubes after growth through Stigma and Style (pistil explants)

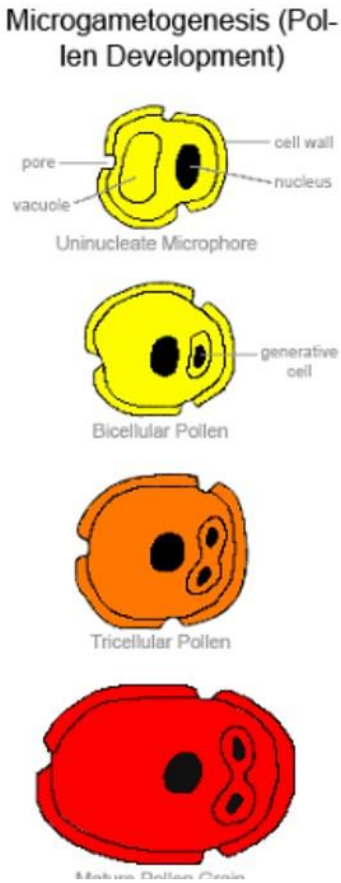
d)


At5g41780 249270_at
Arabidopsis eFP Browser at bar.utoronto.ca
Winter et al., 2007. PLoS One 2(8): e718
i)
Seed/ Silique 5
Seed/ Silique 4
Shoot Apex Inflorescence
Shoot Apex Transition
Shoot Apex Vegetative
Cauline Leaf
Cotyledons
Hypocotyl
2nd Internode
Root
1st Node
Cauline Leaf
24 h Imbibed Seed
Dry seed
Entire Rosette after transition to flowering
Vegetative Rosette
Root
Absolute
1439.03
1295.12
1151.22
1007.32
863.41
719.51
575.61
431.7
287.8
143.9
0.0
Masked
Seed 3 4 5 6 7 8 9 10
Silique + + + - - - - -
Embryo globular heart torpedo walking-stick curled cotyledons green cotyledons
Leaf 1 2 4 6 7 (dist/prox/petiole) 8 10 12 Cauline Senescent
Flowers 9 10 11 12 12 (sep/pet/stam/carp) 15 (flower/pedicel) 15 (sep/pet/stam/carp) mature poller

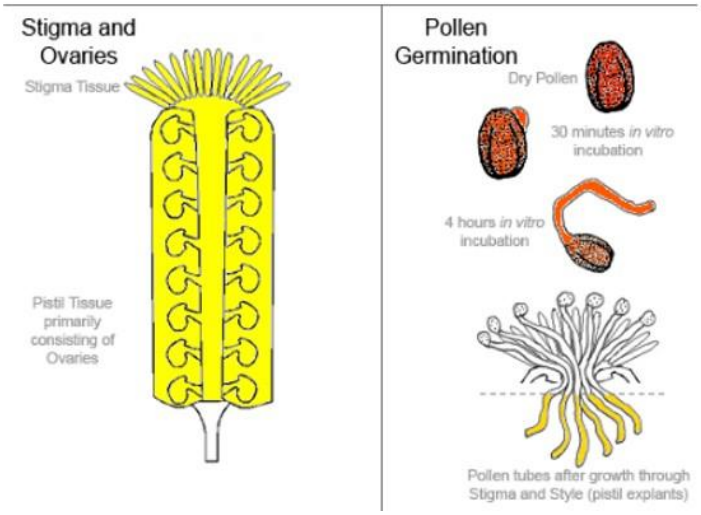
ii)


At5g41780
Cell eFP Browser - bar.utoronto.ca
Cell Wall
Plasma Membrane
Microtubule (Cytoskeleton)
Nucleus
Endoplasmic Reticulum
Cytosol
Mitochondria
Golgi
Vacuole
Chloroplast (Plastid)
Peroxisome
Confidence: min max

iii)

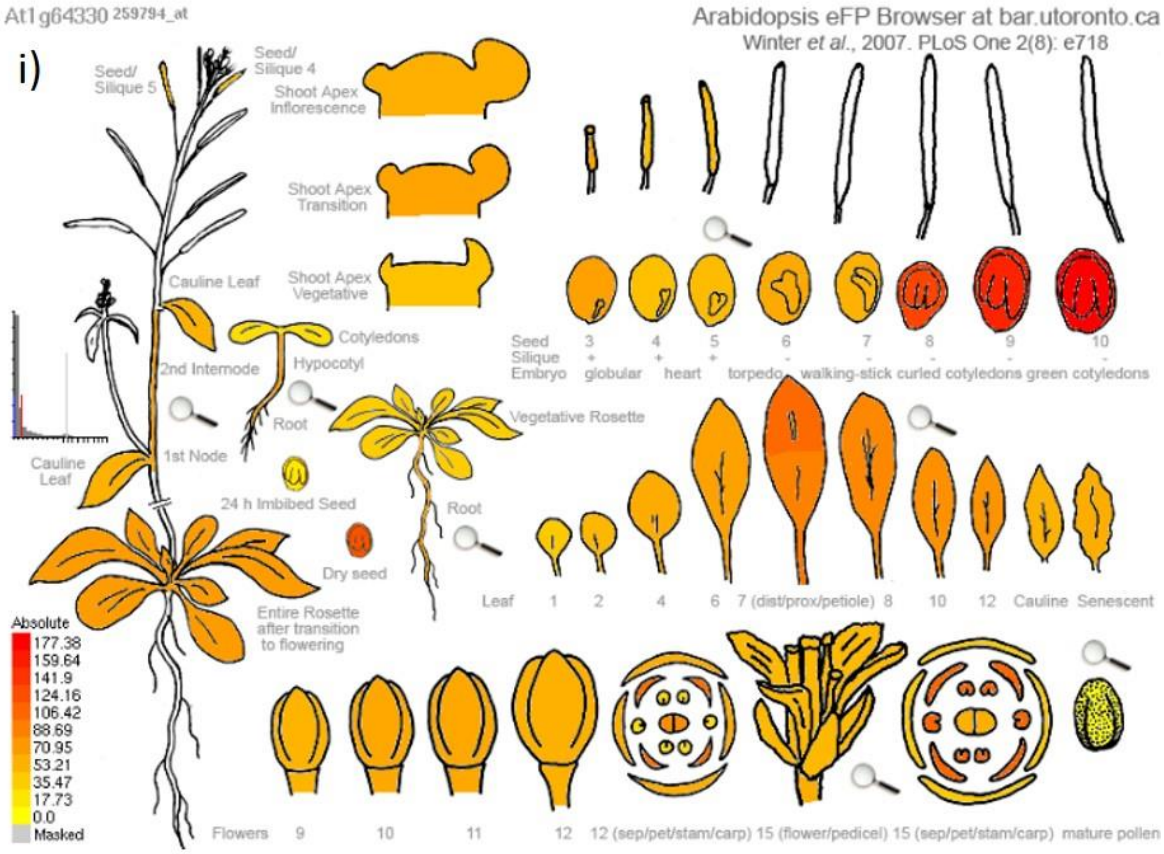

Microgametogenesis (Pollen Development)
pore
cell wall
nucleus
vacuole
Uninucleate Microphore
generative cell
Bicellular Pollen
Tricellular Pollen
Mature Pollen Grain

iv)

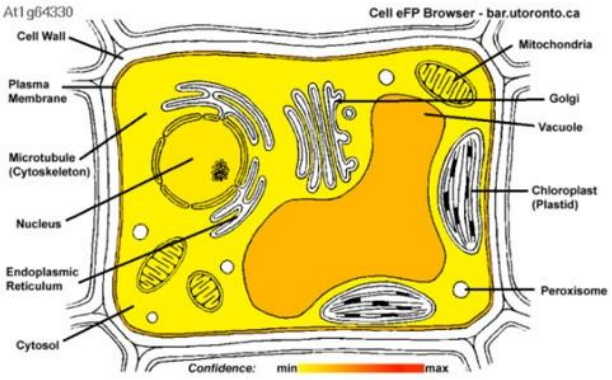

Stigma and Ovaries
Stigma Tissue
Pistil Tissue primarily consisting of Ovaries
Pollen Germination
Dry Pollen
30 minutes in vitro incubation
4 hours in vitro incubation
Pollen tubes after growth through Stigma and Style (pistil explants)

e)

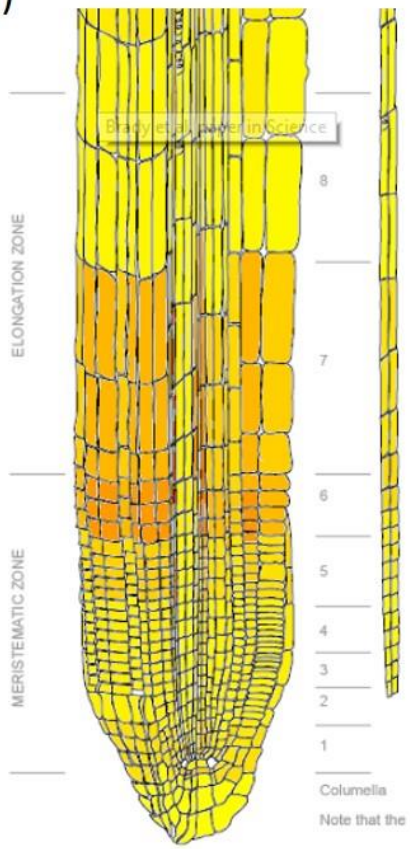
i)


At1g64330 259794_at
Arabidopsis eFP Browser at bar.utoronto.ca
Winter et al., 2007. PLoS One 2(8): e718
Seed/ Silique 5
Seed/ Silique 4
Shoot Apex Inflorescence
Shoot Apex Transition
Shoot Apex Vegetative
Cauline Leaf
Cotyledons
Hypocotyl
2nd Internode
Root
1st Node
Cauline Leaf
24 h Imbibed Seed
Dry seed
Entire Rosette after transition to flowering
Vegetative Rosette
Root
Absolute
177.38
159.64
141.9
124.16
106.42
88.69
70.95
53.21
35.47
17.73
0.0
Masked
Seed 3 4 5 6 7 8 9 10
Silique + + + - - - - -
Embryo globular heart torpedo walking-stick curled cotyledons green cotyledons
Leaf 1 2 4 6 7 (dist/prox/petiole) 8 10 12 Cauline Senescent
Flowers 9 10 11 12 12 (sep/pet/stam/carp) 15 (flower/pedicel) 15 (sep/pet/stam/carp) mature pollen

ii)


At1g64330
Cell eFP Browser - bar.utoronto.ca
Cell Wall
Plasma Membrane
Microtubule (Cytoskeleton)
Nucleus
Endoplasmic Reticulum
Cytosol
Mitochondria
Golgi
Vacuole
Chloroplast (Plastid)
Peroxisome
Confidence: min max

iii)


ELONGATION ZONE
MERISTEMATIC ZONE
8
7
6
5
4
3
2
1
Columella
Note that the

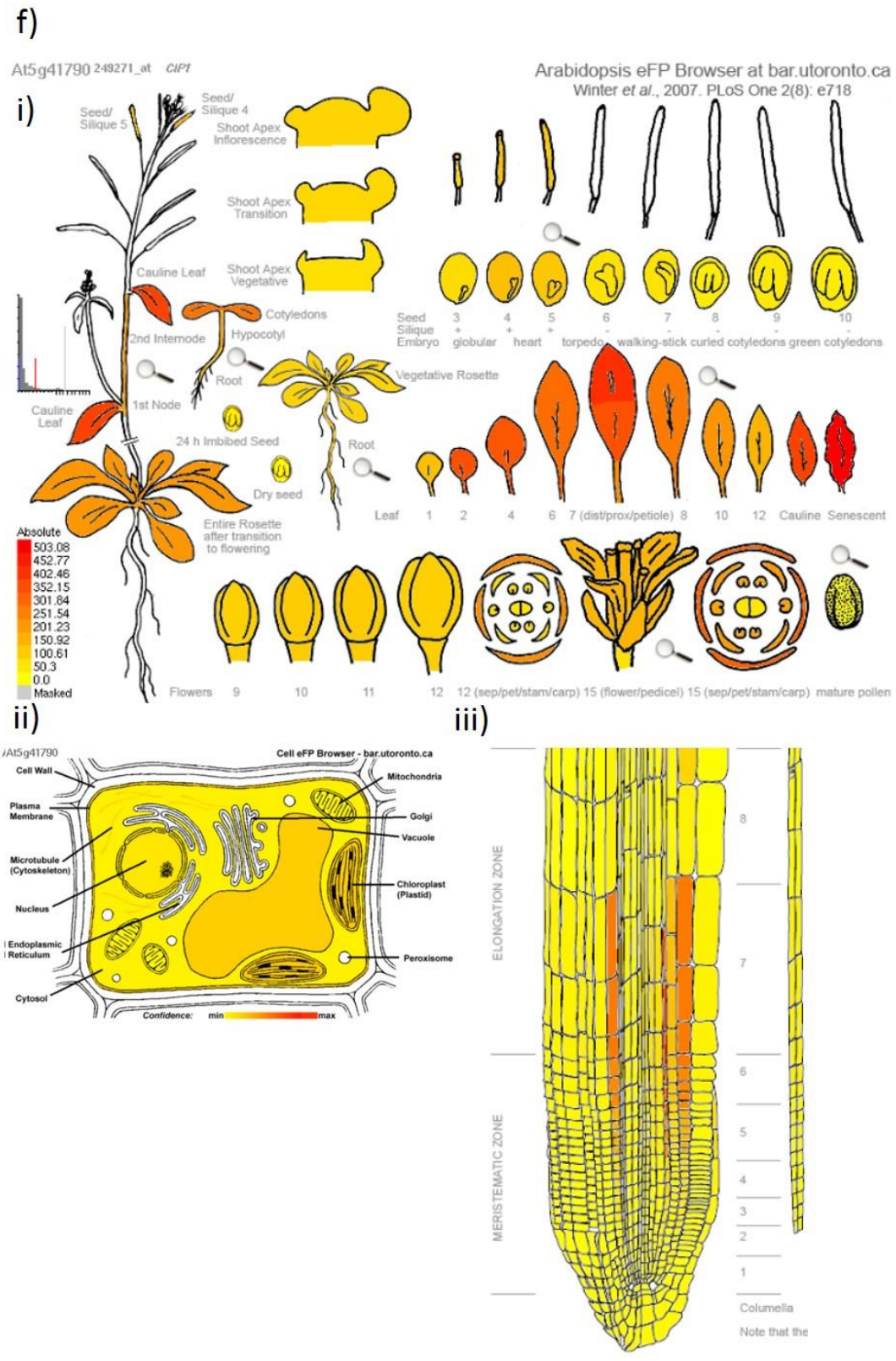


Figure 3.6: Gene expression patterns of the IRQ proteins at the tissue and subcellular levels. The gene expression patterns of the IRQ genes were displayed using the *Arabidopsis* eFP browser at both the tissue and subcellular levels. Expression level is shown based on a colour gradient from low (yellow) to high (red). a) IRQ1, b) IRQ6, c) IRQ2, d) IRQ3, e) IRQ4, f) IRQ5.

Table 3.2: Subcellular analysis of localisation of IRQ proteins summarised from SUBA3 database.

Gene	Prediction	Imaging	Mass Spectroscopy
At3g10880	Cytosol Mitochondrion Nucleus		
At5g05180	Cytosol Mitochondrion Nucleus Plastid		
At1g64320	Cytosol Mitochondrion Nucleus Cytoskeleton		Cytoskeleton (Hamada *et al.*, 2013)
At5g41780	Cytosol Mitochondrion Nucleus Cytoskeleton		
At1g64330	Cytosol Mitochondrion Nucleus Cytoskeleton		Vacuole (Carter *et al.*, 2004; Jaquinod *et al.*, 2007; Nikolovski *et al.*, 2012; Yoshida *et al.*, 2013) Plasma membrane (Elmore *et al.*, 2012)
At5g41790	Cytosol Mitochondrion Nucleus	Cytoskeleton and speckles (Matsui & Deng, 1995)	Vacuole (Carter *et al.*, 2004; Jaquinod *et al.*, 2007; Yoshida *et al.*, 2013) Plasma membrane (Elmore *et al.*, 2012) Plastid (Kleffmann *et al.*, 2004; Kong *et al.*, 2011)

3.7 Phylogenetic analysis of IRQ proteins

3.7.1 Phylogenetic analysis of the IRQ proteins calculated using the IRQ domain alone

In order to answer questions relating to the identity and timing of appearance of the IRQ domain, the emergence of different IRQ proteins, and how well their evolutionary lineage agrees with that of the NET protein superfamily, phylogenetic analysis was carried out across multiple species. Species used in this analysis are shown in evolutionary context in Figure 3.7 and included: *Ostreococcus tauri*, a single celled phytoplankton and one of the smallest known eukaryotes (Palenik *et al.*, 2007); *Physcomitrella patens*, the model moss species occupying a critical position in the evolution of land plants as plants moved from the water and occupied the land (Rensing *et al.*, 2008); *Selaginella moellendorffii;* gymnosperms including *Picea abies* (Norway Spruce) and *Pinus taeda* (Loblolly pine); representatives from the ancient angiosperms, plants which are important in understanding the development of the flowering plants and thought to be the sister lineage to the prolific eudicots and monocots including *Amborella trichopoda,* a monotypic genus of shrub endemic to new Caledonia (Soltis *et al.*, 2008; Zuccolo *et al.*, 2011; Albert *et al.*, 2013; Chamala *et al.*, 2013), *Persea americana,* (avocado) (Chanderbali *et al.*, 2008; Soltis *et al.*, 2009)*, Liriodendron tulipifera*, a magnoliid (Liang *et al.*, 2011)*, Zamia vazquezii,* a critically endangered cycad species*, Nuphar advena* (water lily) (Yoo *et al.*, 2010)*, Aristolochia fimbriata* (Dutchman's pipe), a potentially useful model basal angiosperm (Bliss *et al.*, 2013); representative monocots *Oryza sativa* (rice), *Zea mays* (Maize), and *Brachypodium distachyon* (stiff brome); and a single eudicot species, *A. thaliana*. A search of three non-plant genomes (*Drosophila melanogaster, Saccharomyces cerevisiae,* and *Homo sapiens*) did not produce any homologues, suggesting that the IRQs are a plant specific family.

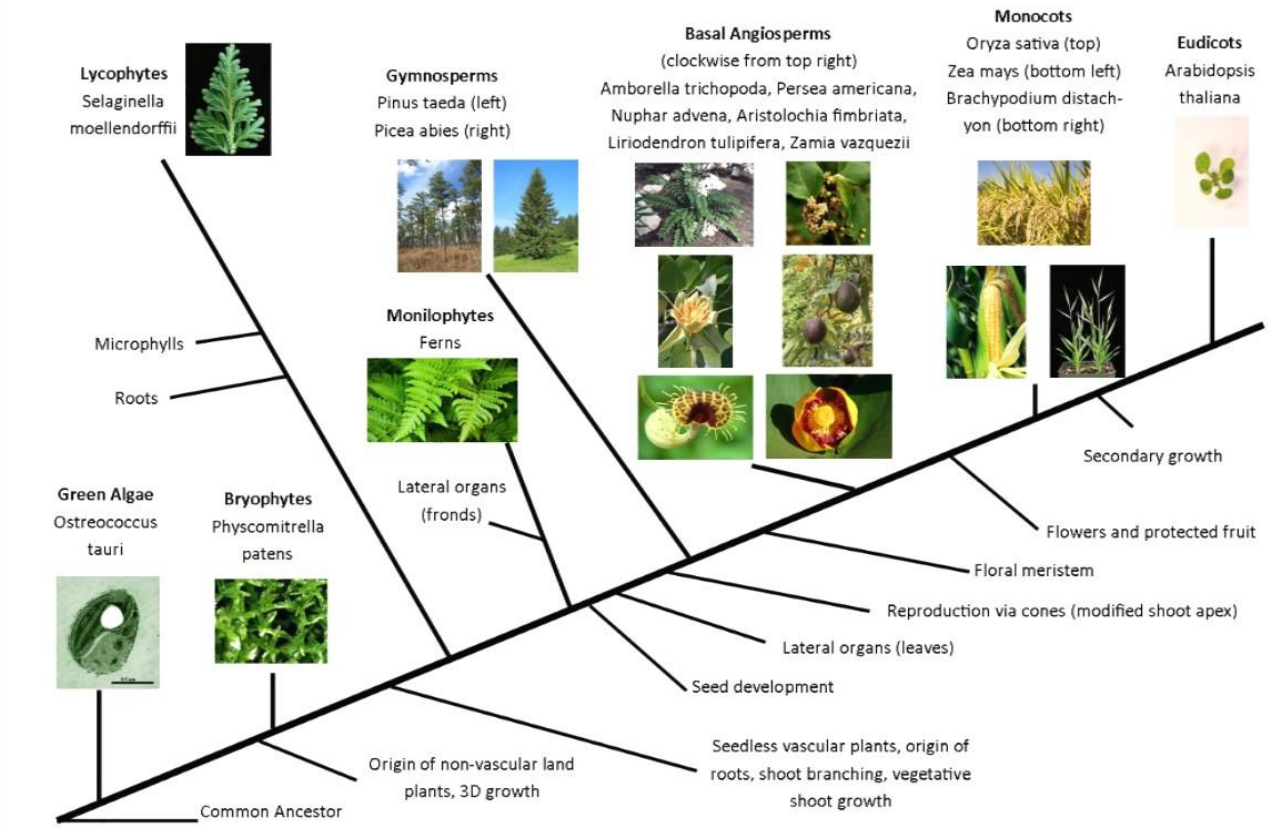


Figure 3.7: Schematic showing the species used in this phylogenetic analysis against major physiological changes during land plant evolution. Names in bold refer to major plant groupings. Plant names underneath are the organisms used during the phylogenetic analysis. Major physiological changes and events during land plant evolution are shown at their appropriate relative places along the evolutionary timeline (major bold line).

The IRQ domains of each IRQ protein were taken in turn and used to search sequence databases as detailed in section 2.8. The putative IRQ domains of any hits returned with at least 40% similarity to the IRQ domain in question were retained, aligned using ClustalX 2.1 (Larkin *et al.*, 2007), and then loaded into MetaPIGA v.20 phylogenetic estimation analysis software (Helaers & Milinkovitch, 2010) using stochastic heuristics for large phylogeny inference with the Metapopulation Genetic Algorithm (metaGA) (Lemmon & Milinkovitch, 2002). The metaGA algorithm uses a maximum likelihood (ML) method for phylogeny inference. GAs mimic evolution by taking a population of individuals and subjecting them to mutation, recombination, reproduction, and selection scenarios. In the case of phylogeny inference GA takes trees with their different topologies and parameters e.g. branch lengths and subjects them to the same process but this time mutations may involve changing the topology of the tree and monitoring the score for the tree. Selection ensures that trees with higher scores after each successive generation are retained, improving the trees over time. The metaGA improves the speed and efficiency with which ML trees are found (Lemmon & Milinkovitch, 2002). The MetaPIGA software increases dataset quality by detecting identical sequences and removing them and also detecting and giving the option to remove highly divergent sequences. Any poorly aligned regions are trimmed away using the trimAl algorithm (Capella-Gutiérrez *et al.*, 2009). MetaPIGA calculations were stopped when the mean relative error of 10 consecutive consensus trees stayed below 5% using trees sampled every 5 generations or the likelihood stopped increasing after 200 iterations. A full list of all the IRQ domains used in this chapter and their alignment can be found in appendix 2. The resulting consensus tree produced was imported into FigTree v1.4.3 (Andrew Rambout, University of Edinburgh) to create figures. Figure 3.8 shows the results of this analysis when using the IRQ domains alone.

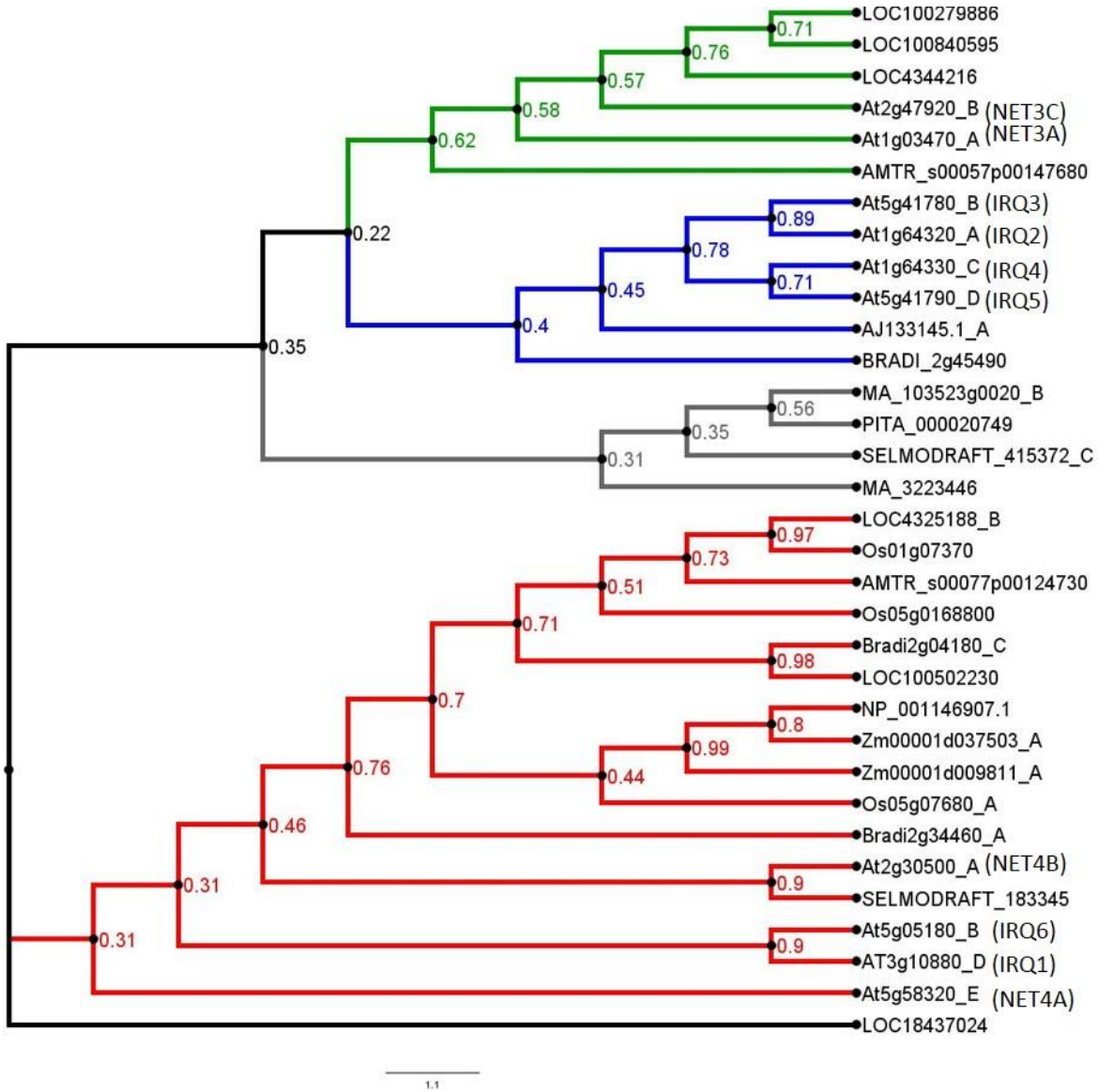


Figure 3.8: Phylogeny of NAB/IRQ and IRQ proteins calculated from the alignment of IRQ domains. Branches are coloured according to groupings of NAB/IRQ and IRQ proteins. Green: NET3 IRQ domains; Blue: IRQ2, IRQ3, IRQ4, and IRQ5 subgroup; Grey: ancestral IRQ group related to green and blue groups; Red: NET4 and IRQ1 and IRQ6 group.

The topology of the tree suggests 4 broad groupings: a NET3 group, a group of four IRQ proteins (IRQ2, 3, 4 and 5), a NET4 related group of proteins, and a NET4 group which also contains two IRQ proteins (IRQ1 and 6).The first appearance of the IRQ domain is in *S. moellendorffii* with sequences most similar to the NET4s. No IRQ proteins appear in the genomes of the organisms tested before the appearance of the NET4 suggesting that the IRQ domain did not appear alone and that the IRQ proteins evolved from the NETs, following a pattern consistent with the evolution of the NETs through evolutionary time (Hawkins *et al.*, 2014). The position of IRQ1 and 6 with NET4A being shown as having the sister domain to these IRQs suggests that IRQ1 and 6 diverged earliest among the IRQ proteins and retained the most similarity to the basal NET4 compared to the remaining IRQs. IRQ1 and 6 do not have a domain with any similarity to the NAB domain so the retention of their respective IRQ domains suggests that the NAB domain was lost in IRQ1 and 6 while the IRQ domain was retained.

Previous work using the NAB domain alone showed that within the clade containing the NET3s there was one orthologue each in *Z. mays* (Zm2g131554) and *B. distachyon* (Bd1g18347; Hawkins *et al.*, 2014). Using the IRQ domain alone the current analysis has identified an additional protein in *O. sativa* (LOC4344216). This protein may not have been identified in the Hawkins *et. al.* (2014) phylogeny because LOC4344216 was not annotated properly in the sequence database, does not possess a NAB domain, or it is of such weak similarity that it was not considered. Its presence within this tree is most likely due to the similarity between itself and the IRQ domains of the NAB/IRQs NET3A and NET3C as shown in Figure 3.8. The presence of a NET3 IRQ domain homologue in *A. trichopoda* is in agreement with the pattern of NET evolution (Hawkins *et al.*, 2014) following the appearance of the proto NET4 in *S. moellendorffii.*

The grouping of IRQs 2, 3, 4, and 5 shows that these IRQ domains are distinct from all other IRQs. Their closeness to the NET3 clade suggests a closer relationship between the IRQ domains in these proteins and the corresponding IRQ domains in the NET3s than to the NET4s. IRQ2 and 3 are suggested by the available affymetrix data to be pollen specific (Figure 3.6c and d) and in this analysis they are shown to be more closely related to one another than to IRQ4 and 5, which are predicted to be expressed weakly in pollen. Within this clade the single examples from the basal angiosperm *P. americana* and the monocot *B. distachyon* are unusual.

BRADI_2g45490 is identified as a golgin and is also annotated in UniprotKB as having a NAB domain. Golgins are a group of coiled-coil tether proteins involved in fusion of vesicles to their target compartment and are effectors of Rab GTPases (Hutagalung & Novick, 2011). Alignment of the IRQ domains in these proteins (Figure 3.9) shows there is a high degree of conservation across the domains, in particular the amino acids 'KREAIRQL'. However, in the BRADI_2g45490 sequence the 'R' of the IRQ has been replaced with a 'K'. As this is the replacement of an amino acid with a positively charged side chain with another of similar charge this does not represent a significant change. The lack of any other monocot studied showing greater specificity to an IRQ rather than a NAB/IRQ protein could be due to the strict criterion used to select sequences, so any sequences of slightly lower similarity that are true IRQ domain containing proteins may be missed. Alternatively, after the divergence between the monocots and eudicots, the BRADI_2g45490 sequence could have diverged and independently had particular amino acid changes which made the IRQ domain sequence more similar to the IRQs it is associated with in Figure 3.8 compared to the earlier ancestral NET4 or NET3 sequences.

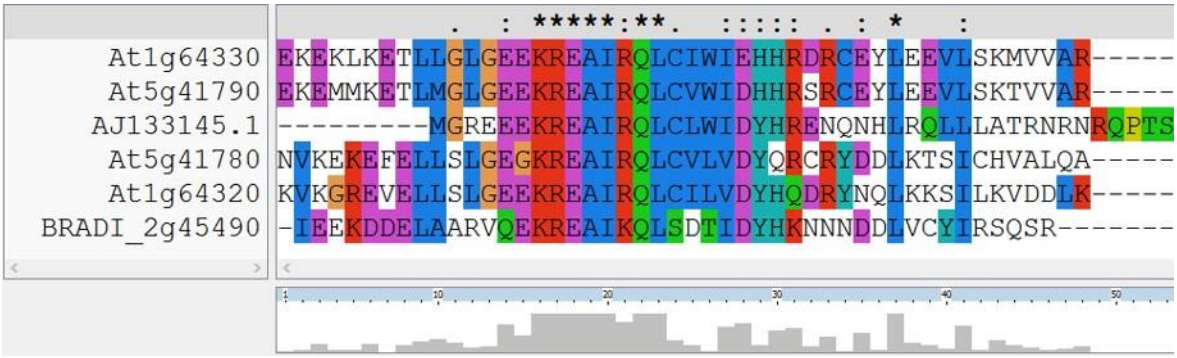


Figure 3.9: Alignment of IRQ proteins with a basal angiosperm and a monocot. The IRQ domains of IRQ2, IRQ3, IRQ4, and IRQ5 were aligned against AJ133145.1 and BRADI_2g45490 using ClustalX 2.1 software with amino acids coloured according the default settings.

The *P. americana* sequence identified (AJ133145.1) codes for a partial sequence designated as a metallothionein-like protein type 2. Metallothionein-like proteins sequester Zn^{2+} ions and are involved in tolerance to high Zn^{2+} levels (Robinson *et al.*, 1996). AJ133145 shows a relatively strong similarity to the core part of the IRQ domain of the other IRQs within this group (Figure 3.8). However, the annotation of this protein as a partial sequence and lack of the availability of a complete sequence for this protein call into question whether this sequence has been annotated correctly and whether it codes for a protein at all.

In summary the phylogenetic analysis using the IRQ domain alone has shown that the IRQ proteins appear to have evolved from the NET proteins, IRQ1 and IRQ6 have IRQ domains most closely related to NET4 proteins so likely diverged first among the IRQ proteins, and no IRQs were found specifically within the gymnosperms suggesting a eudicot/monocot specificity.

3.7.2 Phylogenetic analysis calculated using the full length IRQ protein sequences

Figure 3.9 shows the results of this analysis when using the full length proteins shown in in Figure 3.8. Note the full length IRQ proteins were not used to search the databases in the same way as the IRQ domains; the full length sequences of those proteins identified in Figure 3.8 were aligned and processed in the METAPIGA software to generate Figure 3.9.

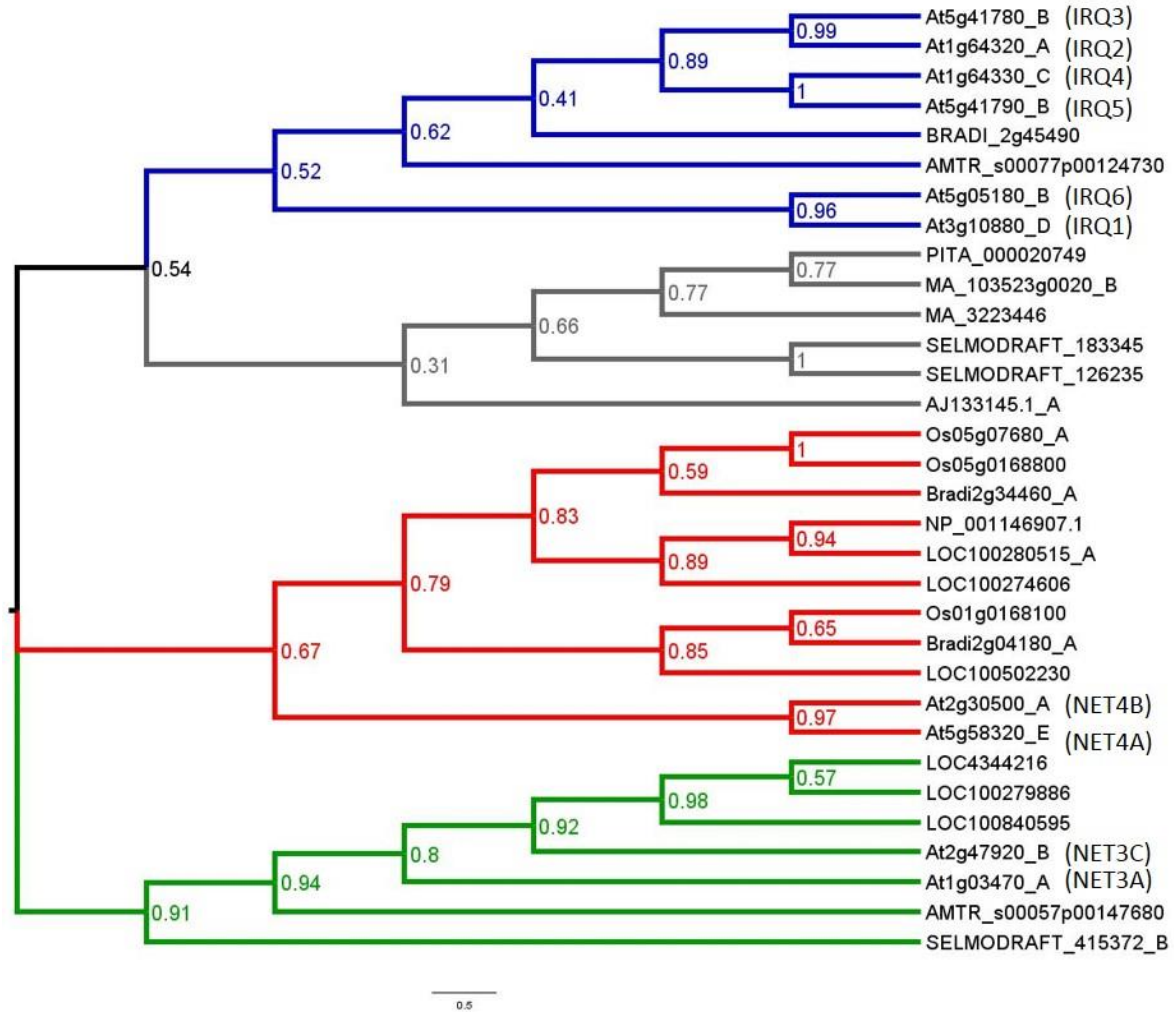


Figure 3.10: Phylogeny of NAB/IRQ and IRQ proteins calculated from the alignment of the full length proteins from Figure 3.8. Branches are coloured according to groupings of NAB/IRQ and IRQ proteins. Green: NET3 IRQ domains; Blue:IRQ1, IRQ6, IRQ2, IRQ3, IRQ4, and IRQ5 subgroup; Grey: ancestral IRQ group related to green and blue groups; Red: NET4 group.

The tree shows the divergence of the IRQ proteins when the full sequence is considered. NET3 and NET4 proteins form separate groups but this time all the IRQ proteins form a separate clade with a branch of NET4 related sequences (Figure 3.9). Within the NET3s the *O. sativa* LOC4344216 protein identified as not being in the Hawkins *et. al.* (2014) paper appears again. Alignment of the full length sequences of NET3A, NET3C and LOC4344216 (Figure 3.11) shows that it possess a highly similar NAB domain as well as the IRQ domain, classifying this protein as a NAB/IRQ and validating its position within both trees.

Within the IRQ proteins the *P. americana* AJ133145.1 sequence that was grouped with these proteins in the IRQ domain alone tree in Figure 3.8 has moved to be with the clade of *S. moellendorffii* and pine sequences when the full length proteins are considered (Figure 3.9). As only a partial sequence was available for this protein, not a great deal of significance can be put on the positioning of this protein in either tree. The BRADI_2g45490 Golgin sequence is again present. Through evolutionary time it is possible that the predicted NAB domain of BRADI_2g45490 has been mutated, possibly giving rise to its position with the tree next to the IRQ proteins rather than either of the NAB/IRQs. According to the criteria used to select sequences for this phylogenetic analysis no other monocot sequences were unique enough from the NET3 and NET4 proteins to be related to the IRQs leaving BRADI_2g45490 as the sole example. The separation between IRQ1 and 6 and the other IRQs by the BRADI_2g45490 and AMTR_s00077p00124730 sequences again underscores how different IRQ1 and 6 are from the other IRQs and the closer relationship between IRQ1 and IRQ6 with NET4 rather than NET3 group proteins.

AMTR_s00077p00124730 is annotated as a NET4A sequence and in the IRQ domain alone tree was found with the NET4s and some of their orthologues (Figure 3.8). In Figure 3.9, using the full length proteins, it is found with the IRQ proteins. An alignment of the full length sequences of AMTR_s00077p00124730 with the NET4s, IRQ1, IRQ6, IRQ4 and IRQ5 (Figure 3.12) shows that this protein is particularly large and has similarity at the IRQ domain and NAB domains. There are several large gaps in the AMTR_s00077p00124730 sequence dispersed between regions of similarity to the NET4s and IRQs and this could indicate a problem with the curation of the *A. trichopoda* database where potentially gene splicing events have not been identified correctly and introns have not been correctly removed leading to their inclusion in the protein sequence. This may mean the position of the AMTR_s00077p00124730 protein within the full length tree is not an accurate reflection of the true evolutionary relationship of this protein to the NAB/IRQ and IRQs.

In summary, comparison of the full length sequences identified in section 3.7.2 shows that the IRQ proteins are more closely related to each other compared to other IRQ domain containing proteins.

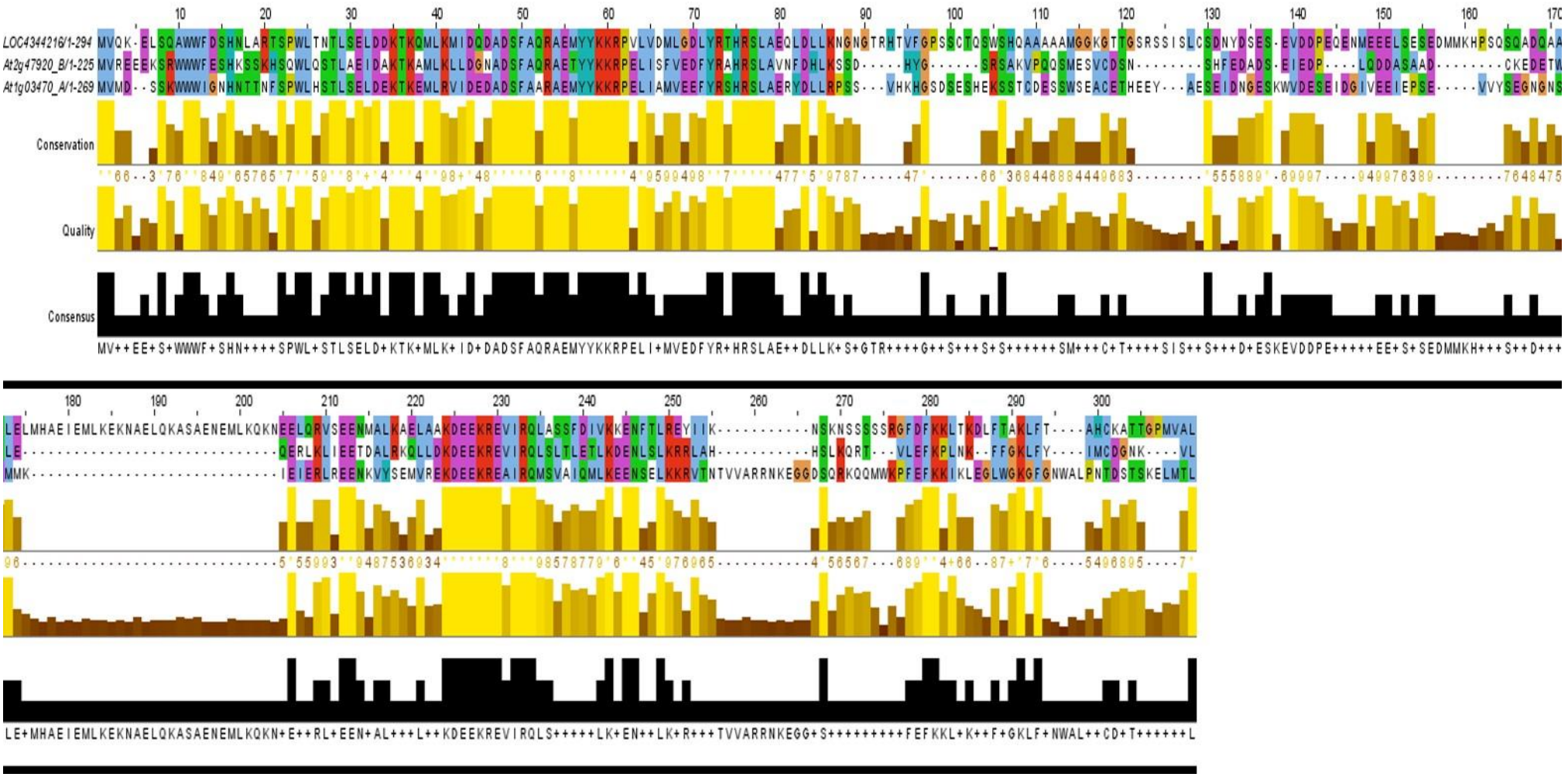


Figure 3.11: Alignment of full length NET3A (At1g03470), NET3C (At2g47920), and the *O. sativa* LOC4344216 sequences. The NET3 sequences and LOC4344216 were aligned in ClustalX 2.1 and exported to Jalview. There is a high degree of conservation at both the N-terminus where the NAB domain is located and at the C-terminus where the IRQ domain is located, verifying the position of this protein with the NET3 sequences in the tree.

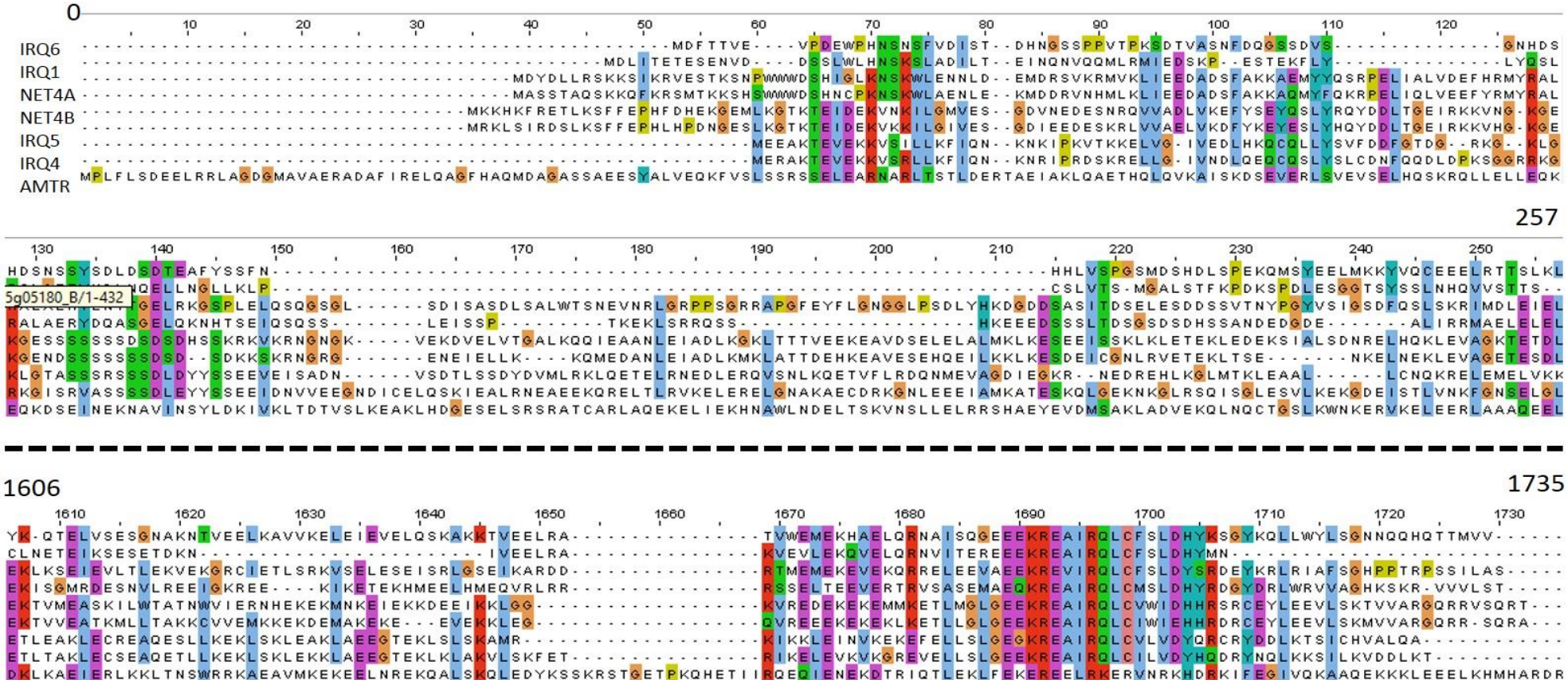


Figure 3.12: Alignment of full length IRQ6, IRQ1, IRQ4, IRQ5, NET4A, NET4B, and AMTR_s00077p00124730 sequences. Sequences were aligned in ClustalX 2.1 and exported to Jalview using colours according to default clustal settings. AMTR stands for AMTR_s00077p00124730. Due to the sizes of the IRQ5 and AMTR_s00077p00124730 sequence sonly the regions corresponding approximately to the NAB domain and IRQ domains are shown. The numbers above the alignment refer to the amino acid number. The dotted line refers to the break between the N-terminal NAB sequences and the C-terminal IRQ sequences.

3.7.3 Eudicot specific phylogenetic analysis of IRQ proteins using the IRQ domain alone

The NET proteins have homologues in monocot species but this was not the case with the IRQ proteins as shown in Figure 3.8 and Figure 3.9, suggesting that the IRQ proteins may be eudicot specific. The eudicots contain at least 70 % of all flowering plant species (Tudge, 2006). The physiological differences between eudicots and monocots e.g. number of flowering parts, pollen physiology, and leaf venation may indicate differences in cellular processes. To see if the IRQ proteins were unique to *A. thaliana* or were present in other eudicots the IRQ domains of the IRQ proteins were used to search Phytozome as described in section 2.8. The three organisms in addition to *A. thaliana* chosen for this analysis were from outside the Brassicaceae and included one member of each of the Fabaceae (*Glycine max*), Vitaceae (*Vitis vinifera*), and Solanaceae (*Solanum lycopersicum*). These organisms were chosen due to their well annotated genomes and to represent a wide range of families as possible within the prolific eudicot clade. Aligned sequences were loaded into METAPIGA as described in section 3.7.1 to produce consensus trees which were visualised in FigTree.

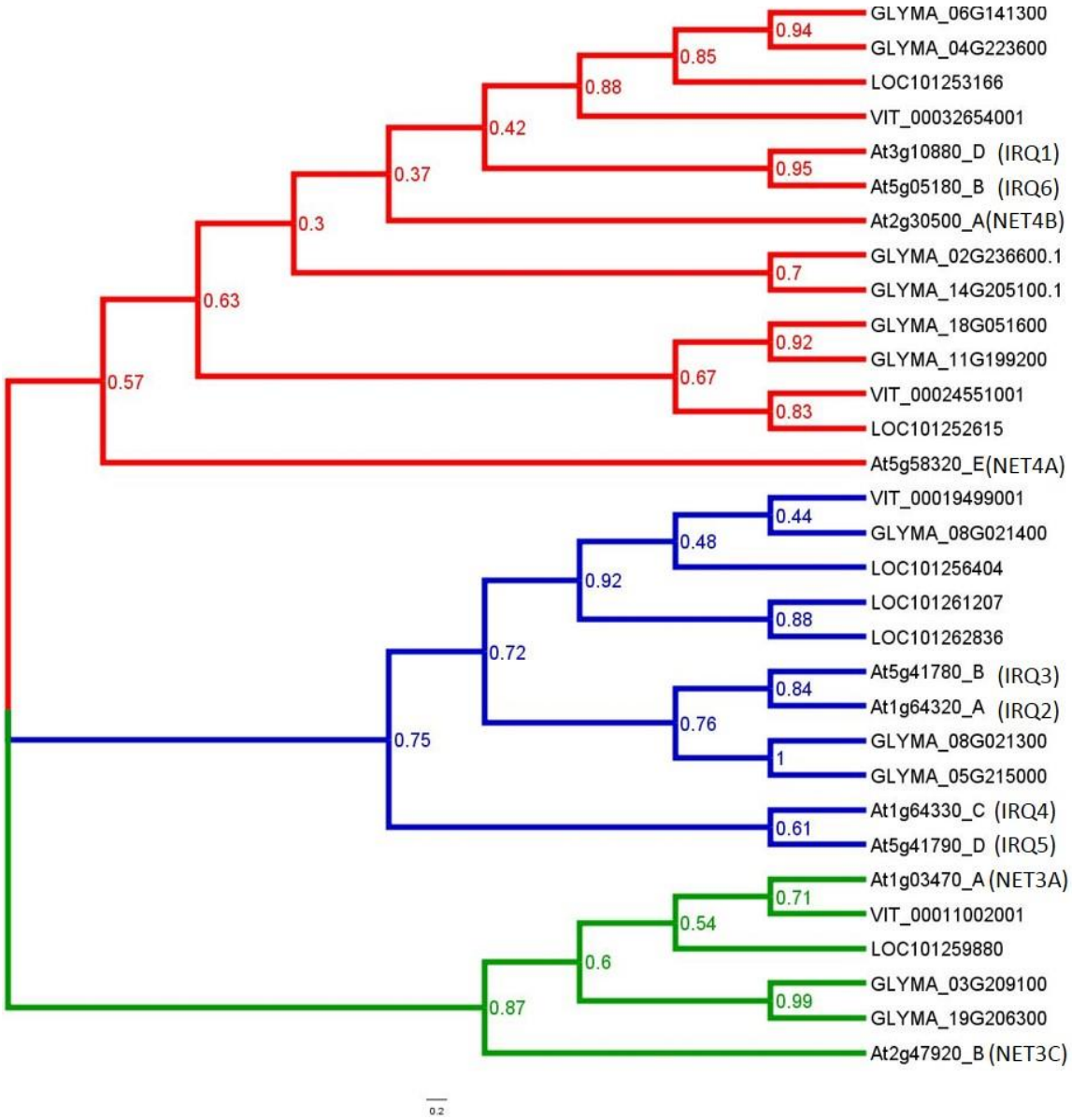


Figure 3.13: Phylogeny of NAB/IRQ and IRQ proteins in four eudicot species calculated from the alignment of the IRQ domains alone. Branches are coloured according to groupings of NAB/IRQ and IRQ proteins. Green: NET3 IRQ domains; Blue: IRQ2, IRQ3, IRQ4, and IRQ5 subgroup; Red: NET4, IRQ1, and IRQ6 group.

The presence of the NET4s with IRQ1 and 6 within the same clade repeats the results shown in Figure 3.9, underscoring the similarity between these IRQ domains in other eudicots and green plants. The separation in this tree between the NET4A and NET4B IRQ domains suggests that NET4A and B have diverged separately within the species studied here (Figure 3.13). This is in comparison to the IRQ1 and 6 sequences in Figure 3.13 which remain closely related. As IRQ1 and 6 could not be separated by the species used in this analysis, it could be tentatively suggested that only one copy of IRQ1 or IRQ6 is present in other eudicots, and any duplication to produce other IRQs similar to IRQ1 or IRQ6 occurred from a single protein (either IRQ1 or IRQ6) rather than the species starting with copies of both IRQ1 and 6. Only one *V. vinifera* and one *S. lycopersicum* sequence is shown branching off from IRQ1 and 6 within Figure 3.13. *Glycine max* appears to have two sequences (GLYMA 06G141300 and GLYMA 04G223600) so could represent a doubling of a single IRQ1 or 6 orthologue. This hypothesis relating to IRQ1 and 6 requires validation by looking at a much broader range and number of eudicot species.

The clade containing IRQs 2, 3, 4, and 5 shows that there is greater similarity between the IRQ2 and IRQ3 IRQ domains than to IRQ4 and IRQ5. Similar to IRQ1 and 6, there are no IRQ domains in other species which are more similar to either proteins within the IRQ2/3 or IRQ4/5 pairs. With the development of angiosperms and their rapid expansion and concomitant increase in the number of flower types (Horn *et al.*, 2014) there was an increase in the complexity of pollination mechanisms (Crepet *et al.*, 1991; Crane *et al.*, 1995; Dodd *et al.*, 1999). As IRQ2 and 3 are indicated to be pollen specific it could be suggested that these putative orthologous sequences are also associated with functions in pollen. The presence of IRQ4 and 5 as a seeming outgroup in this clade may suggest that these two proteins arose slightly earlier than IRQ2 and 3 and their orthologues. This tree overall shows that the divergence between the IRQ domains in these proteins perhaps occurred relatively early on in eudicot evolution.

In summary, the higher level of similarity between the NET4 IRQ domains and the IRQ domains in IRQ1 and 6 compared to the other IRQ domains is repeated in the eudicot species studied here. This lends support to the hypothesis that IRQ1 and 6 were the first IRQ proteins to diverge form the NET4 sequences. IRQ1 and 6 may have arisen from the duplication of a single protein in other eudicot species. Finally, the presence of distinct IRQ isoforms present in these diverse families of eudicots suggests that diversification among the IRQ proteins probably occurred early on in eudicot evolution.

3.7.4 Eudicot specific phylogenetic analysis of IRQ proteins using the full length sequences

Full length sequences from those proteins used to generate Figure 3.13 were aligned, processed in METAPIGA and the consensus tree produced visualised using FigTree as described in section

3.7.1. The tree produced in Figure 3.14 is different to that produced for the IRQ domains alone. The NET4s, NET3s, and their orthologues are more closely related in this tree while IRQ1 and 6, formerly more closely linked to NET4s using only the IRQ domain, form their own separate clade. In this instance *Glycine max* possesses four copies of the NET4s while *V. vinifera* and *S. lycopersicum* possess only one copy of NET4A or NET4B. A similar observation can be made in the NET3s, with two NET3 orthologues in *G. max* but only one NET3 in *V. vinifera* and *S. lycopersicum*. IRQ2, 3, 4, and 5 are again more closely related compared to other IRQs but this time there is a split between their orthologues: there are only single copies of the orthologues for IRQ4 and 5 whereas IRQ2 and 3 have two copies of *G. max* and *S. lycopersicum* but seemingly no orthologues in *V. vinifera.* The single copies of the IRQ4 and 5 orthologues suggest that a gene duplication may have taken place, but only in *A. thaliana*. However, it may be that there are other eudicots not included in this analysis which contain more than one copy of the IRQ4 and 5 orthologues. Also, any conclusions relating to the number of copies must be made with caution as this tree was produced using only the proteins found in the previous analysis using the IRQ domains alone as search terms. There may be other parts in the sequences of these proteins which could increase the number of orthologues present for each NAB/IRQ or IRQ protein in other eudicots.

In summary using the full length proteins shows that the pattern of increased similarity between IRQ proteins compared to other IRQ domain containing proteins is shared in diverse eudicot species.

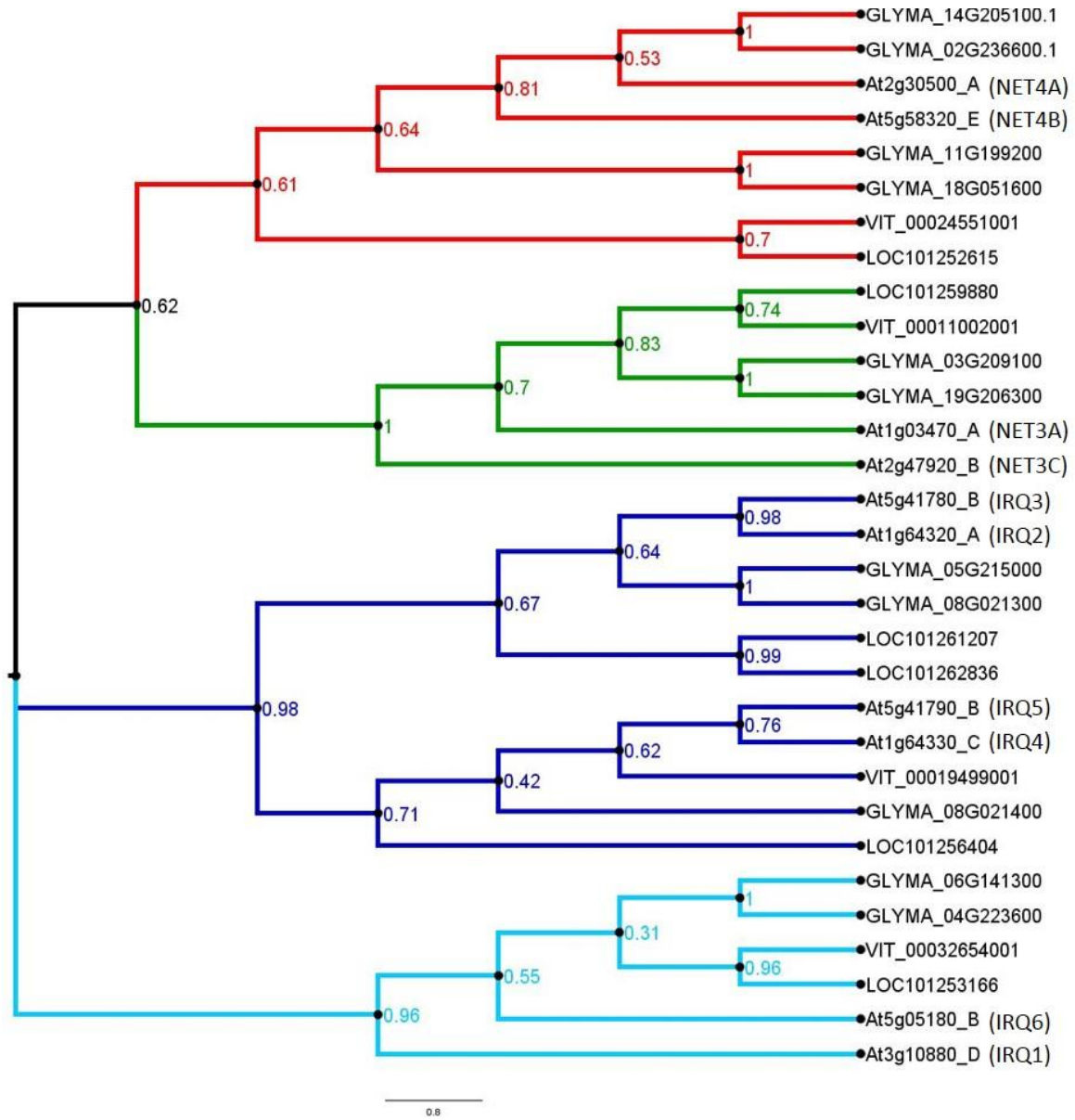


Figure 3.14: Phylogeny of NAB/IRQ and IRQ proteins in four eudicot species calculated from the alignment of the full length proteins from Figure 3.13. Branches are coloured according to groupings of NAB/IRQ and IRQ proteins. Note the groupings have changed from Figure 3.13. Green: NET3 subgroup; Dark blue: IRQ2, IRQ3, IRQ4, and IRQ5 subgroup; Red: NET4 subgroup; Turquoise: IRQ1 and IRQ6 subgroup.

3.7.5 Summary of phylogenetic data

The major findings from the phylogenetic analyses presented above are that the IRQ proteins evolved from the NET proteins and IRQ1 and IRQ6 are likely to have diverged first among the IRQ proteins. The lack of IRQ proteins among gymnosperms and monocots suggests that the IRQ proteins are eudicot specific, and that diversification of IRQ proteins among the eudicots probably occurred early in eudicot evolution.

3.8 IRQ Gene number expansion through evolutionary time in *Arabidopsis thaliana*

Through evolutionary time gene families can contract and expand through mechanisms including gene loss and duplication (Guo, 2013). Approximately 70 % of all angiosperms have undergone some form of polyploidy in their ancestry (Masterson, 1994). Analysis of the *Arabidopsis* genome has shown that 60% of the genome consists of large duplicated segments, with 17% of *Arabidopsis* genes occurring in tandem arrays (The Arabidopsis Genome Initiative, 2000). Despite this duplication *Arabidopsis spp.* gene families have on average shown a net contraction of 0.547 gene lost per gene family, with 989 gene families showing contractions in the analysis by Guo (2013). The *Arabidopsis* genome appears to have undergone a rapid expansion due to whole genome duplication and then been subject to purifying selection (Hoffmann & Palmgren, 2016)

Previous work on the NET superfamily has shown that certain members have undergone duplications, inversions or have potentially been lost over evolutionary time (Hawkins *et al.*, 2014). An isoform of *NET3* lies in close proximity to an isoform of NET1 and it has been shown that this chromosome segment has been duplicated from chromosome 1 to chromosome 4 or *vice versa*. Similarly, *NET2A* and *NET2B* lie in a gene segment that has been duplicated on the same chromosome (chromosome 1). Finally, *NET1A* and *NET1B* lie in a segment that has been duplicated then inverted between chromosomes 3 and 4. To see if any of the IRQ domain containing proteins have undergone any similar processes to the NETs a similar analysis of collinear clusters across the 5 chromosomes was carried out.

Chromosomes 1 and 5 contain one copy of *IRQ4* or *IRQ5* and one copy of *IRQ2* or *IRQ3*, in two pairings: *IRQ2* next to *IRQ4* (Figure 3.15bi), and *IRQ3* next to *IRQ5* (Figure 3.15bii). Using collinear analysis it was found that these genes lie within a section of chromosome which has undergone a tandem duplication and block duplication. *IRQ6* is annotated as having undergone a block duplication (Figure 3.15biii) but the most similar IRQ to *IRQ6, IRQ1*, is not annotated as having been duplicated. Therefore, gene duplications can in part account for the expansion of some of the IRQs. The implications of these duplications for IRQ protein evolution and possible scenarios for the order in which they took place are discussed in section 3.9.

Information on position effects in plants is lacking in plants in comparison to animals and yeast (Singh *et al.*, 2008). It is possible that this might be occur at the IRQ gene loci. A recent example is that of nucleolus organizer regions (NORs) where a large number of rRNA genes are localised. Certain rRNA subtypes on chromosome 2 (NOR2) were silenced while those on chromosome 4 (NOR4) were active (Mohannath *et al.*, 2016). A chromosome 2 position effect was found to due

selective silencing through methylation of the genes. A survey of the IRQ loci using the Methylome mapping tool (Zhang *et al.*, 2006) revealed no significant predicted gene methylation.

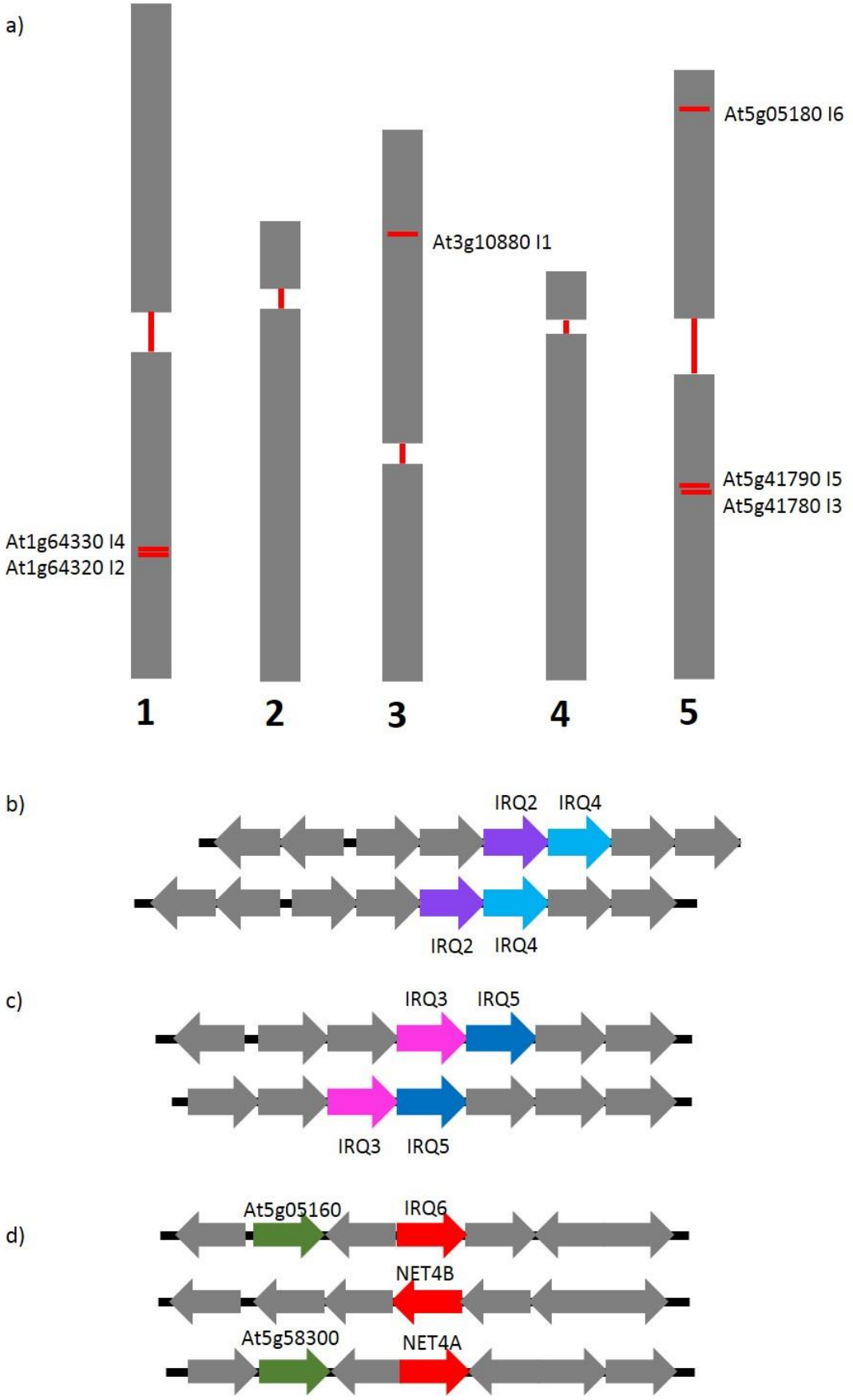


Figure 3.15: Duplication of IRQ genes in the *Arabidopsis* genome. a) Schematic of the approximate positions of the IRQ genes (red bars) mapped onto the five *Arabidopsis* chromosomes (1-5 below each chromosome). I1 indicates IRQ1, I2 indicates IRQ2 etc. b-d) Gene duplication information obtained from the PLAZA tool. Diagram indicating the block duplication and tandem duplications that occurred between IRQ2 and IRQ4 (b) and IRQ3 and IRQ5 (c). d) Diagram indicating the block duplication between IRQ6 and NET4A. Note, the order and orientations of the duplications are not specified here and are discussed in the text.

3.9 Developing a model for the evolution of the IRQ proteins

In order to develop a model for the evolution of the IRQ proteins it is helpful to bear in mind the following relevant information from the analysis presented so far in this chapter:

1. The first protein containing an IRQ domain was a NET4 isoform
2. The IRQ domains of IRQ1 and IRQ6 are more closely related to the NET4 IRQ domains than to NET3 sequences or other IRQ sequences, indicating that they diverged first among the IRQ proteins.
3. The lack of IRQ proteins in gymnosperms and monocots indicates that IRQ proteins are eudicot specific.

As IRQ1 and IRQ6 are likely to have diverged first among the IRQ proteins a fundamental question to answer is whether the IRQ proteins all evolved from IRQ1 and/or IRQ6 or independently via the NET4 or NET3 proteins. The block duplication of IRQ6 not giving rise to IRQ1 is slightly surprising but not unusual; NET3A and NET3B arose due to a block duplication but NET3C did not (Hawkins *et al.*, 2014). IRQ1 could have arisen due to a gene duplication event in an ancestor species to *A. thaliana.*

In support of point 2 above the block duplication involving IRQ6 links this IRQ protein with NET4A and NET4B (Figure 3.15d). In addition, two other proteins, At5g05160 and At5g58300, are also linked to the block duplication of IRQ6 and were not found in the previous linkage analysis with the NET proteins. This could be due more sequencing information being available with the release of the TAIR10 annotation. For example, the TAIR 9 release gave a truncated version of IRQ2 missing 147 amino acids/nucleotides from the N-terminal end. This has been corrected in the TAIR10 release. Alternatively, the most recent version of PLAZA used in this analysis (PLAZA 3.0) was released in 2014, after submission of the Hawkins *et al.* (2014) paper so any improvements in the results would have been missed. At5g05160 encodes Reduced in Lateral growth 1 (RUL1) and At5g58300 encodes for an unknown protein of the Leucine-rich repeat receptor kinase family. Alignment of At5g05160 and At5g58300 shows that there is a high degree of similarity between these two proteins (Figure 3.16). The similarity between these two non-NET/NAB/IRQ proteins suggest that NET4A and At5g58300 were duplicated to create IRQ6 and At5g05160.

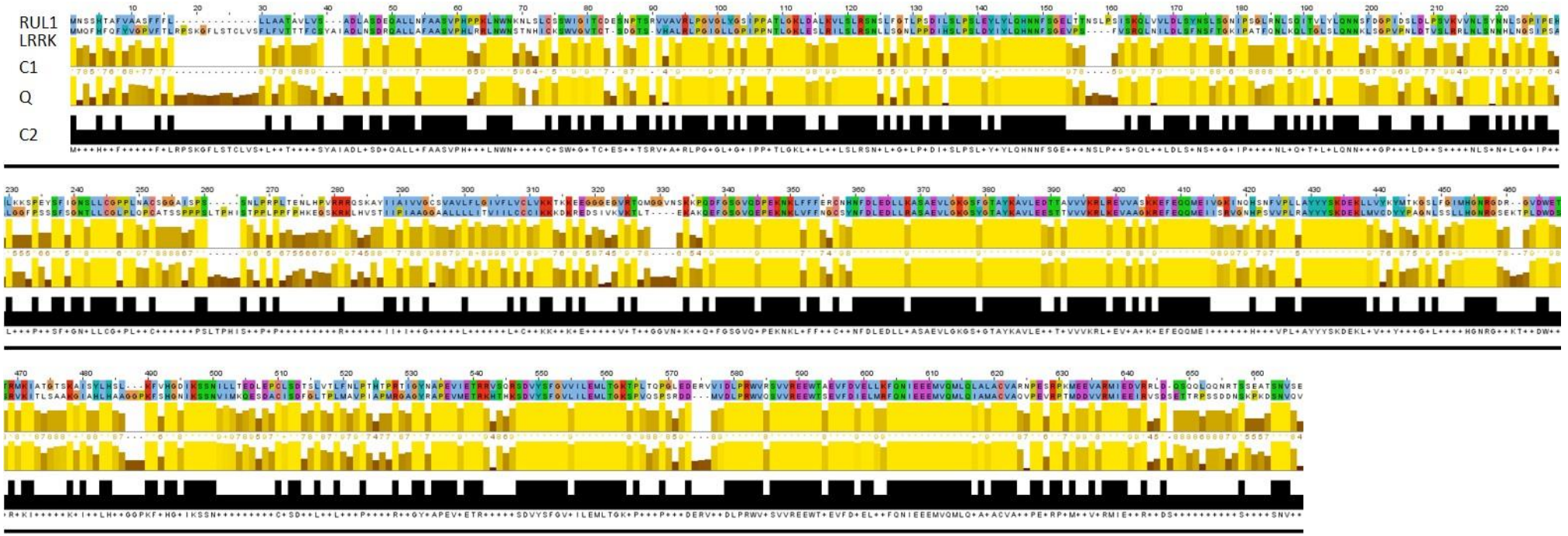


Figure 3.16: Sequence alignment between RUL1 (At5g05160) and LRRK (At5g58300). Sequences for RUL1 and the LRRK family member were aligned using ClustalX 2.1 and exported to Jalview. The colour scheme for amino acids is the default clustal colour scheme. C1=Conservation, Q=Quality, C2= consensus. The high degree of sequence similarity between these two proteins suggests that these two proteins were duplicated as part of a block duplication involving NET4A and IRQ6.

Given the relationship between IRQ1 and IRQ6 and the NET4A and NET4B proteins, is it possible that the rest of the IRQ proteins evolved from IRQ6 and IRQ1? This scenario is shown in Figure 3.17. The phylogenetic analysis and synteny plot have shown that NET4A has been duplicated as part of a block duplication to give IRQ6 (1). IRQ6 may then have given rise to the other IRQs by further duplications and modifications (2). At another time point a gene duplication in an earlier species could have given rise to the IRQ1 protein encoding gene (3).

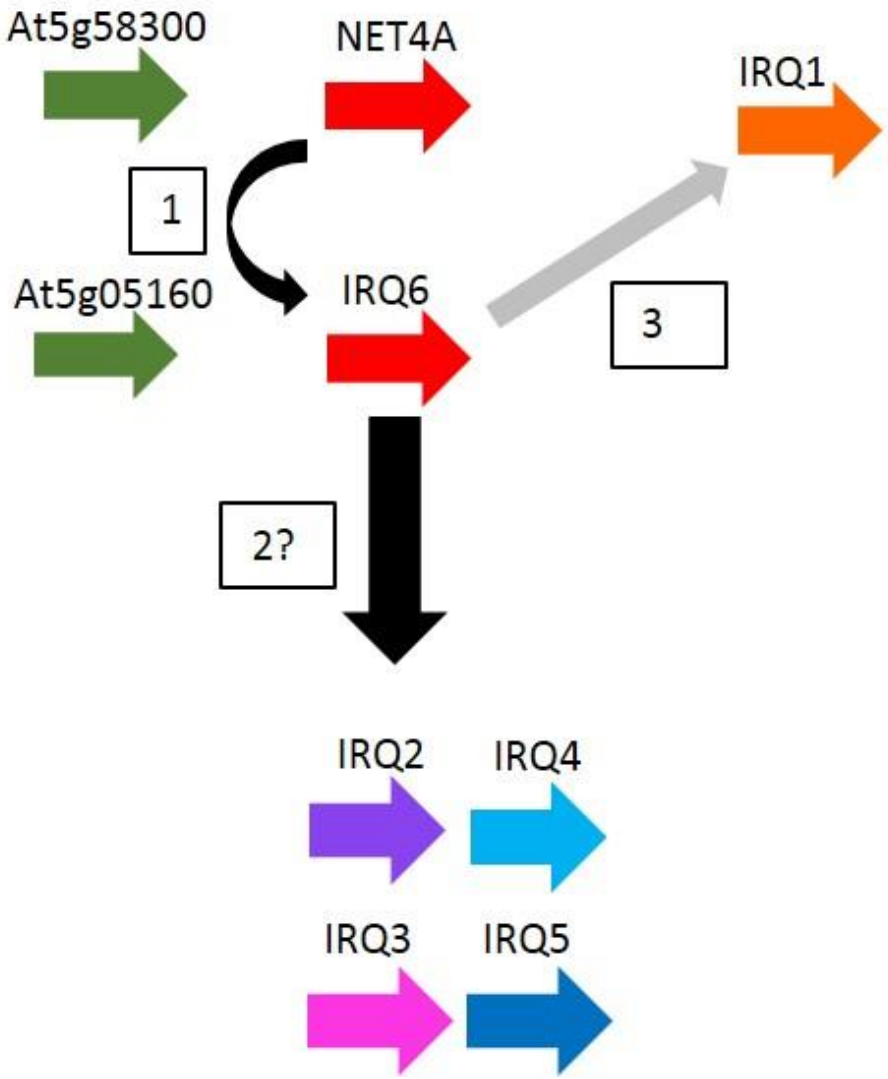


Figure 3.17: Possible scheme of the evolution of IRQ6 from NET4A. PLAZA showed that IRQ6 was part of a duplication involving NET4A. Given that there are no IRQ proteins before the early NET4 in *S. moellendorfii* IRQ6 evolved from NET4A via a chromosome duplication (1). IRQ6 could then have been the progenitor to IRQ2, IRQ3, IRQ4, and IRQ5 proteins (2). In addition, at some other point in evolutionary time during early angiosperm evolution, IRQ6 could have been duplicated to give IRQ1 (3).

Besides an IRQ6 based origin for the rest of the IRQ proteins there are two other possible scenarios for evolution of the remaining IRQs (IRQ2, IRQ3, IRQ4, and IRQ5) which must also be considered giving a total of three possibilities and these are shown in Figure 3.18 and summarised below:

1. The IRQs evolved from IRQ6 or IRQ1 (2A in Figure 3.18)

2. The remaining IRQs evolved separately from IRQ6 via one of the *NET4* genes (2B)
3. The remaining IRQs evolved separately from IRQ6 via one of the *NET3* genes (2C)

However, no further inferences can be made at this point without first examining the direction and orientation of the gene and block duplications identified earlier for IRQ2, IRQ3, IRQ4, and IRQ5 (Figure 3.15bi and ii). This is discussed before returning to the scenarios presented above.

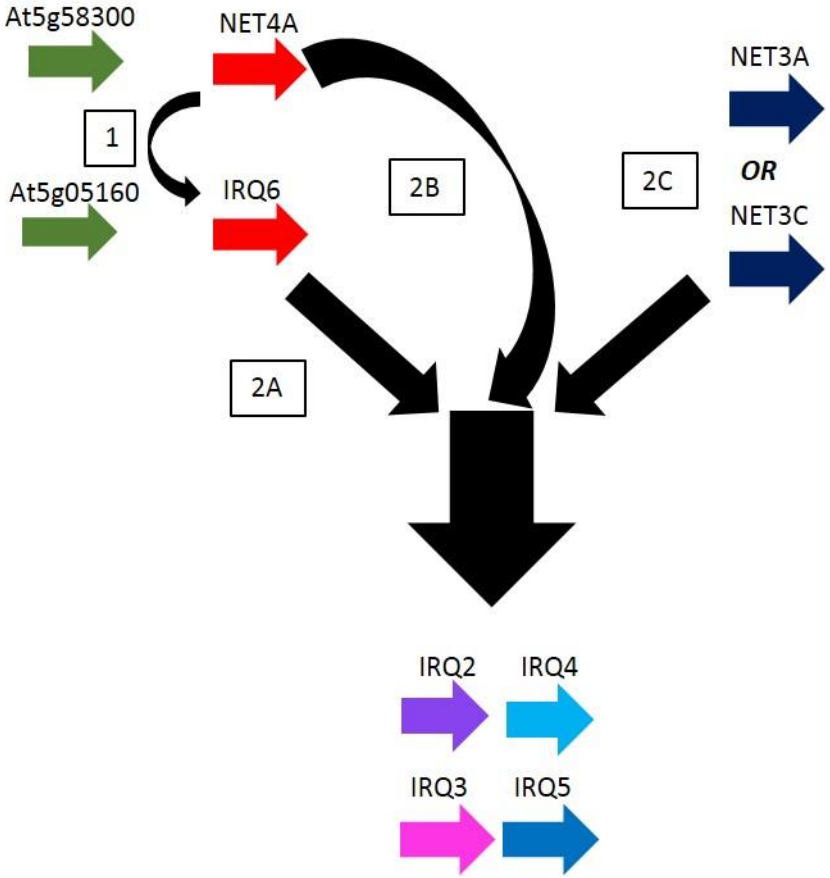


Figure 3.18: Possible origins of the IRQ2, IRQ4, IRQ3, and IRQ5 proteins. Based on the information obtained thus far there are three plausible scenarios to the evolution of the IRQ2, IRQ4, IRQ3, and IRQ5 proteins. For simplicity IRQ1 is not considered here. IRQ6 evolved from NET4A via a block duplication (1). The remaining IRQ proteins could have evolved from IRQ6 in subsequent tandem or block duplications (2A). The second scenario involves the IRQs evolving directly from the NET4A sequence (2B). The final scenario involves the evolution of the IRQ proteins from one of the NET3 proteins (2C). The scenarios are discussed in more detail in the remaining text.

For the IRQ2/4 and IRQ3/5 pairings perhaps the simplest explanation is that the tandem duplication preceded the block duplication (Figure 3.19a). The alternative scenario, where the block duplication occurred first followed by a tandem duplication of the same genes in the same orientation or in the same area in separate chromosomes independently, is very unlikely (Figure 3.19b). The order in which the tandem duplication took place i.e. a proto form of IRQ2 or IRQ3, giving rise to IRQ4 or IRQ5, or *vice versa* may be indicated by the presence of the pseudo-NAB domain found within the N-termini of IRQ4 and IRQ5. Assuming for the sake of argument that IRQ2/4 gave rise to IRQ3/5 the presence of the pNAB suggests that it was IRQ4 that was duplicated to give IRQ2. This can be assumed because it is more likely that the NAB domain in IRQ4 has come from an ancestral NET4 or NET3 and been mutated to give IRQ4 before the pNAB is lost altogether in IRQ2 (Figure 3.20a). The alternative scenario could have occurred where the ancestral NET duplicated to give IRQ2 (Figure 3.20b). This IRQ2 protein was then duplicated to give IRQ4 and the mutated form of the NAB domain was somehow regained. The scenario in Figure 3.20b is far less likely. IRQ5 also possesses a pNAB domain so a similar scenario can be envisaged giving rise to IRQ3 then IRQ2/4. Crucially, the reasoning presented here behind the duplication of IRQ4 or IRQ5 to give IRQ2 or IRQ3 and not the other way around due to the presence of the pNAB domains in IRQ4 and IRQ5 rules out scenario 2A in Figure 3.18. The IRQ2, IRQ3, IRQ4, and IRQ5 proteins must have been derived independently of IRQ6 or IRQ1 from either a NET3 or NET4 protein (scenarios 2B and 2C in Figure 3.18).

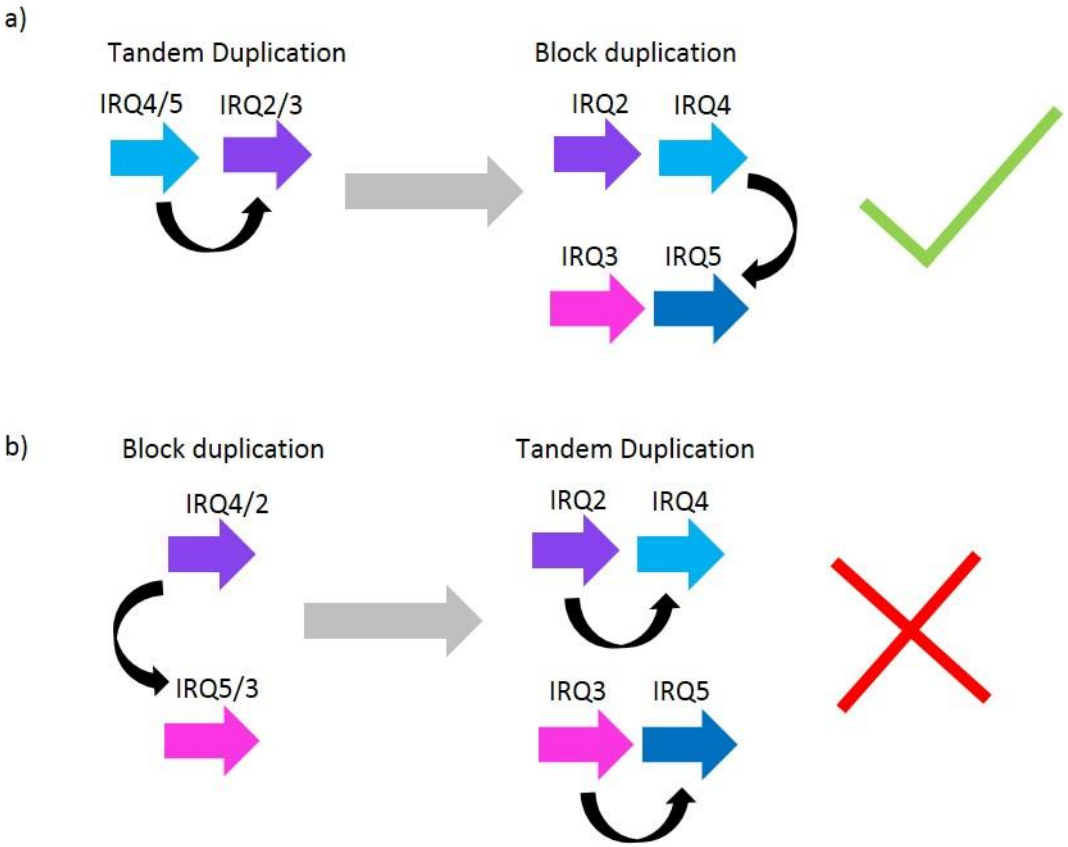


Figure 3.19: Tandem duplication followed by a block duplication is the more likely scenario for the evolution of IRQ2, IRQ3, IRQ4, and IRQ5. The IRQ2, IRQ3, IRQ4, and IRQ5 proteins evolved via a tandem and block duplication. The tandem duplication could have occurred first followed by the block duplication (a) or the block duplication followed by the tandem duplication (b). Scenario (a) is the most likely because the scenario in (b) where the genes tandem duplicate in the same orientation and in the same area in two separate chromosomes is very unlikely in comparison to that presented in (a).

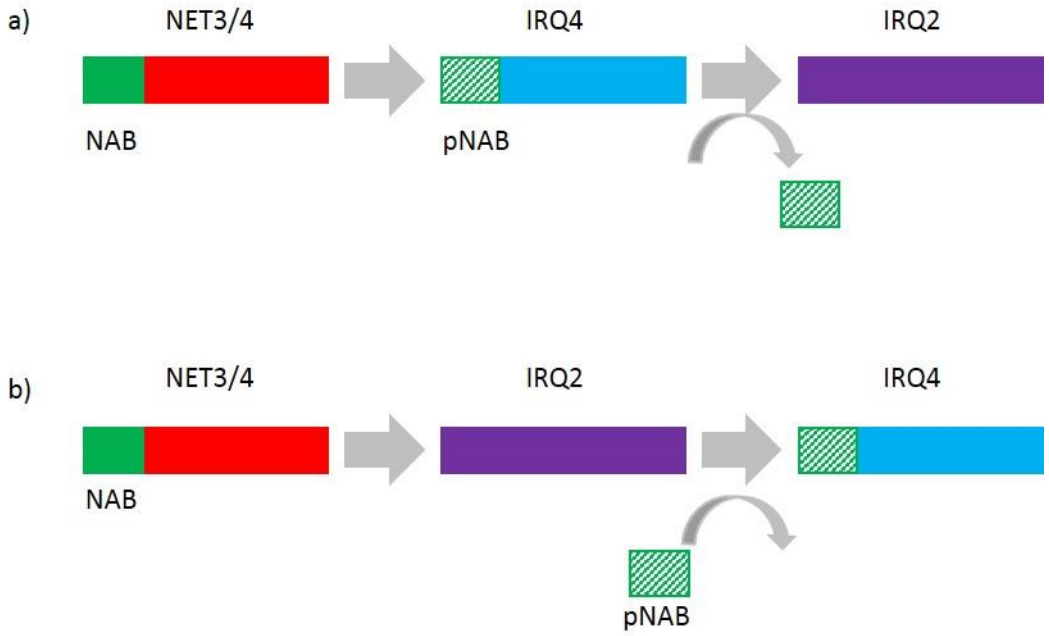


Figure 3.20: IRQ4 is more likely to have evolved before IRQ2. The order of the tandem duplication i.e. whether IRQ2 evolved from IRQ4 or *vice versa* is indicated by the presence of the pNAB domain in IRQ4. IRQ2 does not possess the domain so it follows that either iRQ4 lost the domain (a) to give IRQ2 or that IRQ2 gained the domain to give IRQ4 (b). Although it is not unusual for proteins to gain protein domains it would be highly unlikely to gain a mutated version of a domain from its ancestor (b). It is far more likely for the domain to be changed and altered over time to give the pNAB domain found in IRQ4 before being lost altogether to give IRQ2 (a). The same logic could be used to explain the evolution of IRQ3 from IRQ5.

In considering the block duplication i.e. whether IRQ2/4 gave rise to IRQ3/5 or *vice versa* this may never be known without much more sequence information. Alignment of the IRQ4 and IRQ5 full length protein sequences shows that there are two regions of similarity between them: the N-terminal end where the pNAB is situated and the C-terminus where the IRQ domains are situated (Figure 3.21). IRQ5 is much larger (1586 amino acids for IRQ5 compared to 555 amino acids for IRQ4) and this increased size could be due to an insertion within the sequence, suggesting that IRQ4 evolved first, was copied, and then IRQ5 gained the insertion (Figure 3.22a). Alternatively, IRQ5 evolved first, was copied giving the proto-IRQ4, and a large deletion occurred in this proto-IRQ4 to give the current protein (Figure 3.22b). However, without much more sequencing information from a greater number of samples pre- and post-duplication events in *A. thaliana* it will be difficult to decide which scenario is most likely.

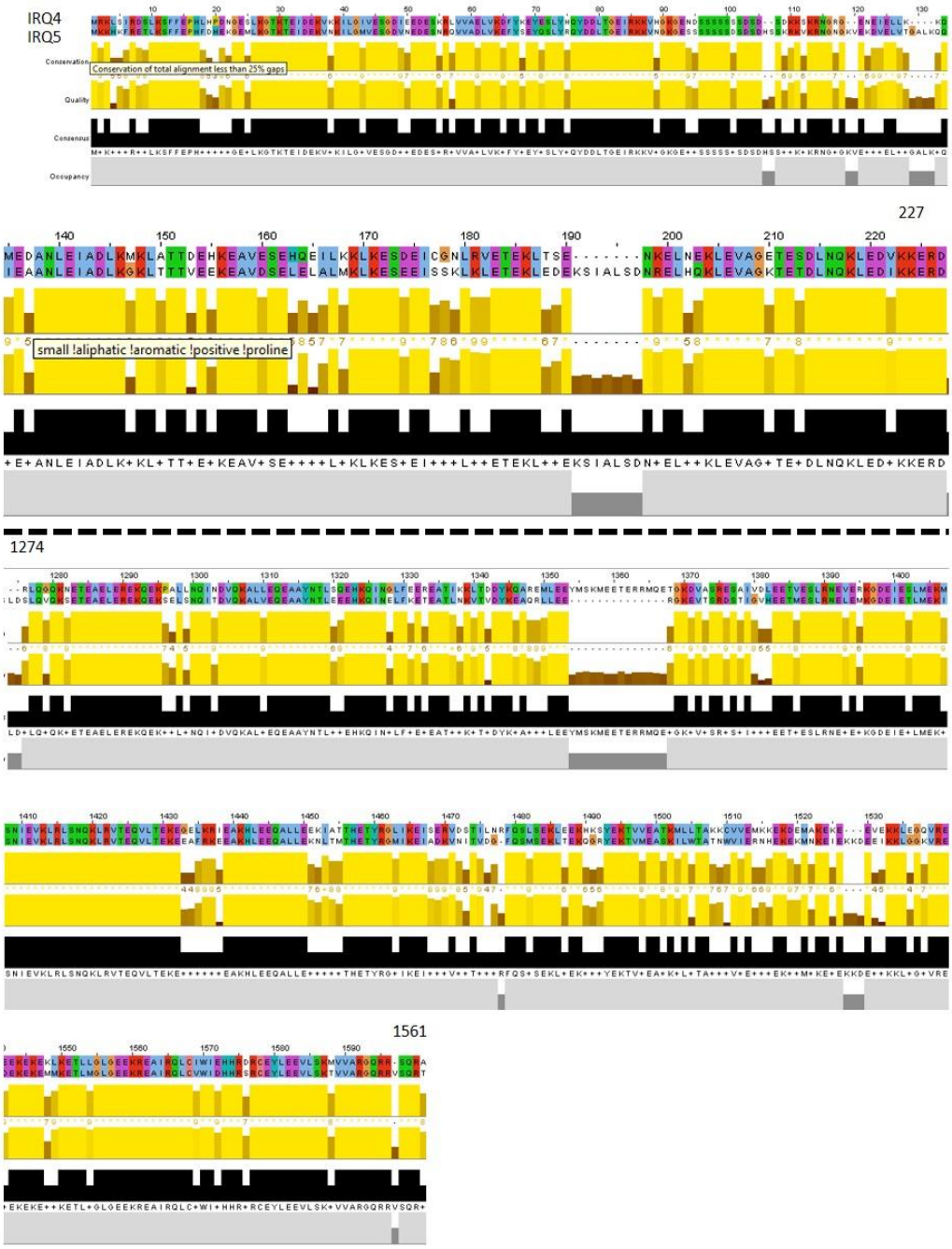


Figure 3.21: IRQ4 and IRQ5 are highly similar at the N-terminus and the C-terminus. The full length IRQ4 and IRQ5 sequences were aligned in ClustalX 2.1 and exported to Jalview for viewing. Amino acids are coloured according to the default clustal colour scheme. IRQ4 and IRQ5 share a high degree of similarity at the N-terminus where the pNAB domain is situated and the C-terminus where the IRQ domain is situated. The dotted line refers to a large section of the IRQ5 protein (between approximately 227-1274 amino acids) which is not shared by IRQ4.

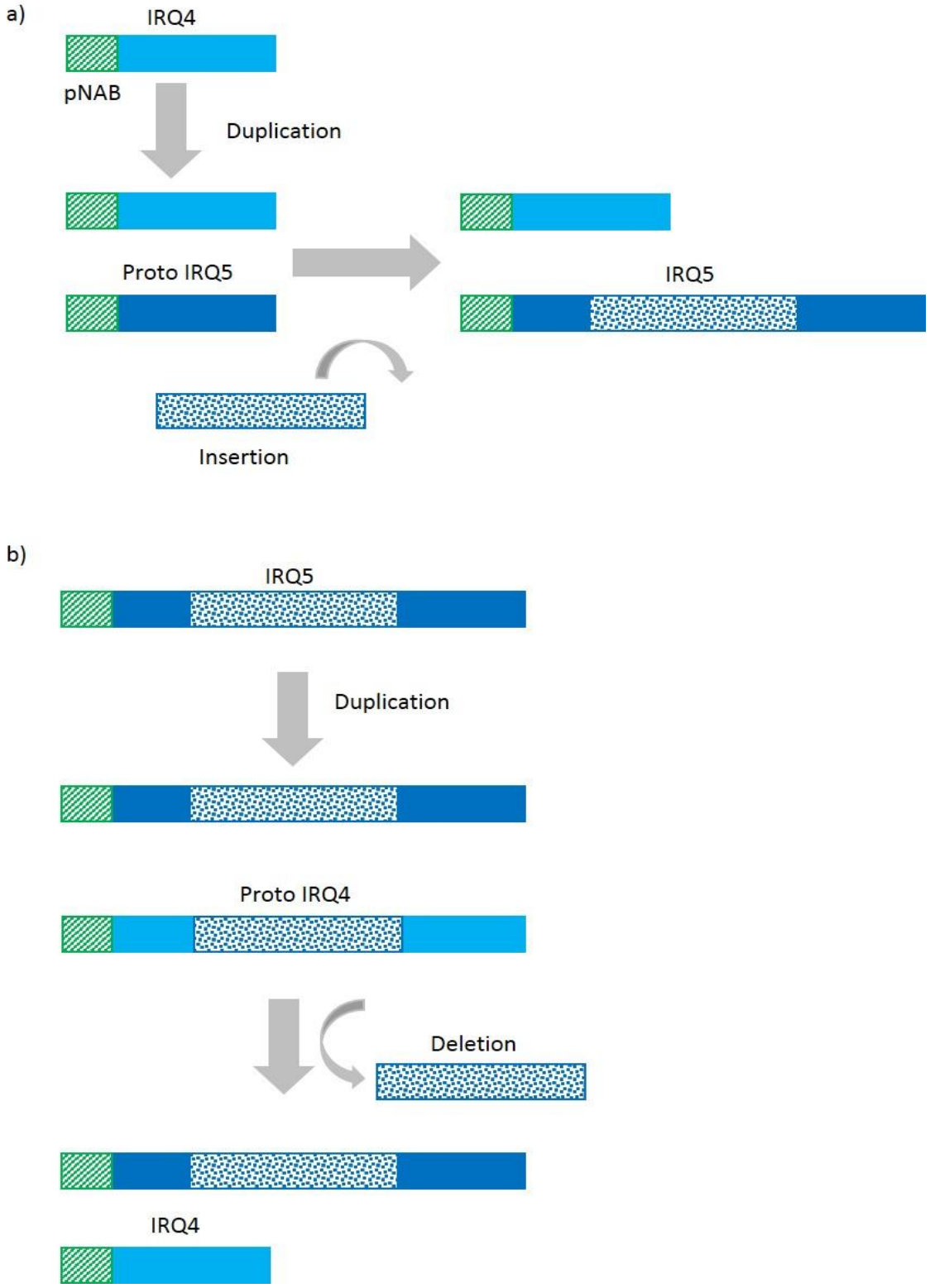


Figure 3.22: Possible scenarios for the evolution of IRQ4 and IRQ5. Whether IRQ4 gave rise to IRQ5 or *vice versa* may be dependent upon the origin of the in IRQ5. a) IRQ4 could have been duplicated giving a proto-IRQ5. This sequence could have then had a single large insertion event (blue stippling) giving the current IRQ5 protein. b) IRQ5 could have been duplicated giving a proto-IRQ4 protein. A deletion could then have removed a large section from the middle of the protein (blue stippling) to give the current version of IRQ4. More sequence information is required for this to be elucidated.

The last outstanding question to address is whether the IRQ4 and IRQ5 proteins were derived from the NET3 or NET4 proteins (scenarios 2B or 2C shown in Figure 3.18). Both NET3 and NET4 proteins possess IRQ and NAB domains. The phylogenetic analysis presented earlier using the IRQ domains alone (Figure 3.8) has shown that IRQ2, IRQ3, IRQ4, and IRQ5 have IRQ domains that are more closely related to NET3A and NET3C rather than NET4A or NET4B. This indicates that IRQ4 or IRQ5 proteins are derived from NET3A or NET3C.

In order to verify this inference from the phylogenetic data a comparison of the NAB domains of the NET3 and NET4 proteins and the pNABs of IRQ4 and IRQ5 was carried out. The first 120 amino acids of NET3A, NET3C, NET4A, NET4B, IRQ4, and IRQ5 were aligned in CLUSTALX (Figure 3.23). This alignment shows the higher degree of similarity between the NAB domains in the NET proteins compared to the pNABs in IRQ4 and IRQ5. An attempt to use this alignment to generate a consensus tree in MetaPIGA failed due to the high degree of similarity between the sequences; the software was unable to distinguish which set of NET proteins are most closely related to the pNABs in IRQ4 and IRQ5.

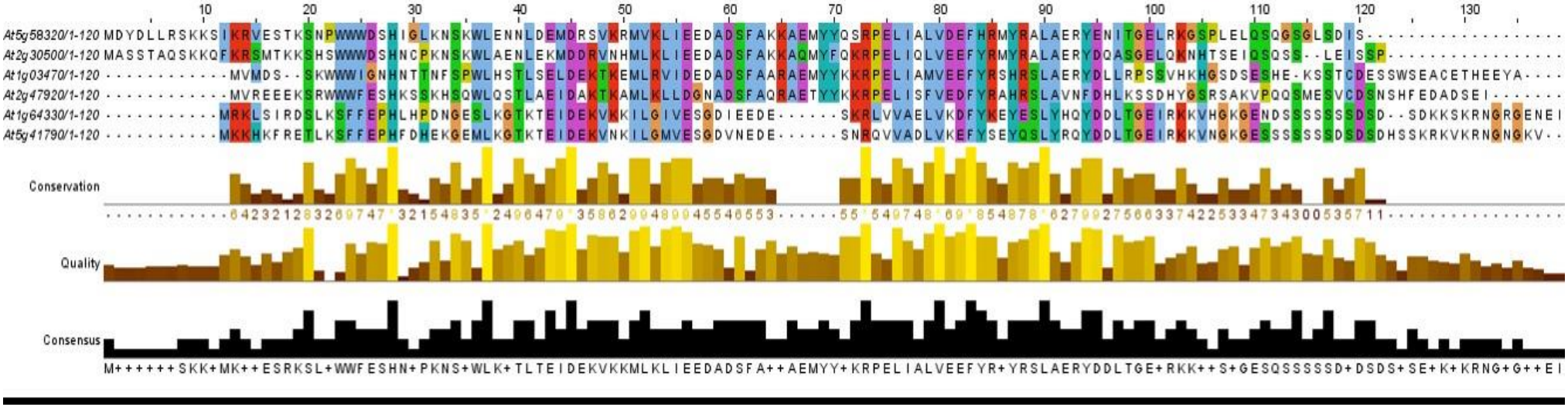


Figure 3.23: There is a high degree of similarity between the NAB and pNAB domains. The first 120 amino acids of NET3A, NET3C, NET4A, NET4B, IRQ4, and IRQ5 were aligned in ClustalX 2.1. Amino acids are coloured according to the default clustal colour scheme. There is a high degree of conservation between the NAB domains of the NET3s and NET4s and the pNABs of IRQ4 and IRQ5.

To overcome this problem pairwise comparisons were conducted using pBLAST between the pNABs of IRQ4 and IRQ5 with the NET3 and NET4 NAB domains. The percentage identities of each pairwise comparison are shown in Table 3.3. Table 3.3 shows that there is a higher percentage identity between the pNAB domains of IRQ4 and IRQ5 and the NET3 proteins compared to the NET4 proteins. The NET1A NAB domain is used as a control as it a NET protein with no IRQ domain. Overall, this agrees with the data from the phylogenetic data and it is reasonable to conclude that IRQ4 or IRQ5 were derived from either NET3A or NET3C.

Table 3.3: Pairwise comparison using pBLAST of pNAB domains in IRQ4 and IRQ5 with the NAB domains of NET3a, NET3C, NET4A, NET4B, IRQ4, IRQ5, and NET1A.

	At1g64330	At5g41790
At1g03470 NET3A	30	30
At2g47920 NET3C	30	27
At5g58320 NET4A	26	26
At2g30500 NET4B	26	26
At1g64330 IRQ4	100	73
At5g41790 IRQ5	71	100
At3g22790 NET1A	27	27

Numbers show number of positive identities between the sequences.

Drawing together the information in section 3.9 it is now possible to derive a scenario whereby the IRQ proteins diverged from the original NET proteins, and this is shown in Figure 3.24. The evolution of the IRQ proteins appears to follow that of the NET proteins where an ancestral NET4 protein arose during the evolution of land plants (Hawkins *et al.*, 2014). As gymnosperms evolved the NET3 proteins appear along with a hybrid NET1/2 protein. As no IRQ proteins were found in gymnosperms it appears that the IRQ proteins are angiosperm specific. During angiosperm evolution the first IRQ proteins appeared with IRQ6 developing from the NET4A sequence and the IRQ4 or IRQ5 evolving from the NET3A or NET3C proteins. As no IRQ domain alone containing proteins were found in in the monocot species studied here, the IRQ proteins appear to be a eudicot specific group of proteins.

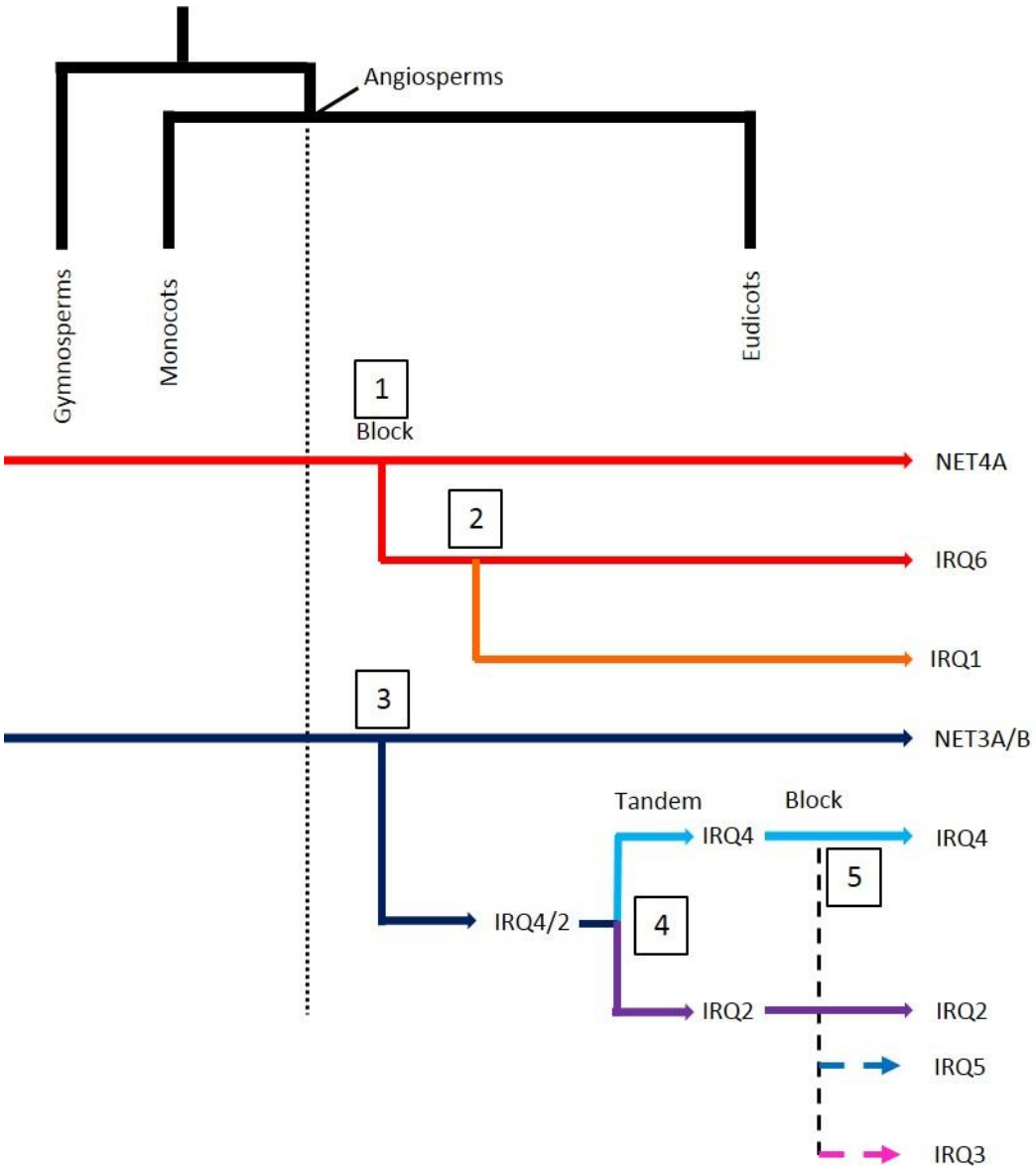


Figure 3.24: **Overview phylogenetic tree of IRQ evolution.** Phylogenetic tree showing the evolution of the IRQ group proteins from the NAB/IRQ proteins. The top of diagram shows phylogenetic tree with the major split between gymnosperms and angiosperms and the split between eudicots and monocots. Reading from left to right indicates moving forward in evolutionary time. In gymnosperms and monocots only NAB/IRQ proteins are present. Sometime after the split between monocots and eudicots a block duplication took place in NET4A to produce IRQ6 (1). At some point after IRQ6 may have been duplicated to give IRQ1 (2). Independently, NET3A or NET3C underwent some form of duplication to give rise to the IRQ4/2 hybrid (3). This hybrid underwent a tandem duplication to produce IRQ4 and IRQ2 (4). A whole section of chromosome including these genes underwent a block duplication to produce IRQ5 and IRQ3 (5). Numbers in parenthesis refer to numbers in boxes. Block above a line indicates a time point when the gene(s) underwent a block duplication. Tandem over a line indicates a time when a gene(s) underwent a tandem duplication. IRQ4 is shown as being derived from NET3A/C first to aid clarity but note that it could also have been IRQ5 instead. The dotted line indicates the evolution of angiosperms. The dashed lines associated with IRQ2, IRQ4, IRQ3, and IRQ5 are related to the block duplication.

3.10 Conclusion

This chapter has used bioinformatics tools and analyses to examine the relationship between the IRQ proteins: their secondary structure, their gene expression and subcellular localisation patterns, and their evolution in relation to the NET proteins.

A comparison of the amino acid sequences of the *Arabidopsis* IRQ domains showed that the domain consists of a central core region flanked by regions of lower similarity. The highly conserved central core region may be required for the functions of the domain while the differing flanking regions may be responsible for different aspects of the interaction e.g. binding affinity or specificity. Notably, comparison of the secondary structure of the IRQ proteins showed that IRQ1 has a truncated IRQ domain, missing the C-terminal flanking region of the domain. This may affect the affinity of this protein for its target.

A comparison of the secondary structure of the full length NAB/IRQ and IRQ proteins showed they have a very similar structure to the NET proteins: proteins with a highly conserved domain (in this case the C-terminal IRQ domain) with the rest of the protein made of coiled-coils of varying length. Coiled coils allow proteins to bind to each other, so these domains could facilitate protein-protein interactions. In addition, IRQ4 and IRQ5 possess a domain with limited similarity to the NAB domain termed the pNAB domain.

Initial cladistics analysis of the full length NAB/IRQ and IRQ proteins showed that they can be split into five groups: 1. NET3, 2. NET4, 3. IRQ1 and IRQ6, 4. IRQ4 and IRQ5, and 5. IRQ2 and IRQ3. Analysis of their gene expression and subcellular localisation patterns using Affymetrix data from the *Arabidopsis* eFP browser and various proteomic and prediction algorithms in SUBA3 generally confirmed these groupings to be correct. IRQ1 and IRQ6 were shown to most strongly localise to meristematic tissues of the shoot and root. IRQ2 and IRQ3 appear to be pollen specific with very low expression seen in any other tissues. IRQ4 and IRQ5 are relatively ubiquitously expressed with IRQ4 showing a greater level of expression in the below ground parts of the plant relative to IRQ5. Subcellular localisation data using both prediction algorithms and subcellular proteomics experiments suggests that both IRQ4 and IRQ5 are localised to the vacuole. Overall, the expression patterns of the IRQ group proteins shows they are expressed across nearly all life stages and tissues in *Arabidopsis thaliana*.

The relationship between NET, NAB/IRQ and IRQ proteins was established using phylogenetic analysis. The IRQ domains of the NAB/IRQ and IRQ proteins were used to interrogate available sequence databases for organisms at important stages of plant evolution. The sequences obtained were aligned and analysed using the MetaPIGA software to create phylogenetic trees.

Full length sequences for all those proteins obtained from the initial screen were also compared to understand if any sequences outside the IRQ domain had a major effect on the relationship between these proteins.

A major finding was that the IRQ proteins evolved from the NET proteins as no other proteins containing an IRQ domain were found earlier in evolutionary time than the NET4 sequence in *S. moellendorfii*. Comparison of all the IRQ domains alone across different species revealed that the IRQ6 and IRQ1 sequences were more similar to the NET4 sequences while IRQ2, IRQ3, IRQ4, and IRQ5 had IRQ domains more similar to the NET3 sequences. This experiment indicated that the IRQ1 and IRQ6 sequences diverged among the IRQ proteins. No IRQ sequences were found in gymnosperms or monocots that did not belong to proteins also containing NAB domains indicating that not only are IRQ proteins angiosperm specific but they are also eudicot specific. Only three monocot species were studied for this analysis so a wider range of monocot species datasets would need to be analysed before making a definitive conclusion with regards to eudicot specificity. A comparison of the full length sequences found in this analysis revealed the higher degree of similarity between the IRQ proteins and the NAB/IRQ proteins,

An investigation of three additional species of eudicots from outside the Brassicaceae including *Glycine max*, *Vitis vinifera*, and *Solanum lycopersicum* showed that the higher degree of similarity between the IRQ1 and IRQ6 IRQ domains and the NET4 IRQ domains holds true in these diverse eudicot species. In addition the existence of multiple IRQ isoforms in these species indicates that the division between the IRQ proteins has occurred early on in angiosperm evolution. To further show the multiplicity of IRQ proteins across the eudicots additional datasets of eudicots need to be studied.

To further understand the gene duplication events which may have led to the diversity of IRQ sequences and to develop a model for the evolution of the IRQ group proteins analysis of collinear clusters across chromosomes in *Arabidopsis* was carried out. A combination of tandem and block duplications occurred to give IRQ2, IRQ4, IRQ3, and IRQ5 while a block duplication linked IRQ6 with NET4A.

Using all the above phylogenetic information generated in this chapter and gene duplication events, a model was developed for IRQ evolution from the NET proteins. In this model IRQ6 arose from a gene duplication event in NET4A. IRQ1 may have arisen from a previous gene duplication event prior to the evolution of *A. thaliana*. At some point during angiosperm evolution, prior to the split between the eudicots and the monocots, NET3A or NET3C gave rise to IRQ4 or IRQ5. IRQ4 or IRQ5 then underwent a tandem duplication followed by a block duplication and these proteins further diversified to give the pollen specific IRQ2 and IRQ3. The exact identities of these

events e.g. whether NET3A or NET3C gave rise to IRQ4 or IRQ5, nor the order of these events e.g. whether IRQ4 and IRQ2 were tandem block duplicated to give IRQ5 and IRQ3 or *vice versa*, cannot be known until a much wider variety of angiosperm species are sequenced and available for interrogation.

Overall this chapter has shown the IRQ proteins as a group evolved relatively recently from the NET proteins and a group are expressed in nearly all tissues and life stages of *A. thaliana*. Although tissue and subcellular localisation algorithms and software are useful indicators of expression the exact localisation of expression can only be verified using experimental techniques. The next chapter takes four of the IRQ proteins (IRQ1, IRQ6, IRQ4, and IRQ5) and uses translational fusions to pinpoint their localisations at the tissue level.

Chapter 4 *In situ* transcriptional localisation analysis of the *IRQ1, IRQ6, IRQ4,* and *IRQ5* genes in *Arabidopsis thaliana*

4.1 Introduction

The previous chapter has demonstrated the relationship between the IRQ, NAB/IRQ, and NET proteins. It also showed Affymetrix data predicting the localisation of expression of the IRQ proteins. It is important that the data obtained from high-throughput microarray experiments is verified on a gene-by-gene basis *in planta*. This chapter will examine the localisation of transcription of the IRQ1, IRQ6, IRQ4, and IRQ5 protein containing genes using the GUS reporter expression system in *Arabidopsis thaliana* plants.

The GUS reporter system uses the GUS coding sequence fused to the promoter of the gene of interest. Activation of transcription creates the GUS gene product, the *Escherichia coli* enzyme β glucoronidase, which catalyses the cleavage of β glucuronides. When X-Gluc (5-bromo-4-chloro-3-indolyl-beta- D-glucuronic acid) is infiltrated into the tissues the β glucoronidase uses this as a substrate, cleaving it to produce a colourless glucuronic acid and a chloro-bromoindigo giving a blue precipitate. As the β glucoronidase is only produced when the gene specific promoter is activated, the blue colouration can be used as a marker for the localisation of transcription (Jefferson *et al.*, 1987).

The initial generation of *IRQ4* and *IRQ5* stably transformed *Arabidopsis thaliana* lines (T_0 plants) were produced previously by Dr. D. Mentlak and Dr T.J. Hawkins. The construct used for *IRQ6* was also produced previously by Dr. D. Mentlak. The construct used for *IRQ1* was produced by Dr. J.T.M. Kroon. All GUS constructs used in this thesis were produced by amplifying 2 kb of promoter sequence upstream of the ATG start codon for each gene using primers listed in appendix 1. This was cloned into the pDONR207 entry vector according to section 2.2.3. For *IRQ4, IRQ5,* and *IRQ6* this was then cloned into the Gateway modified GUS expression vector pBI101G (Jefferson *et al.*, 1987). *IRQ1* was cloned into the commercially available Gateway compatible pBGWFS7 vector which, in addition to GUS, is also fused to GFP allowing fluorescence detection of transcription. A major advantage of the pBGWFS7 vector is it allows monitoring of transcription in live cells and as the GFP product is too large (30 kDa) to be passively transported between cells reduces the problem of the chloro-bromoindigo leaking into non-expressing cells producing false positive results. All the transformed vectors were then transformed into wild type *Arabidopsis thaliana* plants using the floral dip method according to section 2.2.15. T_1 plants were selected on kanamycin (*IRQ4*, *IRQ5*, and *IRQ6*) containing media or by spraying using BASTA (*IRQ1*) as a

selection agent. Positively selected plants were tested for the presence of the GUS containing vector by PCR according to section 2.2.1 and those successful at this step were allowed to self-pollinate to produce the T_2 generation. These T_2 plants were sown and screened using PCR and allowed to self again to increase homozygosity. The T_3 seed produced was then tested for GUS activity according to section 2.5.6 producing 10 lines for *IRQ4* and 10 lines for *IRQ5*. For *IRQ1* the seed used in this thesis was in the T_2 generation and that of *IRQ6* in the T_1 generation. As such, results for *IRQ1* and *IRQ6* should be regarded as preliminary. All lines used had been checked for the presence of their respective *prom*GUS containing vector by PCR. Up to 30 lines were screened for *IRQ1* and 9 lines were screened for *IRQ6*. Expression of transcription was investigated at the following life stages: 5-7 day old seedling, 2 week old seedling, 5-6 weeks old plants, and just germinated seedlings. Appendix 3 describes the expression pattern for all lines tested.

4.1.1Expression pattern of IRQ4

4.1.1.1 Expression pattern in 5-7 day old seedlings of IRQ4 promGUS plants

Seedlings of *IRQ4prom*::GUS were incubated in GUS buffer, vacuum infiltrated for approximately 2-5 minutes at room temperature, then left at room temperature for 15-30 minutes before being washed in an ethanol series to clear the tissues and imaged (according to section 2.5.6).

Expression was evident at the root tip in the epidermis and cortex of the primary root for all lines tested (Figure 4.1). Expression of *IRQ4* weakened moving from the root tip towards the hypocotyl (area between two arrows in Figure 4.1b). In lateral roots, expression occurred at the root cap, then there was a large decrease in expression seen approximately in the meristematic zone, before expression was evident again further up the lateral root (Figure 4.1d). In contrast patchy expression was seen in the epidermis of the main root in the meristematic zone before increasing again in the elongation and differentiation zones of the primary root (Figure 4.1e). No expression was seen in the root cap or below the quiescent centre of the primary root. Whether the expression pattern in the primary root is limited to the epidermis or whether expression is seen in other cell files would require cross sectioning of stained samples. The differences in expression in the meristematic zone of the lateral and primary roots could be due to the potentially differing way in which these two tissues are patterned (Birnbaum, 2016). Additionally, lines 1, 2, and 4 had lateral root cap staining while lines 7 and 9 had staining in the root vasculature. No expression of *IRQ4* was seen in the cotyledons or hypocotyl of seedlings with expression restricted to the roots.

4.1.1.2 Expression pattern in 14 day old seedlings of IRQ4 promGUS plants

Whole seedlings were immersed in GUS buffer and vacuum infiltrated for 5 minutes before being incubated at 37 ^{o}C for 30-60 minutes before being washed in an ethanol series to clear the tissues and imaged (according to section 2.5.6). In all lines tested expression was restricted to the root with the primary root seeming to show expression at the root rip in the cortex, epidermis, and the root vasculature (Figure 4.2d). In lateral roots expression was seen in the root tip from the root cap, into the meristematic zone and into the differentiation and elongation zones before being restricted to the vasculature further up the root (Figure 4.2b). In all the lines tested, the absence of staining that had occurred in the meristematic zone of the lateral roots at 5-7 days had changed in some lateral roots with staining now becoming evident in this region. This could be due to the longer incubation time used for this stage or it could be due to a change in the developmental programme of the lateral roots at this later stage in their development. Of the five lines tested there was again no expression seen in the aerial tissues including cotyledons, true leaves, hypocotyls or trichomes.

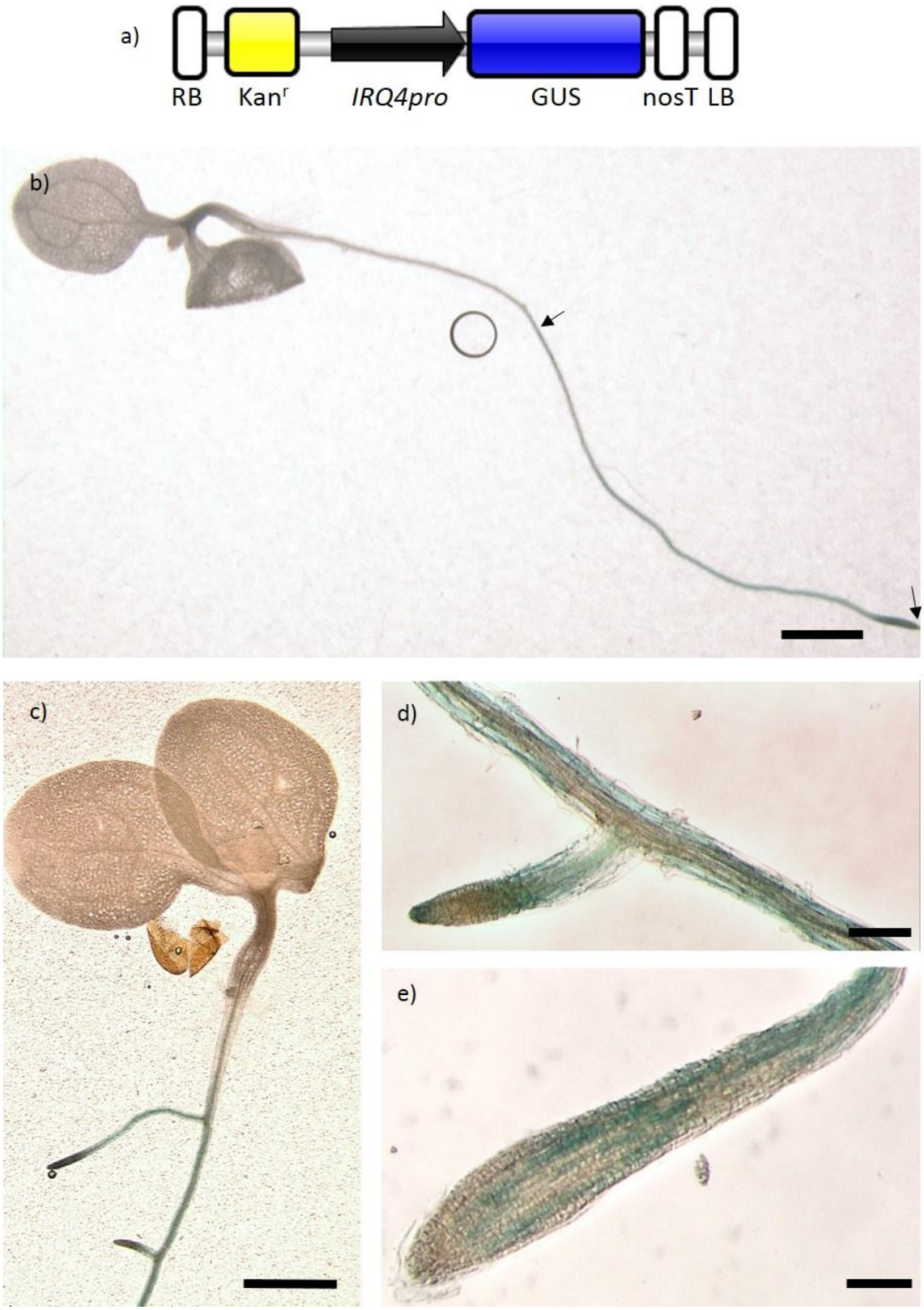


Figure 4.1: Expression profile of *IRQ4* in 7-day old seedlings. a) *IRQ4prom*::GUS construct created using the pBI101G expression vector. b-e) Whole 7-day old seedlings were histochemically stained for GUS activity, washed repeatedly in 70 % ethanol and mounted in 20 % glycerol for imaging. b) Whole seedling showing expression between the two arrows shown in the image. c) Image showing lack of staining in cotyledons but presence of stain in lateral roots. d) Close-up of lateral root showing absence of stain in the meristematic zone above the root tip. e) Close-up of primary root showing absence of staining in the root cap and quiescent centre with staining in the meristematic zone and into the elongation zone of the root. Scale bars: b-c) 1000 µm, d-e) 100 µm.

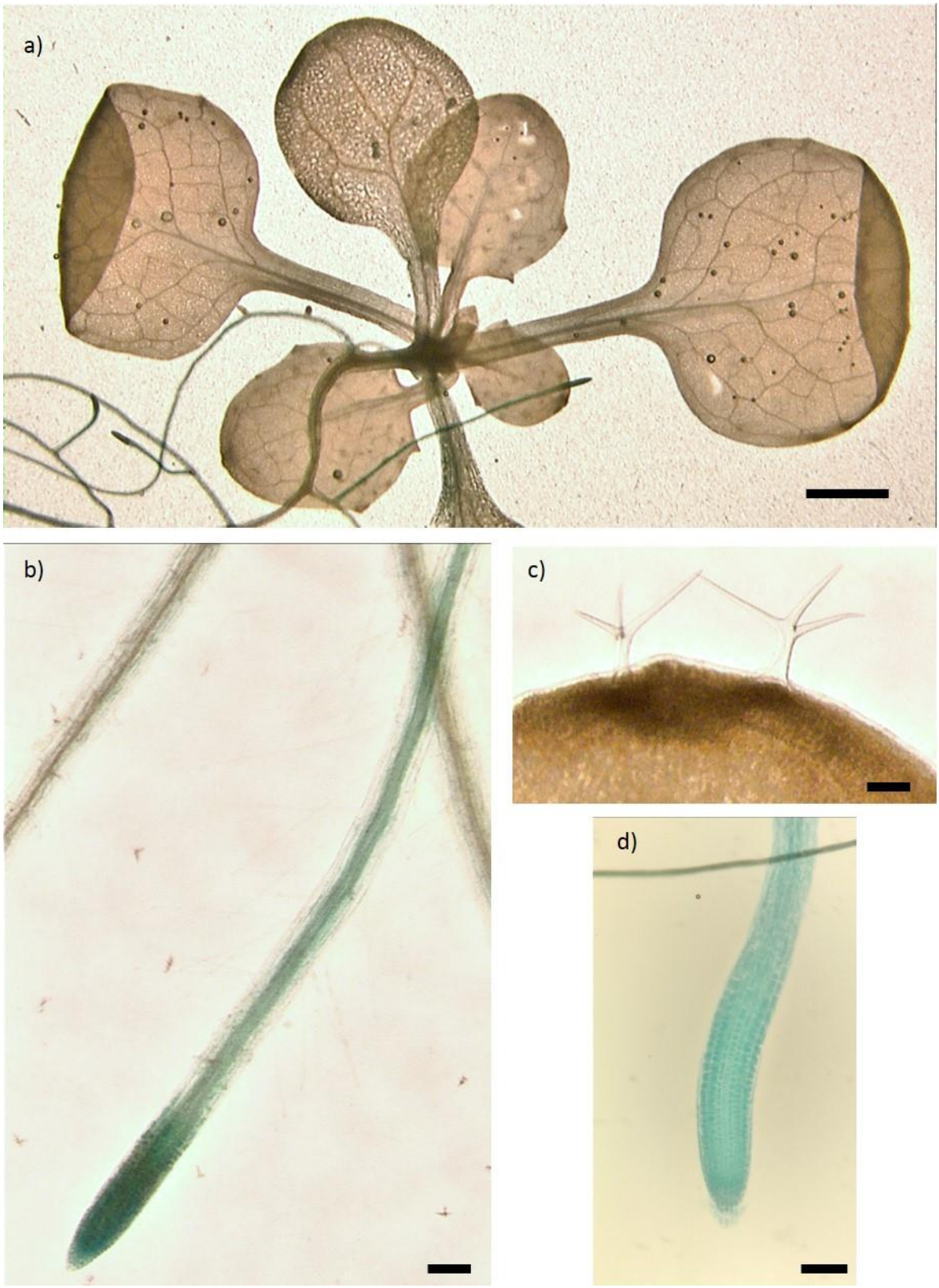


Figure 4.2: Expression profile of *IRQ4* in 14-day old seedlings. Whole 14-day old seedlings were histochemically stained for GUS activity, washed repeatedly in 70 % ethanol and mounted in 20 % glycerol for imaging. a) Histochemical staining revealed no expression in the cotyledons at 14 days or in any of the younger leaves. b) Image of a lateral root showing GUS staining throughout the meristematic zone and into the elongation zone with prominent staining of the vasculature. c) Close-up of the trichomes on a young leaf showing no expression at this stage. e) Close-up of primary root showing staining throughout the root similar to b). Scale bars: a) 1000 μm, b-d) 100 μm.

4.1.1.3 Expression pattern in 5-6 week old IRQ4 promGUS plants

Whole inflorescences, cauline leaves, and developing and older siliques were immersed in GUS buffer and vacuum infiltrated for 5 minutes before being incubated at 37 ^{o}C for 60 minutes before being washed in an ethanol series to clear the tissues and imaged (according to section 2.5.6). All five lines tested showed expression at the receptacle at the base of the silique (Figure 4.3b) and the stigma (Figure 4.3c). Only line 3 showed expression in the vasculature of the petals. Two of the lines tested (lines 1 and 6) showed staining in the trichomes of the more mature leaves. Trichomes are single cell extensions arising from the epidermal surface of the leaf (Hulskamp *et al.*, 1994). They function as in both storage and defence capacities, e.g. storage of anthocyanins and defence compounds (Wagner *et al.*, 2004). For example the sequestration of cadmium and formation of calcium crystals in secretory trichomes (Choi *et al.*, 2001). As *IRQ4* has been identified in the vacuolar proteome (discussed in section 3.6) the expression of *IRQ4* could reflect a role of *IRQ4* in a storage associated function in the trichome.

4.1.1.4 Expression pattern in just germinated seedlings of IRQ4 promGUS plants

To look at the potential expression pattern in very young plants seeds were grown on 1xMS plates as per section 2.5.6 but 1-2 days after being vernalised were taken from the plates, vacuum infiltrated for 1-2 minutes then incubated at room temperature for 15-30 minutes before being washed in an ethanol series to clear the tissues and imaged (according to section 2.5.6). Problems with germination meant that only two lines were tested.

Expression in the primary root showed a similar pattern to that of lateral roots in 5-7 day old seedlings: expression was evident near the root cap, a decrease in expression was seen in the meristematic zone then an increase in the differentiation and elongation zones (Figure 4.4a and b). However, no expression was seen in the very root tip as seen in 5-7 day old plants (compare Figures 4.4d with 4.4a and b). Again no expression was seen in the cotyledons or hypocotyls.

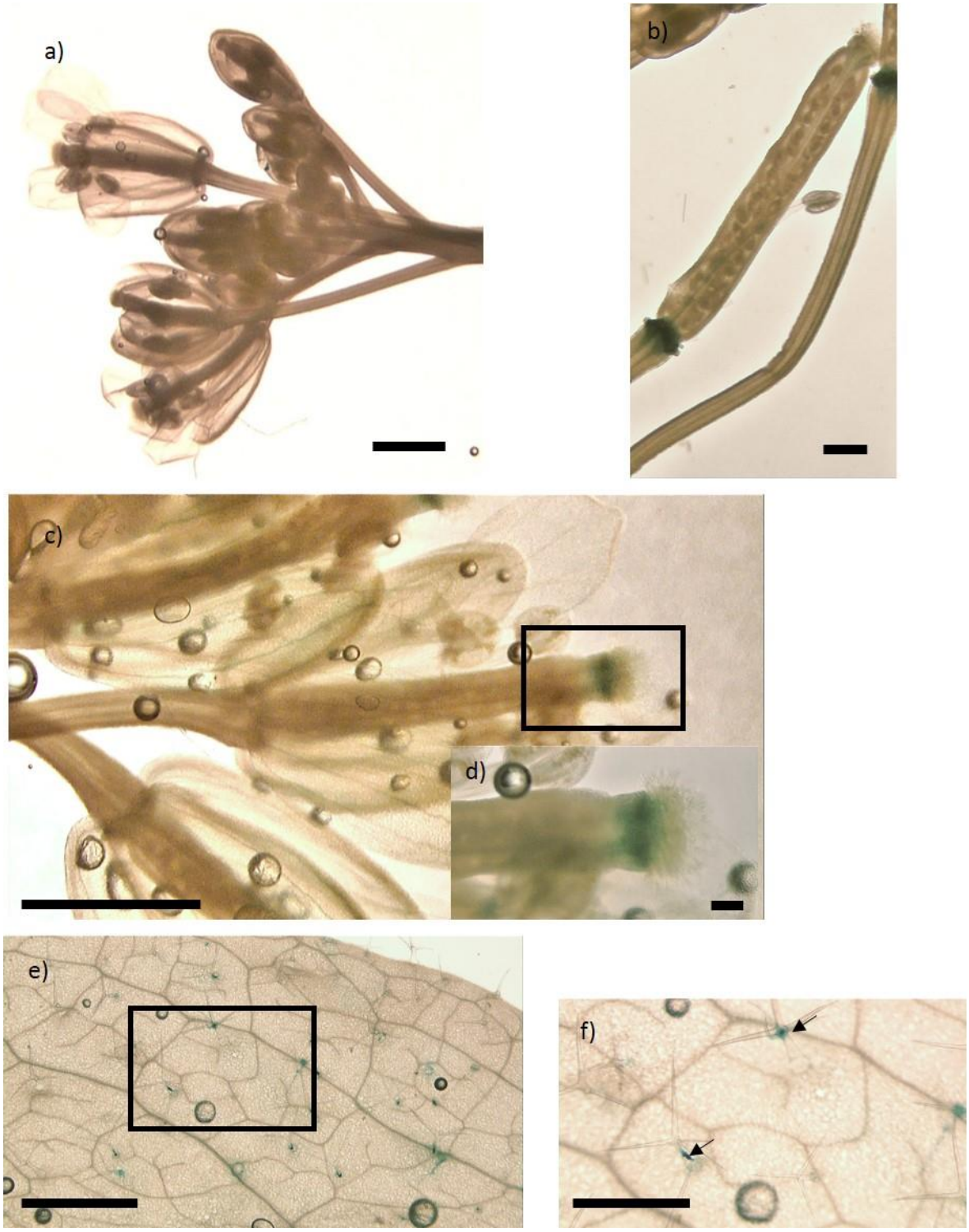


Figure 4.3: Expression profile of *IRQ4* in 5-6 week old plants. Leaves, flowers and siliques of 5-6 week old plants were histochemically stained for GUS activity, washed repeatedly in 70 % ethanol and mounted in 20 % glycerol for imaging. a) Whole inflorescence showing no staining. b) Developing silique only showing staining at the receptacle at the base of the silique. c) Close up of flower (c) showing expression at the stigma, shown in more detail in the inset of boxed area (d). e) Staining of trichomes in mature rosette leaf, shown in more detail in f, (boxed area in (e), arrows indicate stained trichomes). Scale bars: a-c) 1000 μm, d) 100 μm, e) 1000 μm, f) 500 μm.

Figure 4.4: Expression profile of *IRQ4* in seedlings one day after germination. Seedlings were histochemically stained for GUS activity, washed repeatedly in 70 % ethanol and mounted in 20 % glycerol for imaging. a-b) Whole seedlings showed expression at the root tip, a decrease in expression between the two arrows in a) and b) then a slight increase before showing weak staining throughout the rest of the root. No expression was seen in the cotyledons. Scale bars: a-b) 250 µm.

4.1.1.5 Summary of the expression of IRQ4

Analysis of the expression patterns of *IRQ4* show the transcription of *IRQ4* to be generally localised to the root, with the cortex and epidermis being most prevalent at both the 5-7 and 14 day old stages. In addition, at 14 days staining was evident in wider tissues of the root including the root cap and vasculature. This could reflect a differing role for IRQ4 in the root depending on the growth stage of the plant.

4.1.2 Expression pattern of IRQ5

4.1.2.1 Expression pattern in 5-7 day old seedlings of IRQ5 promGUS plants

Previous work by Ren *et al.* (2016) using their *proCIP1*:GUS construct had shown that *IRQ5* was expressed throughout the hypocotyl, cotyledons, cotyledon vasculature, rosette leaf vasculature and throughout the leaves, and petal vasculature. Their construct contained only 1549 bp of the upstream sequence and was cloned into pCAMBIA1301. In contrast the IRQ *prom*GUS construct used here consisted of 2kb of upstream promoter sequence and was cloned into a different GUS vector (pBI101G). It would be of interest to see if the results produced by Ren *et al.* (2016) could be replicated using the more extensive promoter region cloned here as additional promoter elements important for the localisation of *IRQ5* may be missing when a shorter promoter sequence is used.

Initially, seedlings of all lines of *IRQ5prom*::GUS were incubated in GUS buffer then vacuum infiltrated for approximately 2-5 minutes at room temperature, then incubated at 37^0C for 15 minutes before being washed in an ethanol series to clear the tissues and imaged (according to section 2.5.6). However, Ren *et al.* (2016) used a different method of histochemical staining including much longer incubation times of overnight at 37 ^{o}C. Three of the lines initially tested (lines 1, 4, and 5) were taken and the incubation times in GUS buffer were extended to 1-2 hours at 37 ^{o}C for each stage described (5-7 day old, 2 week, 5-6 week, just germinated seedling). The extended incubation time was to see if longer incubation times with the GUS buffer and conditions used here would make any difference to the result. A mixture of results is presented below. Lines 3 and 8 failed to express at 5-7 days.

Expression across all 8 lines was strongest within the root. Within the primary root, most lines showed expression in the cortex, epidermis, and root vasculature (Figure 4.5d-g). At the primary root tip, expression was absent from the root cap and the cells approximately below the quiescent centre. Expression was stronger at the bottom or the meristematic zone before weakening again through the meristematic zone (area between the arrows in Figure 4.5g) before increasing again in the differentiation and elongation zones. Expression was absent in emerging lateral roots (Figure 4.5f) but more developed lateral roots had an expression pattern similar to 5-7 day old seedlings of *IRQ4prom*::GUS expression at the lateral root tip, absence of expression in the meristematic zone, before expression was seen in the differentiation and elongation zones (compare Figure 4.1d to Figure 4.5f).

Expression was relatively high at the boundary area between the hypocotyl and primary root with a relatively sharp cessation of expression moving from the root into the hypocotyl (Figure 4.5d and e). This indicates that there appears to be some form of barrier to expression in the hypocotyl but not in the leaf tissues at this developmental stage. Lines 1, 4, and 5 showed expression possibly corresponding to stipule areas near the SAM region (Figure 4.5d and e). Weak staining was seen in the non-vasculature of cotyledons in lines 6 and 9 (Figure 4.5b). Expression was absent in the vasculature of all 8 lines tested, including lines 1, 4, and 5 treated with longer incubation times.

The absence of strong expression in the aerial parts of the plant is interesting given the results of Ren *et al*. (2016) which showed stronger expression in these tissues. This could reflect the difference in promoter regions selected for this study compared to Ren *et al.* (2016) or could be a reflection of the much longer incubation times used by Ren *et al.* (2016) compared to that used here.

4.1.2.2 Expression pattern in 14 day old seedlings of IRQ5 promGUS plants

Whole seedlings were immersed in GUS buffer and vacuum infiltrated for 5 minutes before being incubated at 37 °C for 30-60 minutes before being washed in an ethanol series to clear the tissues and imaged (according to section 2.5.6). In the repeat experiment, lines 1, 4, and 5 were vacuum infiltrated for 5 minutes then incubated in GUS buffer at 37 °C for 1-2 hours before imaging.

The majority of lines showed expression throughout the primary root (Figure 4.6b and d) with lines 1, 2, 5 and 9 showing expression in the primary root vasculature. In the lateral roots a similar expression pattern to 5-7 day old seedlings of *IRQ5* occurred: expression at the lateral root tip, absence of expression in the meristematic zone, before expression was seen in the differentiation zone (Figure 4.6c).

Expression in aerial parts was limited to lines 1, 5, and 9. Weak expression was seen in the cotyledons of lines 3 and 4 (Figure 4.6a). *IRQ5* was expressed in trichomes of true leaves in lines 1 and 5 (Figure 4.6e), showing expression at an earlier stage than *IRQ4* (5-6 week old mature leaves in *IRQ4* Figure 4.3) compared to 2- week old plants for *IRQ5*.

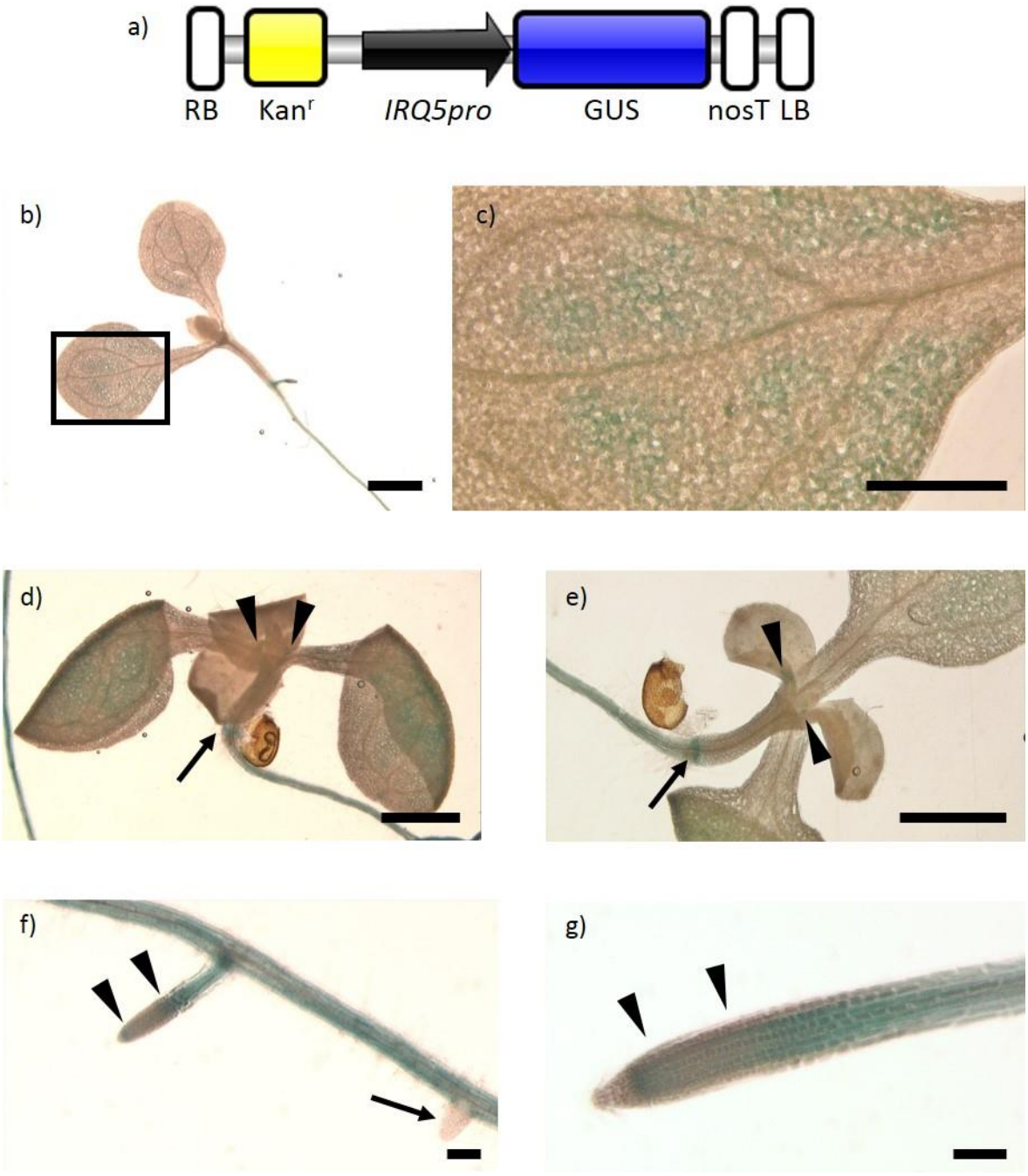


Figure 4.5: Expression profile of *IRQ5* in 7-day old seedlings. a) *IRQ5prom*::GUS construct created using the pBI101G expression vector. b-g) Whole 7-day old seedlings were histochemically stained for GUS activity, washed repeatedly in 70 % ethanol and mounted in 20 % glycerol for imaging. b) Cotyledons showed a diffuse staining indicating expression in this tissue which can be seen more clearly in c) which is a magnification of the box in b). d-e) Image showing expression in cotyledons. Expression begins at the boundary between the hypocotyl and the root and that IRQ5 is only being expressed within the root and the hypocotyl ((arrows in d) and e)). Expression can also be seen at the boundary between the hypocotyl and primary root (arrow heads in c) and d)). f) Lateral root showing absence of stain in the meristematic zone above the root tip (area between arrowheads). Also note the complete absence of expression in the newly emerging lateral root (arrow). g) Primary root showing absence of staining in the root cap and a reduction in expression in part of the meristematic zone (area between two arrowheads). Expression can be seen in the more mature parts of the meristematic zone and into the elongation zone of the root. Scale bars: bi) 1000 µm, bii) 500 µm, c-d) 1000 µm, e-f) 100 µm.

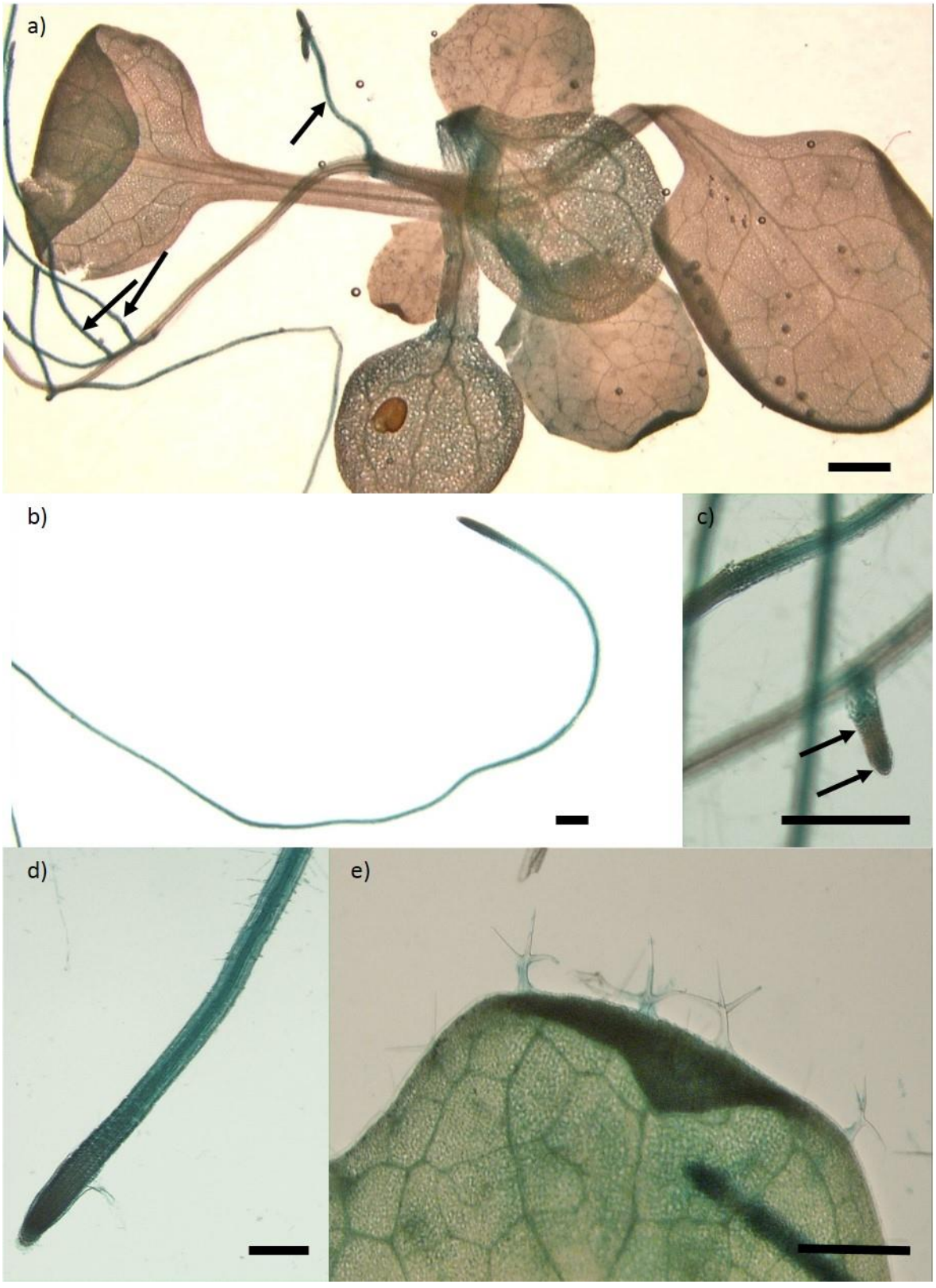


Figure 4.6: Expression profile of *IRQ5* in 14-day old seedlings. Whole 14-day old seedlings were histochemically stained for GUS activity, washed repeatedly in 70 % ethanol and mounted in 20 % glycerol for imaging. a) Expression was seen in the cotyledons but was much reduced or absent in some true leaves. Expression was strong in the lateral roots labelled by the arrows. b) Expression was seen in the lower part of the primary root from the root tip and far into the elongation and differentiation zones. c) Expression was absent in the meristematic zone in emerging lateral roots as shown by the area between two arrows. d) Expression was seen throughout the primary root tip. e) *IRQ5* was expressed in trichomes of young true leaves. Scale bars: a-b) 1000 µm, c-e) 500 µm.

4.1.2.3 Expression pattern in 5-6 week old IRQ5 promGUS plants

Whole inflorescences and developing and older siliques from all lines were immersed in GUS buffer and vacuum infiltrated for 5 minutes before being incubated at 37 °C for 30-60 minutes before being washed in an ethanol series to clear the tissues and imaged (according to section 2.5.6). In most of the lines tested (except line 3) expression was seen in the anthers, the stigma and at the base of the silique (Figure 4.7a-d). Lines 1, 4, and 5 also showed expression in the petal vasculature. Only lines 1 and 4 showed expression in the filament (Figure 4.7c) while only lines 5 and 9 showed expression in the style. Only line 4 showed expression at the base of the flower.

4.1.2.4 Expression pattern in just germinated seedlings of IRQ5 promGUS plants

To look at the potential expression pattern in very young plants seeds were grown on 1xMS plates as per section 2.5.6 but 1-2 days after being vernalised were taken from the plates, vacuum infiltrated for 1-2 minutes then incubated at 37 °C for 15-30 minutes before being washed in an ethanol series to clear the tissues and imaged (according to section 2.5.6).

Expression was seen in the primary root with no expression seen in the hypocotyl or cotyledons (Figure 4.8a-c). In the youngest of seedlings (annotated seedlings in Figure 4.8c) expression was seen throughout the root but in slightly older seedlings expression was more limited to the root tip region (Figure 4.8a and b).

4.1.2.5 Summary of expression of IRQ5

Overall, this analysis has shown that *IRQ5* is expressed more widely than *IRQ4* with expression seen throughout the root during earlier stages of development and especially at 14 DAG. The results generated here differ to that of Ren *et al.* (2016) in that *IRQ5* is predominantly root localised here while in the Ren *et al.* (2016) show staining throughout the plant in both the root system and the aerial tissues of the plant. This could be because of the far longer incubation times (overnight) used by Ren *et al.* (2016) leading to bleed through of the stain to non-expressing adjacent cells. In addition, the promoter used in the experiments is larger and may incorporate more of the elements involved in localisation of transcription. Therefore, the promoter construct used here is potentially a more accurate reflection of the transcription of the protein.

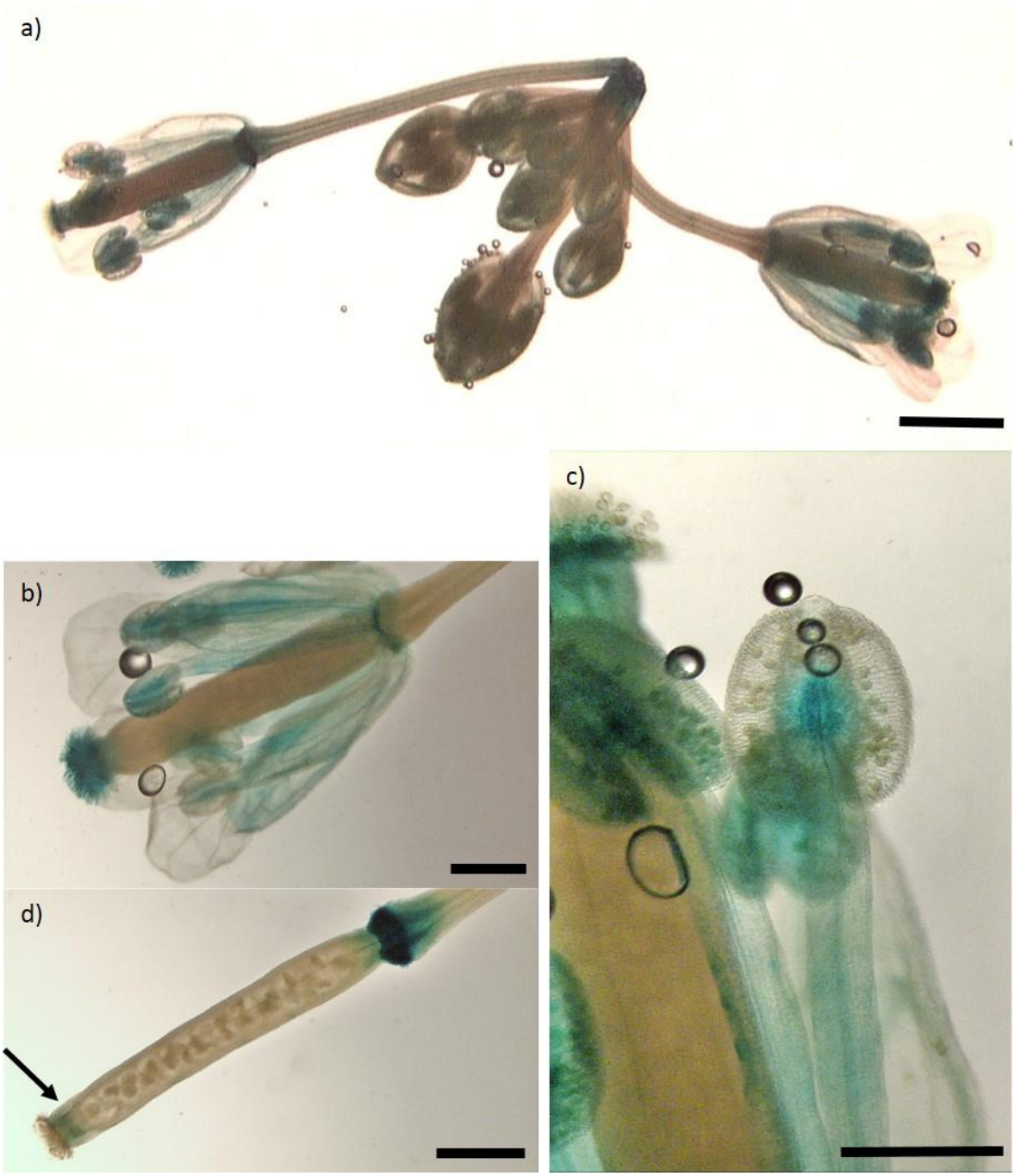


Figure 4.7: Expression profile of *IRQ5* in 5-6 week old plants. Flowers and siliques of 5-6 week old plants were histochemically stained for GUS activity, washed repeatedly in 70 % ethanol and mounted in 20 % glycerol for imaging. a) *IRQ5* expression in an inflorescence. b) Expression of *IRQ5* can be seen in the vasculature of the petals, on the stigma and in the filament of the anther which can be seen more clearly in c). c) Anther showing expression in the filament and absence of staining in the pollen grains. d) Expression was also seen in the receptacle at the base of the developing silique. Very faint staining was also visible at the tip of the silique (arrow) Scale bars: a) 1000 µm, b-c) 500 µm, d) 250 µm.

Figure 4.8: Expression profile of *IRQ5* in seedlings one day after germination. Seedlings were histochemically stained for GUS activity, washed repeatedly in 70 % ethanol and mounted in 20 % glycerol for imaging. a-c) Whole seedlings showing expression at the root tip (areas between the two arrows in a-c). Expression was also seen higher up the root and to the approximate start of the hypocotyl (star in c). Scale bars: a-c) 250 μm.

4.1.3 Expression pattern of IRQ1

4.1.3.1 Expression pattern in 5-7 day old seedlings of IRQ1 promGUS plants

Seedlings of *IRQ1prom*::GUS were incubated in GUS buffer, vacuum infiltrated for approximately 2-5 minutes at room temperature then left at 37 ^{o}C for 15-30 minutes before being washed in an ethanol series to clear the tissues and imaged (according to section 2.5.6).

Expression seemed to be localised to very specific regions: at or very near to emerging lateral roots (Figure 4.9d-f) or within specific cell files in the centre of the root tip of lateral roots distal to the root cap and in an area approximately between the meristematic and differentiation zones (Figure 4.9f and g). Expression seemed to be specific for lateral roots as the staining was only generally visible on the epidermis on the side of the root to which the lateral root was present (Figure 4.9d). Staining was present in the locations where lateral roots had emerged and were mature, indicating that *IRQ1* may have some function beyond lateral root formation and may be perhaps involved in maintain the lateral root. No expression was visible in the primary root tip. No expression of *IRQ1* was seen in the cotyledons or hypocotyl tissue (Figure 4.9a and b). Expression in all nine lines was only found in the roots.

As confirmation of the GUS localisation, the expression of *IRQ1* was also examined in 5-7 day old seedlings under CLSM set to detect GFP (according to section 2.5.3). The expression of the GFP mirrored the blue stain in the GUS assays (Figure 4.10 a-d) with GFP detected in accumulations at or near lateral roots that had emerged or were forming. GFP was also seen in the region of the root tips of lateral roots in the same manner as the GUS assay (Figure 4.10 c and d).

4.1.3.2 Expression pattern in 14 day old seedlings of IRQ1 promGUS plants

Seedlings of *IRQ1prom*::GUS were incubated in GUS buffer then vacuum infiltrated for approximately 2-5 minutes at room temperature, then left at 37 ^{o}C for 15-30 minutes before being washed in an ethanol series to clear the tissues and imaged (according to section 2.5.6).

Similar to the 5-7 day old seedlings, expression was mainly localised to the roots but there was a limited amount of expression in the SAM region in 11 of the 20 lines tested (Figure 4.11c). Within the roots, expression of *IRQ1* mirrored the 5-7 day plants with accumulations of staining at the locations where lateral roots were emerging or were present (Figure 4.11a and b). Lines 9 and 19 had limited expression in non-vascular tissue in the cotyledons and lower expression in the vascular tissue in true leaves. No expression was seen in the trichomes (Figure 4.11d and e).

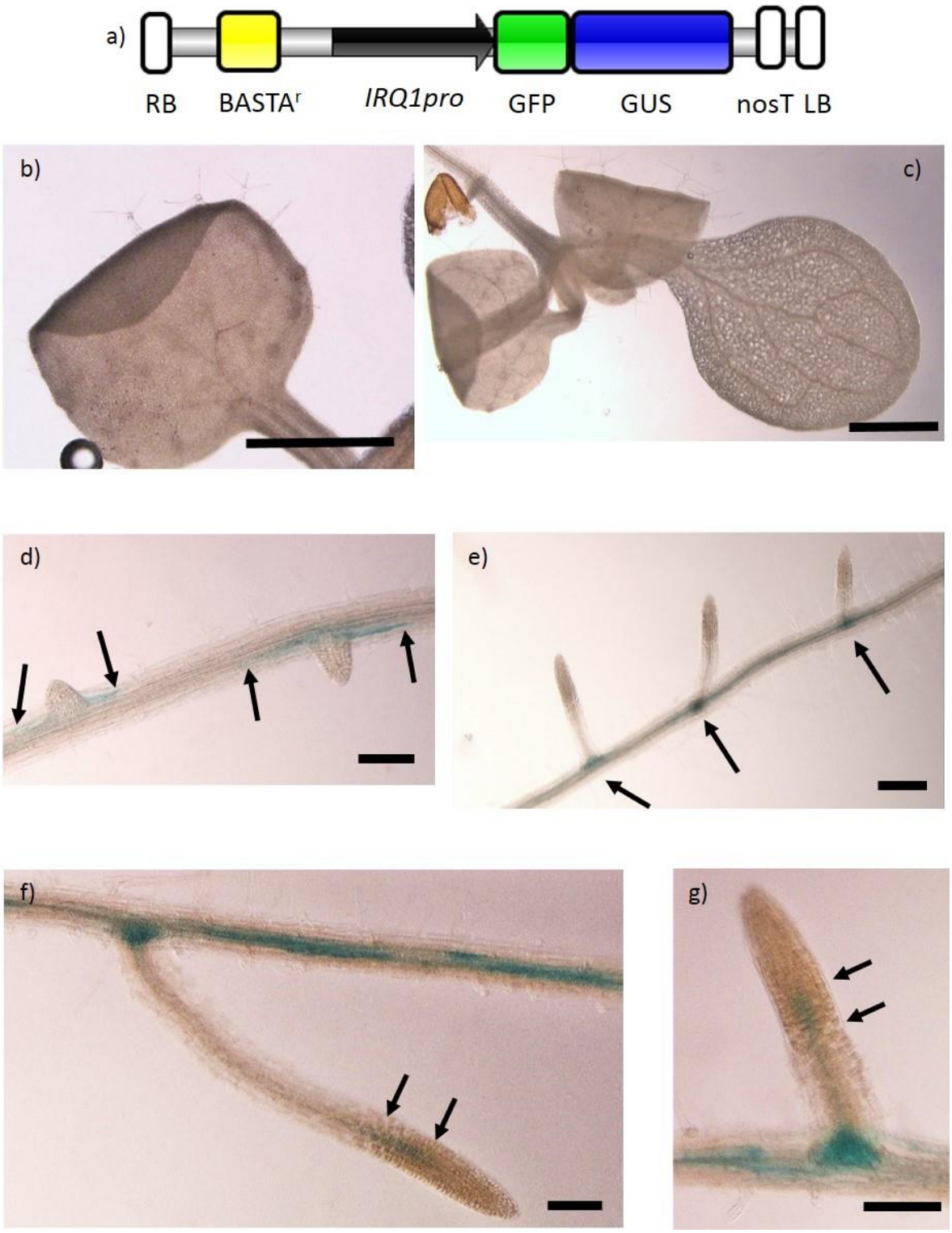


Figure 4.9: Expression profile of *IRQ1* in 5-7 day old seedlings. a) *IRQ1prom*::GUS construct created using the pBGWFS7 expression vector. b-g) Whole 7-day old seedlings were histochemically stained for GUS activity, washed repeatedly in 70 % ethanol and mounted in 20 % glycerol for imaging. b-c) Cotyledons showing no expression in the cotyledons or trichomes. d) Expression maxima occurred at areas where lateral roots are emerging but not in the lateral roots themselves (areas between arrow in (d) and at arrows in (e)). f-g) Expression was also seen in the central files of the lateral roots as shown by the areas between the two arrows in (f-g). Scale bars: b) 250 µm, c) 1000 µm, d) 250 µm, e) 100 µm, f) 1000 µm, g) 100 µm.

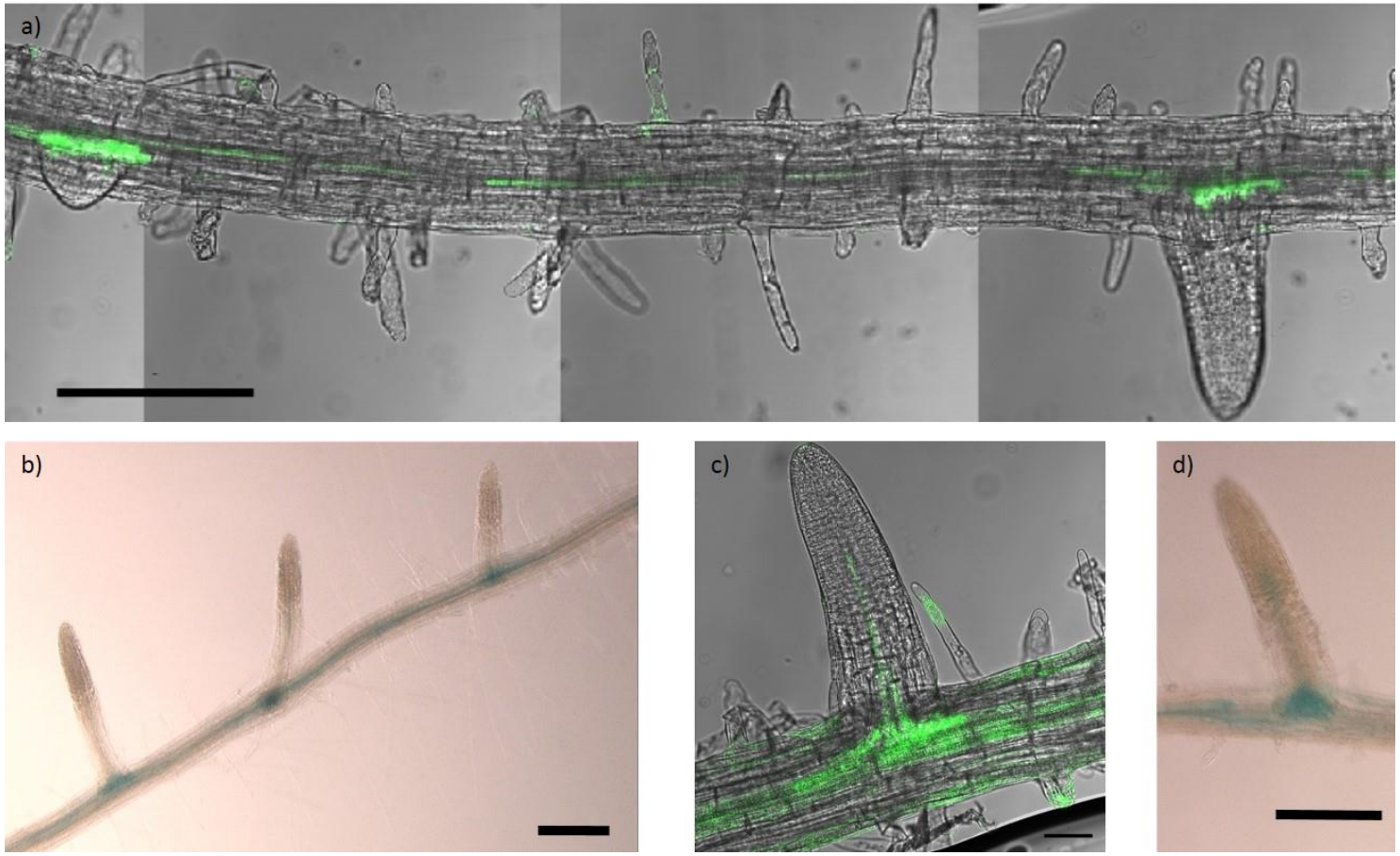


Figure 4.10: Expression profile of *IRQ1* in 7-day old seedlings using CSLM. a) Tile scan of GFP expression showing accumulations of GFP in the same places as occurred for GUS staining. An example of the GUS staining is shown in (b). c) Magnified image of an early lateral root showing accumulation of GFP expression at the base and some GFP in central file of root. A similar image of GUS staining is shown in (d). All CSLM images have the bright-field image overlaid to show cell files. Scale bars: a) 50 µm, b) 250 µm, c) 50 µm, d) 200 µm.

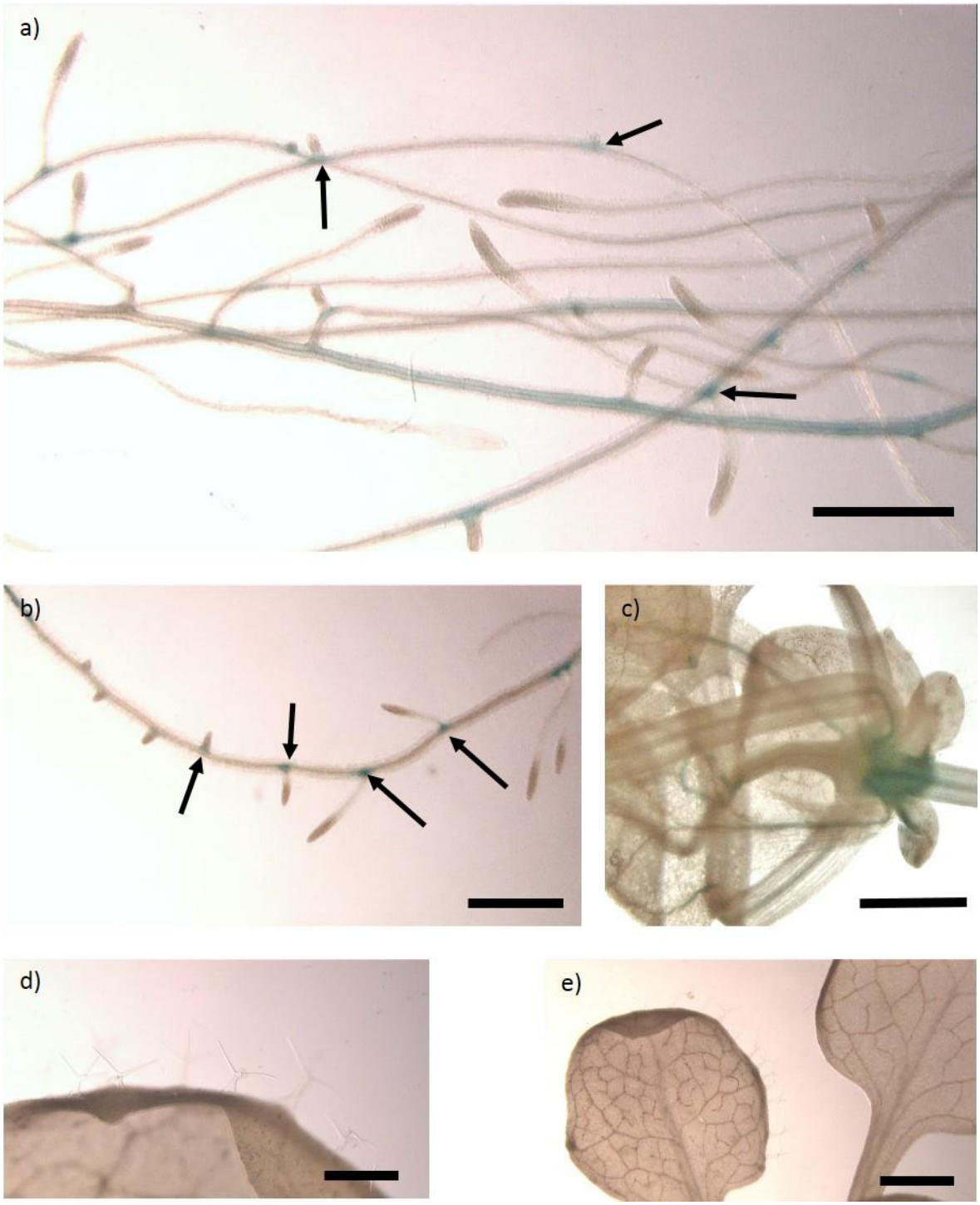


Figure 4.11: Expression profile of *IRQ1* in 14-day old seedlings. Whole 14-day old seedlings were histochemically stained for GUS activity, washed repeatedly in 70 % ethanol and mounted in 20 % glycerol for imaging. a-b) Expression occurs where lateral roots have emerged or are merging. Arrows indicate some examples. C) Expression in some lines was evident in the shoot apical meristem region. d-e) Histochemical staining revealed no expression in the trichomes and cotyledons at 14 days or in any of the younger leaves. Scale bars: a) 1000 μm, b) 1000 μm, c-d) 1000 μm.

4.1.3.3 Expression pattern in 5-6 week old IRQ1 promGUS plants

Whole inflorescences and developing and older siliques from all lines were immersed in GUS buffer and vacuum infiltrated for 5 minutes before being incubated at 37 °C for 30-60 minutes before being washed in an ethanol series to clear the tissues and imaged (according to section 2.5.6).

Of the 13 lines tested in this assay, 11 showed expression at the receptacle at the base of the silique (Figure 4.12c) while five showed expression at the receptacle at the base of the flower. Lines 9, 22, and 23 showed expression in the anthers of flowers and in the vasculature of the petals (Figure 4.12a and b).

4.1.3.4 Expression pattern in just germinated seedlings of IRQ1 promGUS plants

To look at the potential expression pattern in young very young plants seeds were grown on 1xMS plates as per section 2.5.6 but 1-2 days after being vernalised were taken from the plates, vacuum infiltrated for 1-2 minutes then incubated at 37 °C for 15-30 minutes before being washed in an ethanol series to clear the tissues and imaged (according to section 2.5.6).

In contrast to earlier results for *IRQ1* expression, *IRQ1* was seen in the root tip in an expression pattern very similar to the root tip of lateral roots (Compare Figures 4.13a and b with Figures 4.9f and g and Figure 4.10c and d). The distribution of staining was almost identical with it being limited to the centre of the root, within the region between the meristematic and differentiation zones. *IRQ1* was also expressed weakly in the hypocotyl vasculature in lines 11 and 19 (Figure 4.13c).

4.1.3.5 Summary of the expression of IRQ1

IRQ1 is almost exclusively expressed in the roots and has a very striking and specific localisation, with expression being seen at lateral root primordia. *IRQ1* accumulates in the primary root where the lateral root primordia are developing and remains after their emergence and subsequent growth. *IRQ1* is also expressed in two specific cell files just above the meristematic zone of the root. This results coupled with the lateral root primordia accumulations strongly suggest that *IRQ1* is expressed at in xylem pole pericycle cells, which are cells involved in the initiation of lateral roots (Atkinson *et al.*, 2014).

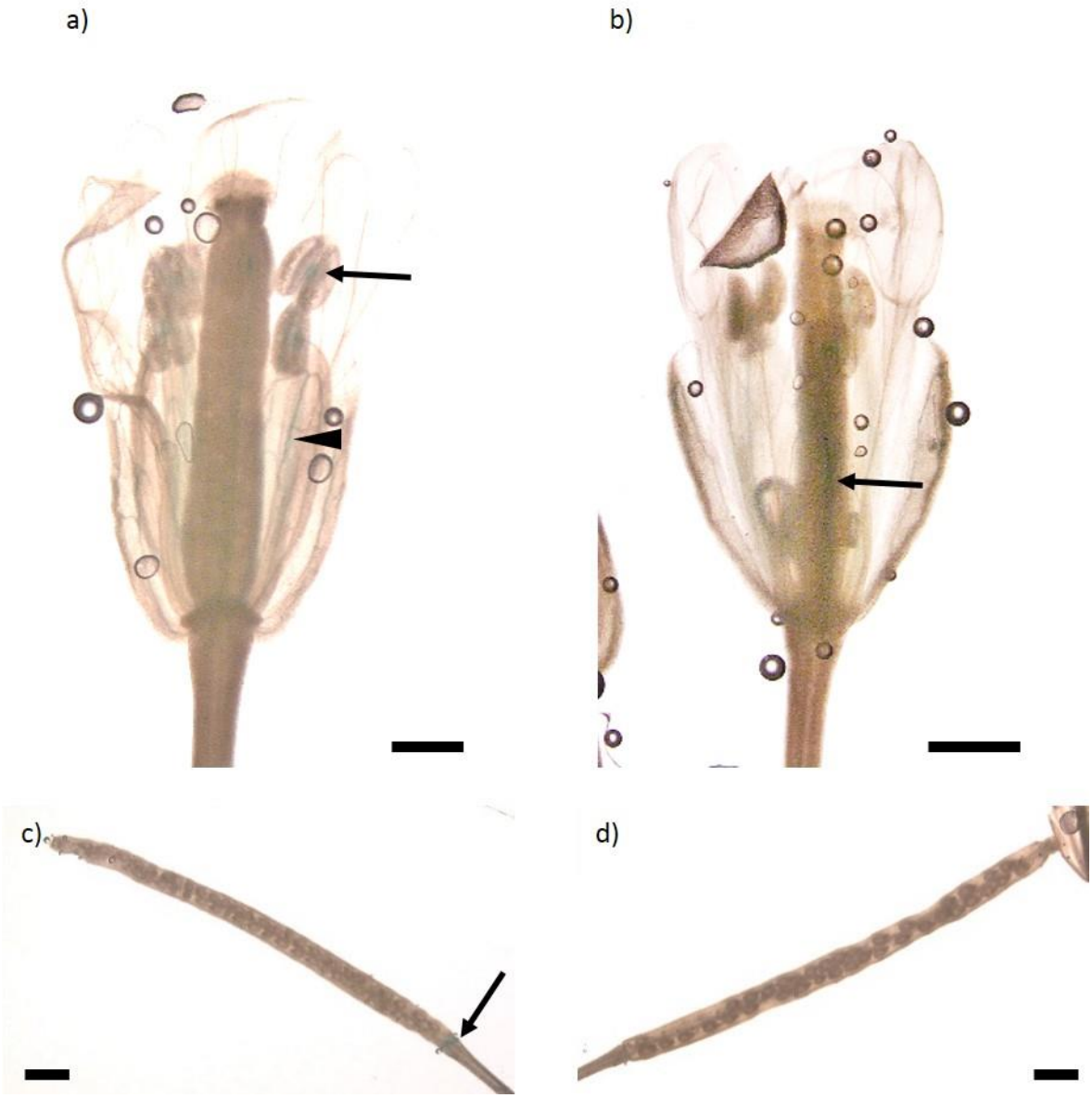


Figure 4.12: Expression profile of *IRQ1* in 5-6 week old plants. Flowers and siliques of 5-6 week old plants were histochemically stained for GUS activity, washed repeatedly in 70 % ethanol and mounted in 20 % glycerol for imaging. a-b) Whole inflorescence showing weak expression in the anthers as indicated by the arrow. Expression is also weakly seen in the vasculature of the petals as shown by the arrow head in (a) and arrow in (b). c-d). No staining is visible in the silique except in the receptacle at the base of the silique (arrow in (c)). Scale bars: a-b) 500 μm, c-d) 1000 μm.

Figure 4.13: Expression profile of *IRQ1* in seedlings one day after germination. Seedlings were histochemically stained for GUS activity, washed repeatedly in 70 % ethanol and mounted in 20 % glycerol for imaging. a) Whole seedlings showing expression at the root tip as indicated by the arrows. bi) Whole seedling showing expression at the root tip. bii) shows a close up image of the region in the box in (bi) showing the central cell file expression. c) Weak expression was seen in the vasculature of the hypocotyl (arrow in c). No expression was seen in the cotyledons at this stage. Scale bars: a-bi) 250 μm, bii) 100 μm, c) 100 μm.

4.1.4 Expression pattern of IRQ6

4.1.4.1 Expression pattern in 5-7 day old seedlings of IRQ6 promGUS plants

Seedlings of *IRQ6prom*::GUS were incubated in GUS buffer then vacuum infiltrated for approximately 2-5 minutes at room temperature then left at 37 ^{o}C for 15-30 minutes before being washed in an ethanol series to clear the tissues and imaged (according to section 2.5.6).

IRQ6 showed a much greater variety of expression compared to *IRQ1*, as expression was seen in both aerial and root tissues. *IRQ6* was strongly expressed in emerging lateral roots in all lines tested except line 9 (see arrow in Figure 4.14d). *IRQ6* was expressed in the vasculature of the primary root until approximately between the meristematic and differentiation zones (Figure 4.14e). In the lateral roots expression was similar to the primary root in that the vasculature was stained but there was also a region of the lateral root where most cell files seemed to be stained (area between two arrows in Figure 4.14f). Similar to the primary root, the level of expression in the lateral roots decreased through the meristematic zone towards the root tip.

IRQ6 was expressed in the vasculature of the cotyledons in all seven lines tested (Figure 4.14b). A low level of staining was also present in the non-vascular tissue of the cotyledons. Expression was also seen in the stipules (arrowheads in Figure 4.14b-d). *IRQ6* was expressed relatively strongly at the base of the hypocotyl at the region where the hypocotyl meets the primary root (Figure 4.14d). Staining was visible in the hypocotyl itself but despite the relatively high concentration of ferricyanide and ferrocyanide used (5 mM) this could represent leakage of stain between cells.

4.1.4.2 Expression pattern in 14 day old seedlings of IRQ6 promGUS plants

Seedlings of *IRQ6prom*::GUS were incubated in GUS buffer then vacuum infiltrated for approximately 2-5 minutes at room temperature then left at 37 ^{o}C for 15-30 minutes before being washed in an ethanol series to clear the tissues and imaged (according to section 2.5.6).

At 14 days the expression of *IRQ6prom*::GUS plants more closely matched that of *IRQ1* with no expression in the true leaves in any of the lines tested (Figure 4.15a) nor in the trichomes of true leaves (Figure 4.15b). Within the primary root, *IRQ6* was expressed at the regions corresponding to the emergence of lateral roots (Figure 4.15c-e) but the expression of *IRQ6* was more widespread within the lateral root and less specific compared to *IRQ1* (Figure 4.15c and f). *IRQ6* did not seem to express in the region corresponding to the meristematic zone in the lateral roots, but did express in the root cap (area between the two arrows in Figure 4.15f).

4.1.4.3 Expression pattern in 5-6 week old IRQ6 promGUS plants

Whole inflorescences and developing and older siliques from all lines were immersed in GUS buffer and vacuum infiltrated for 5 minutes before being incubated at 37 °C for 30-60 minutes before being washed in an ethanol series to clear the tissues and imaged (according to section 2.5.6).

In all lines tested, *IRQ6* was expressed in the petal vasculature, filaments, style, and receptacle at the base of the silique (Figure 4.16a-c). Line 1 also showed some expression at the receptacle at the base of the flower but this was not present in any of the other lines tested.

4.1.4.4 Expression pattern in just germinated seedlings of IRQ6 promGUS plants

To look at the potential expression pattern in very young plants, seeds were grown on 1xMS plates as per section 2.5.6 but 1-2 days after being vernalised were taken from the plates, vacuum infiltrated for 1-2 minutes then incubated at 37 °C for 15-30 minutes before being washed in an ethanol series to clear the tissues and imaged (according to section 2.5.6).

In all lines tested expression of *IRQ6* was seen in the root with some lines showing expression at the root tip only (Figure 4.17a) and others showing expression at the root tip and weaker expression throughout the rest of the root (Figure 4.17b). No expression was seen in the cotyledons or hypocotyl in any of the lines tested.

4.1.4.5 Summary of expression of IRQ6

IRQ6 has an over-lapping expression pattern with *IRQ1*, being predominantly expressed in the roots and also giving a distinct localisation at the emerging lateral roots, especially in seedlings 14 days old. However, in contrast to *IRQ1, IRQ6* has a more widespread expression pattern within younger seedlings of the root, being expressed more widely in the lateral roots in seedlings 5-7 days old compared to *IRQ1*. As discussed in chapter 3, IRQ6 is the result of a tandem duplication of NET4A and is therefore likely to have evolved before IRQ1. After the duplication of IRQ6 to give IRQ1, IRQ6 could have diversified in function resulting in the different localisation patterns between these two proteins. Alternatively IRQ1, could have a specific function not fulfilled by IRQ6.

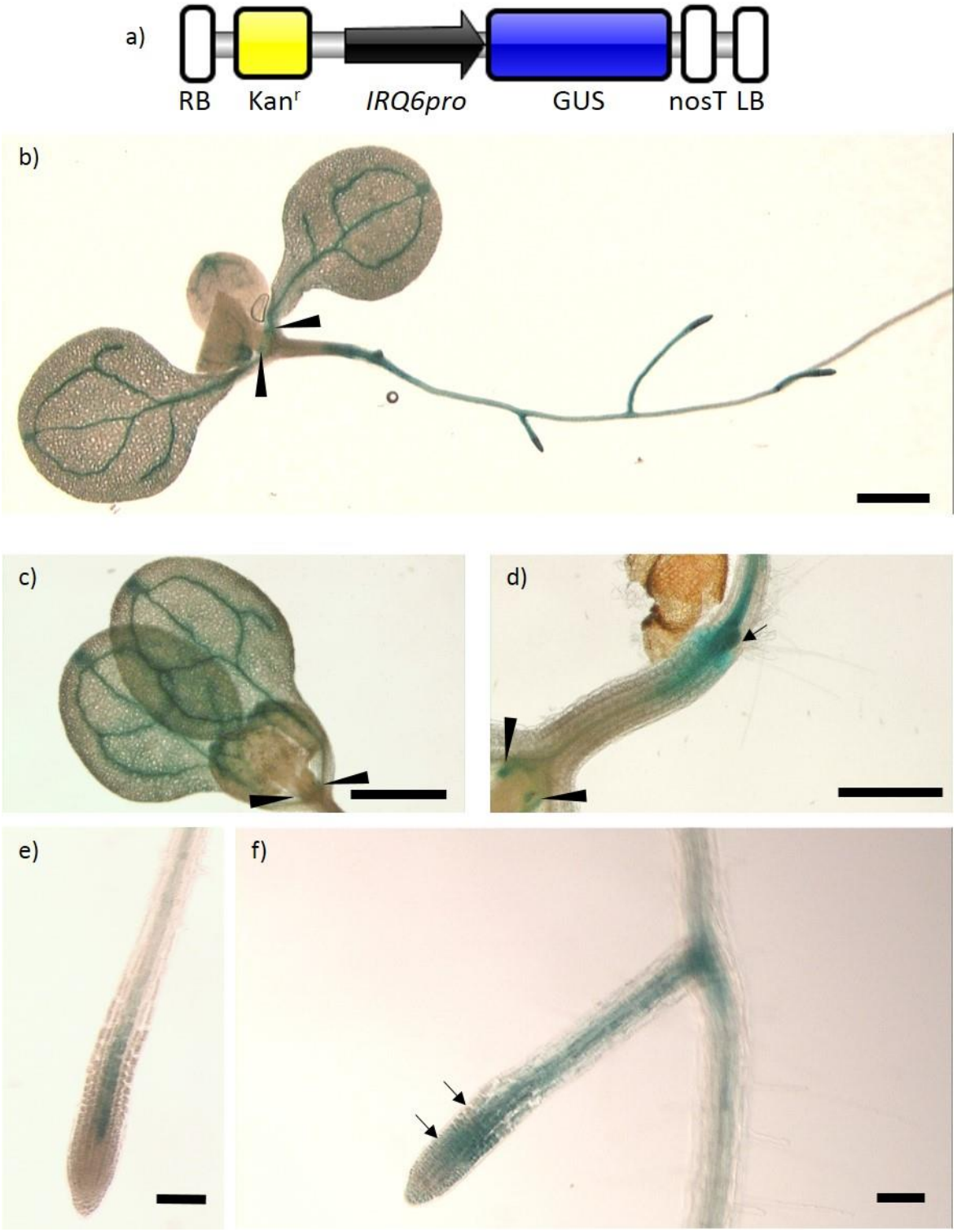


Figure 4.14: Expression profile of *IRQ6* in 7-day old seedlings. a) *IRQ6prom::*GUS construct created using the pBI101G expression vector. b-f) Whole 7-day old seedlings were histochemically stained for GUS activity, washed repeatedly in 70 % ethanol and mounted in 20 % glycerol for imaging. b) Whole seedling showing expression in the vasculature of the cotyledons (also see c), the stem just above the root, and into the root (also see d). Expression can be seen in the lateral root initial (arrow in d) and in more developed lateral roots (f). Expression appears to be localised to the central cell files in the primary root tip (e). Scale bars: b) 1000 µm, c) 1000 µm, d) 500 µm, e-f) 100 µm 1000 µm

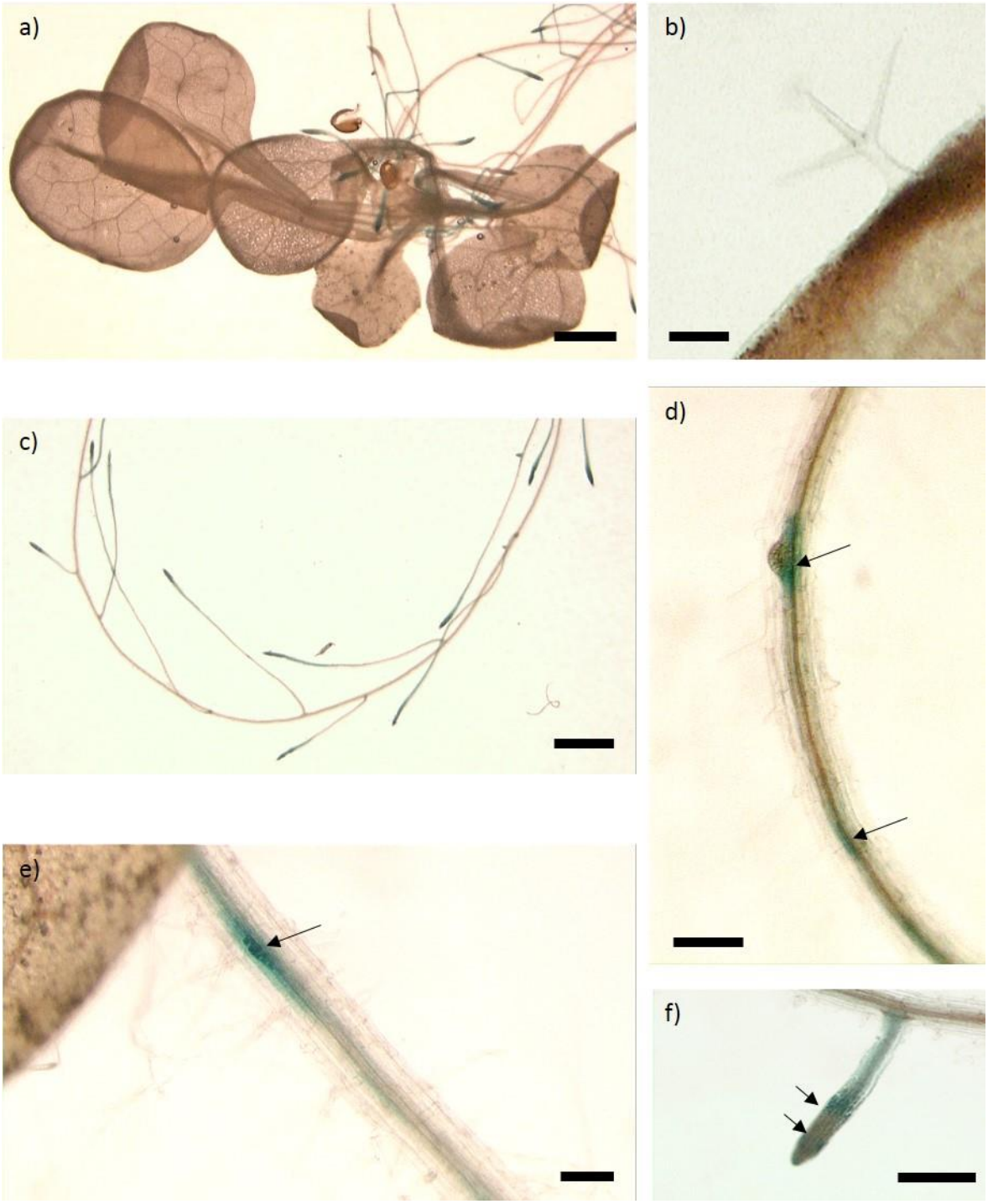


Figure 4.15: Expression profile of *IRQ6* in 14-day old seedlings. Whole 14-day old seedlings were histochemically stained for GUS activity, washed repeatedly in 70 % ethanol and mounted in 20 % glycerol for imaging. a) Histochemical staining revealed no expression in the cotyledons at 14 days or in any of the younger leaves. b) Close-up of a trichome showing no expression. c-f) Expression can be seen in the lateral roots (c,f) with expression absent in the region between the two arrows in f). d-e) Expression accumulates in the cells located around the point of emergence of the lateral root as indicated by the arrows. Scale bars: a) 1000 μm, b) 100 μm, c) 1000 μm, d-e) 100 μm, f) 500 μm.

Figure 4.16: Expression profile of *IRQ6* in 5-6 week old plants. Leaves, flowers and siliques of 5-6 week old plants were histochemically stained for GUS activity, washed repeatedly in 70 % ethanol and mounted in 20 % glycerol for imaging. a-b) Whole inflorescence showing expression in the flower and filaments of the flowers. The boxed region in b) is expanded in c) showing clearer filament staining and some expression in the vasculature of the petals. d) No expression is visible in the silique but is present at the receptacle at the base of the silique and the tip of the silique. Scale bars: a-bi) 1000 μm, bii) 500 μm, c) 1000 μm.

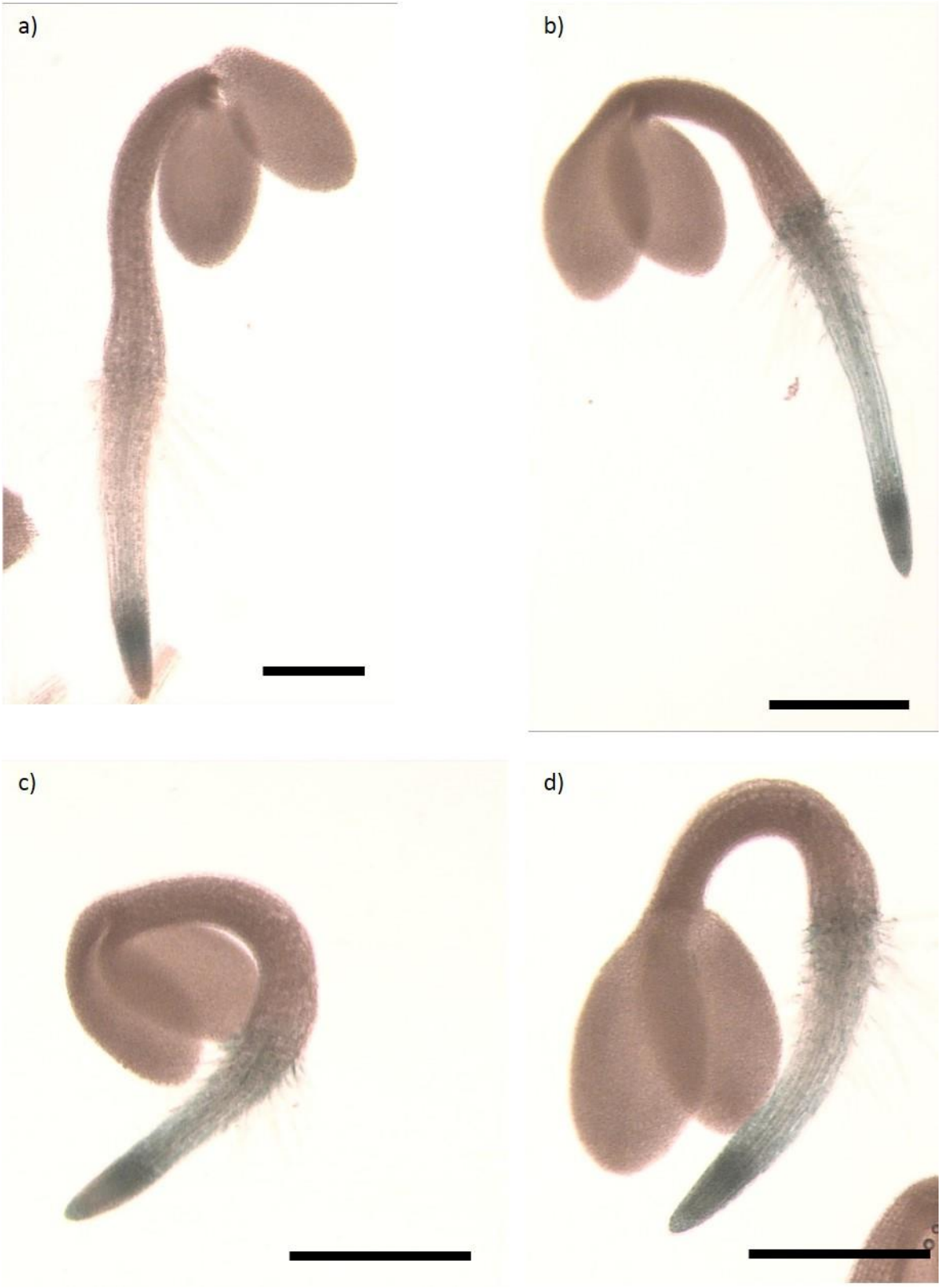


Figure 4.17: Expression profile of *IRQ6* in seedlings one day after germination. Seedlings were histochemically stained for GUS activity, washed repeatedly in 70 % ethanol and mounted in 20 % glycerol for imaging. a-d) Whole seedlings showing expression at the root tip (a-d) and slightly lower expression throughout the young root (b-d). Some lines showed no expression apart from the root tip. Scale bars: a-d) 500 μm.

4.2 Conclusion

In this chapter an analysis of the localisation of transcription of *IRQ4*, *IRQ5*, *IRQ1*, and *IRQ6* was experimentally determined using the *prom*GUS expression system in *Arabidopsis thaliana* plants. A schematic summarising the expression of the IRQ proteins tested here is shown in Figure 4.18.

As has been discussed in chapter 3 IRQ4 and IRQ5 are closely related due to their shared N-terminal pNAB domains and their C-terminal IRQ domains. IRQ4 is generally localised to the root and, in contrast to the data obtained from eFP browser, is seemingly not expressed in the aerial tissues of the plant apart from the trichomes on more mature leaves (Figure 4.18). Expression was also not found within the flowers to the extent of that predicted by microarray experiments. However, expression of IRQ4 in the roots was shown to be in approximately the correct location with an increase in expression seen at the top of the meristematic zone and into the elongation zone as predicted by the eFP browser (Figure 3.6ei and Figure 3.6eiii). Discrepancies between the data obtained given by microarray experiments and the experimental data obtained here are likely due to the high-throughput nature of the microarray experimental setup. In the *prom*GUS system used here single genes are looked in isolation over several growth stages. Therefore, *prom*GUS fusions generally provide more accurate data on localisation with the microarray data providing more of a guide to the localisation of expression.

IRQ5 is much more widespread in expression, being present in both the roots and cotyledons, true leaves and the hypocotyl. The localisation shown here for *IRQ5* generally supports earlier work on *IRQ5* transcriptional localisation (Ren *et al.*, 2016) and also the microarray data from chapter 3. The roots of 5-7 day old seedlings displayed expression similar to levels expected from the microarray experiments. However, expression was weaker in the leaves of the seedlings compared to that predicted by the microarray data. Expression in the leaves at 14 days was again weaker than that predicted by microarray data but expression in roots was much stronger than predicted at this life stage. At 5-6 weeks expression of *IRQ5* was weaker in the petals than predicted but was as expected in the anthers.

There were differences in the extent to which staining in IRQ5 plants occurred in above ground and below ground tissues when comparing the results presented here with that of Ren *et al.* (2016). These are likely due to the differing size of the fragment used when cloning the promoter region of IRQ5. Ren *et al.* (2016) used 1549 bp upstream of the start codon whereas the promoter region used to create the IRQ5 construct used here incorporated 2 kb of sequence upstream of the start site. The increased area used here could have included elements which reduced the extent to which *IRQ5* was expressed in the aerial tissues. In addition, Ren *et al.* (2016) used a different method to that used here with incubation times in GUS buffer overnight compared to

the shorter times used in this thesis. Shorter times are likely to be representative of the true localisation of transcription as the stain is less likely to escape and bleed through into cells in which the gene is not expressed giving false positive results.

Both IRQ4 and IRQ5 are known to have mRNA which is mobile (Thieme *et al.*, 2015) suggesting that although transcription may be active in a restricted area the site of translation and the location where the encoded protein is active may be different to the site of transcription. This is perhaps not surprising when considering the importance of coordinated above and below ground responses to abiotic and biotic stimuli e.g. to cytokinins (Werner & Schmulling, 2009).

Both IRQ1 and IRQ6 showed a relatively similar localisation pattern. IRQ6 was less restricted in localisation compared to IRQ1. Microarray data had predicted that both *IRQ1* and *IRQ6* were strongly expressed at the shoot apical meristem (SAM) and the hypocotyl (Figure 3.6bi). This study has shown that *IRQ1* is expressed at the SAM at 7 DAG but *IRQ6* transcription was not localised to this region (Figure 3.18). As has been discussed in chapter 3, NET4A underwent a tandem duplication to produce IRQ6 so is likely to have evolved before IRQ1. Subsequently IRQ6 could have duplicated to produce IRQ1. Duplicated genes often have overlapping but divergent functions compared to their ancestors (Taylor & Raes, 2004) therefore IRQ1 could have evolved to perform a function in the SAM as well as the lateral root primordia.

Expression of *IRQ1* was predicted to be weakly expressed in leaves and flowers by microarray data and this was consistent with the results from the GUS assays. *IRQ6* was predicted to have stronger expression in leaves compared to *IRQ1 b*ut only vasculature staining was visible at 5-7 DAG and no staining was visible in 14 day old plants. *IRQ6* was expression in flowers matched that of the microarray data, with some expression seen in the anthers.

Both *IRQ1* and *IRQ6* expression localised strongly to emerging lateral roots. This is in contrast to the microarray data which did not suggest a root localisation for the expression of these gene products. *IRQ1* seemed to accumulate at sites where no lateral root was yet visible. In addition, *IRQ1* localises to particular cell files in lateral roots that appear to be xylem pole pericycle cells. These cells are able to de-differentiate and are the founder cells of lateral root primordia (Malamy & Benfey, 1997). IRQ1, and to some extent IRQ6, could be involved in the formation and development of lateral root primordia.

This chapter has shown that the IRQs analysed here show an array of different localisations of transcription within the plant with *IRQ1* and *IRQ4* being more specific in their localisation compared to *IRQ6* and *IRQ5* respectively. Although *prom*GUS studies are able to provide information the localisation of transcription of a protein and to what extent it is expressed at different stages of plant development, the localisation of transcription may not represent the site

of action of the protein. In addition, *prom*GUS studies are unable to provide information on the subcellular localisation of a protein. One way in which to study this is to use fusion proteins to fluorescent reporters. This is the subject of the next chapter where the localisation and behaviour of IRQ4 is studied in more detail.

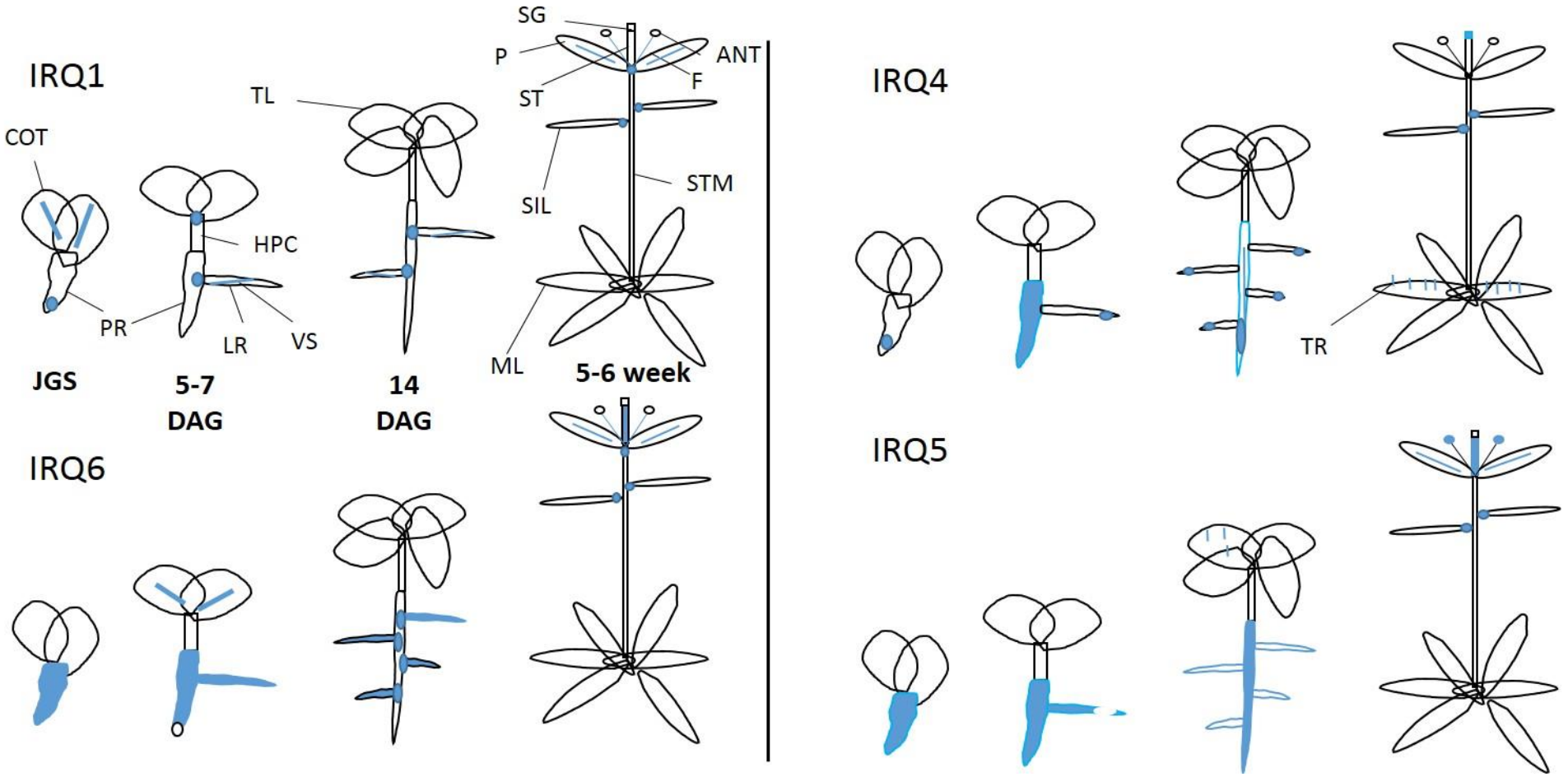


Figure 4.18: Summary of GUS staining in *IRQ1, IRQ6, IRQ4,* and *IRQ5*. The pattern of GUS staining observed in just germinated seedlings (JGS) and seedlings 5-7 and 14 days after germination (DAG) and 5-6 week old plants are represented by the respective schematic. Single blue lines represent staining in the vasculature of the root or leaf as indicated. ML= mature Leaf, SIL= silique, P= petal, STM= stem, ANT= anther, F= filament, SG=stigma, ST= style, PR= primary root, LR- lateral root, VS= Vasculature, COT= cotyledon, HPC= hypocotyl, TL= true leaf, TR=trichome, JGS= Just germinated seedling, DAG= Days after germination.

Chapter 5 *In vitro* and *in vivo* analysis of IRQ4 localisation

5.1 Introduction

The previous chapter has shown that *IRQ1* and *IRQ6* are mainly transcribed in the root, and that *IRQ4* and *IRQ5* are also transcribed in the root but that *IRQ5* has a wider distribution including the aerial tissues of the plant. As the site of transcription of an mRNA encoding a protein is not the site of translation and activity of that protein, it is important to use translational fusions to proteins of interest to study their localisation further. This chapter describes experiments which aimed to study the localisation and behaviour of translational fusions of IRQ4. IRQ4 was selected for further study because it is an unknown protein of unknown function but it is known that its mRNA is cell-cell mobile (Thieme *et al.*, 2015), indicating that perhaps the IRQ4 protein is present in tissues other than the one in which it is transcribed.

Throughout these assays the construct of interest was introduced into plant cell culture or intact plants using an *Agrobacterium tumefasciens* mediated system. Transformation using *A. tumefasciens* is a highly efficient method for both the transient or stable transformation of a wide variety of crop and other plant species (Gelvin, 2003, 2010). Stable transformation uses the ability of *A. tumefasciens* to infect plants and integrate its own DNA within the plant genome via a type iv protein secretion system. The DNA to be integrated is contained within a large plasmid, either the tumour inducing (Ti) or rhizogenic (Ri) plasmid, within the bacterium. The region with the Ti or Ri plasmid is termed the T-DNA. Transfer of the T-DNA into the plant cell is aided by virulence genes. During stable transformation the gene of interest is cloned into the T-DNA and in a separate replicon the virulence genes are contained, this constitutes a binary system. The functionality of the virulence genes is reduced to prevent the formation of tumours detrimental to the plant. The DNA of interest is integrated within the genome of the plant, possibly using a non-homologous end joining (NHEJ) mechanism, and the protein of interest is subsequently expressed. In transient expression systems the T-DNA is not integrated into the plant genome.

5.2 Initial cloning of the IRQ4 full length protein and visualisation in *Nicotiana benthamiana* leaf epidermal cells

The full length IRQ4 protein (1-555 aa) was initially cloned by Dr. D. Mentlak into the pMDC83 destination vector which creates a fusion to GFP at the C-terminal end of the protein constitutively expressed under the control of a dual CaMV35S promoter (Figure 5.1a). The IRQ4 fusion protein was expressed in leaf epidermal cells of *N. benthamiana* using an *Agrobacterium*

mediated transient expression method as described in section 2.2.13. After 2-3 days of incubation the GFP-fusion was imaged using confocal laser scanning microscopy (CLSM).

The IRQ4-GFP fusion seemed to localise to punctate structures (Figure 5.1b). Using time lapse imaging these punctae were found to move rapidly across the cell (Figure 5.1c and d). The punctate structures showed a variety of behaviours including touching but not merging, merging then splitting, or not touching at all (Figure 5.1d for examples). There are a number of different compartments that these structures could be, including prevacuolar compartments/multivesicular bodies (PVC/MVB), autophagasomes, or another as yet unknown endosome compartment.

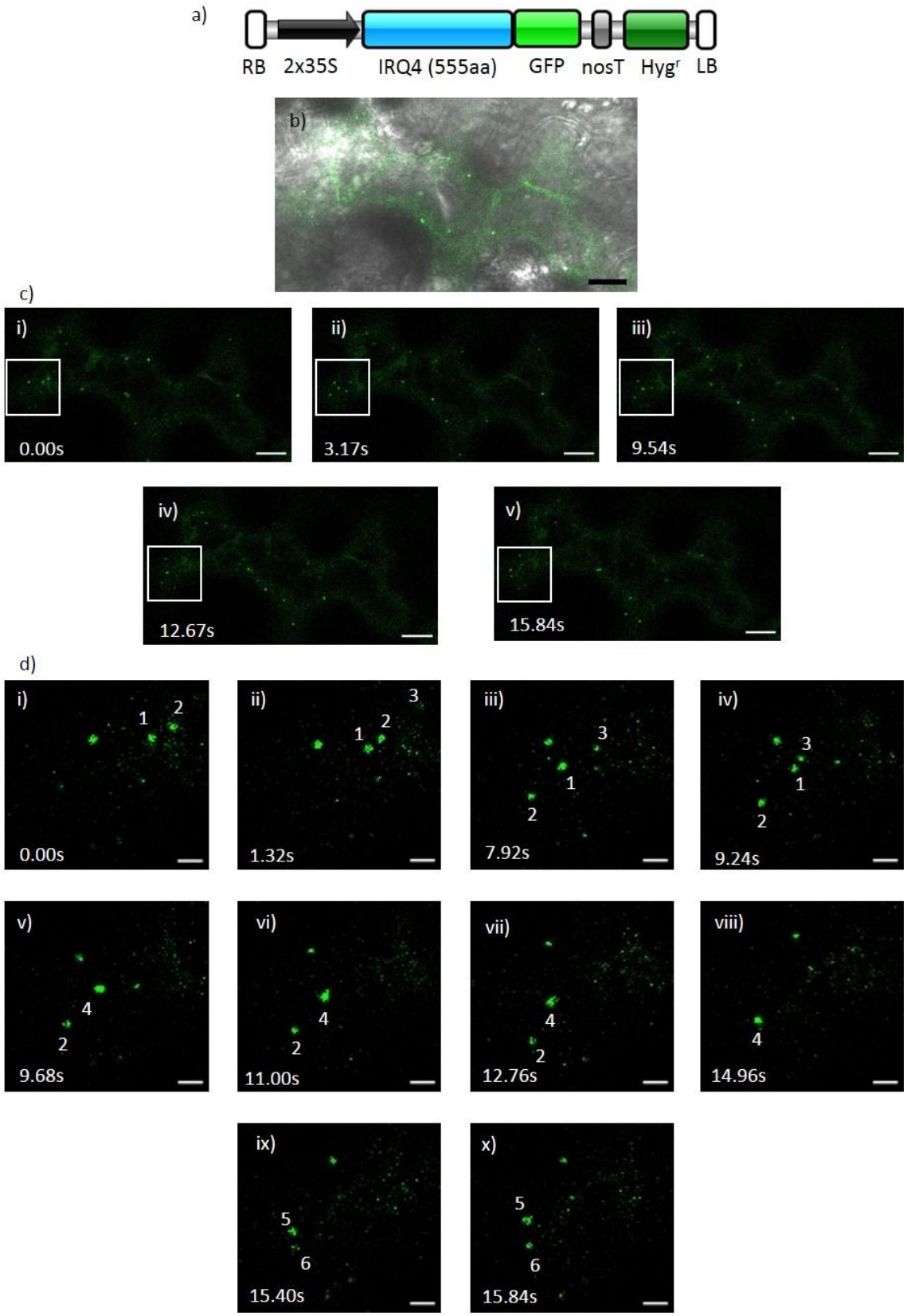


Figure 5.1: Transient expression of IRQ4-GFP fusion protein in N. benthamiana leaf epidermal cells. a) IRQ4-GFP construct made using the pMDC83-GFP expression vector. b) Merged image of DIC and GFP fluorescence image of *N. benthamiana* expressing IRQ4::GFP showing punctae. c) Time series from *N. benthamiana* leaf segment showing overview of cell during image acquisition. Relative time shown in seconds at bottom left (ci-v). White box indicates area shown in di-x over same time period. d) Punctae of IRQ4-GFP show different patterns of behaviour; sometimes coming close together but not merging or touching (1 and 2, di-ii), and other times coming together (1 and 3, dii-iv), seeming to merge to form a hybrid compartment (4, dv-viii) before splitting again to form two separate compartments (5 and 6, dix-x). Scale bars: b) 10 μm, ci-v) 10 μm, di-x) 2 μm.

5.3 Analysis of IRQ4 expression in stably transformed BY2 cell culture

Unfortunately very few cells were labelled using the transient transformation technique so in order to study protein expression in more detail the same IRQ4-GFP construct used in the transient assays was stably transformed into cultures of *Nicotiana tobacum* Bright Yellow 2 cells (BY2) according to section 2.2.14. Stable expression in BY2 cells removes some of the variability that is associated with transient expression systems, the cells are relatively easy to transform and the cells are larger than other cell culture systems such as *Arabidopsis* allowing easier visualisation (Geelen & Inze, 2001; Nagata *et al.*, 2004).

Seven day old cultures of BY2 cells expressing IRQ4-GFP were subcultured according to section 2.6.2 and 2-3 days after were mounted in BY2 media and visualised using CSLM. In interphase BY2 cells expression was seen at the membrane of what appeared to be the vacuole (arrowheads in Figure 5.2b). In addition, punctate structures similar to that seen in the transient expression system were also seen (compare Figure 5.1d with structures at arrows in Figure 5.2bi). The punctae also moved rapidly around the cell as visualised by time lapse imaging (Figure 5.2c-e). The range of behaviours exhibited by the punctae were also similar including seeming to merge and splitting quickly thereafter (Figure 5.2d), or just touching, not merging, but moving together across the cell before disappearing from the plane of view (Figure 5.2e). None of the punctae in the images merged with the larger potentially vacuole type compartments in the time lapse images taken. This does not preclude this from happening as time lapse images were taken in a single plane so the punctae could move out of the plane of vision before merging with the larger membranes.

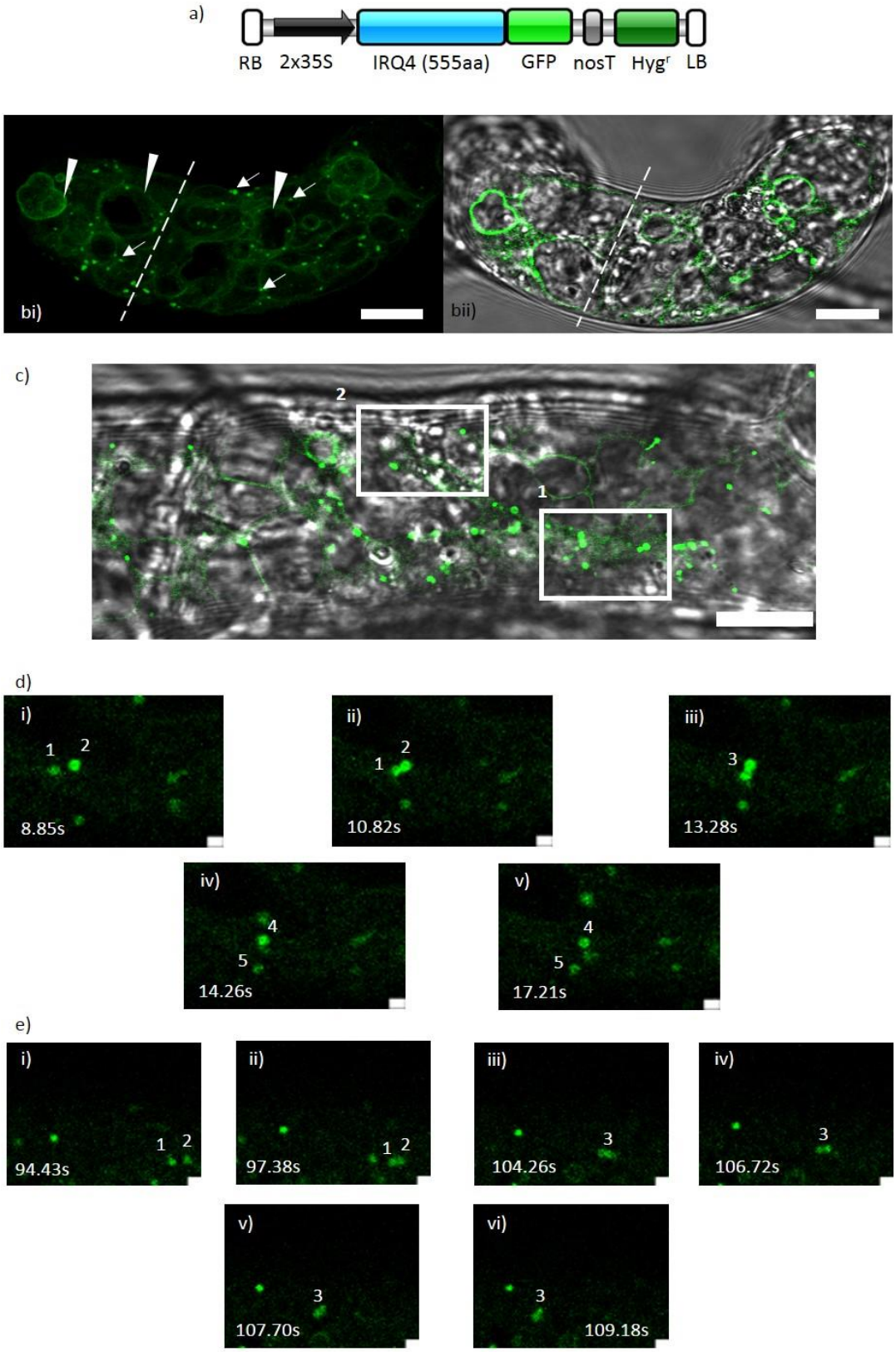


Figure 5.2: Localisation of the IRQ4-GFP fusion protein stably expressed in BY2 interphase cells. a) IRQ4-GFP construct made using the pMDC83-GFP expression vector. b) Maximum projection of BY2 cell showing both punctae (arrows) and slightly fainter vacuole localisation (arrowheads). Dotted line indicates cell wall between two cells. bii) is a merged image of brightfield image with IRQ4-GFP. c) Overview of cell used for time series in d) and e). Boxed regions correspond to areas imaged for screenshots in d (Box 1) and e (Box 2). d) Punctae of IRQ4-GFP come close together touching but not merging (1 and 2, di-ii), seeming to form a hybrid compartment (3, diii), before splitting and forming separate compartments (4 and 5, div-v). e) Some punctae touch but do not merge (1 and 2, ei-ii) but remain together and move together without fusing (3, diii-vi). Relative time shown in seconds at bottom left (di-v, ei-vi). Scale bars: b-c) 10 μm, d-e) 1μm.

5.3.1 Colocalisation analysis of IRQ4 with endomembrane markers

To identify the membrane system IRQ4 localises to, the IRQ4-GFP stably expressing BY2 cells were transformed with markers for the endoplasmic reticulum (HDEL-RFP), Golgi (ST-RFP), and the vacuole (VAMP711-MCherry) according to section 2.2.14. Only HDEL-RFP was successfully transformed into the IRQ4-GFP expressing cells. Seven day old cultures of the co-expressing cells were sub-cultured according to section 2.6.2 and after 2-3 days were mounted on a slide in BY2 media and imaged using CSLM.

IRQ4-GFP associates with the ER (Figure 5.3a and b) with the punctae seemingly aligned with the ER network. The large volume of the vacuole relative to the internal volume of the cell, pushes the ER to the periphery of the cell along with the IRQ4-GFP punctae and other cell contents. Therefore, the images shown in Figure 5.3 can only show association rather than direct colocalisation without further experiments. The ER network is extensive and potentially has connections to nearly every organelle in the cell including the plasma membrane, Golgi, and mitochondria (Hawes *et al.*, 2015). It is therefore not surprising that the IRQ4-GFP punctae seem to associate with this organelle. To more specifically localise the protein to a particular organelle the IRQ4-GFP expressing cells need to be colabelled with organelle markers for the Golgi and vacuole.

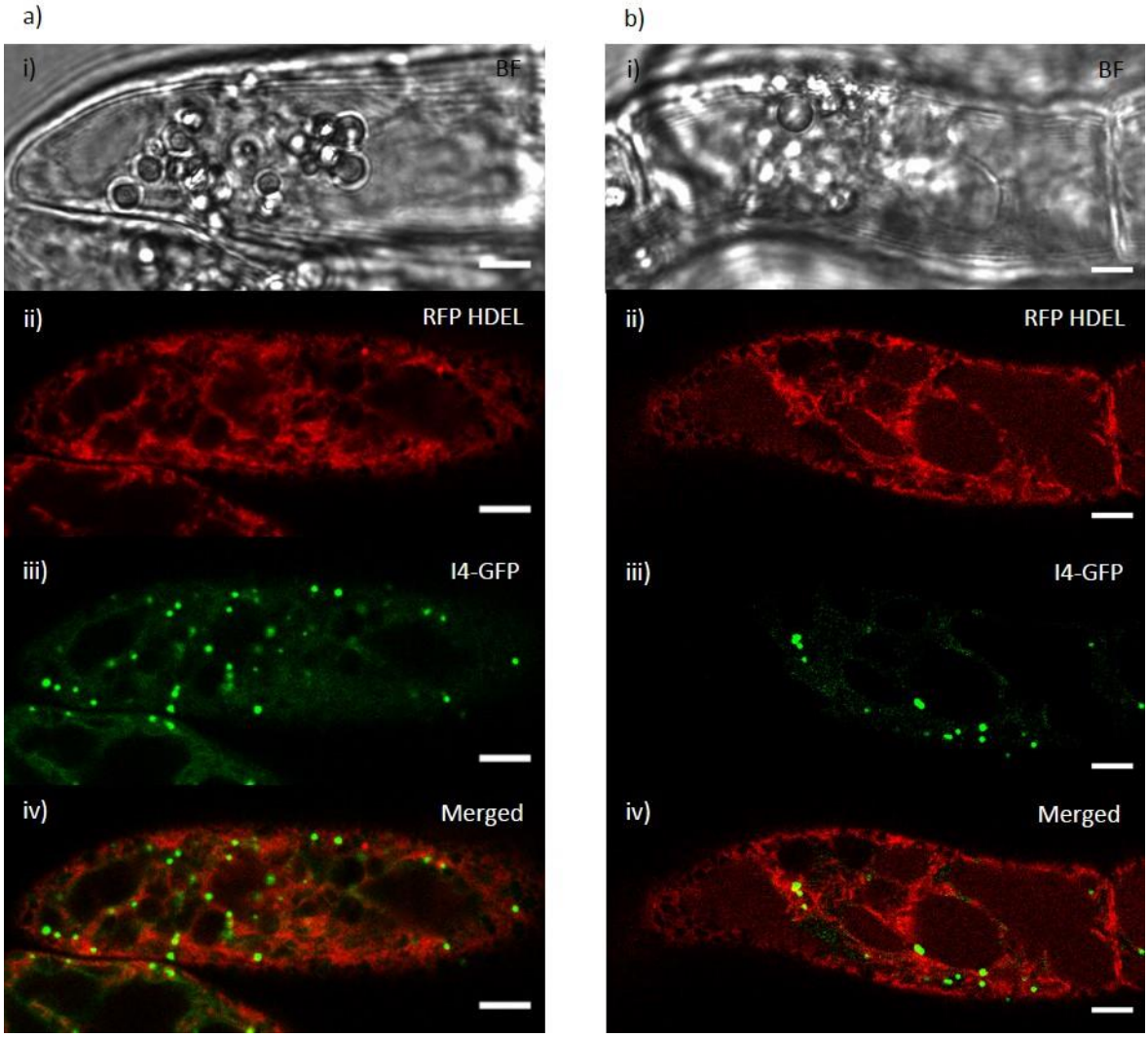


Figure 5.3: Co-expression of stably expressed IRQ4-GFP with HDEL-RFP in BY2 interphase cells. a) and b) Single BY2 cells co-expressing IRQ4-GFP with RFP-HDEL. ai) and bi) Brightfield images of cells, aii) and bii) HDEL-RFP, aiii) and biii) IRQ4-GFP, aiv) and biv) merged images. Scale bars: ai-iv) and bi-iv) 10 μm.

5.3.2 Drug studies to assess the localisation and behaviour of IRQ4 in stably expressing BY2 cell culture

The previous results have suggested that IRQ4 has two localisation patterns in BY2 cell culture: a large vacuole-like endomembrane compartment and smaller punctae. Another tool to assess the localisation of proteins within the endomembrane system is to use specific inhibitors of transport through the endomembrane system such as Brefeldin A (BFA) and wortmannin. BFA inhibits the GTPase activity of ADP-ribosylation factor 1 (Arf1p), blocking the formation of COPI coated vesicles at the Golgi leading to the loss of distinct Golgi cisternae in a *cis*-to-*trans* direction, changes in morphology of the ER and the formation of ER-Golgi hybrid compartments (Ritzenthaler *et al.*, 2002). BFA has no direct effect on the structure of the vacuole. Wortmannin is an inhibitor of phosphatidylinositol 3-kinases (PI3K), and to some extent PI4Ks, and halts both the generation of autophagasomes causing them to disappear from the cell (Merkulova *et al.*, 2014)

and causes the homotypic fusion of prevacuolar compartments (PVC) causing them to swell (Wang *et al.*, 2009). Treatment with wortmannin has no effect on Golgi bodies (Tse *et al.*, 2004).

As has been noted, the IRQ4-GFP labelled punctae move rapidly across the cell (Figure 5.2). Directed movement of endomembrane compartments within plants cells is controlled by either the actin or microtubule cytoskeleton. The movement of organelles and the bulk flow of cytoplasm is mostly achieved through the action of the actin and myosin motor proteins (Woodhouse & Goldstein, 2013; Tominaga & Ito, 2015). To determine which of the cytoskeletal elements is responsible for the movement of the punctae, cytoskeletal disrupting drugs can be used. Latrunculin B inhibits actin filament formation by binding to actin monomers, preventing the formation of filamentous F-actin and the eventual depolymerisation of the actin cytoskeleton (Morton *et al.*, 2000). Oryzalin is a dinitroaniline herbicide and inhibits the formation of microtubules by binding to plant tubulin in such a way that it cannot polymerise (Strachan & Hess, 1983).

5.3.2.1 Treatment of stably expressing BY2 cells with BFA

Seven day old cultures of IRQ4-GFP expressing cells were sub-cultured according to section 2.6.2 and after 2-3 days were treated with either DMSO (control) or 10 µM BFA for 1 hour (according to section 2.5.4. Cell samples were taken every 15 minutes for 60 minutes, mounted on a slide and imaged using CSLM. Two independently transformed cell lines were tested, images from one representative line is shown in Figure 5.4.

In the control BY2 cells there was no change in localisation over the time of the experiment meaning the localisation of the protein is not affected by the level of DMSO used (Figure 5.4a). Treatment with 10 µM BFA had no effect on the localisation of the IRQ4 protein with images similar to the control treatment (Figure 5.4b). None of the punctae were seen accumulating in BFA bodies over the time of the experiment. IRQ4 does not appear to localise to the Golgi.

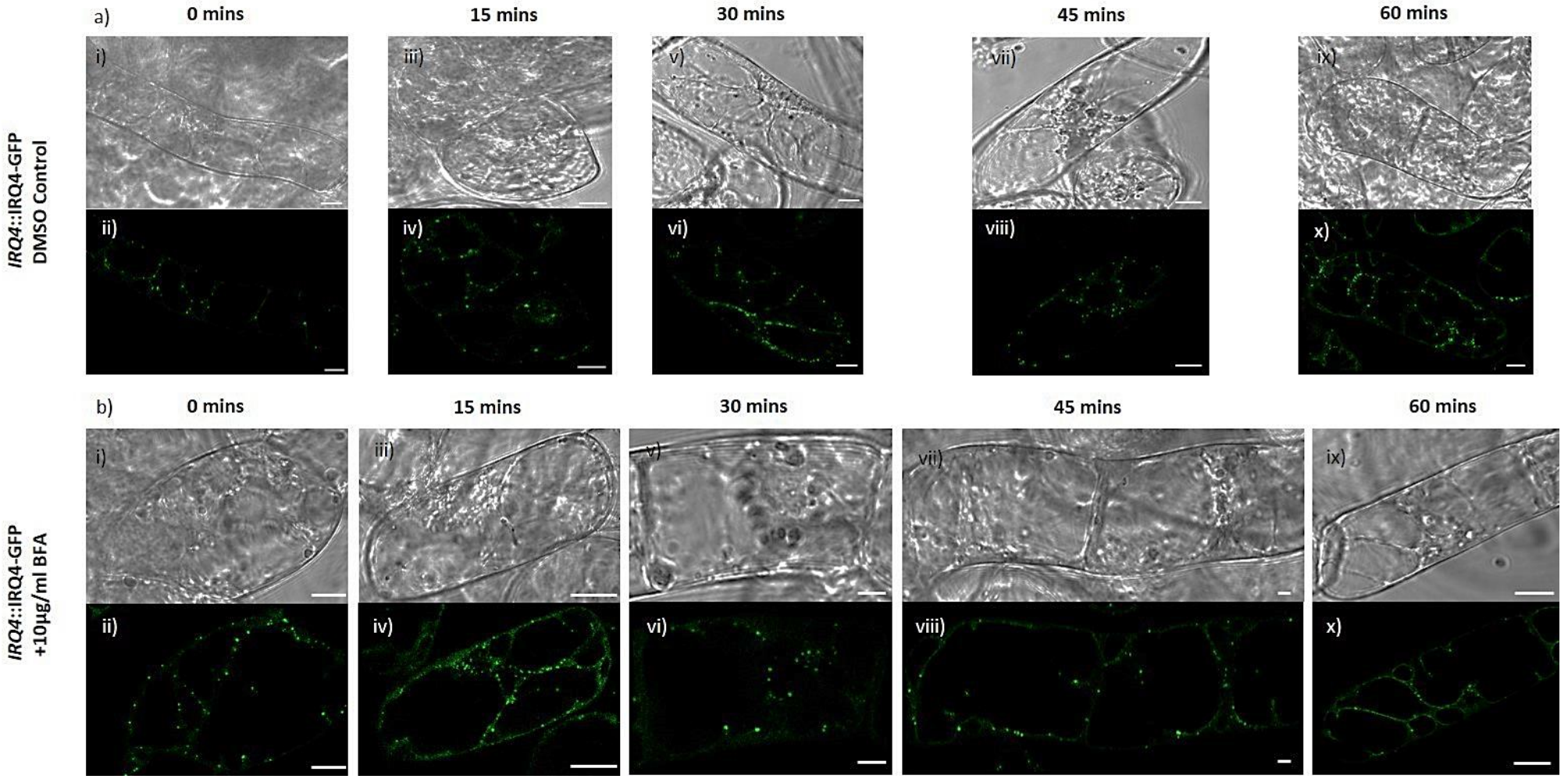


Figure 5.4: BFA treatment has no effect on the localisation of IRQ4-GFP construct. Stably expressing IRQ4-GFP BY2 cells were treated with either DMSO a) or with 10µg/ml BFA b). Brightfield images (a-i,-iii, -v, -vii,-ix; bi, -iii, -v, -vii, -ix) with their corresponding GFP images (aii, -iv, -vi, -viii, -x; bii, -iv, -vi, -viii, -x). Images were taken every 15 minutes for 1 hour. Time points are indicated above images. Scale bars: ai-x) and bi-x) 10 µm.

5.3.2.2 Treatment of stably expressing BY2 cells with wortmannin

Seven day old cultures of IRQ4-GFP expressing BY2 cells were sub-cultured according to section 2.6.2 and after 2-3 days were treated with either DMSO (control) or 30 µM wortmannin for three hours (according to section 2.5.4). Cell samples were taken after three hours and mounted on a slide and imaged using CSLM. In this experiment two independently transformed cell lines are shown in Figure 5.5.

Control treated cells of both lines 7 and 23 show the punctate localisation of the IRQ4-GFP construct as seen earlier in interphase BY2 cells (Figure 5.5) meaning the localisation of the protein is not affected by the level of DMSO used or the length of time of the treatment. In both lines treatment with 30 µM wortmannin for three hours caused the punctae to become bigger, forming swollen vesiculated compartments which cluster together. This result is similar to that obtained by Wang *et al.* (2009) using GFP-BP-80 expressing BY2 cells as a marker for the PVC. Therefore, the result shown in Figure 5.5 indicates that the IRQ4 punctae represent PVC.

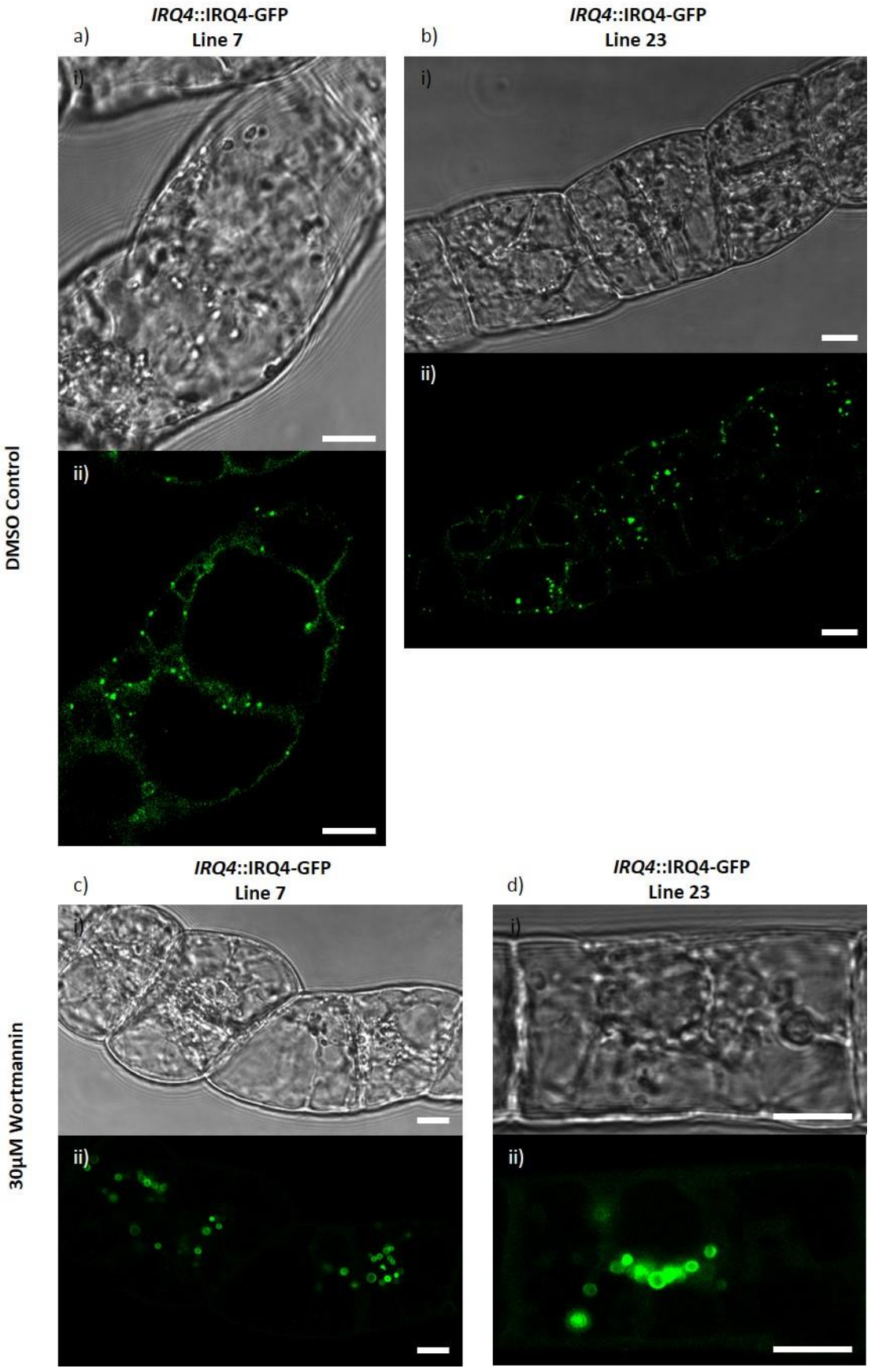


Figure 5.5: Effect of wortmannin treatment on the localisation of the IRQ4-GFP construct. Stably expressing IRQ4-GFP BY2 cells treated with either DMSO a-b) or with 30 µM wortmannin c-d). Brightfield images (ai, bi, ci, di) are shown with their corresponding GFP images (aii, bii, cii, dii). Images were taken after 3 hours of DMSO or wortmannin treatment. Two independently transformed BY2 lines, line 7 and line 23, are shown. Scale bars: ai-iv) and bi-iv) 10 µm.

5.3.2.3 Treatment of stably expressing BY2 cells with oryzalin

Seven day old cultures of IRQ4-GFP expressing BY2 cells were sub-cultured according to section 2.6.2 and after 2-3 days were treated with either DMSO (control) or 37.5 µM oryzalin for 1.5 hours (according to section 2.5.4). Cells were imaged using CLSM in a single focal plane for 200 frames according to section 2.5.3. The speed of the punctae within each cell was calculated using the spot tracking plugin in the ICY software (Chaumont *et al.*, 2012) as described in section 2.7.2. The average speed of punctae in the control and treated cells and a student's t test was calculated to compare the means according to section 2.7.2.

The mean speed of punctae in oryzalin treated cells was not significantly different in treated cells (0.35 μms^{-1}, s.d.=0.15, 95 % CIs: 0.32 to 0.38) compared to the DMSO treated control cells (0.32 μms^{-1}, s.d.= 0.17, 95 % CIs:0.24 to 0.39; t(22)=1.15, p=0.26; Figure 5.6a). The behaviour of the punctae when treated with oryzalin did not change showing that the punctae do not depend on the microtubule cytoskeleton for their movement.

5.3.2.4 Treatment of stably expressing BY2 cells with latrunculin B

Seven day old cultures of IRQ4-GFP expressing BY2 cells were sub-cultured according to section 2.6.2 and after 2-3 days were treated with either DMSO (control) or 10 µM latrunculin B for one hour according to section 2.5.4. Cells were imaged using CLSM in a single focal plane for 200 frames according to section 2.5.3. The speed of the punctae within each cell was calculated using the spot tracking plugin in the ICY software (Chaumont *et al.*, 2012) as described in section 2.7.2. The average speed of punctae in the control and treated cells and a student's t test was calculated to compare the means according to section 2.7.2.

The mean speed of PVCs in latrunculin B treated cells was significantly lower (0.23 μms^{-1}, s.d.=0.09, 95 % CIs: 0.17 to 0.30) compared to the DMSO treated control cells (0.47 μms^{-1}, s.d.= 0.11, 95 % CIs: 0.39 to 0.55; t(19)=5.10, p<0.001; Figure 5.6b). The behaviour of the punctae when treated with latrunculin B changes from mostly directed movement in control cells to a prevalence of undirected, Brownian motion in treated cells (Supplemental movies S1 (control) and S2 treated)). This indicates that the movement of the punctae is dependent upon the actin cytoskeleton.

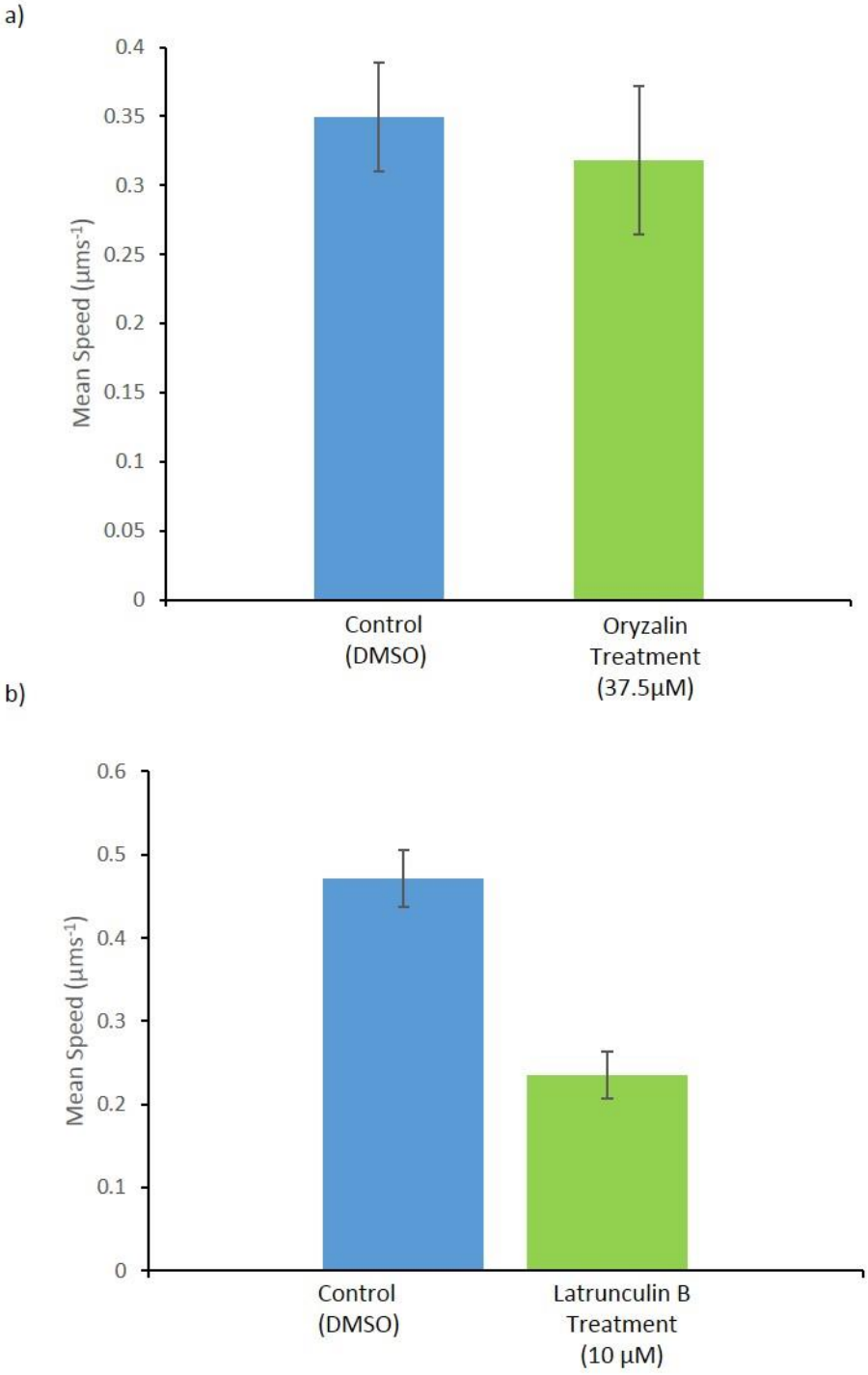


Figure 5.6: Effect of cytoskeletal disrupting drugs on the movement of IRQ4 punctae. a) BY2 expressing the IRQ4-GFP construct were treated with either 34.5 µM oryzalin or DMSO (control) for 1.5 hours. b) BY2 cells expressing the IRQ4-GFP construct were treated either with 10 µM latrunculin B or DMSO (control) for 1 hour. In a) and b) cells were imaged using CLSM in a single focal plane for 200 frames according to section 2.5.3. The speed of the punctae within each cell was calculated using the spot tracking plugin in the ICY software (Chaumont *et al.*, 2012) as described in section 2.7.2. The average speed of punctae in the control and treated cells and a student's t test was calculated to compare the means according to section 2.7.2. Oryzalin does not have a significant effect on the speed of the punctae whereas disrupting the actin cytoskeleton using latrunculin B causes a significant decrease in their movement. Bars on graph indicate S.E.M.

5.4 In vivo analysis of the localisation of the full length IRQ4 protein expressed under the control of the IRQ4 promoter in *Arabidopsis thaliana* plants

To verify the localisation pattern seen in the BY2 cells stably expressing transgenic *Arabidopsis* plants were created using the full length IRQ4 protein fused to GFP under the control of its own promoter. Whereas the construct used to create the BY2 IRQ4-GFP lines utilised a dual 2X35S CaMV promoter, using constitutive overexpression can sometimes create false localisation and potentially the creation of protein aggregates. Using a constitutive promoter also means that in intact plants expression is driven in cell types in which the protein may not normally be expressed. Expressing the protein under its own promoter reduces the potential for artefacts and is more likely to represent the true localisation of the wild type protein. The original clone used was created by Dr. D. Mentlak. Briefly, the cDNA clone of the full length protein plus approximately 2 kB of upstream promoter sequence was cloned into the pDONR207 Gateway compatible vector via a BP clonase reaction. The construct was then cloned into the pMDC107 destination vector via an LR reaction which creates a C-terminal fusion of the protein to GFP with expression driven by the endogenous promoter (Figure 5.7a).

Wild type *Arabidopsis* were grown until flowering then transformed by the *Agrobacterium* mediated floral dip method as described in section 2.2.15. T_0 seed were sown and selected on hygromycin as described in section 2.6.4. Positive transformants were screened using fluorescence microscopy as described in section 2.6.4 and checked for the presence of the *promIRQ4*::IRQ4-GFP insert by PCR as described in section 2.2.1 using primers listed in appendix 1. The cycle of growth and selection was completed two more times until homozygous T_3 seed was obtained and this seed was used throughout the project for imaging.

Plants were grown vertically on solid 1XMS plates as described in section 2.6.3 for 5-7 days before being mounted in liquid 1XMS and imaged via CLSM. No expression was seen in the aerial tissues including the cotyledons or the hypocotyl, indicating that although the mRNA encoding IRQ4 is capable of moving between ground and aerial tissues (Thieme *et al.*, 2015) the transcript appears not to be translated in the aerial tissues. Subsequent analyses focused on the root tissues. Stacks of images were taken in the regions of the root indicated in Figure 5.7b and used to create maximum projections (Figure 5.7c-d). IRQ4 localises to the meristematic zone of the root and is absent from the root cap. Expression is seen throughout the meristematic zone and into the differentiation and elongation zones (Figure 5.7ci-iv). The intracellular localisation of the IRQ4 protein in the root tip is very similar to that seen in the BY2 cell cultures in section 5.3. IRQ4 localises to a large membranous compartment which takes up a large volume of the cell (see inset in Figure 5.7di). However, the punctae seen in the BY2 cells are largely absent from these cells.

Further into the meristematic zone in the area corresponding to Figure 5.7bii) the localisation changes to become much more similar to that seen in the BY2 cells with localisation seen at the large vacuolar type compartment but also punctae are visible (Figure 5.7dii). Further up the root at the end of the meristematic zone and into the differentiation zones (areas corresponding to Figure 5.7biii and iv) the IRQ4-GFP localises solely to punctae (Figure 5.7diii and iv).

In the lateral root, the change in localisation with the change in the cell type of the root can be seen in a single image (Figure 5.8b). At the tip of the lateral root the vacuole type localisation can be seen (area between arrows in Figure 5.8b). Moving towards the primary root the localisation changes with the punctate localisation becoming more prominent (arrowheads in Figure 5.8b). Punctae are also visible near the root tip in the approximate region of the meristematic zone in the lateral root (arrowheads in Figure 5.8b).

As there were punctae present the dynamic behaviour of the punctae was studied in the roots. Single cells in the late meristematic zone were imaged using time lapse microscopy and the behaviour of the puncta examined. The same patterns of behaviour that were observed using the IRQ4-GFP construct in *N. benthamiana* leaf epidermal cells and transgenic BY2 cells were seen in the root cells: punctae appeared to merge, form hybrid compartments and split back into smaller compartments (Figure 5.9b and c). This finding verifies the behaviour of the dots in the heterologous systems of *N. benthamiana* leaf epidermal cells and *N. tobacum* cell cultures and indicates that the localisation is unlikely to be an artefact due to overexpression as the same behaviour was observed using a construct which utilises expression under the promoter of the gene in question, making it more likely the localisation pattern reflects the behaviour of the wild type protein.

Taken together, the above localisation data obtained from the *Arabidopsis* roots suggests that the localisation of IRQ4 changes depending on the zone of the root being imaged. This also indicates that the localisation of IRQ4 may change depending of the maturity of the tissue being examined.

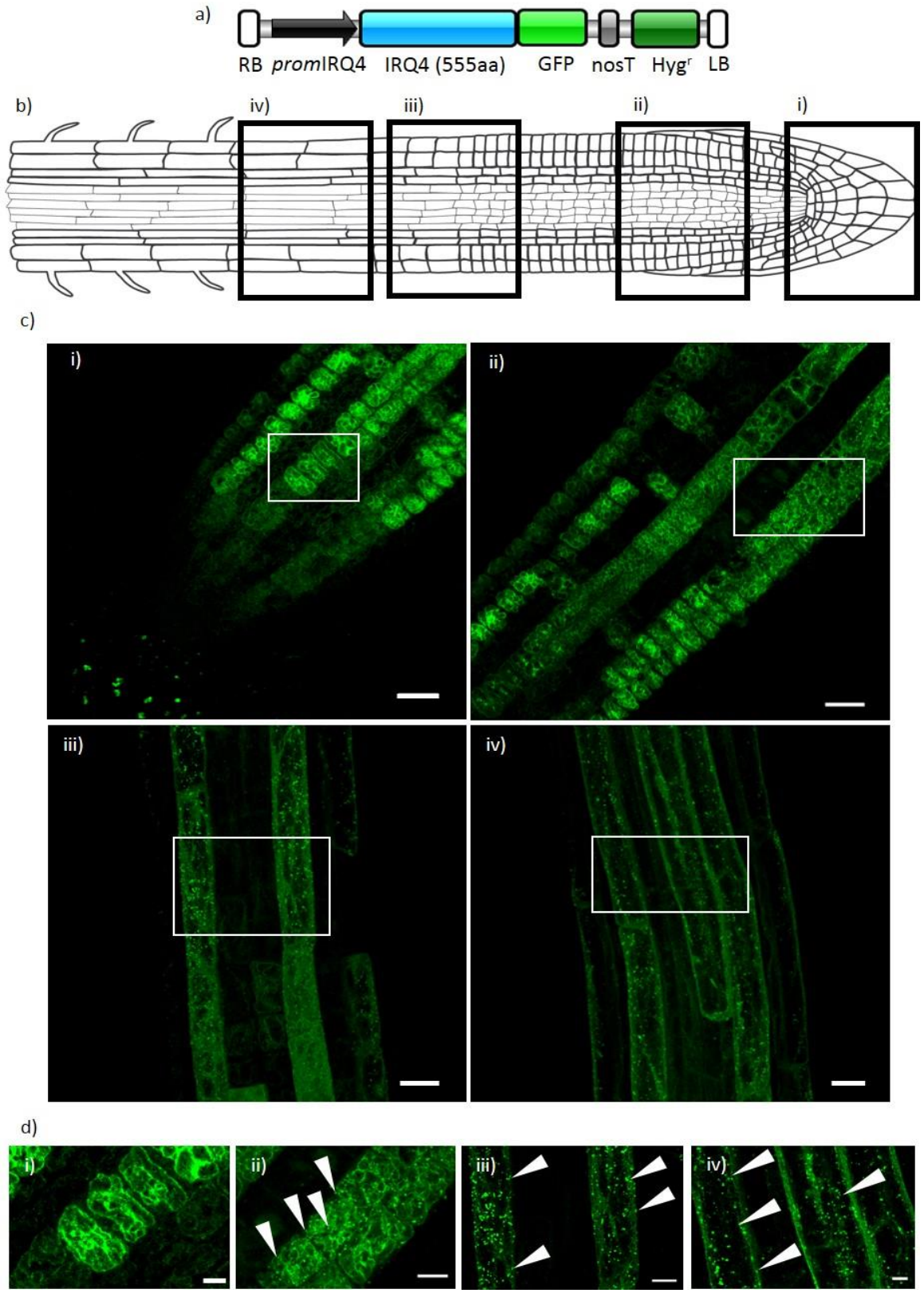


Figure 5.7: Localisation of IRQ4 in the primary roots of transgenic *Arabidopsis thaliana* expressing IRQ4 under its own promoter. a) *promIRQ4*::IRQ4-GFP construct created using the pMDC107-GFP expression vector. b) Schematic of root showing approximate locations (black boxes) where images in c) were taken i) Root rip, ii) meristematic zone, iii) meristematic/differentiation zone, iv) differentiation/elongation zone. Numbering reflects numbering in c) i.e. bi)=ci), bii)=cii) etc. c) Expression of *prom*IRQ4::IRQ4-GFP construct in areas of root corresponding to the schematic in b). White boxes indicate area used to show magnified images in d). d) Magnified images showing changes in localisation with progression from the root tip to the elongation zone. Numbering corresponds to boxes in c i.e. di)= box in ci), dii) = box in cii) etc. Arrowheads dii-iv) indicate appearance of punctate structures. Schematic in b) modified from Plant Illustrations (2017). Scale bars: c) 20 μm, di, iv) 5 μm, dii, iii) 10 μm.

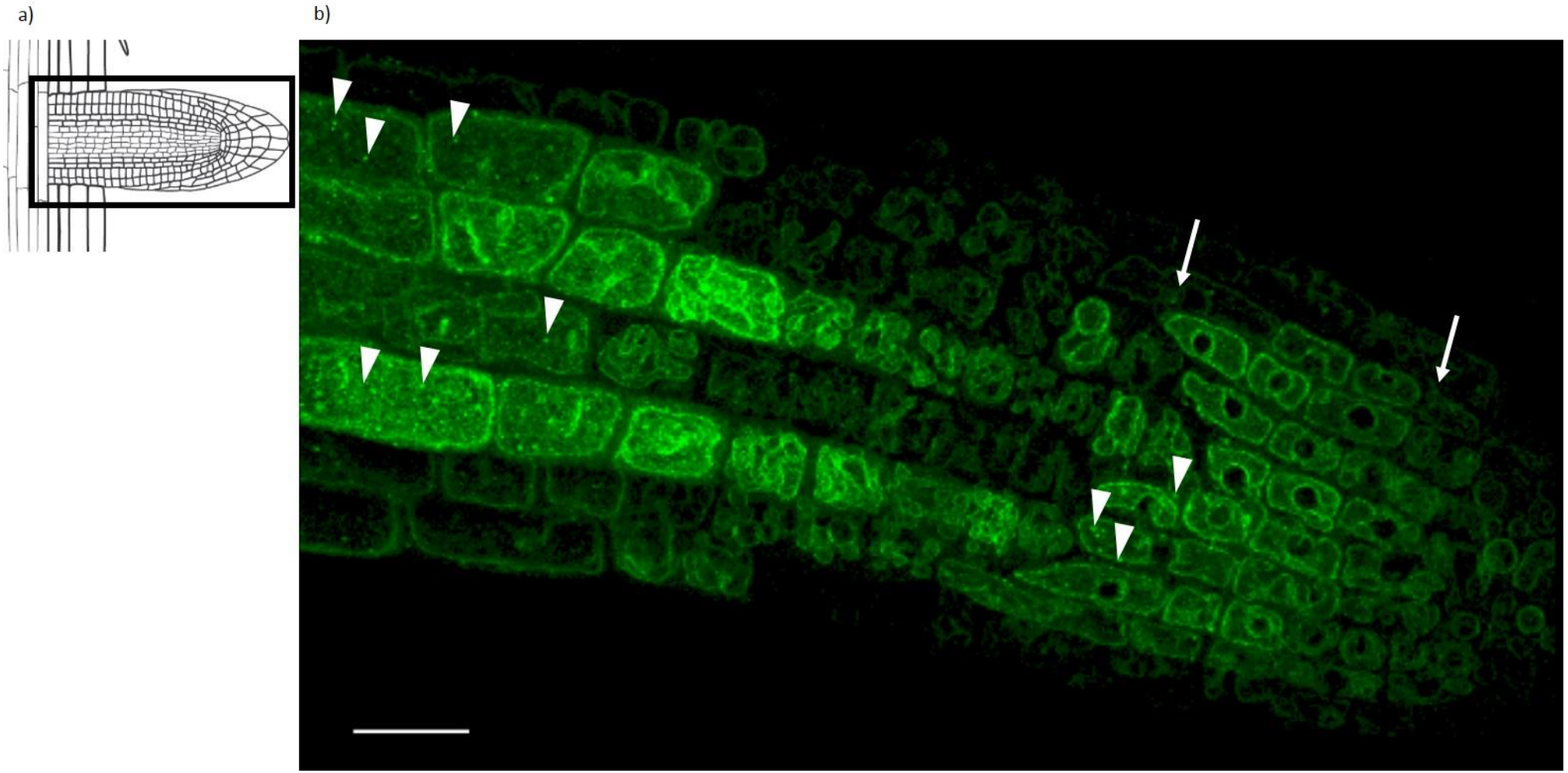


Figure 5.8: Localisation of IRQ4 in lateral roots of transgenic *Arabidopsis thaliana* expressing IRQ4 under its own promoter. a) Schematic showing area of lateral root (black box) used for imaging in b). b) Expression of *promIRQ4::*IRQ4-GFP construct in the lateral root. Area between arrows shows where vacuole localisation is most common. Arrowheads indicate punctate structures. Schematic in a) modified from Plant Illustrations (2017). Scale bar: b) 20 μm.

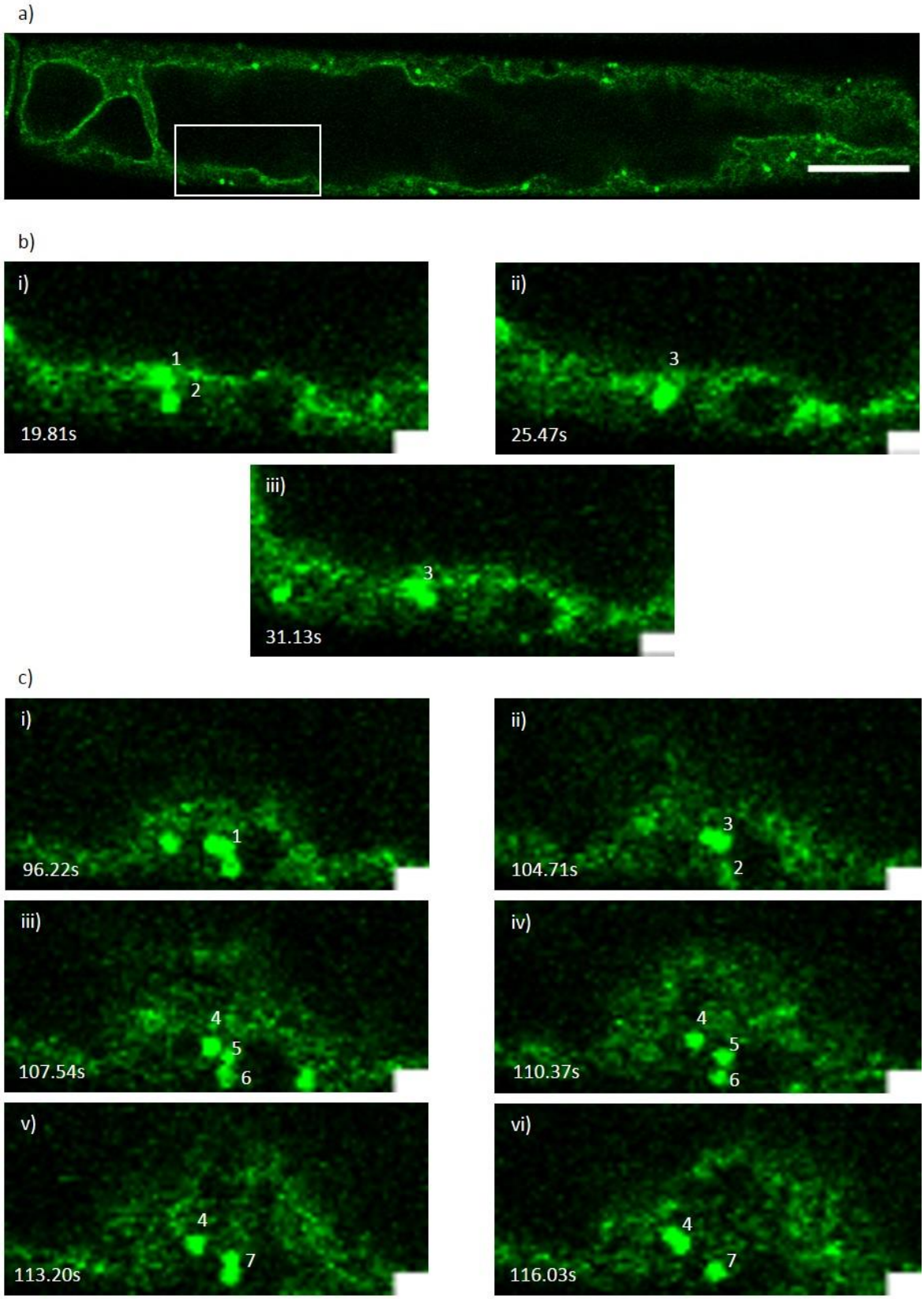


Figure 5.9: Dynamic behaviour of punctae in *Arabidopsis thaliana* plant roots expressing the *promIRQ4*::IRQ4-GFP construct. a) Overview of single plant root epidermal cell used for time series images in b-c). White box indicates area used for images in b-c). b) Punctae (i; 1 and 2) will meet and seem to partially fuse and form hybrid compartments (ii, iii; 3). c) Punctae are seen fusing together forming hybrid compartments (i; 1) before one compartment breaks away (ii; 2). The single compartment (ii; 3) seems to break up into 3 separate compartments (iii-iv; 4, 5, 6) before compartments 5 and 6 form another hybrid compartment, 7 (v-vi). Scale bars: a) 10 μm, b-d) 1 μm.

5.4.1 Analysis of *promIRQ4*::IRQ4-GFP localisation in *Arabidopsis thaliana* plant roots coexpressing endomembrane markers

To identify the endomembrane compartments which the *promIRQ4*::IRQ4-GFP localises to, the stably expressing plants were dipped with plasmids containing markers for the endoplasmic reticulum (HDEL-RFP), the Golgi apparatus (ST-RFP), and the vacuole (VAMP711-MCherry). The T_3 generation of *promIRQ4*::IRQ4-GFP plants were grown until flowering and transformed using the floral dip method as described in section 2.2.14. T_0 seed were grown on selective media (Kanamycin for ST-RFP and hygromycin for HDEL-RFP coexpressing plants) or were sown on soil and sprayed with BASTA (For VAMP711-MCherry containing plants) according to section 2.6.4. All plants were screened for fluorescence according section to 2.6.4 and presence of the insert was confirmed by PCR using primers listed in appendix 1. One more round of selection was undertaken until T_2 seed was generated. This seed was used for imaging experiments. The *promIRQ4*::IRQ4-GFP plants dipped with VAMP711-MCherry containing plasmid were not transformed successfully so were not used in this analysis.

5.4.1.1 Analysis of IRQ4 localisation in Arabidopsis thaliana plants coexpressing promIRQ4::IRQ4-GFP with ST-RFP Golgi marker

Plants were grown vertically on solid 1XMS plates as described in section 2.6.3 for 5-7 days before being mounted in liquid 1XMS and imaged via CLSM. ST-RFP was expressed in all cell files but only in the lower meristematic zone of the root and not in the upper meristematic zone or into the elongation zone. IRQ4 punctae were found in lower abundance in this zone of the root (Figure 5.10) but where they were present they did not colocalise or associate with the ST-RFP labelled Golgi compartments.

5.4.1.2 Analysis of IRQ4 localisation in Arabidopsis thaliana plants coexpressing promIRQ4::IRQ4-GFP with HDEL-RFP ER marker

Plants were grown vertically on solid 1XMS plates as described in section 2.6.3 for 5-7 days before being mounted in liquid 1XMS and imaged via CLSM. The HDEL-RFP construct was not strongly expressed within the root and only at the root tip in the lower meristematic zone of the root. The inset in Figure 5.11b shows the IRQ4 fusion protein that localises to the larger vacuole type compartment does not colocalise with the HDEL-RFP labelled ER. Within roots, the ER can have a different appearance compared to aerial tissues with thicker strands being formed. These fibres are typically induced due to stress. Even careful removal of the plant from the agar and placement on the slide for imaging caused the fibres to appear. To remove this effect plants would need to be grown and imaged in such a way that they were not disturbed during the imaging process. This

could be achieved by growing the plant in capillary tubing and imaging using lightsheet microscopy, a technique which allows the plant to be imaged without removing it from the capillary in which it is growing.

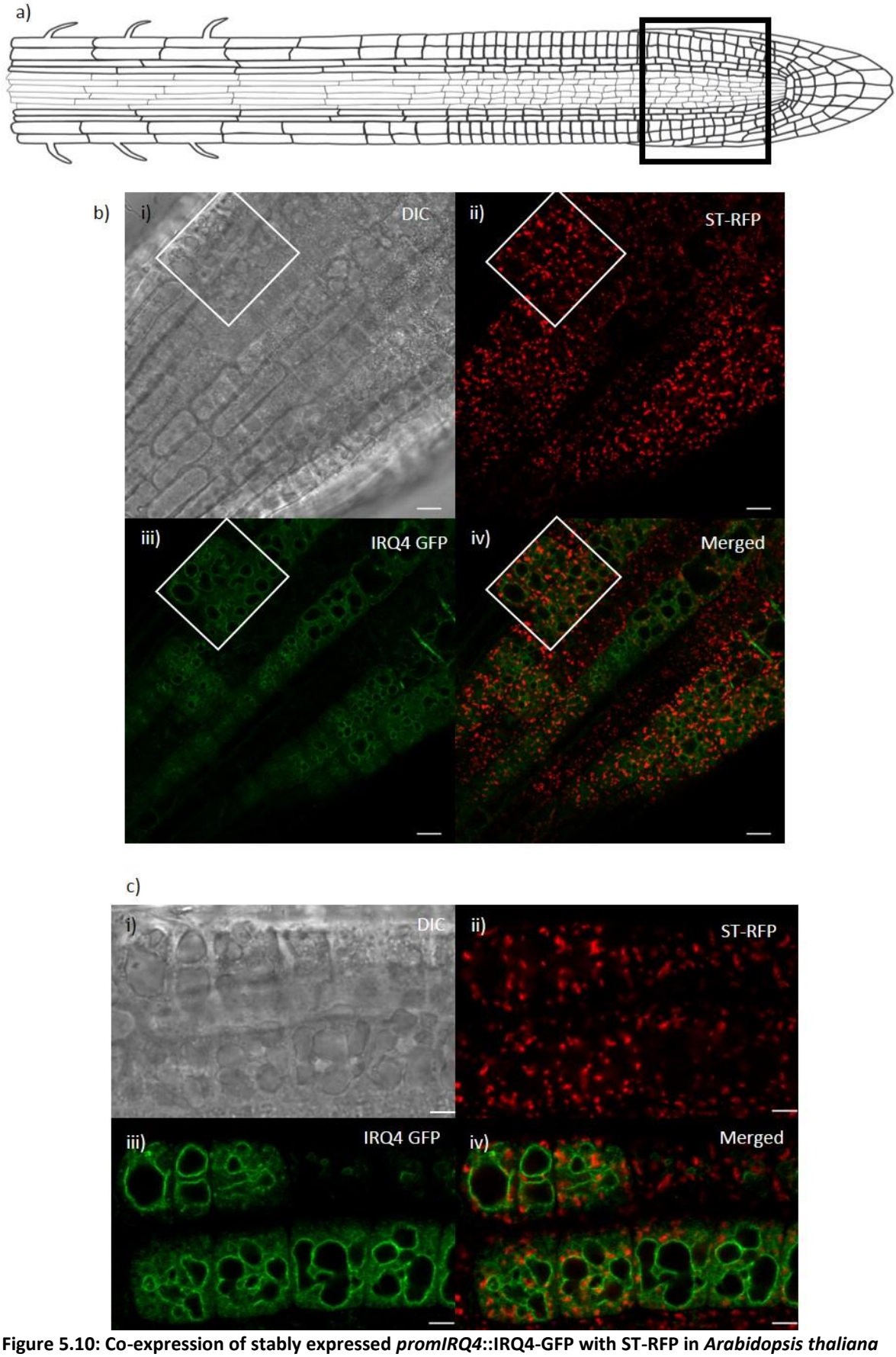


Figure 5.10: Co-expression of stably expressed *promIRQ4*::IRQ4-GFP with ST-RFP in *Arabidopsis thaliana* roots. 5-7 day old seedlings were grown vertically and mounted in water for imaging. a) Schematic showing area of lateral root (black box) used for imaging in b). b) Maximum projection of root and c) magnification and expansion of area bounded by white box in b). bi) and ci) brightfield images, bii) and cii) ST-RFP alone, biii) and ciii) *promIRQ4*::IRQ4-GFP alone, biv) and civ) merged image of ST-RFP and *promIRQ4*::IRQ4-GFP. Scale bars: bi-iv) 20 μm and ci-iv) 10 μm.

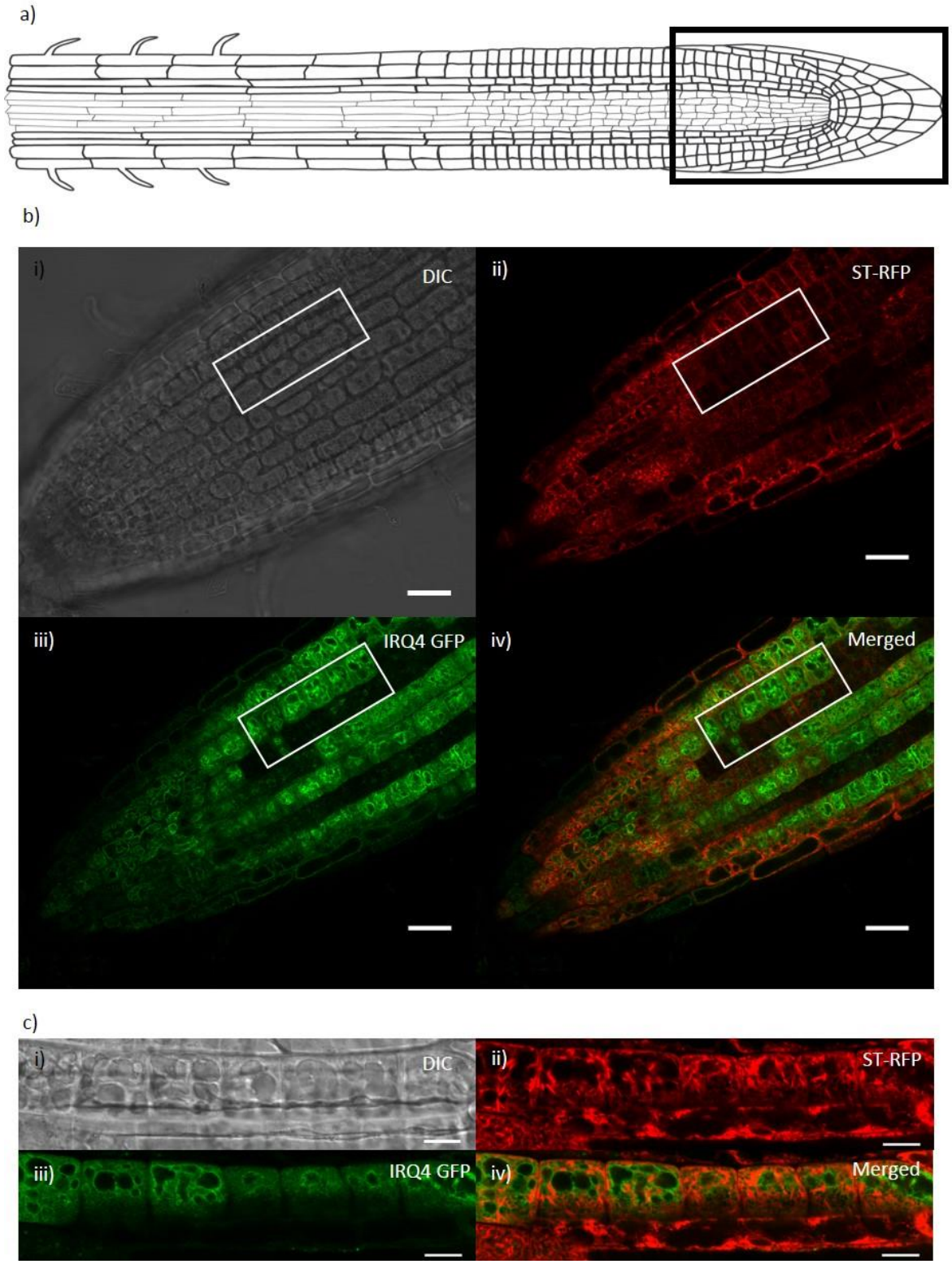


Figure 5.11: Co-expression of stably expressed *promIRQ4*::IRQ4-GFP with HDEL-RFP in *Arabidopsis thaliana* roots. 5-7 day old seedlings were grown vertically and mounted in water for imaging. a) Schematic showing area of lateral root (black box) used for imaging in b). b) Maximum projection of root and c) magnification and expansion of area bounded by white box in b). bi) and ci) brightfield images, bii) and cii) HDEL-RFP alone, biii) and ciii) *promIRQ4*::IRQ4-GFP alone, biv) and civ) merged image of HDEL-RFP and *promIRQ4*::IRQ4-GFP. Scale bars: bi-iv) 20 µm and ci-iv) 10 µm.

5.4.1.3 Localisation of IRQ4 in promIRQ4::IRQ4-GFP stably expressing Arabidopsis thaliana plants labelled using the vacuole staining dye neutral red

As the plants co-expressing the vacuole marker VAMP711-MCherry had failed to be transformed neutral red was used to stain the vacuole. Neutral red is a lipophilic dye which in its unprotonated form penetrates the plasma membrane and the tonoplast membrane. Due to protonation it then becomes trapped within acidic compartments and accumulates in the vacuole (Dubrovsky *et al.,* 2006). Staining is relatively fast and its low toxicity makes it useful for vacuole staining.

Plants were grown vertically on solid 1XMS plates as described in section 2.6.3 for 5-7 days before being treated with 4 µM neutral red dissolved in water (pH 5.8) for 30 minutes and mounted in the neutral red solution for imaging by CLSM. The neutral red dye accumulates in the vacuoles (Figure 5.12ai). Some neutral red is found in the cytoplasm but the majority is found in the vacuole using this particular pH for staining. The merged image shows the *promIRQ4*::IRQ4-GFP labels the tonoplast on the neutral red stained vacuoles. The tonoplast localisation can be more clearly seen in the insets in Figure 5.12b and c.

Using the localisation data from the co-expression studies and the dye it is possible to state that the IRQ4 punctae or larger vacuole type compartment do not colocalise with the Golgi. Both localisation patterns (punctate and vacuole-type) can only be stated to associate with the ER. The use of neutral red dye to stain the vacuole lumen has shown that IRQ4 localises to the tonoplast of the vacuole in the root.

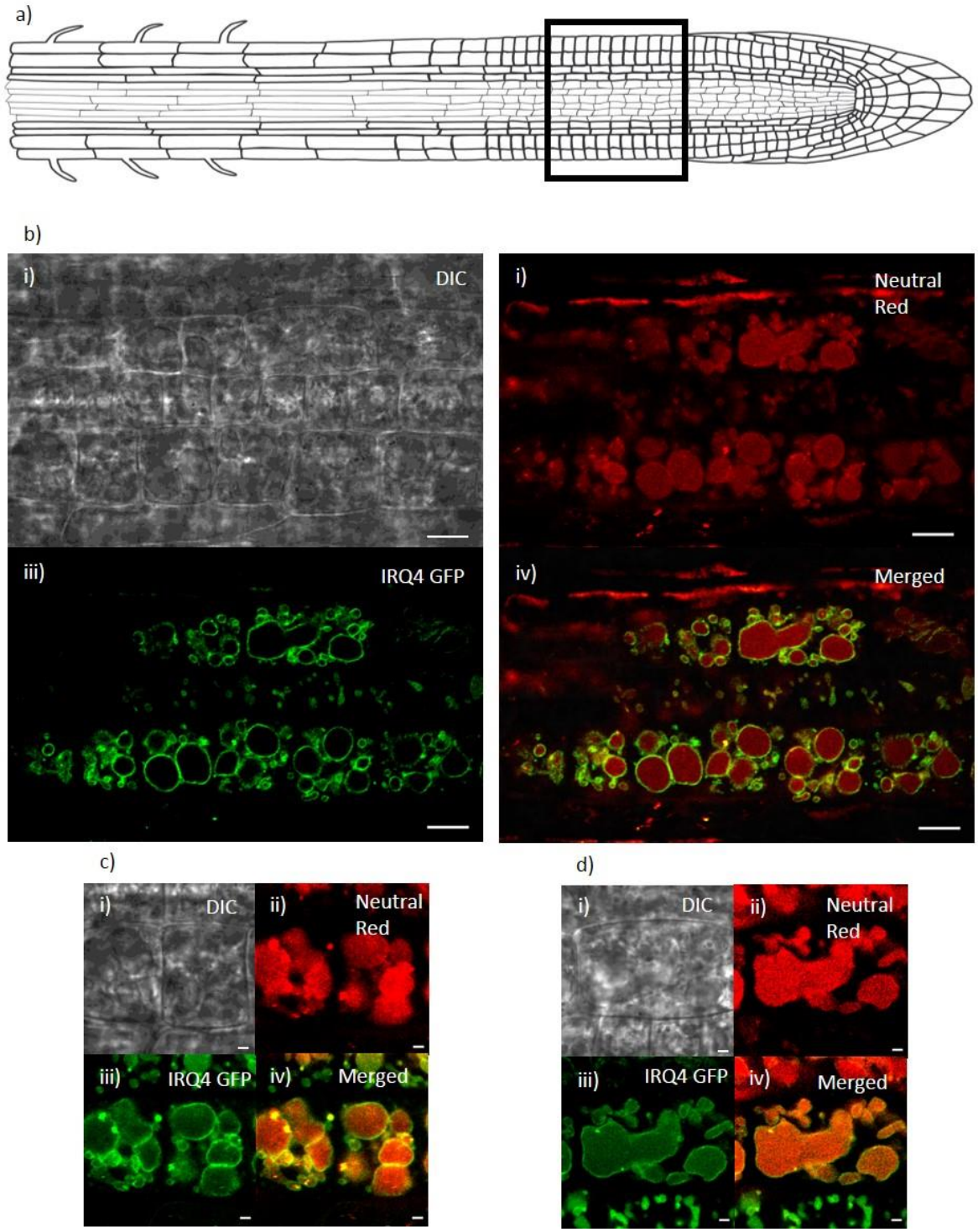


Figure 5.12: Labelling of stably expressing *promIRQ4*::IRQ4-GFP plants stained with neutral red dye in *Arabidopsis thaliana* roots. 5-7 day old seedlings were grown vertically, stained with 4µM neutral red in water (pH 5.8) for 30 mins and mounted in the neutral red solution for imaging. a) Schematic showing area of lateral root (black box) used for imaging in b). b) Maximum projection of root meristematic zone. c) and d) Magnification of cells from a meristematic region of roots. bi), ci), and di) brightfield images; bii), cii), and dii) neutral red dye alone; biii), ciii), and diii) *promIRQ4*::IRQ4-GFP alone; biv), civ), div) merged image of neutral red and *promIRQ4*::IRQ4-GFP. Scale bars: bi-iv) 10 µm, ci-iv) and di-iv) 2 µm.

5.4.2 Drug studies to assess the localisation and behaviour of IRQ4 in *promIRQ4*::IRQ4-GFP stably expressing *Arabidopsis thaliana* plants

To verify the results obtained in BY2 cells expressing the IRQ4-GFP construct the *promIRQ4*::IRQ4-GFP stably expressing *Arabidopsis thaliana* plants were subjected to treatment with the same drugs including BFA, wortmannin, oryzalin, and latrunculin B.

5.4.2.1 Treatment of stably promIRQ4::IRQ4-GFP stably expressing Arabidopsis thaliana plants coexpressing ST-RFP with BFA

Plants were grown vertically on solid 1XMS plates as described in section 2.6.3 for 5-7 days before being transferred to 1XMS liquid media treated with either DMSO (control) or 10 µM BFA for one hour (according to section 2.5.4). During treatment plants were kept in normal growing conditions as described in section 2.6.3. Images were taken every 15 minutes. For the washout images, plants which had been treated with DMSO or BFA for 60 minutes were transferred to liquid 1xMS containing no DMSO or BFA and left shaking for 1 hour before imaging (washout, WO).

IRQ4 does not change localisation during or after treatment with BFA (Figure 5.13). The IRQ4 PVC have a similar distribution after 60 minutes in BFA and after WO compared to the beginning and compared to the DMSO control. The latter finding indicates that the *promIRQ4*::IRQ4-GFP construct is not affected by the time course of the experiment or the level of DMSO used. The ST-RFP construct was used as a control to show that the concentration of BFA used caused BFA-compartment formation within the root cells (Takáč *et al.*, 2011). The ST-RFP can be seen to form aggregates after 45 minutes, forming BFA compartments and in the merged images the IRQ4 does not preferentially localise to these BFA bodies (Figure 5.13d-e). Figure 5.13f shows that after WO the ST-RFP returns to a normal construct and that the washout treatment was sufficient to remove the BFA from the cell.

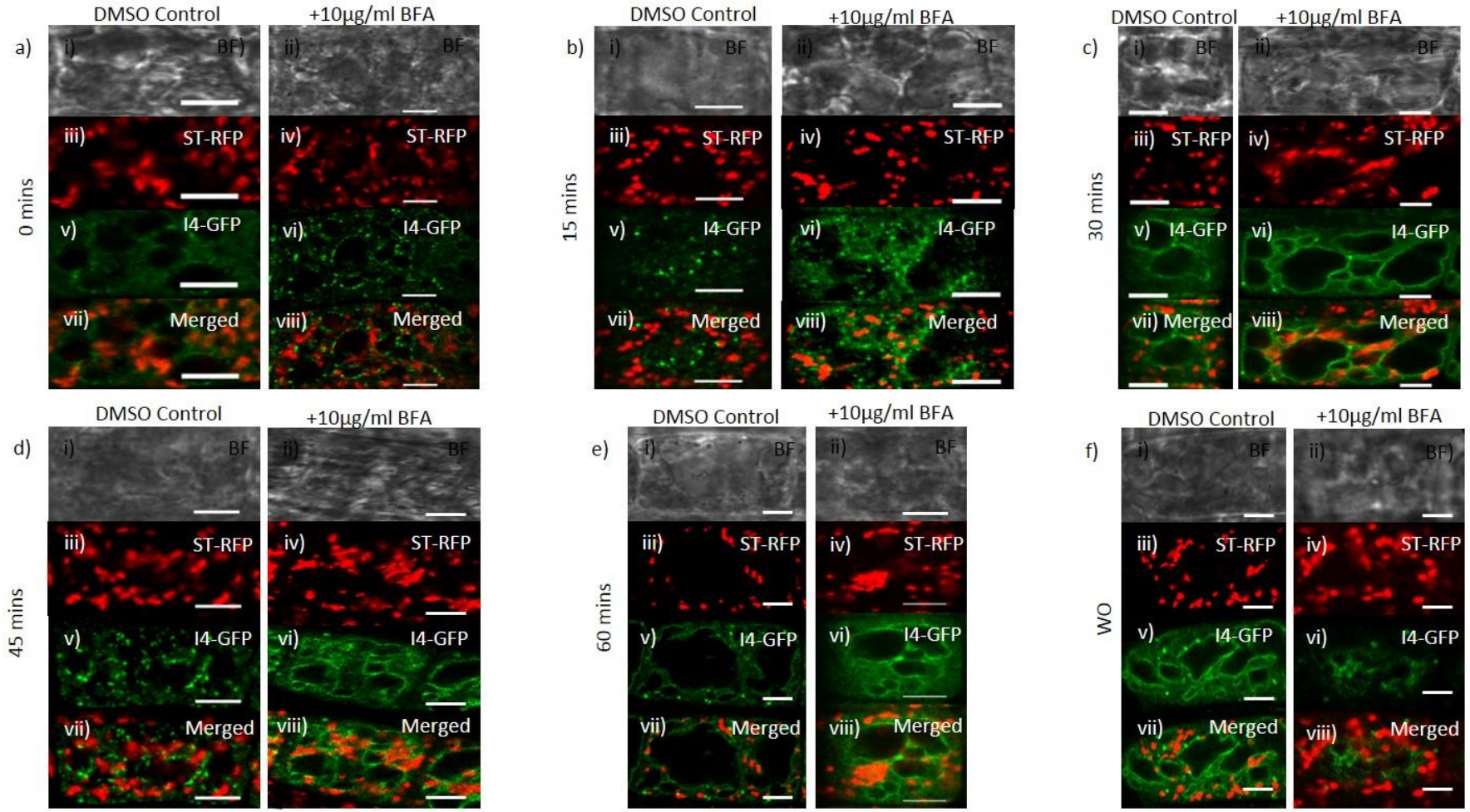


Figure 5.13: BFA treatment has no effect on the localisation of the *promIRQ4*::IRQ4-GFP in stably expressing *Arabidopsis thaliana* plants. 5-7 day old stably expressing *promIRQ4*::IRQ4-GFP +ST-RFP seedlings were treated with either DMSO (a-, b-, c-, d-, e-, f-i, iii, v, vii) or with 10µg/ml BFA (a-, b-, c-, d-, e-, fii, iv, vi, viii). Brightfield images (a-, b-, c-, d-, e-, fi, ii) with their corresponding ST-RFP (a-, b-, c-, d-, e-, f- iii, iv), I4-GFP (a-, b-, c-, d-, e-, f-, v, vi), and merged images (a-, b-, c-, d-, e-, f- vii, viii) Images were taken every 15 minutes for 1 hour. Time points are indicated to the side of images. Scale bars: a-, b-, c-, d-, i, iii, v, vii) 5 µm; e-, f-, i, iii, v, vii) 10 µm; a-, b-, c-, d-, e-, f-, ii, iv, vi, viii) 5 µm.

5.4.2.2 Treatment of stably expressing promIRQ4::IRQ4-GFP Arabidopsis thaliana plants with wortmannin

Plants were grown vertically on solid 1XMS plates as described in section 2.6.3 for 5-7 days before being transferred to 1XMS liquid media treated with either DMSO (control) or 30 µM wortmannin for three hours (according to section 2.5.4). Plants were incubated under normal growth conditions (section 2.6.3) during the incubation with the drug. Plants were then imaged using CLSM. To see if the behaviour of the punctae was different when treated with wortmannin in more mature tissues compared to less developed tissues images were taken at the root tip/meristematic zone (Figure 5.14b-, c-i, iii) and in the elongation/differentiation zone (Figure 5.14b-, c-ii, iv).

The DMSO control shows that the length of time the treatment took place over and the level of DMSO used had no effect on the appearance of the punctae (Figure 5.14bi-v). Treated plants showed a similar phenotype to that of the BY2 cells shown in Figure 5.4.2.2c and d with swollen vesicular compartments which cluster together. The behaviour of the punctae was the same in the treated plants between younger root cells and those cells which were more developed which may indicate that all the punctae labelled by the *promIRQ4*::IRQ4-GFP construct are PVC/MVB regardless of tissue developmental stage.

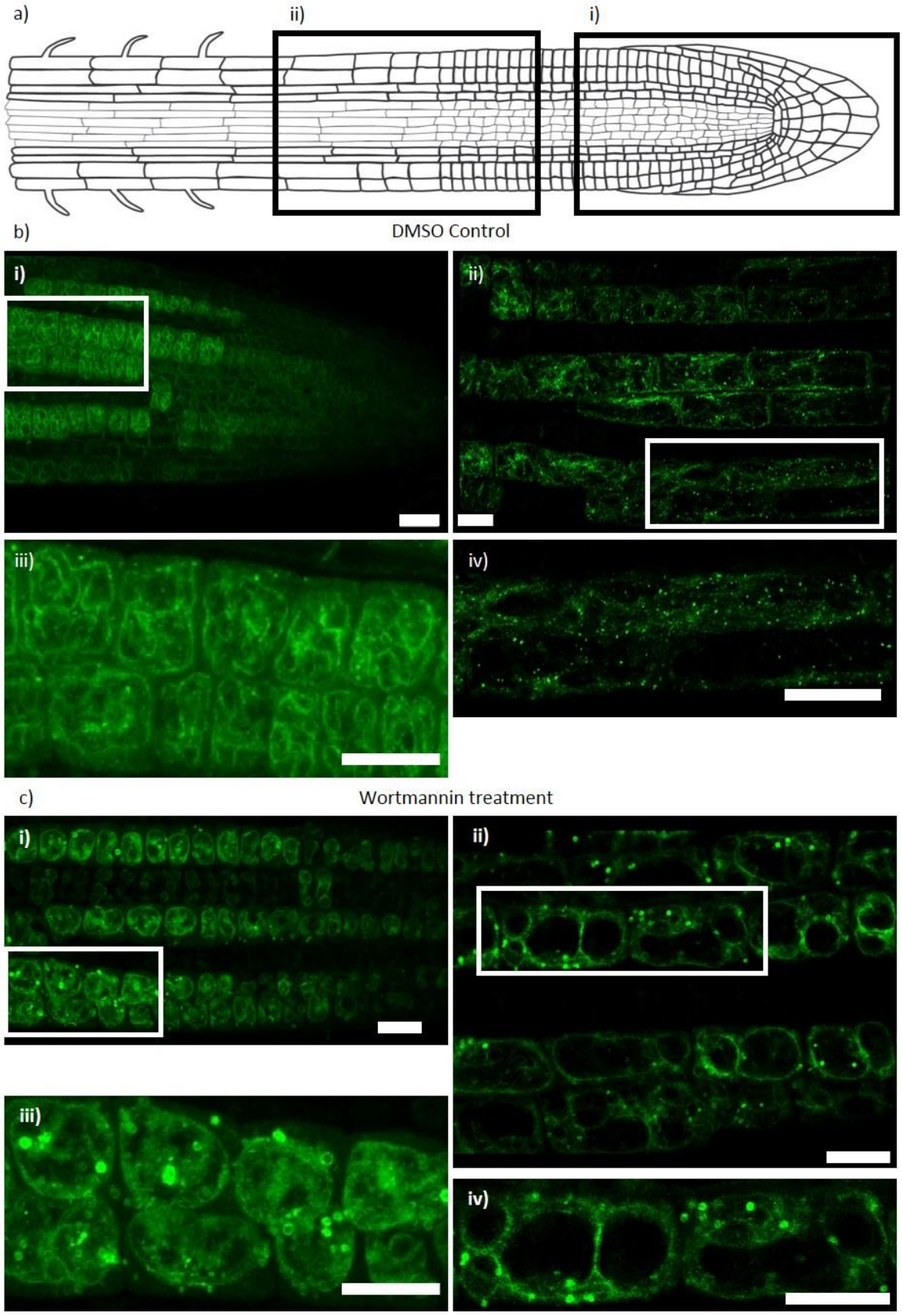


Figure 5.14: Effect of wortmannin treatment on the localisation of the *promIRQ4*::IRQ4-GFP construct. Five to seven day-old *Arabidopsis* seedlings were treated with either DMSO (Control; b) or 30 µM wortmannin (c) for 3 hours before being mounted in water and the roots imaged via CSLM. Images for bi, iii, ci, ciii were taken from region i) as indicated in a). Images for bii, iv, cii, iv were taken from the region indicted in aii). Areas in white boxes are shown magnified below their respective images. Scale bars: bi, bii, ci, and cii) 20 µm; biii, biv, ciii, civ) 10 µm.

5.4.2.3 Treatment of stably expressing promIRQ4::IRQ4-GFP Arabidopsis thaliana plants with oryzalin

Plants were grown vertically on solid 1XMS plates as described in section 2.6.3 for 5-7 days before being transferred to 1XMS liquid media treated with either DMSO (control) or 37.5 µM oryzalin for 20 mins. Plants were incubated under normal growth conditions (section 2.6.3) during the incubation with the drug. Plants were imaged using CLSM with single cells being selected from the late meristematic zone into the differentiation zone. Cells were imaged in a single focal plane for 200 frames according to section 2.7.2. The speed of the PVC/MVB compartments within each cell was calculated using the spot tracking plugin in the ICY software (Chaumont *et al.*, 2012) as described in section 2.7.2. The average speed of the control and treated cells and a student's t test was calculated to compare the means according to section 2.7.2.

The mean speed of PVCs in oryzalin treated cells was not significantly different (0.33 μms^{-1}, s.d.=0.16, 95 % CIs: 0.26 to 0.40) compared to the DMSO treated control cells (0.32 μms^{-1}, s.d.=0.18, 95 % CIs: 0.25 to 0.40; t(17)=-0.10, p=0.92; Figure 5.15a). This verifies the results obtained in the BY2 cells in section 5.3.2.3 that showed the movement of the PVC/MVB compartments is not dependent upon an intact microtubule cytoskeleton.

5.4.2.4 Treatment of stably expressing promIRQ4::IRQ4-GFP Arabidopsis thaliana plants with latrunculin B4

Plants were grown vertically on solid 1XMS plates as described in section 2.6.3 for 5-7 days before being transferred to 1XMS liquid media treated with either DMSO (control) or 4 µM latrunculin B for one hour (according to section 2.5.4). Plants were incubated under normal growth conditions (section 2.6.3) during the incubation with the drug. Plants were imaged using CLSM with single cells being selected from the late meristematic zone into the differentiation zone. Cells were imaged in a single focal plane for 200 frames according to section 2.7.2. The speed of the PVC compartments within each cell was calculated using the spot tracking plugin in the ICY software (Chaumont *et al.*, 2012) as described in section 2.7.2. The average speed of the control and treated cells and a student's t test was calculated to compare the means according to section 2.7.2.

The mean speed of PVCs in latrunculin B treated cells was significantly lower (0.31 μms^{-1}, s.d.=0.01, 95 % CIs: 0.27 to 0.35) compared to the DMSO treated control cells (0.58 μms^{-1}, s.d.=0.10, 95 % CIs: 0.53 to 0.63; t(45)=8.86, p<0.001; Figure 5.15b). The behaviour of the PVCs when treated with latrunculin B changes from mostly directed movement in control cells to a prevalence of undirected, Brownian motion in treated cells (Supplemental movies S3 (control) and

S4 (treated)). This result supports the drug studies in the heterologous BY2 system and indicates that the movement of the PVCs is dependent upon the actin cytoskeleton.

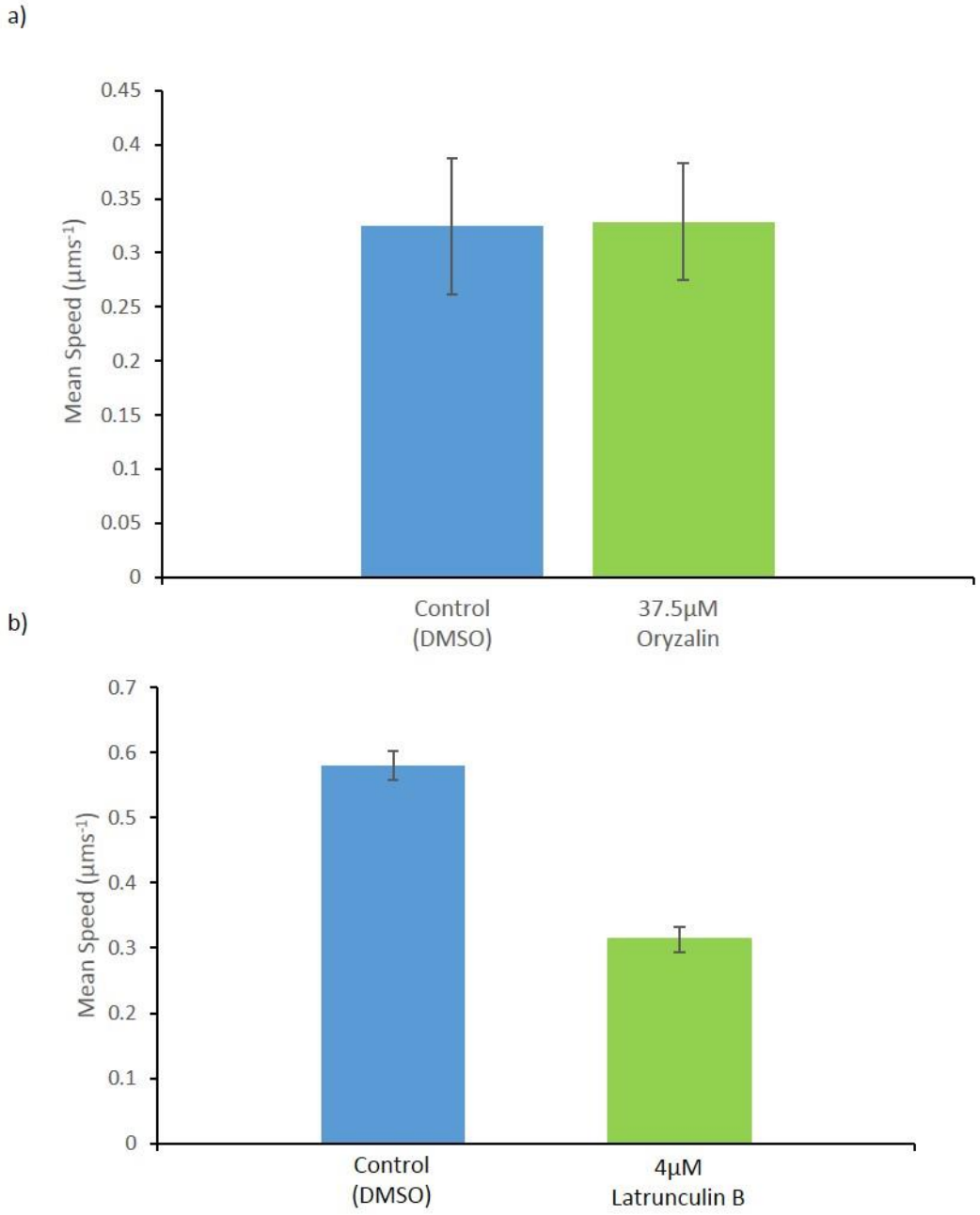


Figure 5.15: Effect of oryzalin and latrunculin B treatment on the speed of punctae in Arabidopsis plants expressing IRQ4. a) 5-7 day old seedlings stably expressing the *promIRQ4*::IRQ4-GFP construct were treated with either 34.5 µM oryzalin or DMSO (control) for 1.5 hours. b) 5-7 day old seedlings stably expressing the *promIRQ4*::IRQ4-GFP construct were treated either with 4 µM latrunculin B or DMSO (control) for 1 hour. In a) and b) cells were imaged using CLSM in a single focal plane for 200 frames according to section 2.7.2. The speed of the punctae within each cell was calculated using the spot tracking plugin in the ICY software (Chaumont *et al.*, 2012) as described in section 2.7.2. The average speed of punctae in the control and treated cells and a student's t-test was calculated to compare the means according to section 2.7.2. Similar to the results obtained in the heterologous BY2 system, treatment with oryzalin does not have a significant effect on the speed of the punctae. However, disrupting the actin cytoskeleton using latrunculin B causes a significant decrease in their movement. Bars on graph indicate S.E.M.

5.5 Analysis of IRQ4 Expression in *Arabidopsis thaliana* using polyclonal antibodies

Experiments in sections 5.3 and 5.4 have demonstrated that IRQ4 localises to the vacuole and a prevacuolar-type compartment. To confirm the localisation found using the fluorescent fusion proteins polyclonal antibodies were raised to IRQ4. A protein fragment of interest was cloned and expressed in bacterial cells before being purified and used to inoculate a host species. Antibodies are generated by the animal in response to the inoculation. These antibodies are polyclonal as different types of antibodies are generated by the host species which are capable of binding to different epitopes of the protein. Once purified this blood sera can be used as a primary antibody to bind to the protein of interest in appropriately fixed plant tissue. For visualisation a secondary antibody, either a fluorochrome-conjugated or gold-conjugated secondary antibody can be used. The subcellular localisation of IRQ4 was further characterised through immunogold labelling and Transmission Electron Microscopy. The creation of the antibody also allowed the translational profile of IRQ4 to be examined using Western blotting.

5.5.1 Identification of the IRQ4 antigen fragment

To generate polyclonal antibodies animals, typically mice or rabbits, are inoculated with an antigen designed to target a specific protein. The antigen causes an immune response in the animal leading to the production of antibodies against the target protein. The antibodies are harvested form the blood serum of the animal. The specificity of an antibody is critical as non-specific binding causes an increase in background signal in Western blots and imaging through binding to non-target proteins.

To ensure that the antibody was as specific as possible the sequences of the IRQ proteins were aligned in ClustalX 2.1 and exported to Jalview for viewing. The region shown in Figure 5.16a corresponds to a region with the lowest similarity between IRQ4 and the other IRQ proteins. The high degree of homology between IRQ4 and IRQ5 as shown by the dot plot shown in Figure 5.16b meant the region chosen had to be as specific to IRQ4 as possible. As the N-termini of IRQ4 and IRQ5 have some limited similarity to the NAB domain and the C termini contain the IRQ domain these conserved areas were avoided. Alignment of IRQ4 and IRQ5 showed there was only one region (Figure 5.17a) which satisfied the criteria used for antigen design in this project: an area of low hydrophobicity to increase the solubility of the antigen during inoculation and illicit an immune response in the animal, at least 50 amino acids in length, and contain no more than 1 run of 8 amino acid residues which are conserved between IRQ4 and IRQ5. The hydropathy plot of both the full length IRQ4 (Figure 5.17d) show the low hydrophobicity of the IRQ4$^{125-248}$ fragment chosen (Figure 5.17b and c) and therefore its suitability for raising an antiserum.

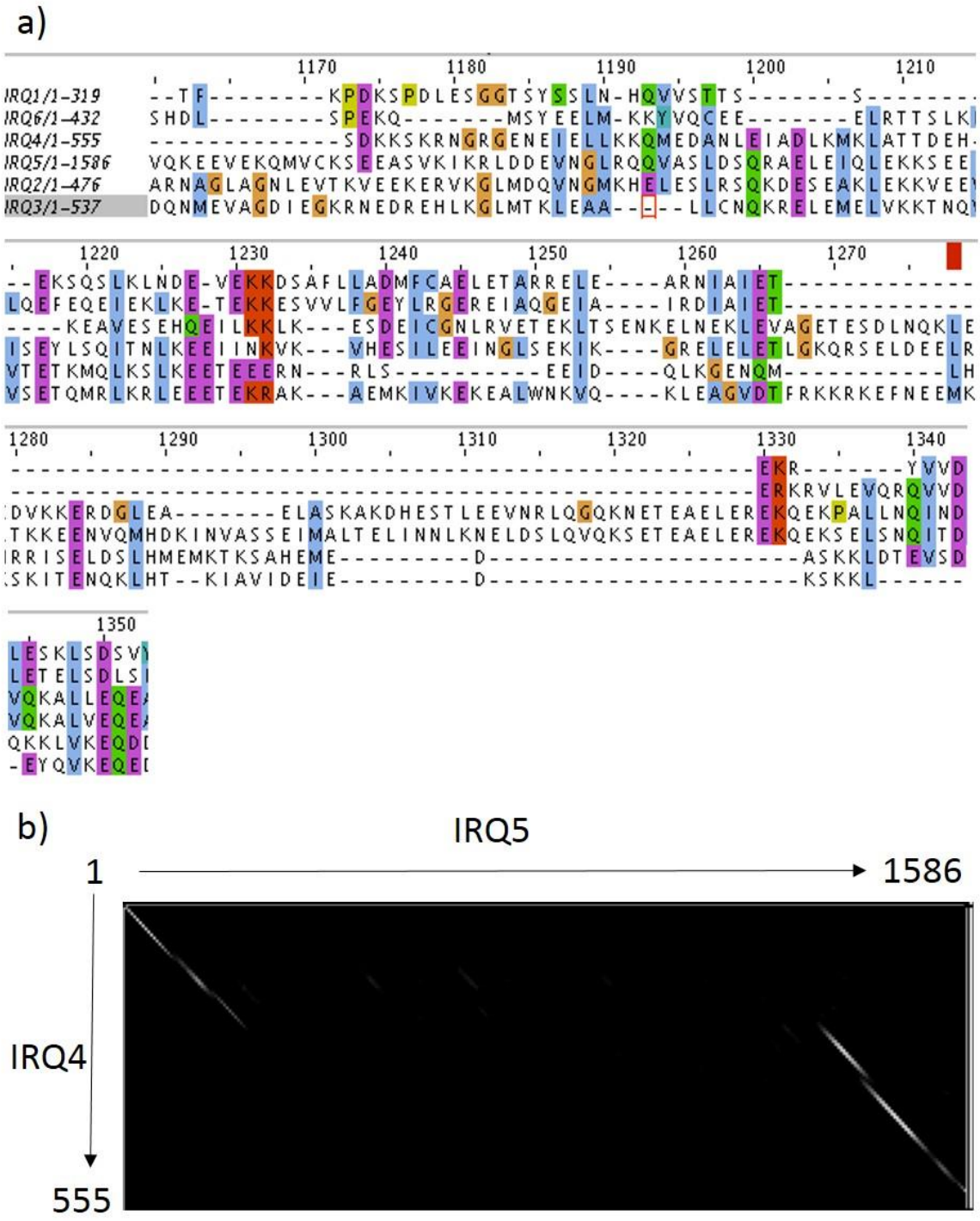


Figure 5.16: Identifying the IRQ4 antigen fragment. a) Alignment of the IRQ protein sequences highlighted a region which avoided the pNAB domains of IRQ5 and IRQ5 and also the IRQ domains of the IRQ proteins. b) Dot plot showing the similarity in amino acid composition between IRQ4 and IRQ5. White dots represent similar amino acid positions. There is a large central region within IRQ5 which does not share any similarity to IRQ4.

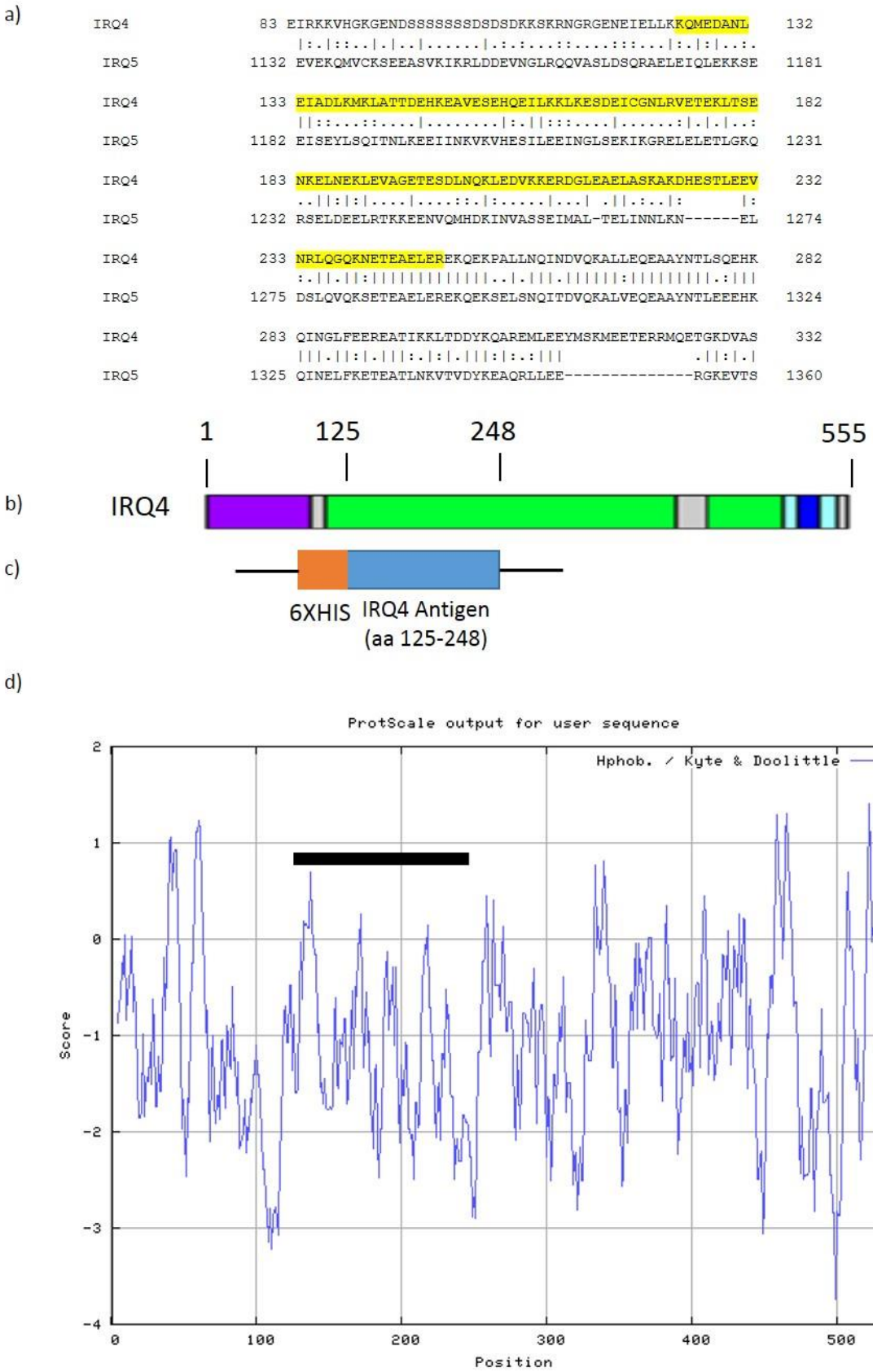


Figure 5.17: Ensuring specificity of the IRQ4 antigen fragment compared to IRQ5. a) Alignment of IRQ4 and IRQ5 showing the region selected for the creation of the IRQ4 antigen between 125-248 amino acids. b) Schematic of the IRQ protein showing the location of the IRQ4 antigen fragment. c) Schematic of the IRQ4$^{125\text{-}248}$ construct fused to the 6XHIS tag. d) Hydropathy plot showing the region of the IRQ4$^{125\text{-}248}$ antigen fragment (black bar). The generally negative hydropathy values indicate the suitability of the fragment for raising an antiserum.

5.5.2 Cloning of the IRQ4$^{125-248}$ for recombinant protein production

The IRQ4$^{125-248}$ fragment was amplified from the full length IRQ4 sequence in the pMDC83 vector using primers listed in appendix 1. The fragment was then cloned into the pDONR207 entry vector using the Gateway system (according to 2.2.3). The fragment was then cloned into from the pDONR207 vector into the pGAT4 expression vector for expression in *E. coli*. The pGAT4 expression vector used here has been modified by Dr. Tijs Ketelaar (Wageningen University & Research Centre) using the pET-17 backbone with the addition of the multiple cloning site of pET-28a and a Gateway cassette to make it Gateway compatible. The vector contains an N-terminal 6xHistidine tag, allowing purification of the antigen fragment using nickel-nitrilotriacetic acid (Ni-NTA) beads.

The 6xHis-IRQ4$^{125-248}$ antigen fragment was expressed in *E. coli* Rosetta 2 cells as a histidine-tagged fusion protein using auto-induction according to the method described in section 2.4.1. The 6XHis-IRQ4$^{125-248}$ antigen fragment was purified using a column containing nickel nitrilotriacetic acid (Ni-NTA) agarose beads (Qiagen) in section 2.4.1 and dialysis using buffer exchange to remove urea as described in section 2.4.2. The Bradford assay (Bio-Rad) was used to quantify the protein.

The purified fragments were run on a 15 % SDS-PAGE gel to confirm the molecular weight of the protein (Figure 5.18). The predicted molecular weight of the fragment was 14.2 kDa. Shown in the Figure 5.18 is the successful induction showing an enlarged band at approximately 16 kDa in the induced sample (AIM) compared to the non-induced (T_0) sample. The induced protein is slightly larger than that of the predicted molecular weight due to the presence of the 6XHis tag. After purification both sets of pooled fractions (F_{12} and F_{36}) produce a clear band of the same size corresponding to the size of the enlarged band in the induced sample (AIM; Figure 5.18).

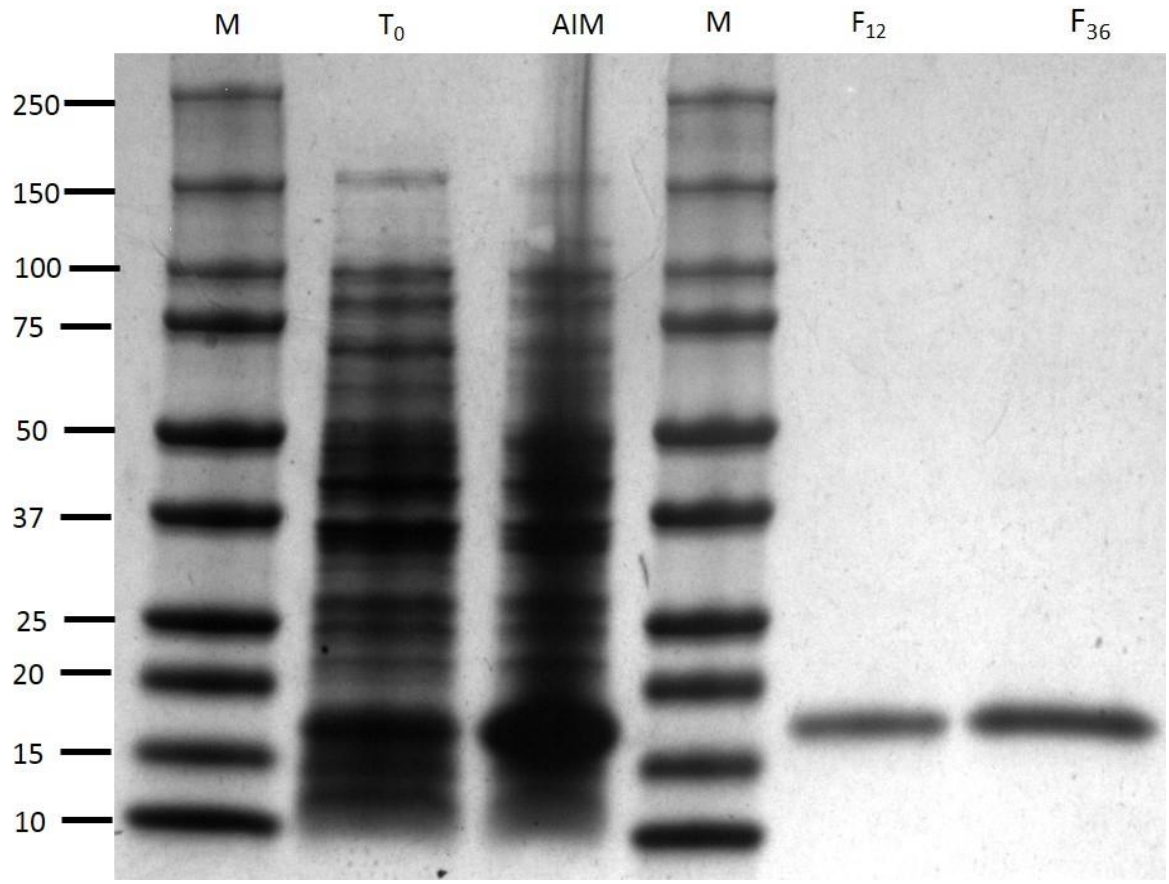


Figure 5.18: Purification of the IRQ4 antigen fragment (6XHis-IRQ4$^{125-248}$). The IRQ4 antigen fragment was expressed using an autoinduction method as a 6xHistidine tagged fusion protein and purified from E. coli Rosetta 2 cells using a nickel-nitrilotriacetic acid column, followed by overnight dialysis into PBS. Shown here is a crude *E.coli* extract before induction (T_0) and after induction (AIM). The pooled fractions (1 and 2, and 3 to 6) are shown (F_{12}, F_{36}) using 1µg of protein per lane. M= PageRuler Plus Prestained Protein ladder (Thermo Scientific).

5.5.3 Production and validation of the anti-IRQ4 polyclonal antibodies

Polyclonal antibodies to the IRQ4$^{125-248}$ antigen fragment were raised in mice. As variation can occur in the individual immune response of each host organism, three different mice (M1, M2, and M3) were inoculated with the same protein fragment (as described in 2.4.4). The specificity of the antibodies raised in these mice was assessed using Western blot analysis. Total protein extracted from 15-day old vertically grown wild-type and *promIRQ4*::IRQ4-GFP *Arabidopsis* plants was separated on a 10 % SDS-PAGE gel and transferred to a nitrocellulose membrane (according to 2.4.3 and 2.4.8). The sera derived from the terminal bleeds were initially used to probe the membrane at dilutions of 1:1 00 and 1:500 (Figure 5.19). In addition, an anti-GFP antibody at a dilution of 1:200 (Chromotek) was used as a control.

Both M1 and M2 do not produce a clear band and give a smear at both dilutions. M2 produces a large band at approximately 150 kDa but this is far too large for IRQ4 so may represent a specific binding. M3 produces a clear band between 75 and 50 kDa, approximately corresponding to the

65 kDa of IRQ4. There is a large band produced at approximately 150 kDa in the 1:100 dilution similar to M1 and M2 but this band disappears at higher dilutions of 1:500 and :1000. The 65 kDa band is retained at these higher dilutions demonstrating the specificity of the antibody.

To further confirm the utility of the M3 terminal bleed a further western blot was undertaken using total protein extracted from 15-day old WT Col-0 and *promIRQ4*::IRQ4-GFP plants as described in section 2.4.6. The extracts were separated on a 10% SDS-PAGE gel and transferred to a nitrocellulose membrane (according to 2.4.3 and 2.4.8).

Using a dilution of 1:500 the anti-IRQ4 antibody recognised a band at approximately 65 kDa in WT plant extracts (Figure 5.20). In the *promIRQ4*::IRQ4-GFP extracts the antibody detected both the 65 kDa band but also another band at approximately 95 kDa. This would correspond to the IRQ4 protein fused to GFP as the GFP is approximately 30 kDa in size. Using the anti-GFP antibody (Chromotek) alone on the *promIRQ4*::IRQ4-GFP plant extracts produces a band at 95 kDa, matching the higher band produced when using the M3 terminal bleed. This indicates that the higher band corresponds to the IRQ4 protein and also that the protein expressed as a GFP fusion is actually IRQ4. The lack of bands in the anti-GFP only lane tested in WT shows that the anti-GFP antibody does not recognise any other proteins and is specific for the GFP tagged protein.

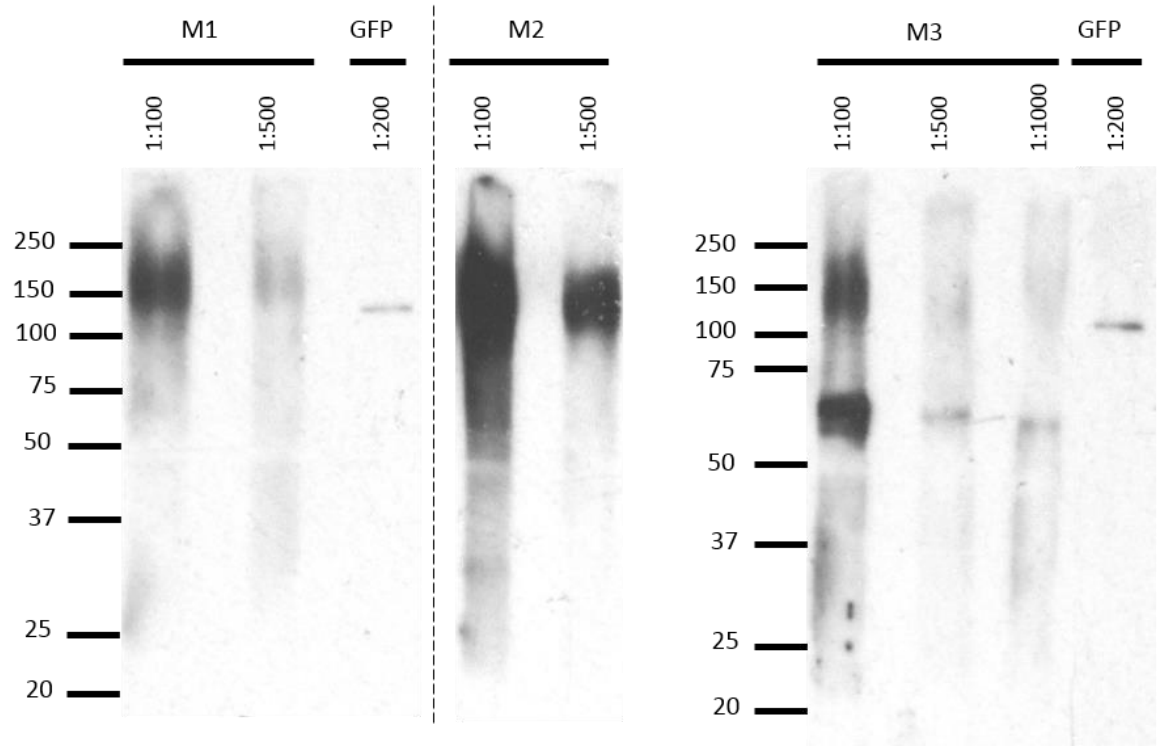


Figure 5.19: Western blot analysis of the mouse terminal bleeds. Total protein extracted from 15-day old wild-type *Arabidopsis* and *promIRQ4*::IRQ4-GFP plants was separated on a 10% SDS- PAGE gel and transferred to a nitrocellulose membrane for Western blot analysis. The membrane was cut into strips and probed with various dilutions of the M1, M2, and M3 mouse terminal bleeds (M1 and M2: 1:100 and 1:500, M3: 1:100, 1:500, 1:1000), followed by a secondary anti-Mouse/HRP antibody (1:1500; Dako). For the GFP lanes the anti-GFP antibody (Chromtek) was used at a dilution of 1:200 against the *promIRQ4*::IRQ4-GFP plant extracts followed by a secondary anti-rat/HRP antibody (1:1500). The M3 sera produced two one bands, one at 150 kDa and one at 65 kDa at the 1:100 dilution. This higher band disappeared at higher dilutions leaving a single clear band at 65 kDa, a size which corresponds to the predicted molecular weight of IRQ4. The M1 and M2 sera produced non-specific banding approximately 150 kDa so were not used in further experiments. The anti-GFP used as a control produced a band at approximately 115 kDa. The Precision Plus Protein Dual Colour (Bio-Rad) was used as a molecular weight standard. The dashed line indicates where the gel was cut to remove an empty lane from the image.

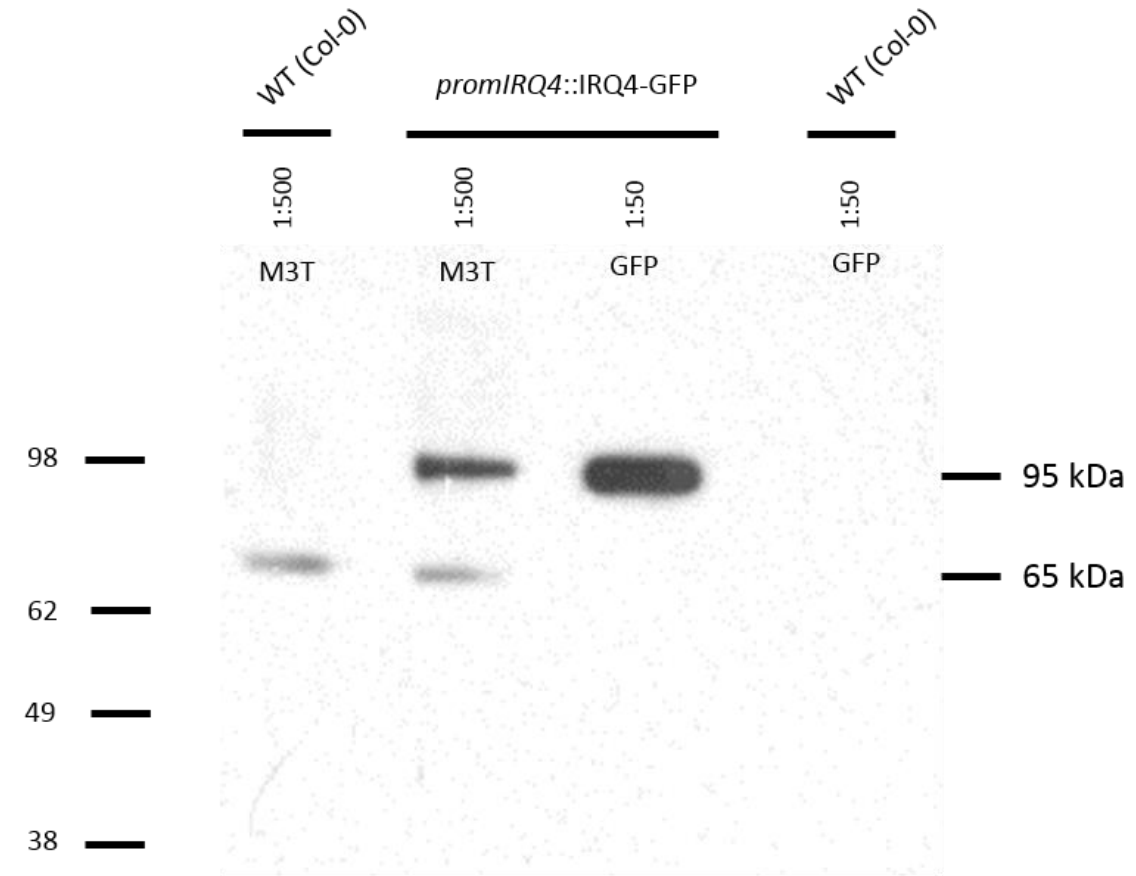


Figure 5.20: Further analysis of the M3 terminal bleed using western blotting. Total protein extracted from 15-day old wild-type Arabidopsis and *promIRQ4*::IRQ4-GFP plants was separated on a 10% SDS- PAGE gel and transferred to a nitrocellulose membrane for Western blot analysis. The membrane was cut into strips and probed either with the M3 terminal bleed at a dilution of 1:500, followed by a secondary anti-Mouse/HRP antibody (1:1500; Dako), or with an anti-GFP primary antibody at 1:50 (Chromtek) followed by a secondary anti-rat/HRP antibody (1:1500; Dako). The plant tissue used for each lane is indicated above the black lines. The primary antibody and dilution of each used for each lane is indicated below the black line. The M3 sera recognises a single band at 65 kDa, the predicted size of IRQ4, in the WT plant extracts. In the first lane of the *promIRQ4*::IRQ4-GFP plant extract the M3 recognises the 65 kDa wild type protein band as before but also a higher band at approximately 95 kDa. This corresponds to the IRQ4 protein with an additional GFP tag of approximately 30 kDa. The anti-GFP only lane using the *promIRQ4*::IRQ4-GFP plant tissue extract produced only the single 95 kDa band. The anti-GFP antibody does not produce non-specific bands as demonstrated by the lack of bands in the anti-GFP only lane using WT plant extract. The Precision Plus Protein Dual Colour (Bio-Rad) was used as a molecular weight standard.

To further validate that the specificity of the M3 sera for IRQ4, Western blot analysis was performed with the M3 pre-bleed sera obtained prior to inoculation with the purified IRQ4 antigen fragment (Figure 5.21). Equal amounts of total protein (50 µg) extracted from 15-day old vertically grown wild-type Arabidopsis plants (as described in 2.4.6) were separated on a 10 % SDS-PAGE gel and subsequently transferred onto a nitrocellulose membrane for Western blot analysis (as described in 2.4.3 and 2.4.8). Equal loading of the sample was confirmed through staining the membrane with ponceau. Probing the membrane with the M3 pre- and terminal bleed at a dilution of 1:500 produced a single band at approximately 65 kDa which could only be

detected using the terminal bleed and not the pre-bleed sera. This demonstrates that this band is detected by polyclonal antibodies produced in response to the IRQ4 antigen fragment.

Taken together Figures 5.19, 5.20, and 5.21 suggest that the M3 terminal bleed is specific for a protein of the correct approximate molecular weight of IRQ4 at 65 kDa. M3 was therefore used for all further experiments.

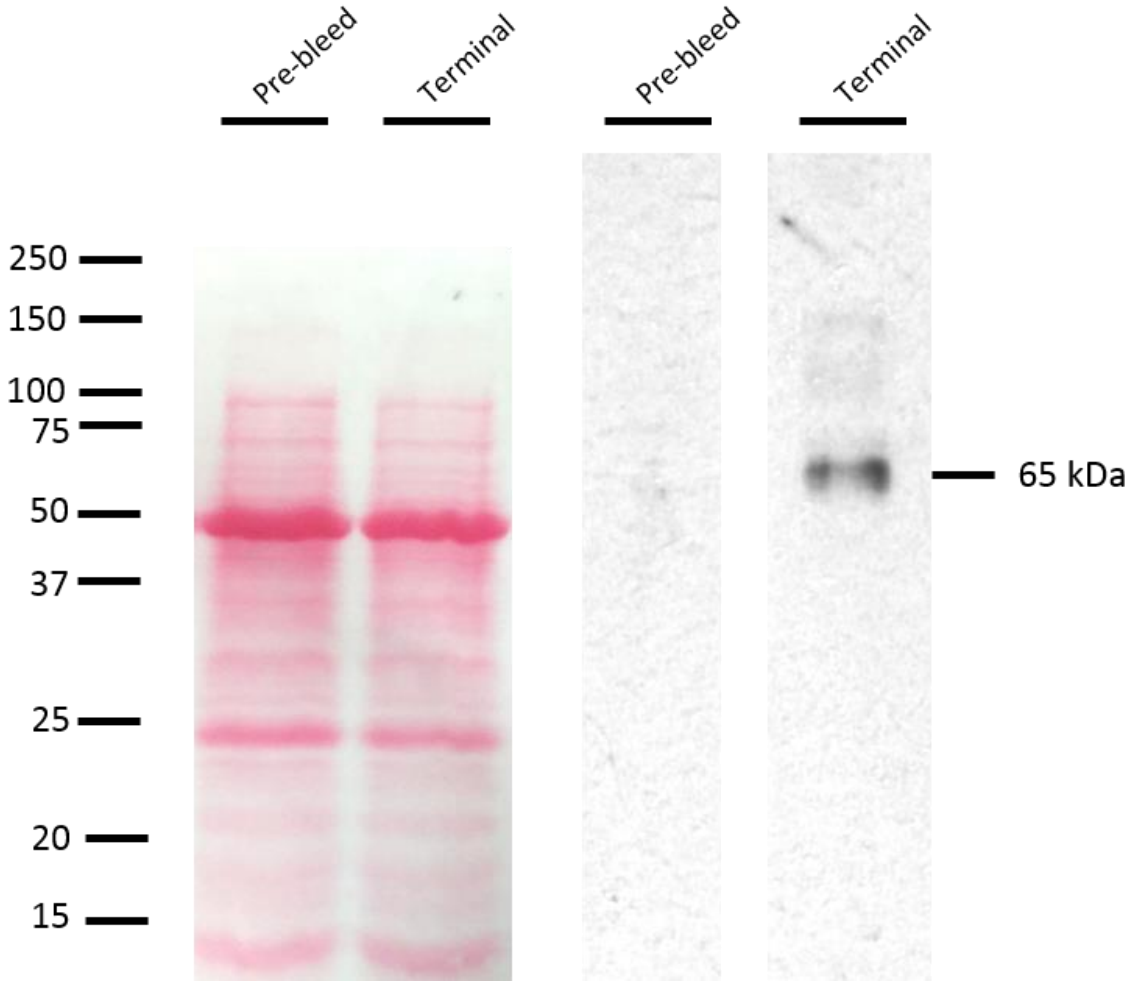


Figure 5.21: Western blot analysis of the M3 pre- and terminal-bleeds. Total protein extracted from 15-day old wild-type Arabidopsis and *promIRQ4*::IRQ4-GFP plants was separated on a 10 % SDS- PAGE gel and transferred to a nitrocellulose membrane for Western blot analysis. The membrane was cut into strips and probed either with the M3 pre-bleed sera or M3 terminal bleed sera at a dilution of 1:500, followed by a secondary anti-Mouse/HRP antibody (1:1500; Dako). The M3 terminal bleed only detected the 65 kDa band in the terminal-bleed sera and not the pre-bleed sera. This shows that the band produced is due to the inoculation of the mouse with the IRQ4 antigen. The Precision Plus Protein Dual Colour (Bio-Rad) was used as a molecular weight standard.

5.5.4 Analysis of IRQ4 expression in different *Arabidopsis* tissue using Western blot analysis

Using the specific antibodies generated in the previous section they were used to probe different tissues in *Arabidopsis* for the expression of the protein. Equal amounts of total protein (50 µg) extracted from 5-7 day old seedling root and cotyledon, 4-6 week old stem, mature leaf, flowers, and siliques (as described in 2.4.6) and were separated on a 10 % SDS-PAGE gel and subsequently transferred onto a nitrocellulose membrane for Western blot analysis (as described in 2.4.3 and 2.4.8).

Probing the membrane with the anti-IRQ4 antibody (1:500 dilution) showed that IRQ4 was present in the roots and of seedlings and flowers but was absent from the stem, mature leaf, and silique samples (Figure 5.22). There are bands present in the seedling root and seedling cotyledon lanes, as well as flowers. The presence of the protein in the roots verifies the localisation for IRQ4 found in the *prom*GUS expression lines and GFP fusions in stably expressing *Arabidopsis* plants. No GFP expression was seen in the cotyledon leaves and the *prom*GUS studies revealed IRQ4 was not expressed in these tissues at this stage of development. IRQ4 is known to have mRNA that is mobile (Thieme *et al.*, 2015) so it could be that the mRNA is transcribed in the roots and moves to the cotyledons. The protein may be highly regulated in these tissues and subsequently quickly degraded in the leaves but with enough to be detected by the sensitive Western blot. The expression of IRQ4 in the flowers backs up the bioinformatics data and *prom*GUS analysis as it shows expression is visible in flowers (see Figure 3.6ei and 4.3).

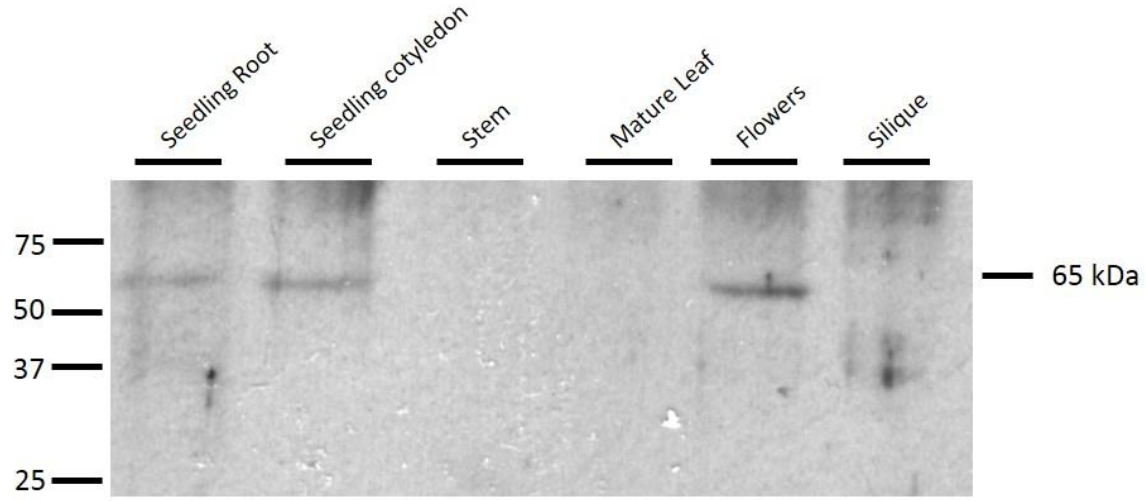


Figure 5.22: Western blot to show IRQ4 protein expression in different *Arabidopsis* tissues. Total protein (50 µg) was extracted from various tissues (above lanes) and separated on a 10 % SDS-PAGE gel and transferred to a nitrocellulose membrane. The membrane was probed with the anti-IRQ4 antibody (1:500 dilution). Indicated on the right is the 65 kDa for IRQ4. The Precision plus Protein Dual Colour (Bio-Rad) was used as a molecular weight standard.

5.5.5 Localisation of IRQ4 in Arabidopsis root tips using immunogold labelling and transmission electron microscopy

The native localisation of IRQ4 was assessed using the polyclonal antibody against IRQ4 generated in section 5.5.3 using immunogold labelling and visualisation using TEM. Five to seven day old *A. thaliana* root tissue was used using cells from the meristematic regions as this is an area where the GUS expression studies and *promIRQ4*::IRQ4-GFP plants had shown IRQ4 is present at this developmental stage. The native IRQ4 protein was detected using the IRQ4 antibody as the primary antibody in ultrathin sections of high pressure frozen and freeze substituted (HPFFS) sections of *Arabidopsis* root tissue. The primary antibody was detected using 5nm colloidal gold-conjugated goat anti-mouse antibodies. HPFFS was used as a fixation method it preserves more of the structural detail and also does not rupture the tonoplast membrane, which can occur when aldehyde fixatives are used during chemical fixation (Wilson & Bacic, 2012). Samples were embedded in resin and ultrathin sections (50-70 nm) were prepared using a diamond knife. Samples were first blocked in BSA (Bovine Serum Albumin) and then incubated with the IRQ4 primary antibody at a dilution of 1:25. Finally, the colloidal gold-conjugated secondary antibody was applied at a dilution of 1:20.

Labelling the cells using the anti-IRQ4 antibody showed gold particles present at the plasma membrane and also at double membrane structures. At the plasma membrane gold particles were found where membrane compartments were either fusing with the plasma membrane or where endocytosis was occurring, and these structures represent invaginations of the PM. Gold labelling was found in compartments at the PM containing either a single membrane (Figure

5.23a-b, i-j) or multiple membranes (Figure 5.23e-h, m-p). Some of the compartments where the gold was found appeared to contain multiple compartments (Figure 5.23e-h, m-p) so could represent pre-vacuolar compartments/multivesicular bodies (PVC/MVB). The double membrane structures not associated with the plasma membrane were highly convoluted (Figure 5.23e-f) and could represent autophagasomes. No labelling of the vacuole was found in these samples. Dual labelling of PVC/MVBs and autophagasomes is possible as fusion between the two compartments has been observed in yeast and mammalian cells (Lamb *et al.*, 2012; Muller *et al.*, 2015) and also plant cells (Zhuang *et al.*, 2015). Some components of the ESCRT machinery and Rab GTPases also have dual functions in autophagy pathways and PVC/MVB vacuolar trafficking (Cui *et al.*, 2016). Overall the results from the immunogold labelling using the anti-IRQ4 antibody verify the localisation to the PVC/MVB seen in the stably expressing BY2 cells and *promIRQ4*::IRQ4-GFP *Arabidopsis* plants.

5.6 Statistical analysis of the number of PVC/MVB when plants are grown under autophagy inducing conditions

The localisation of IRQ4 to autophagasome like structures found using the anti-IRQ4 antibody lead to an investigation into whether any of the previously classified PVC/MVB seen in the *promIRQ4*::IRQ4-GFP plants were autophagasomes. Autophagy can be induced by transferring seedlings from normal media to media lacking nitrate (Bassham, 2015). Transfer to starvation conditions causes an increase in autophagasome number as has been demonstrated in roots using the autophagy marker Atg8 fused to GFP (Wang *et al.*, 2016a).

Stably expressing *Arabidopsis* plants containing the *promIRQ4*::IRQ4-GFP construct were grown vertically for 5 days using normal media according to section 2.6.3. Seedlings were then transferred to either media containing nitrate or media lacking nitrate. Seedlings were left for 48 hours before being imaged using CLSM and the number of punctae counted. The number of punctae is expressed as the number per unit area to control for differences in cell size.

The number of punctae in plants transferred to N starvation conditions was significantly lower (3.16 punctae per unit area, s.d.= 1.60, 95 % CIs: 2.38 to 3.96) compared to control plants transferred to media containing nitrate (5.98 punctae per unit area, s.d.=1.63, 95 % Cis: 5.17 to 6.79; t(36)=5.23, p<0.001; Figure 5.24). In general, the overall fluorescence of plants that had been transferred to media lacking nitrogen was also lower.

The starvation conditions appears to have caused a disruption to the localisation of IRQ4 from the PVC/MVB. PVC/MVB are known to fuse with autophagasomes to form hybrid structures (Cui *et al,* 2016). During this process the GFP from the IRQ4 GFP-fused construct may be severed leaving the

IRQ4 protein to remain in the hybrid structure. As IRQ4 does not possess a transmembrane domain it is also possible that IRQ4 does not bind directly to the PVC/MVB membrane itself but is associated with another protein by binding using its coiled-coil domains. This other protein may be removed or degraded during the autophagic process, not allowing IRQ4 to bind. Alternatively the type of autophagy induced by nitrogen starvation may not involve IRQ4. As introduced in section 1.3 different types of autophagy exist in plants so it is possible that IRQ4 has a role in constitutive autophagic processes rather than induced autophagy. The IRQ4 GFP-fused construct does not seem to be degraded in the vacuole as fluorescent protein fusions which are transported to the vacuole via autophagic processes can lead to an accumulation of the fluorophore within the vacuole, as was seen with RFP cleaved from a Cyt b5-RFP construct (Toyooka *et al.*, 2009).

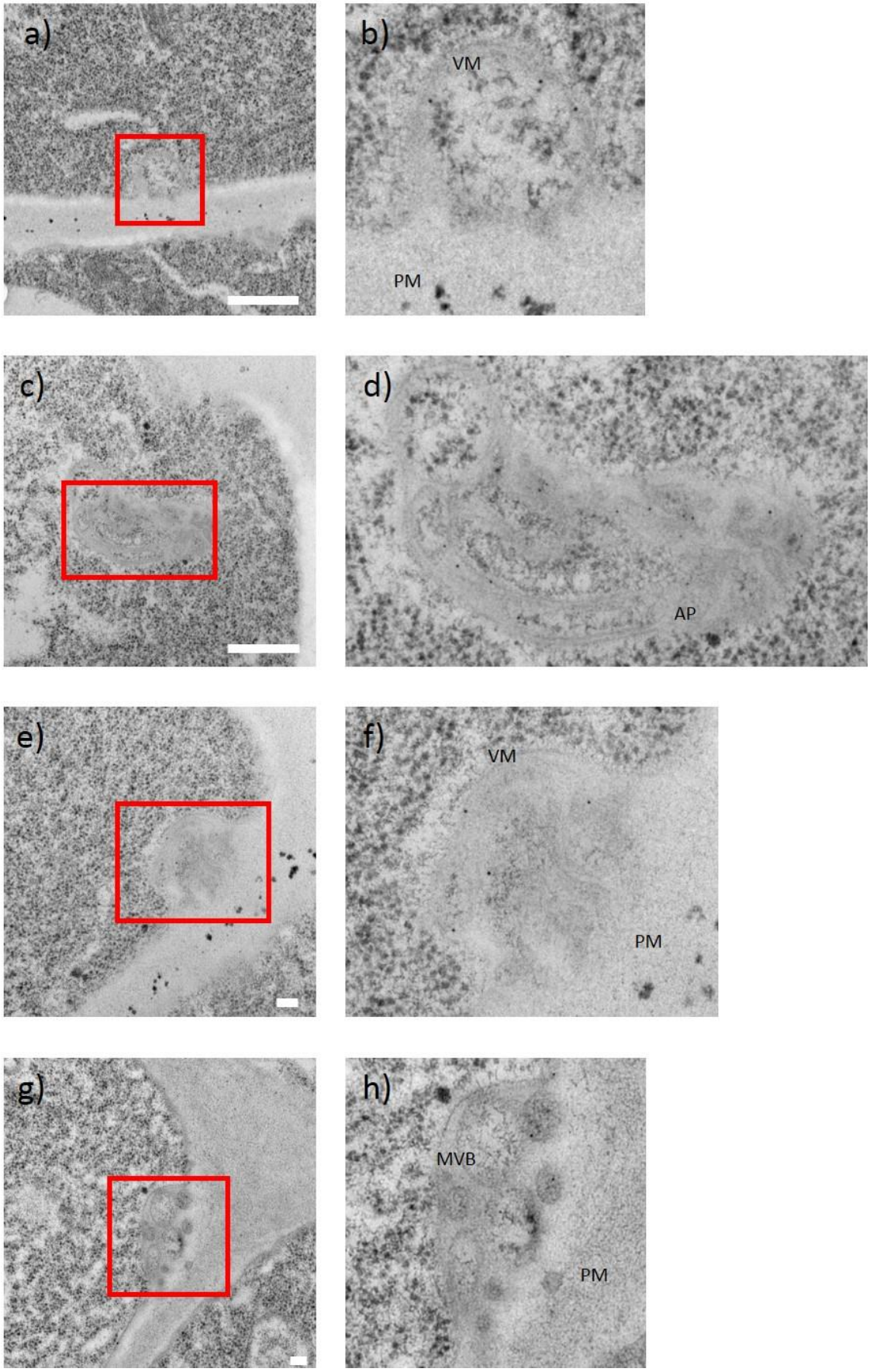


Figure 5.23: *In* situ localisation of IRQ4 in *Araboidopsis* root tissue using immunogold labelling and TEM. Several structures containing multiple membranes (MVB) were labelled by the anti-IRQ4 antibody and also structures which likely correspond to autophagasomes (AP). In addition, vesicle membranes (VM) at the plasma membrane (PM) were also labelled. Examples of cells in the *Arabidopsis* root were imaged at high magnification (a, c=x60000, e=x80000, f=70000) and digital zoom was used to enhance the images shown in b, d, f, and h which correspond to a, c, e, and g respectively. Scale bars: a, c=0.5 µm, e, f=0.1 µm.

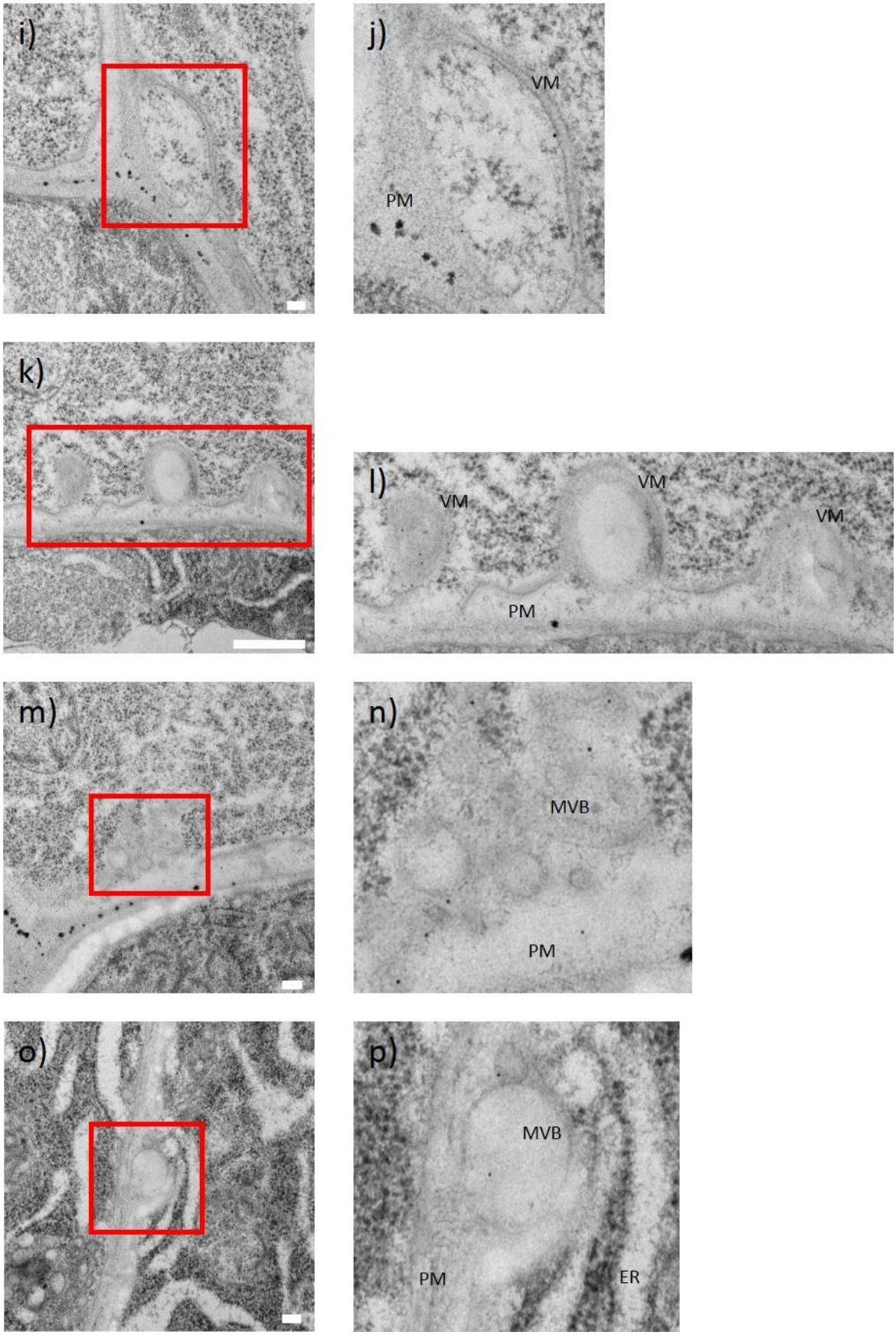


Figure 5.23 continued: i=x80000, k=x60000, m,l=x70000 magnification. Images in j, l, n, and p are enlargements of the red boxes shown in the corresponding images I, k, m, and o respectively. Scale bars: i, m, n=0.1 µm, k=0.5 µm.

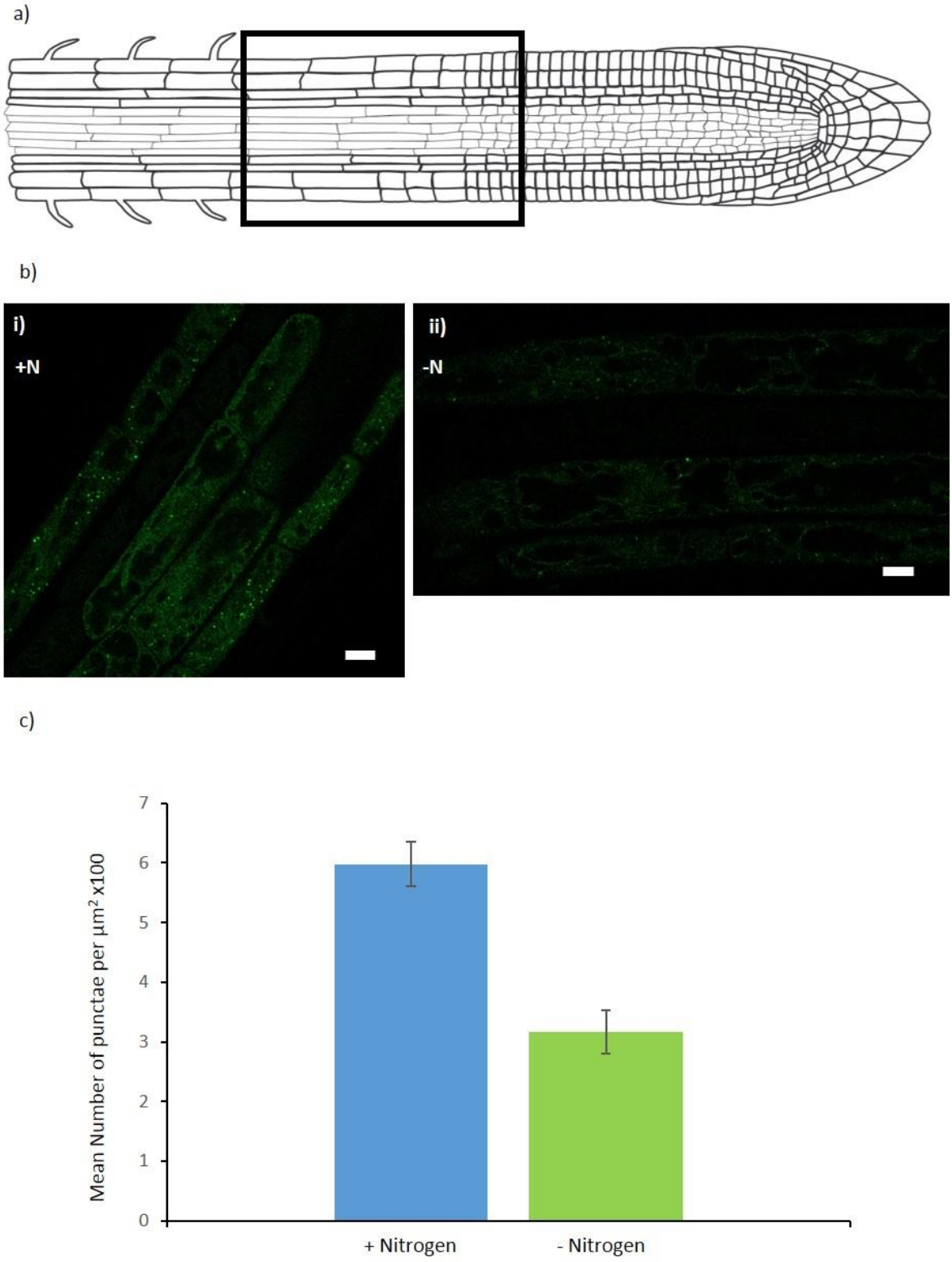


Figure 5.24: Number of PVC/MVB compartments in *Arabidopsis* plants containing the *promIRQ4*::IRQ4-GFP construct grown on autophagy inducing media. Five day old seedlings were grown on plates containing media with nitrate then were either transferred to media containing nitrate (+ Nitrogen, control) or media lacking nitrate (- Nitrogen). Seedlings were grown for 48 hours then mounted in water and imaged on the confocal. a) Schematic of primary root showing the area (black box) between the meristematic and elongation zones from which images used in this figure were taken. b) Overview of an area of root where single cells were analysed from plants transferred to nitrate containing media (+N, bi) or lacking nitrate (-N, bii). c) Statistical analysis of the mean number of PVC counted per unit area multiplied by 100 in cells from nitrate containing media (+Nitrogen) and cells from media lacking nitrate (-Nitrogen). Error bars show +/- standard error of the mean.

5.7 Conclusion

This chapter has described several different approaches to assess the *in situ* expression and subcellular localisation of the IRQ4 protein in the heterologous BY2 system and the native *Arabidopsis* environment.

Conducting transient expression assays in *N. benthamiana* leaves showed that constitutively expressed IRQ4 under the control of a 2X35S promoter localised to punctate structures which moved rapidly around the cell. These punctate structures showed a variety of behaviours including seeming to fuse, forming hybrid compartments then separating.

To support the transient expression studies, the IRQ4-GFP construct was stably transformed into *N. tobacum* BY2 cells. In this system the punctae were again prevalent, displaying similar behaviours to those seen in the transient expression system with rapidly moving punctae seeming to fuse and split. Weak localisation was also seen at the vacuole. Colocalisation analysis in the BY2 cells demonstrated that the punctae associate with the ER but do not necessarily colocalise with the ER. Drug studies in the BY2 cells using latrunculin B and oryzalin revealed that the movement of the punctae is dependent upon the actin and not the microtubule cytoskeleton. Treatment of the BY2 cells with endomembrane trafficking disrupting drugs BFA and wortmannin revealed that the IRQ4 punctae are potentially PVC/MVBs.

To verify the data obtained in the heterologous systems, the localisation of IRQ4 was assessed by transforming *Arabidopsis thaliana* plants with the *promIRQ4*::IRQ4-GFP construct with IRQ4 under the control of its own promoter. IRQ4 was found to be expressed in the roots, verifying the GUS data obtained in chapter 4. In addition, the subcellular localisation was analysed and IRQ4 was found to localise more strongly to the tonoplast membrane and also gave punctae verifying the results obtained in the *N. benthamiana* leaves and in the BY2 cells. The localisation in the primary and lateral roots was found to change with the maturity of the tissue with the vacuolar localisation being more prominent in the meristematic regions and then the punctate localisation becoming more prominent in the differentiation and elongation zones of the root. Co-expression of the *promIRQ4*::IRQ4-GFP construct with the ER marker HDEL-RFP showed a similar result to the BY2 cells showing that IRQ4 does not colocalise with the ER. In addition the Golgi marker ST-RFP was successfully co-expressed in *promIRQ4*::IRQ4-GFP expressing *A. thaliana* plants and showed that the IRQ4 does not colocalise with the Golgi. Staining using the pH dependent dye showed that IRQ4 localises to the tonoplast membrane in *A. thaliana* roots, verifying the Affymetrix data shown in chapter 3.

The same drug studies that were conducted on the BY2 cells were repeated on the stably expressing *A. thaliana* plants. BFA treated plants co-expressing the *promIRQ4*::IRQ4-GFP and ST-RFP constructs showed that IRQ4 does not colocalise with Golgi and does not aggregate with the Golgi in BFA bodies. Plants expressing *promIRQ4*::IRQ4-GFP alone when treated with wortmannin showed the same result as the BYs cells with the formation of larger vesiculated compartments similar to that seen by Wang *et. al.* (2009) when compartments labelled with the PVC marker BP-80 reacted in the same manner to wortmannin treatment.

Finally, to verify the localisation of the translational fusions polyclonal antibodies to IRQ4 were generated and tested using western blotting and the localisation of the construct was tested using immunogold labelling. Western blotting showed that IRQ4 is expressed in the seedling roots and cotyledons, and flowers. The distribution of the gold particles was mostly found at double membrane structures, possibly corresponding to autophagasomes, MVBs or at membranous structures either undergoing an exo- or endocytotic process.

Overall this chapter has shown that IRQ4 protein is expressed in the root tissue and its subcellular localisation is at the tonoplast or PVC/MVB or potentially autophagasomes. IRQ4 is appears to be involved in the late secretory pathway. The next chapter aims to explore the potential interacting partner of IRQ4 in order to ascertain a function for the protein in this pathway.

Chapter 6 Phenotypic analysis of the IRQ proteins

6.1 Introduction

The experiments in chapters 3, 4, and 5 have described the evolutionary relationship between the NAB/IRQ and IRQ proteins, the transcriptional localisation of *IRQ1, IRQ6, IRQ4*, and *IRQ5*, and the translational localisation of IRQ4. Phylogenetic analyses have demonstrated that IRQ4 and IRQ5 are more closely related to each other than to other IRQ proteins in *Arabidopsis*. *prom*GUS fusions to *IRQ4* and *IRQ5* have given a root localisation for transcription of these genes . The subcellular localisation of IRQ4 using translational fusions to GFP driven by the native IRQ4 promoter has shown that IRQ4 is translated in the root, and within the cell is localised to the vacuole and PVC, linking IRQ4 to the late endocytic pathway. Using a polyclonal antibody raised against IRQ4 for immunogold labelling in *Arabidopsis* roots has shown a potential localisation of IRQ4 to autophagasomes, which could link to IRQ4 to autophagic processes.

The function of IRQ4 is yet to be determined. Using mutant plants in which the gene of interest is partially or fully knocked out to create loss of function mutants and conducting phenotypic analyses is a well-established technique to ascertain protein function. This chapter describes experiments using this reverse genetic approach using mutant lines of *IRQ4* and *IRQ5*, both single mutant lines and double mutant lines. As IRQ4 and IRQ5 proteins have sequence similarity at both the N-terminal and C-terminal ends of the proteins and are expressed in some of the same tissues, double mutants of *IRQ4* and *IRQ5* were generated as there may be functional redundancy. These lines were then tested to identify if a phenotype could be determined. The mutants analysed arose through either T-DNA insertional mutagenesis and were purchased through NASC (Nottingham Arabidopsis Stock Centre), or were generated by Dr. J. Kroon using Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) technology using the Cas9 enzyme. Phenotypic analysis of the mutant lines focussed on primary root growth as IRQ4 has been shown to be predominantly expressed in this tissue, including an examination of vacuole morphology.

As CRISPR technology is relatively new to the plant field and the method used to generate the mutants plants used here is not widely known, a brief explanation of how CRISPR works is first given. This is followed by the verification of the knockout status of the lines created and bought using PCR, RT-PCR, and western blotting. Finally, phenotypic analysis of the lines is described.

6.2 CRISPR Technology

Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)/Cas (CRISPR associated) technology is an adapted version of RNA guided archaeal and bacterial immune systems which

degrade exogenous substrates (Jinek *et al.*, 2012). There are a wide variety of such systems and modifications but the CRISPR/Cas9 system is the most widely used (Mali *et al.*, 2013; Voytas, 2013; Jinek *et al.*, 2014; Nishimasu *et al.*, 2014; Sternberg *et al.*, 2014) and applied to plants (Xing *et al.*, 2014). The expansion of the technology has been extensive since its introduction to the academic community with a variety of potential applications across a wide spectrum of scientific fields, as demonstrated by the numbers of patents filed in relation to CRISPR/Cas (Mali *et al.*, 2013; Egelie *et al.*, 2016; Wright *et al.*, 2016).

CRISPR/Cas technology can overcome certain problems in the generation of mutants in plants. Approximately 12 % of genes have no T-DNA insertion available with only 8 % represented by a single gene (O'Malley & Ecker, 2010). Another issue is that generation of plants with mutations in multiple genes using T-DNA lines involves the laborious crossing of mutant lines which is highly time consuming. Plants with multiple mutations are important to removing the functional redundancy often present in multi-gene families or when studying genes which epistatic relationships in their genetic pathway. CRISPR/Cas overcomes these problems by allowing the generation of single or multiple gene knockouts in a more precise and much shorter time-frame, even in a single generation.

CRISPR/Cas systems use the small guide RNA (sgRNA)/Cas complex endonuclease activity to produce double stranded breaks at sites in the genome specified by sgRNAs. These double stranded breaks cause activation of the DNA repair system in the host, typically utilising the non-homologous end joining pathway (Voytas, 2013). The repair pathway is often error prone and leads to the creation of indel mutations in the sequence.

Agrobacterium-mediated techniques used to create transgenic plants include *in planta* transformation and embryonic callus based transformation. In *Arabidopsis* transformation of plants is usually achieved using the floral dip method (Clough & Bent, 1998) where the egg cell is the target of the T-DNA insertion.

Using a binary plant vector where the promoter of the egg cell specific EC1.2 gene (Steffen *et al.*, 2007; Sprunck *et al.*, 2012) drives the expression of Cas9 in Arabidopsis could efficiently lead to homozygous or biallelic mutants for multiple target genes in Arabidopsis in as early as the T1 generation.

6.2.1 Generation of the CRISPR double mutant lines for *irq4* and *irq5*

In order to create a double mutant for the phylogenetically related and potentially redundant IRQ4 and IRQ5 genes a CRISPR/Cas9 strategy was designed (see section 2.2.18) based on a toolkit

using a binary vector set which has the pCAMBIA or pGreen backbone and a sgRNA module vector set (Xing *et al.*, 2014).

The two targets for CRISPR/Cas9 in IRQ4 and IRQ5 are shown below in Figure 6.1.

T1 IRQ4

Class1.0 gRNA mmMM 4/4 + strand

TTCTTCTTCGAGCTCAGATT**CGG**

T2 IRQ4

Class1.0 gRNA mmMM 3/3 + strand

AGCTGTTGAATCAGAGCACC**AGG**

T3 IRQ5

Class1.0 gRNA mmMM 5/3 + strand

ATTCTGACCATTCTTCTAAG**AGG**

T4 IRQ5

Class1.0 gRNA mmMM 5/4 + strand

CAGGGAACTGCATCAGAAAC**TGG**

Figure 6.1: Target sites in the *IRQ4* and *IRQ5* genes specified by the CRISPRsearch tool. Class number indicates specificity of Cas targeting with values between 0 and 1 being considered specific.

The resulting CRISPR/Cas9 construct was transformed into *Agrobacterium tumefasciens* strain GV3101/pMP90 (according to section 2.2.7.2) and used in floral dip plant transformation of *Arabidopsis thaliana* Columbia-0 (according to section 2.2.15).

T1 seed was obtained followed by Hygromycin screening for positive independent transformants and T1 CRISPR/Cas9 generated genomic editing events. Also T2 seed was obtained and T2 plants analysed for the presence and genetic status of CRISPR/Cas9 generated gene editing events.

Two independent double mutants were obtained (*irq4/5CR.1* and *irq4/5CR.2*) and analysed in the T1 and T2 generations. For *IRQ4* a bi-allelic deletion of 52 amino acids between the two target

sequences at the 5’ end appears to have occurred in two independent lines but the sequence has been left in its original frame or in-frame (Figure 6.3). The gene may therefore produce a protein product but this may get degraded at the RNA stage or after translation. If a protein product is produced the protein may not be functional when missing such a large proportion of the N-terminal end.

a)

IRQ4 Mutation

IRQ4 Wild type protein sequence

MRKLSIRDSLKSFFEPHLHPDNGESLKGTKTEIDEKVKKILGIVESGDIEEDESKRLVVAELVKDFYKEYESLYHQ
YDDLTGEIRKKVHGKGENDSSSSSSDSDSDKKSKRNGRGENEIELLKKQMEDANLEIADLKMKLATTDEHK
EAVESEH***QEILKKLKESDEICGNLRVETEKLTSENKELNEKLEVAGETESDLNQKLEDVKKERDGLEAELASK***
AKDHESTLEEVNRLQGQKNETEAELEREKQEKPALLNQINDVQKALLEQEAAYNTLSQEHKQINGLFEERE
ATIKKLTDDYKQAREMLEEYMSKMEETERRMQETGKDVASRESAIVDLEETVESLRNEVERKGDEIESLMEK
MSNIEVKLRLSNQKLRVTEQVLTEKEGELKRIEAKHLEEQALLEEKIATTHETYRGLIKEISERVDSTILNRFQSL
SEKLEEKHKSYEKTVVEATKMLLTAKKCVVEMKKEKDEMAKEKEEVEKKLEGQVREEEKEKEKLKETLLGLGE
EKREAIRQLCIWIEHHRDRCEYLEEVLSKMVVARGQRRSQRA

Protein sequence in *irq4/5CR.1* and *irq4/5.CR2*.

MRKLSIRDSLKSFFEPHLHPDNGESLKGTKTEIDEKVKKILGIVESGDIEEDESKRLVVAELVKDFYKEYESLYHQ
YDDLTGEIRKKVHGKGENDSSSSSSDQEILKKLKESDEICGNLRVETEKLTSENKELNEKLEVAGETESDLNQ
KLEDVKKERDGLEAELASKAKDHESTLEEVNRLQGQKNETEAELEREKQEKPALLNQINDVQKALLEQEAAY
NTLSQEHKQINGLFEEREATIKKLTDDYKQAREMLEEYMSKMEETERRMQETGKDVASRESAIVDLEETVES
LRNEVERKGDEIESLMEKMSNIEVKLRLSNQKLRVTEQVLTEKEGELKRIEAKHLEEQALLEEKIATTHETYRG
LIKEISERVDSTILNRFQSLSEKLEEKHKSYEKTVVEATKMLLTAKKCVVEMKKEKDEMAKEKEEVEKKLEGQV
REEEKEKEKLKETLLGLGEEKREAIRQLCIWIEHHRDRCEYLEEVLSKMVVARGQRRSQRA

b)

IRQ5 Mutation

Protein sequence in *irq4/5CR.1*

MKKHKFRETLKSFFEPHFDHEKGEMLKGTKTEIDEKVNKILGMVESGDVNEDESNRQVVADLVKEFYSEYQ
SLYRQYDDLTGEIRKKVNGKGESSSSSSSDSDSDHSSSDAVPCYHSVLLTFRPLTSQFQVSTYC***

Figure 6.2: Mutations found in *IRQ4* and *IRQ5* genes in the double CRISPR lines. a) The WT protein sequence is shown. Red highlighting indicates the part of the sequence deleted by the CRISPR process. Underneath is the sequence found in the *irq4/5CR.1* and *irq4/5CR.2* mutant plants. The blue highlighting indicates the sequence of the protein fragment used to generate the anti-IRQ4 antibody (see section 5.5.1). As part of the fragment is still present in the mutated CRISPR lines the antibody may therefore still be able to recognise the protein product if one is produced and is stable. b) The only part of the sequence that was produced when sequencing the *irq4/5.CR1* line is shown. This indicates a major disturbance to the gene in this line.

For *IRQ5* the two independent lines have had different alterations. i*rq4/5CR.1* has a bi-allelic alteration in the genomic target area and has been replaced with an undetermined sequence (Figure 6.2b). The coding sequence has been disturbed and a functional protein product will not be produced for *IRQ5*. For the *irq4/5CR.2* line a PCR product could not be obtained suggesting a large and dramatic change in the genomic target area. Therefore, *irq4/5CR.2* is highly unlikely to produce a functional protein product.

In summary, CRISPR/Cas9 technology was used to generate double mutants for the *IRQ4* and *IRQ5* genes. Two independent lines were produced. Both lines were bi-allelic for *IRQ4* and gave a deletion of 52 amino acids which fell in frame and may be capable of producing a protein product. For *IRQ5* a large deletion occurred in the target area for the *irq4/5CR.1* line while the *irq4/5CR.2* line could not be sequenced for the target site. In both cases a large disturbance has occurred in the *IRQ5* sequence and no functional protein product is likely to have been produced.

6.3 Generation and validation of the single T-DNA Insertion lines, and validation of the double CRISPR mutant lines for *irq4* and *irq5*.

6.3.1 Identification of T-DNA Lines for IRQ4 and IRQ5

T-DNA insertional mutagenesis involves the random insertion of foreign DNA into the genome of the organism which can both disrupt gene function and also act as a marker for subsequent analysis via PCR (Krysan *et al.*, 1999). The Salk Institute Genomic Analysis Laboratory (SIGnAL) and German Plant Genomics Program Kölner Arabidopsis T-DNA lines (GABI-Kat; Rosso *et al.*, 2003) provide access to a large population of plant lines containing T-DNA insertions. Precise locations for 88,000 T-DNA insertions have been determined with the identification of mutations in approximately 21,700 genes (Alonso *et al.*, 2003).

The SIGnAL T-DNA Express Arabidopsis Gene Mapping Tool was used to identify candidate mutant lines for phenotypic analysis of *IRQ4* and *IRQ5*. To ensure maximum gene disruption and increase the probability of obtaining a fully knocked out gene, lines were chosen which had insertions in the exon of the gene in question, as insertions within the intergenic regions can be spliced out (Wang & Wang, 2008). A single SALK T-DNA homozygous knockout line (SALK_078284) was identified for *IRQ4*. For *IRQ5* two insertion candidates were identified: a SAIL T-DNA homozygous knockout line (SAIL_293_E12) and one GABI-Kat insertion line (GABI_865F04). The SALK line used by Ren *et al.* (2016; SALK_070302) was not used in this study due to the location of the T-DNA within the intron. However, as a phenotype was shown with this single line, this line has been obtained and work using this line is being continued within the laboratory. As part of the process

to check the position of the insertion the sequence give for the insertion point for the SAIL_293_E12 line did not align with the IRQ5 sequence when compared using the alignment function of BLAST. This SAIL line was not used for further analysis so may represent a mis-annotation in the database. Two T-DNA insertion lines in total for *IRQ4* and *IRQ5* were used in this analysis and the positions of the T-DNA insertions, the primers used to genotype them, and the expected sizes of the amplified fragments with each primer combination are listed in Table 6.1. The SALK_078284 IRQ4 line was renamed *irq4.1* and the GABI_865F04 line for *IRQ5* was renamed *irq5.1.*

Table 6.1: Details of T-DNA insertion lines selected for *IRQ4* and *IRQ5*.

Names of the insertion lines used are given along with the locations of the insertions within the gene. Primer pairs used for the genotyping reactions are given along with expected band sizes for each type of reaction. See appendix 1 for list of primer sequences.

Line	Insertion point in gene	Insertion location	Primers for wildtype PCR	Wildtype fragment (bp)	Primers for insert PCR	Insert fragment (approx. bp)
IRQ4						
SALK_078284 (*irq4.1*)	Exon 2	1206 (1337)	IRQ4284Fw+ IRQ4284Rv	946	LBa1.1+ IRQ4284_R	455
IRQ5						
GABI_865F04 (*irq5.1*)	Exon 2	1771 (1873)	IRQ5_865_F2+ IRQ5_865_R2	1273	IRQ5_865_F2 + 8409_GB_LB	1031

As the lines received are a segregating T3 population, homozygotes must be identified by PCR. PCR reactions for genotyping were carried out according to section 2.2.1. Using the tools available at SIGnAL and BLAST the positions of the T-DNA insertions were identified within each gene (Table 6.1). Seeds from all lines were first grown on vertical plates for 5-7 days before being transferred to soil and grown under standard conditions (according to section 2.6.3). Samples of young leaves were taken and DNA extracted for genotyping according to section 2.2.16 and the DNA was used in the genotyping reaction. Primers were designed for two different PCR reactions: a wild type (WT) PCR reaction and an insert PCR reaction. In the WT PCR reaction forward and reverse primers were designed to be approximately 500 bp (for *IRQ4*) or 954 bp (for *IRQ5*) upstream and 500 bp (for *IRQ4*) or 320 bp (for *IRQ5*) downstream of the insertion site respectively. For the insert PCR reaction the primers used were either the forward or reverse primer from the WT reaction in

combination with either the SALK left border primer LBa1.1 for the *irq4.1* line or the CAP2 08409 primer for the *irq5.1* line. The LBa1.1 and CAP2 08409 primers were designed to only bind to sequences within the T-DNA of their respective lines. Each plant sample was tested using both WT and insert PCR reactions. If the line had no T-DNA insertion meaning it was azygote, in the WT reaction a band of approximately 950 bp (for *IRQ4*) or 1300 (for *IRQ5*) would be seen on the gel while no band would be present in the insert reaction. If the line was heterozygote for the insertion, an approximately 950 bp (for *IRQ4*) or 1300 bp (for *IRQ5*) band would be present in the WT reaction and another band of approximately 500 bp (for *IRQ4*) or 1050 bp (for *IRQ5*) would be present in the insert reaction. If the line was homozygote for the insertion no band would be present for the WT reaction but a band of approximately 500 bp (for *IRQ4*) or 1050 bp (for *IRQ5*) would be present in the insert reaction. Typical results from a set of genotyping reactions showing azygote and homozygote samples can be seen in Figure 6.3 for *irq5.1*. No heterozygote plants were found for the *irq5.1* line. Plants which were identified as homozygote and azygote for both *irq4.1* and *irq5.1* lines were allowed to grow and set seed which was collected according to section 2.5.4. In order to ensure the gene was completely inactivated reverse transcriptase PCR (RT-PCR) was carried out to check for the presence of transcript in the *irq4.1* and *irq5.1* lines. In addition, T-DNA insertion lines and CRISPR mutants were analysed for the presence of the IRQ4 protein by using the polyclonal anti-IRQ4 antibody (see section 5.5 for its generation) in a western blot to ensure no protein was translated.

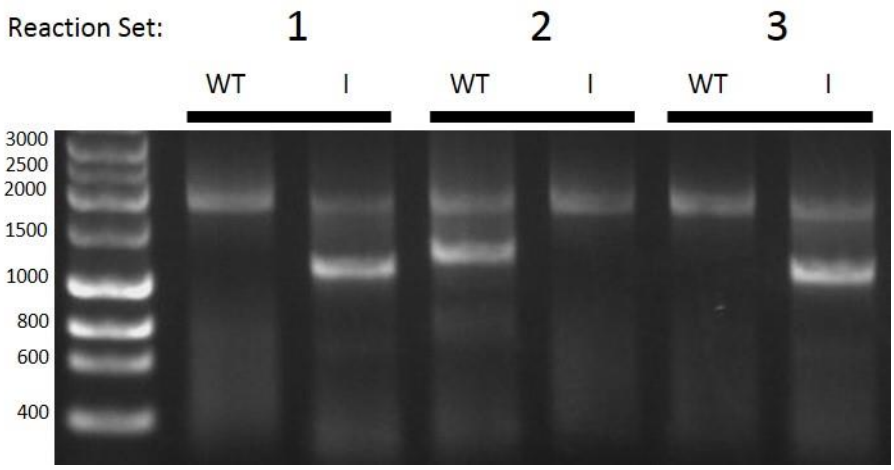


Figure 6.3: Typical result of a PCR used to genotype the T-DNA insertion mutants. Genomic DNA was extracted from young leaves (in this case the *irq5.1* line is shown) and presence of the insert was tested using two different PCR reactions. WT reactions used the IRQ5_865_F2/ IRQ5_865_R2 primers to generate an approximately 1300 bp fragment, while insert PCR reactions used the IRQ5_865_F2/CAP2 0849 primers to generate an approximately 1050 bp fragment. Azygote plants containing no insertion had the 1300 bp band in the WT reaction but no band in the insert reaction (reaction set 2). Homozygote plants contained no band in the WT reaction but contained a 1050 bp band in the insert reaction (reaction sets 1 and 3). No heterozygotes were found for the *irq5.1* line. WT reaction primer sets were designed to span the insertion location while insert reaction primers were designed to amplify from the left border of each of the insertions (in both *irq4.*1 and *irq5.*1). As a positive control to ensure the PCR reactions had worked the *mago* Fw/Rv primers were used which produce a band of approximately 1868 bp. Hyperladder I (Bioline) was used as a molecular weight marker.

6.3.2 Confirmation of lack of transcriptional activity in the *irq4* and *irq5* single mutant lines using RT-PCR

Before continuing with phenotypic analysis confirmation that no transcript was present was required. In RT-PCR RNA is isolated from plant tissue (see section 2.2.10) and used as a template to synthesise cDNA using reverse transcriptase (see section 2.2.17). In all the analysis that follows RNA was isolated from 14-day old vertically grown seedlings. As the RNA isolated represents the total RNA present within the tissue sample, standard PCR reactions can be used on the transcribed cDNA on various parts of the sequence using specific primers to confirm the absence of transcript. In general for each line primers were designed to amplify a section of sequence at the 5' end, another set of primers around the insert, and a final set of primers were designed to amplify a section of cDNA at the 3' end of the sequence. These upstream, insert, and downstream RT-PCR reactions would clarify if there was any truncated sequences (upstream and insert reactions) or whether a full transcript is synthesised (downstream reaction). Primers used for the reactions and the fragment sizes expected from each primer combination are shown in Table 6.2. Locations of the insertions and relative positions of the primers within in each gene are given in Figure 6.4.

Table 6.2: Primers used to verify knockout status of *IRQ4* and *IRQ5* insertion lines using RT-PCR with expected band sizes.

Line	5' End Primers	5' End Band Size	Flanking Primers	Flanking Band Size	3' End Primers	3' End Band Size
IRQ4						
SALK_078284 (*irq4.1*)	IRQ4_F1+ IRQ4_R2	489	IRQ4284Fw+ IRQ4_R1	609	IRQ4_F3+ IRQ4284Rv	209
IRQ5						
GABI_865F04 (*irq5.1*)	IRQ5_F1+ IRQ5_R2	556	IRQ5_865_F2+ IRQ5_865_R2	1273	IRQ5_F2+ IRQ5_R1	533

Primers were designed to amplify sequences at the 5' end of the gene upstream of the insertion site, across the T-DNA insert location, and at the 3' end of the gene downstream of the T-DNA insert location. Band sizes indicate the expected size of the amplified fragments.

Three levels of checking were used to confirm that there was an absence of WT cDNA or genomic contamination within sample used for RT-PCR. The 5' end primers for the *irq4.1* and *irq5.1* mutants were designed to cover the region spanning an intron. Introns are spliced out after transcription meaning that the fragments derived from genomic DNA contaminated samples

would be larger than those derived from cDNA samples. The insertion primers were designed to cover the area of the insertion. If genomic DNA contamination was present then a band would be visible on the gel. Finally, primers designed to cover the intron spanning region of an independent gene, *NETWORKED 1A* (*N1A*; At3g22790), were used as a further check to ensure that no genomic DNA contamination occurred. Genomic DNA produces bands of 763 bp while pure cDNA produces a band of 588 bp for *N1A*. In addition, for checking *irq5.1* plants, reactions used *EF-1α* or Elongation Factor 1-α (*EF*-1α; At5g60390) or *N1A*. *EF-1α* is a well-known referencing gene used to check for transcript levels in RT-PCR due to its ubiquitous expression pattern (Czechowski *et al.*, 2005).

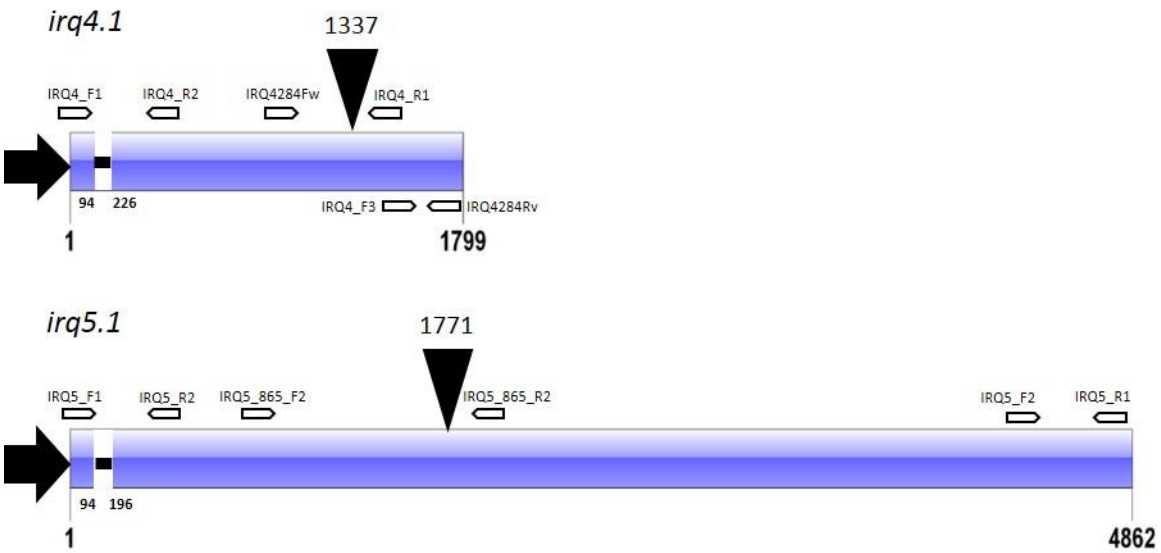


Figure 6.4: Schematic of *IRQ4* and *IRQ5* genes with locations of the T-DNA insertion and primers used in RT-PCR analysis. Promoter regions are indicated by the large black arrows to the left. Exons are labelled by the large central blue bars. Introns are marked by the smaller black bars. Numbers underneath each of the gene schematics indicate the number of base pairs marking the 5' end and 3' ends of the sequence. Numbers also mark the position of the single intron in each gene: between base pairs 94 and 226 in *IRQ4*, and between base pairs 94 and 196 in *IRQ5*. Primer locations are indicated by the white arrows. Names of primers are shown above or to the side of the arrows. The precise sizes of the fragments produced are shown in Table 6.2. The large, black arrowheads above each gene schematic indicate T-DNA locations with the position of the insertion given above each arrowhead.

6.3.2.1 Confirmation of lack of transcriptional activity in the irq4.1 T-DNA insertion line

Figure 6.5 shows that a band was present at approximately 489 bp for the 5' reaction in *irq4.1* sample. No band was seen in the insertion reaction nor at the 3' end for *irq4.1*. The azygote and WT Col-0 control samples produced bands of the correct size for cDNA samples in all four reactions: 5', insert, 3', and *N1A* control.

As an additional control and to illustrate the size difference between the fragments produced by the 5' and *N1A* PCR reactions in genomic compared to cDNA derived samples, a genomic DNA

sample was used with separate reactions using all four primer combinations for the *irq4.1* mutants: 5', insert, 3', and *N1A*. For the 5' reactions a single band was produced which is approximately 620 bp representing the genomic sized band for the 5' reaction including the intron. Similarly the insert and 3' reactions produced the correctly sized bands of approximately 609 bp and 209 bp respectively. The *N1A* reaction producing a band at approximately 763 bp which is the correct size for the fragment if it is derived from genomic DNA containing the intron.

The *N1A* reaction in the *irq4.1* sample produced a band at approximately 588 bp which represents the size of band expected if no genomic contamination is present in the sample. The lack of double band in the *N1A* lane for *irq4.1* indicates that there is no mixture of genomic DNA and cDNA. This leads to the conclusion that the *irq4.1* sample is not contaminated with genomic DNA. The lack of any bands in the insert and 3' RT-PCR reactions shows that the transcript has been disturbed by the T-DNA insertion. Any RNA derived from the 5' end may be degraded as it is incomplete leading to the production of no protein product and indicating that the *irq4.1* is a transcriptional null mutant of the *IRQ4* gene.

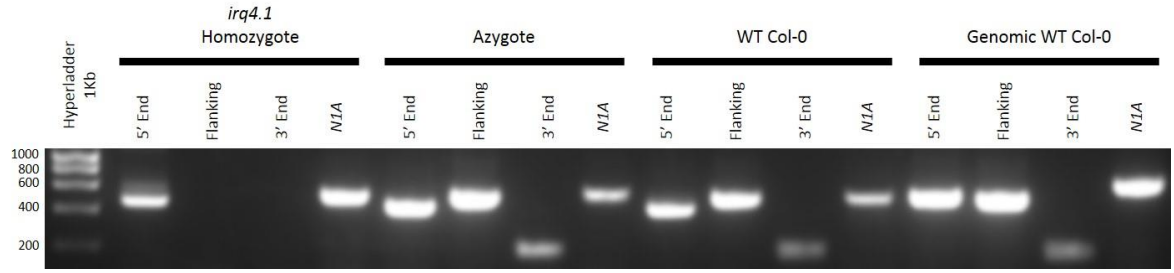


Figure 6.5: RT-PCR analysis of transcriptional activity of *IRQ4* in the *irq4.1* mutant line. RNA was extracted form 14 day old plants homozygote for the *irq4.*1 insertion and azygote plants as well as WT Col-0 seedlings. This RNA was used as a template to synthesise cDNA using the Go-Script reverse transcriptase enzyme (Promega). In addition, a genomic control sample from WT Col-0 plants was also used to indicate the size of bands expected from genomic DNA using the 5' end and *N1A* primer sets. Using the 5' end primer set a band was produced in the *irq4.1* homozygote line but no band was produced using the insert or 3' end primers indicating disruption to the transcription of the gene by the presence of the T-DNA insertion. The lack of band in the reaction using the insertion primers shows that no genomic contamination was present in the cDNA sample. Bands for the 5', flanking, and 3' end PCR reactions were present in the azygote and WT samples. The genomic DNA control sample produced bands in for all the reactions, including a band of increased size for the 5' end and *N1A* primer sets as expected for intron containing DNA. The sizes of the 5' end and *N1A* bands for the *irq4.1*, azygote, and WT plants were all smaller than the bands for the 5' end and *N1A* reactions produced using the genomic DNA as a template indicating no genomic DNA contamination in the *irq4.1* homozygote, azygote, and WT Col-0 cDNA samples.

6.3.2.2 Confirmation of lack of transcriptional activity in the irq5.1 T-DNA insertion line

The *irq5.1* lines show bands in the 5' and 3' end reactions but no band over the insert (Figure 6.6). The lack of band in the insert reaction shows there was no genomic contamination present. The azygote control shows bands in the 5', insert, and 3' reactions. The Col-0 shows identical bands to

the azygote control with bands present in the 5', insert, and 3' end reactions. The 187 bp *EF-1α* band is present across all samples and has the same intensity indicating that there was equal loading of the samples.

To confirm the presence of the insert within the *irq5.1* line a further PCR reaction was set up using one set of primers designed to flank the insert sequence (Control Primer Set 1 (CP1)) and a second set of primers including the CAP2 08409 (8409_GB_LB) primer specific for the GABI-Kat insert sequence and an *IRQ5* specific primer (Control Primer Set 2 (CP2)). Primers used and band sizes expected for each reaction are shown in Table 6.3. Figure 6.7a shows the locations of each primer set on the *IRQ5* gene. If a band is present with the CP1 primers then no insert is present or there is genomic contamination. If a band is seen in the reaction with the CP2 primers then the insert is present. The CP2 reaction for the homozygote line gives a very clear and strong band showing that the insert is present within this line (Figure 6.7b). The CP1 reaction does not give a band showing the insert has disrupted the sequence and no genomic contamination is present. The presence of the CP1 band in the azygote and absence in the CP2 lane shows that the azygote does not contain the insert. The *N1A* control bands show that there is no genomic contamination and that there was approximately equal loading of the samples.

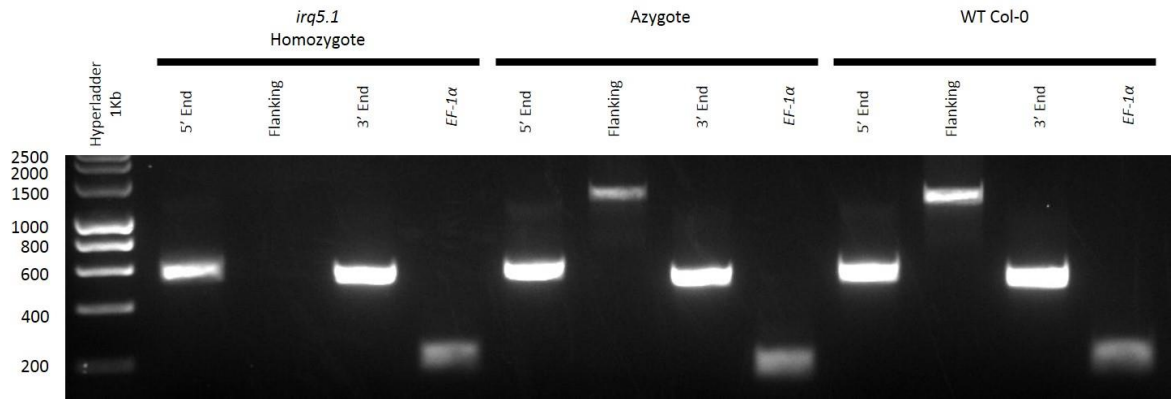


Figure 6.6: RT-PCR analysis of transcriptional activity of *IRQ5* in the *irq5.1* mutant line. RNA was extracted form 14 day old plants homozygote for the *irq5.*1 insertion and azygote plants as well as WT Col-0 seedlings. This RNA was used as a template to synthesise cDNA using the Go-Script reverse transcriptase enzyme (Promega). For the *irq5.1* homozygote line bands were produced for both the 5' end and 3' end primer sets but no band was produced for the flanking primer set. This indicates that the T-DNA insertion is present and that there was no genomic contamination. All the primer sets (5', flanking, and 3') produced bands of the correct size for in the azygote and WT Col-0 reactions. The *EF-1α* reactions all produced bands showing that equal loading occurred in all samples.

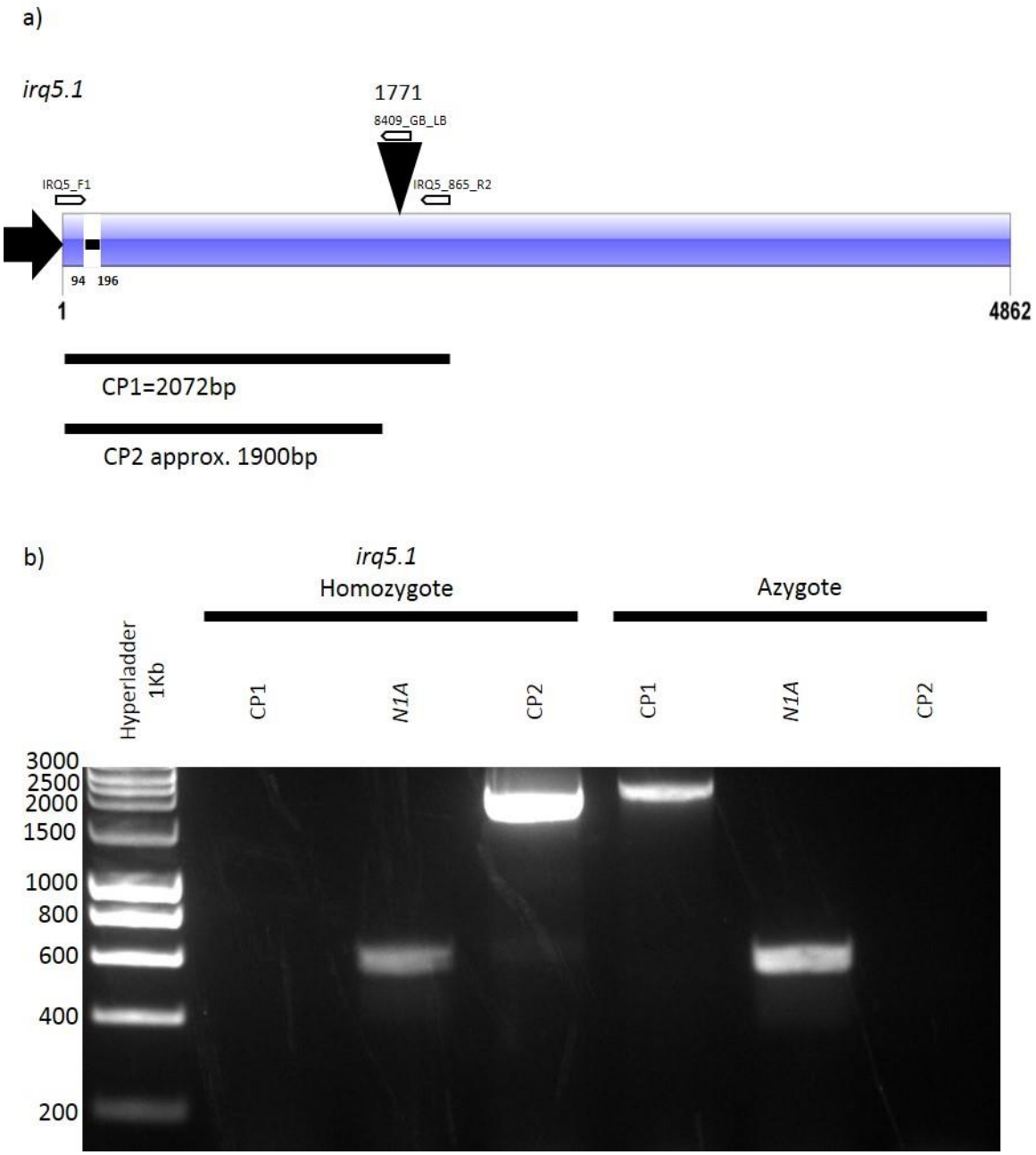


Figure 6.7: RT-PCR analysis to check for the presence of the insertion within the *irq5.1* line. a) To confirm the presence of the insertion within the *irq5.1* line primers were designed to amplify either around the T-DNA insertion (Control Primer Set 1 (CP1)) or designed to amplify from within the T-DNA insertion (Control Primer set 2 (CP2)). Amplification of a band using the CP1 primers means there is no insertion. Presence of a band using the CP2 primers means there is an insertion within the line. Samples which had bands in both lanes would be classed as heterozygote for the insertion. *N1A* primers were used to check for genomic contamination. b) RT-PCR using the CP1, CP2, and *N1A* primer sets to check for the presence of the insert within the *irq5.1* line. The homozygote contains no band in the CP1 lane showing that there is no genomic contamination, a band of the correct size for cDNA in the *N1A* lane, and a strong band in the CP2 lane showing presence of the insert. The azygote line produces a band in the CP1 lane showing there is no insert present, the *N1A* lane contains a band of the correct size indicating no genomic contamination, and there is no band in the CP2 lane confirming the lack of insert within the line.

Table 6.3: Primers used to verify the presence of an insert in the *irq5.1* line.

Line	Control Primer Set 1= CP1	Expected band size (bp)	Control Primer set 2= CP2	Expected band size (approx. bp)
IRQ5				
GABI_865F04 (*irq5.1*)	IRQ5_F1+ IRQ5_865_R2	2072	IRQ5_F1+8 409_GB_LB	1900

Two sets of control primers were used: Control Primer Set 1 (CP1) were designed to amplify a fragment from the beginning of the sequence and across the insertion point. Control Primer Set 2 used the insertion specific 8409_GB_LB primer paired with the IRQ5_F1 primer to check for the presence of the insert.

In summary, analysis of the T-DNA insertion lines using RT-PCR has shown two lines, *irq4.1* and *irq5.1,* have insertions contained within the sequences of *IRQ4* and *IRQ5* respectively that disrupt their transcriptional activity.

6.3.3 Confirming the lack of transcriptional activity of *IRQ4* in the *irq4.1* single T-DNA line and *irq4/5CR.1* and *irq4/5CR.2* CRISPR mutant lines using western blotting.

The RT-PCR analysis of the *irq4.1* mutant showed no fragments were amplified at the insert position or at 3' end position of the sequence (see Figure 6.4a for locations and primers used) but a band was detected at the 5' end. The likelihood of this truncation being able to produce a viable protein is low but it is still a possibility. To rule this possibility and to further verify the double *irq4/5CR.1* and *irq4/5CR.2* CRISPR mutants a western blot analysis was carried out using the anti-IRQ4 antibody. The sequence used to generate the fragment of IRQ4 that was used to generate the antibody (see section 5.5) partially overlaps by 159 bp with the fragment amplified by the 5' reactions used in the RT-PCR. If any of the truncated transcript was not degraded and translated into protein the epitopes present with the fragment may allow the anti-IRQ4 antibody to bind to them and truncated protein would appear on a western blot.

Total protein was extracted from 15-day old vertically grown *irq4.1*, *irq4/5CR.1*, *irq4/5CR.2*, and (for *irq4.1* genotype) azygous plants as described in section 2.4.6. A single sample from two independent parent plants was used for the *irq4.1* line. The extracts were separated on a 10% SDS-PAGE gel and transferred to a nitrocellulose membrane (according to 2.4.3 and 2.4.8). Using the anti-IRQ4 antibody at a dilution of 1:500 the antibody detected a single band of approximately 65 kDa in the WT Col-0 and azygote samples corresponding to the IRQ4 protein (Figure 6.8). No bands were present in the *irq4.1* lanes showing that none of the *IRQ4* transcript was translated into protein and that the fragment that was produced in the RT-PCR was likely degraded and did not produce functional protein. Within the three lanes containing the CRISPR generated mutants a band was present at approximately 60 kDa (shown in Figure 6.9 by ΔIRQ4), slightly below the 65 kDa band expected for IRQ4. This shows that altered N-terminal sequence found during the CRISPR mutant analysis (see section 6.2.1 and Figure 6.2) produced a truncated mRNA product which was still able to be translated producing a truncated protein product. Although not a complete knockout for the *IRQ4* transcript these plants were still used for phenotyping as a truncated protein may change the ability of the IRQ4 protein to fold properly into its correct tertiary structure and as a result not perform its function within the cell. Also, the *IRQ5* gene was a true knockout in the CRISPR lines so these would be informative as additional independent knockout lines to supplement *irq5.1*.

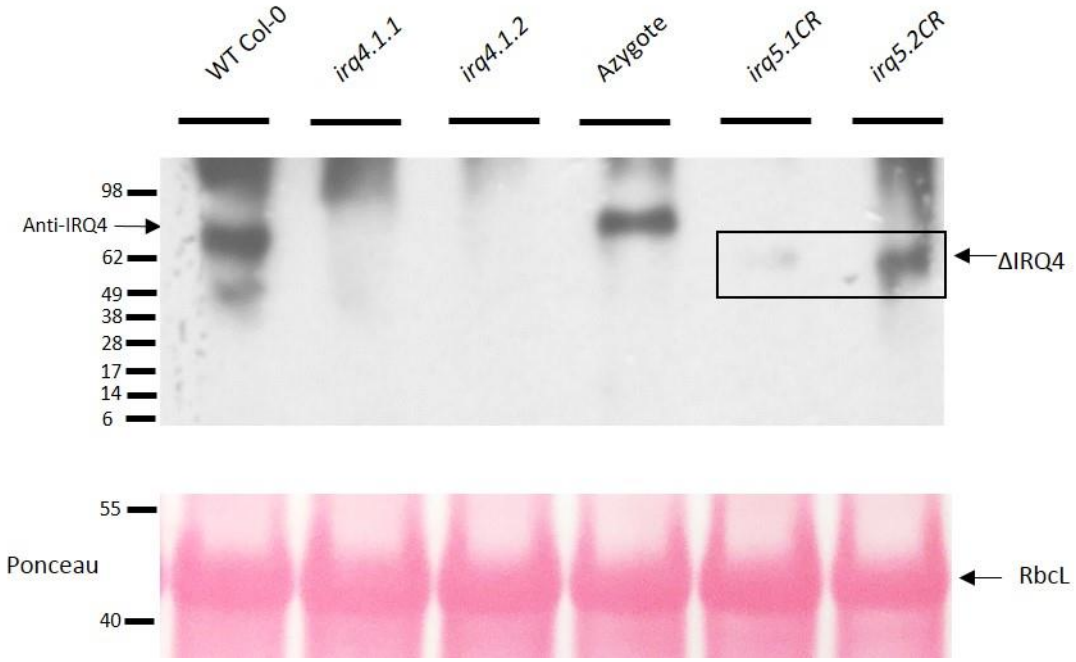


Figure 6.8: Western blot of mutant lines to check for absence of IRQ4 protein. Total protein extracted from 15-day old wild-type Col-0 *Arabidopsis,* the *irq4.1.1*, *irq4.1.2* parent mutant lines, *irq4.1* azygote mutant plants, and *irq5/4CR.1* and *irq4/5CR.2* mutant lines and was separated on a 10% SDS- PAGE gel and transferred to a nitrocellulose membrane for Western blot analysis. The membrane was cut into strips and probed with various dilutions of the anti-IRQ4 antibody at a dilution of 1:500. This was followed by a secondary anti-Mouse/HRP antibody (1:1500; Dako). The WT Col-0 and azygote lines produce a band at approximately 65 kDa, the correct size for WT IRQ4 protein. The *irq4.1.1* and *irq4.1.2* parent lines produce no band showing an absence of protein in the plant line and demonstrating the transcriptional null nature of the plant line. The CRISPR *irq4/5CR.1* and *irq4/5CR.2* lines produce a band (fainter in the *irq4/5CR.1* line than the *irq4/5CR.2* line) at approximately 62 kDa (shown by the ΔIRQ4). This shows that the CRISPR lines are not transcriptional null mutants and a truncated protein product is produced.

In summary the RT-PCR analysis has shown that the two T-DNA insertion lines (*irq4.1* and *irq5.1*) have their respective transcripts disrupted by the presence of the insert. In addition, western blotting using the anti-IRQ4 antibody has confirmed that no IRQ4 protein is produced in the *irq4.1* lines but a truncated product for IRQ4 is present in the *irq4/5CR.1* and *irq4/5CR.2* CRISPR lines. Having confirmed some form of disruption in transcript or of protein produced in these lines the *irq4.1*, *irq5.1* T-DNA insertion lines, and *irq4/5CR.1,* and *irq4/5CR.2* CRISPR lines were used for phenotypic testing.

6.3.4 Complementation of the *irq4.1* line by transformation with the *promIRQ4*::IRQ4-GFP construct

Validation of any potential phenotype requires that the phenotype in question is caused by the knockout of that particular gene. For the *irq4.1* line this was achieved by transforming the mutant line with the *promIRQ4*::IRQ4-GFP construct. Plants of the *irq4.1* line were grown on soil for 4-6 weeks until ready for flowering and were transformed by floral dipping as described in section 2.2.14. T_0 seed was selected according to section 2.6.4 and positives were transferred to soil. Positives were verified by PCR using a reverse GFP primer (83_GFP_R) and a gene specific primer (IRQ4_F3; Figure 6.9a). These primers were used to confirm that the construct was present in the plants. Figure 6.9b) shows a typical genotyping reaction showing the presence of the insert. Lines which were positive were allowed to grow and set seed. To check the localisation of the construct in the *irq4.1* background remained the same in these complemented lines, seeds of the plants were sterilised and sown on plates according to section 2.6.3. After 5-7 days of being grown vertically under standard conditions plants were imaged via CSLM. Figure 6.9d shows the typical localisation pattern of IRQ4 at the tonoplast membrane, similar to that seen in section 5.4. These lines were therefore confirmed to be complemented lines of *irq4.1* and were used for phenotypic analysis.

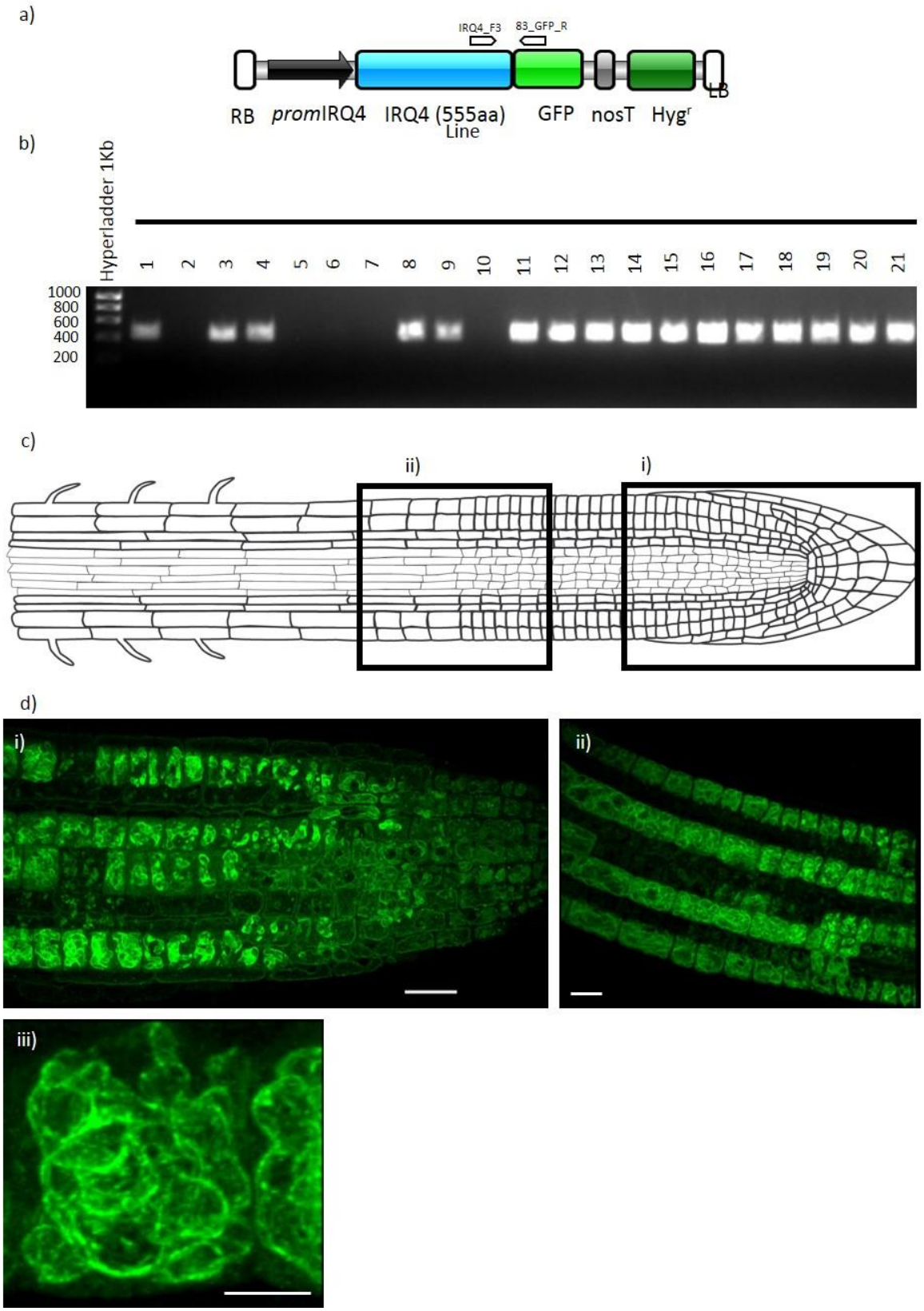


Figure 6.9: Complementation of the *irq4.1* line. a) The *promIRQ4*::IRQ4-GFP construct created using the pMDC107-GFP expression vector labelled with the locations of the primers used to genotype the complimented lines. b) Example PCR genotyping gel used to check for the presence of the insert in the dipped *irq4.1* line. c) Schematic of root showing approximate locations (black boxes) where images in d) were taken i) Root rip, ii) differentiation/elongation zone. Numbering reflects numbering in d) i.e. ci)=di), bii)=cii). d) Expression of *promIRQ4*::IRQ4-GFP in the *irq4.1* background showing the same localisation patterns seen for the *promIRQ4*::IRQ4-GFP construct dipped in the WT Col-0 background. i) Root tip, ii) differentiation/elongation zone, iii) single cell showing tonoplast localisation. Scale bars di-ii)=20 µm, diii)=5 µm

6.4 Analysis of Primary Root Growth in *irq4* mutant lines

Chapter 5 has shown that IRQ4 is localised to the late endocytic pathway, including the PVC/MVB, the vacuole, and to autophagasomes within the roots of *A. thaliana* plants. Roots were therefore studied to ascertain whether any discernible phenotype could be seen in this plant tissue and to assess whether IRQ4 has a functional role in root growth. Phenotypic tests focussed firstly on vacuole morphology as IRQ4 may have a structural role at the vacuole and the loss of IRQ4 may lead to defects in vacuolar shape or organisation. Due to the localisation of IRQ4 at autophagasomes and the disruption of the *promIRQ4*::IRQ4-GFP construct under nitrogen starvation conditions two tests were undertaken to assess root growth under autophagy inducing conditions and ascertain whether the autophagy response was impaired within these plants and whether this led to any gross changes in primary root growth.

6.4.1 Assessing the morphology of the lytic vacuole in unstressed *irq4* mutant lines

As discussed in chapter 1 a continuous flow of membranes and substances are brought to and taken away from the vacuole. If there any changes to the type or quantity of substances taken or removed to or from the vacuole vacuolar morphology defects may arise. For example mutations of RabG3f are inhibited in trafficking to the vacuole and have enlarged PVCs and altered lytic vacuole morphology (Cui *et al.*, 2014).

Using the pH indicating dye 2',7'-Bis-(2-carboxyethyl)-5-(6)-carboxyfluorescein, acetoxymethyl ester (BCECF-AM) allows the visualisation of the vacuole as the dye accumulates in the acidic lytic compartment and is an established dye for vacuole visualisation (Feraru *et al.*, 2010; Nodzynski *et al.*, 2013; Viotti *et al.*, 2013; Kolb *et al.*, 2015; Scheuring *et al.*, 2015). Initially the dye is membrane permeable but after crossing the membrane cellular esterases cleave the acetoxymethyl ester, trapping the dye within the compartment. Visualisation of the dye is possible due to its two excitation peaks at 440 nm and 495 nm. The 440 nm peak is pH insensitive and the 495 nm peak is pH sensitive. The ratio between these two intensities allows the intracellular pH to be calculated (Boens *et al.*, 2006). As only structural information was required the dye was only excited at 488 nm allowing the vacuole to be visualised using CSLM.

Five-day old seedlings of the F3 generation of *irq4.1*, *irq5.1*, *irq4/5CR.1,* and *irq4/5CR.2* CRISPR, azygous, and WT Col-0 lines were incubated for one hour in 10 µM BCECF-AM, and then briefly washed for 10 minutes before imaging using CLSM. Five roots were analysed per genotype and a representative example of each is shown in the maximum projections shown in Figure 6.10. No significant differences were observed between the WT Col-0 lines, azygous or mutant lines as the vacuoles appeared to occupy the majority of the cell volume in those cells imaged at or very near

the epidermis. This indicates that loss of IRQ4 and/or IRQ5 does not cause any gross morphological changes to the lytic vacuole either during its development in the meristematic zone or in more mature vacuoles in the differentiation and elongation zones.

Due to the imaging conditions used it may be that vacuoles in deeper cell files of the root not imaged here may be disrupted. To assess this, lines of the mutant plants could be transformed with a marker for the vacuole such as VAMP711 (yellow fluorescent protein-Vesicle Associated Membrane Protein 711; Geldner *et al.*, 2009) to allow the vacuole to be imaged at greater depth. The transformed lines could be imaged less invasively over time to track the development of earlier lytic compartments in more detail and compare between the genotypes.

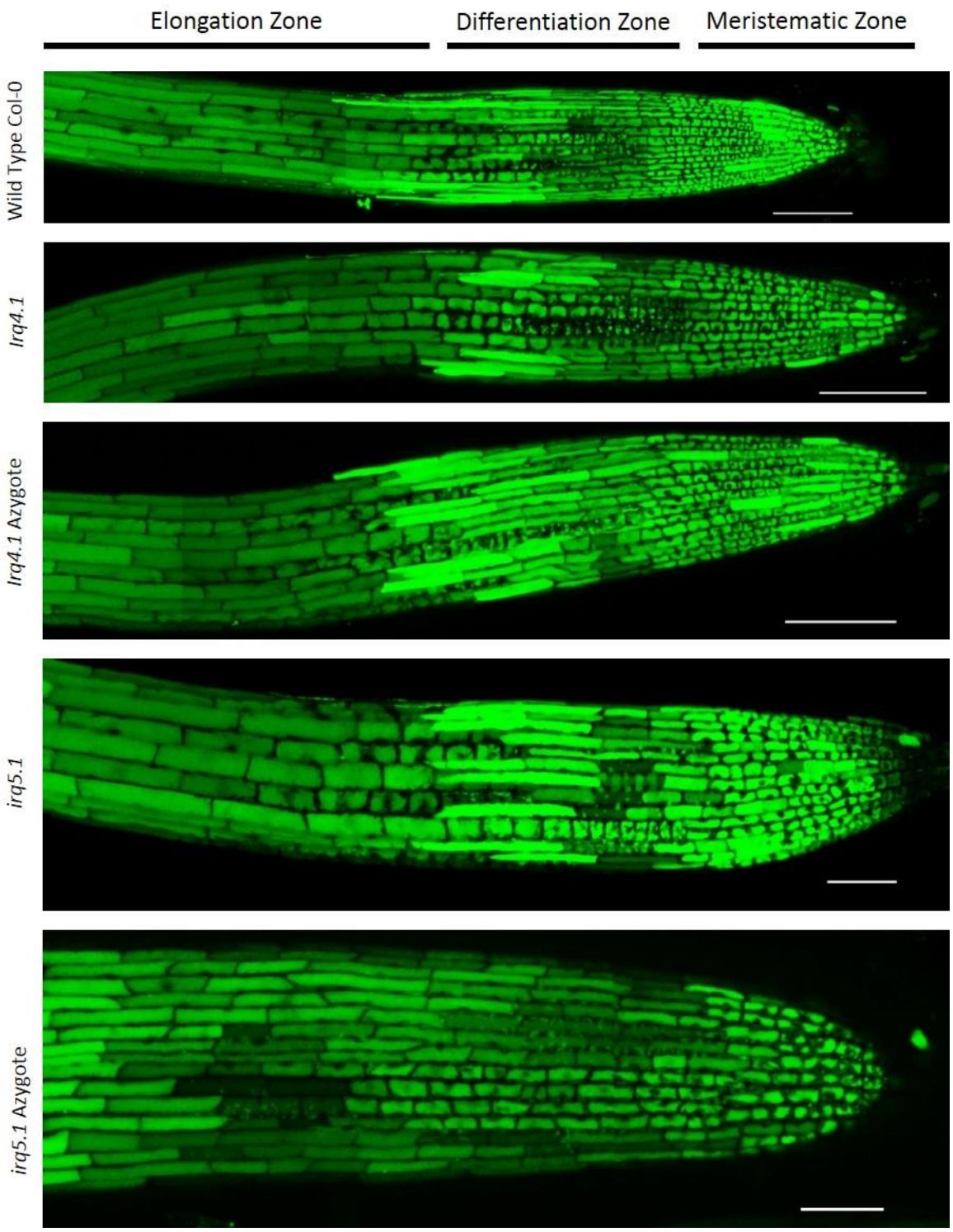


Figure 6.10: Comparison of lytic vacuole morphology in T-DNA insertional and CRISPR generated mutants, azygote, and WT Col-0 lines. Seedlings of the relevant lines (indicated to the left of each image) grown for 5-7 days were stained with the pH sensitive dye BCECF-AM (10 µM) for one hour. BCECF-AM accumulates in the lytic vacuole before and can be imaged by exciting the dye at 488 nm by CSLM. No phenotypic differences were observed in vacuole morphology between all the lines. Scale bars: WT Col-0, *irq4.1*, *irq4.1* Azygote=100 µm; *irq5.1*, *irq5.1* Azygote=50 µm.

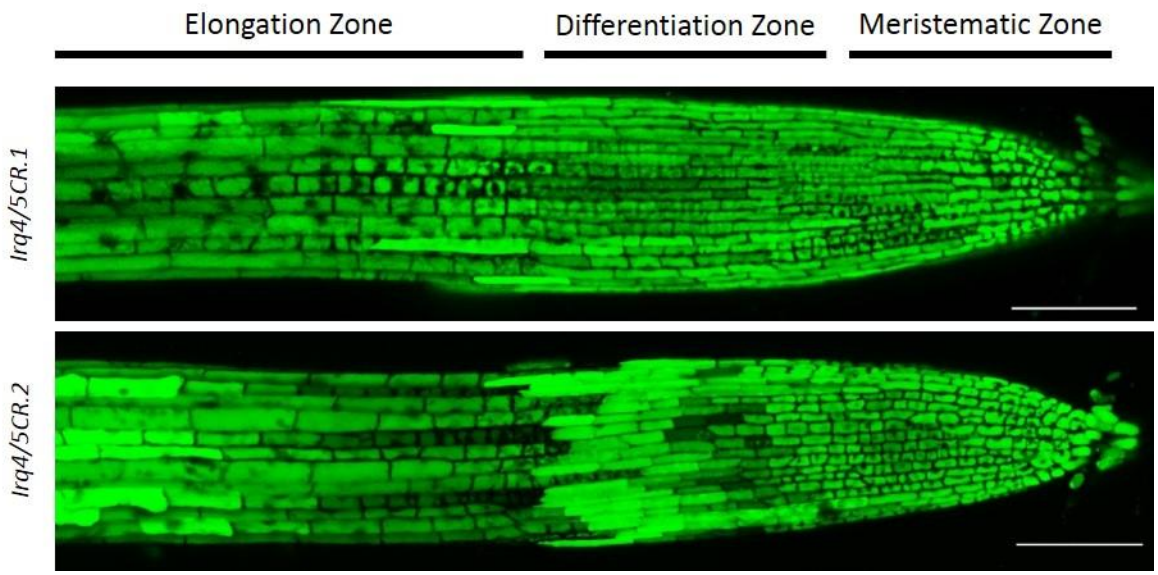


Figure 6.10 continued: Comparison of lytic vacuole morphology in T-DNA insertional and CRISPR generated mutants, azygote, and WT Col-0 lines. Scale bars: *irq4/5CR.1*, *irq4/5CR.2*= 100 µm

6.4.2 Analysis of plant growth under autophagy inducing conditions

The localisation of IRQ4 to autophagasomes and the significant decrease in IRQ4-labelled MVB/PVC described in chapter 5 may link IRQ4 to autophagy in some way. Two different test were used to assess primary root growth: nitrogen starvation and salt stress.

A common phenotype of Arabidopsis mutants defective in autophagy is hypersensitivity to nutrient stress e.g. carbon or nitrogen starvation which leads to earlier chlorosis in cotyledons and rosette leaves and/or earlier senescence (Doelling *et al.*, 2002; Hanaoka *et al.*, 2002b; Chung *et al.*, 2010). Growth of mutants defective in autophagy are significantly shorter under nitrogen starvation (Wang *et al.*, 2016a). Plants with autophagy related phenotypes are also hypersensitive to salt stress with chlorosis of leaves occurring faster in autophagy deficient mutant lines (Liu *et al.*, 2009; Luo *et al.*, 2017).

6.4.2.1 Analysis of primary root growth under nitrogen starvation conditions in irq4 and *irq5 mutants*

The nitrogen starvation assay followed the protocol of Wang *et al.* (2016). Seedlings of the F3 generation of *irq4.1*, *irq5.1*, *irq4/5CR.1,* and *irq4/5CR.2* CRISPR, azygous, complemented *irq4.1* and WT Col-0 lines were sterilised and sown on plates containing either normal nitrogen containing media (+N) or on media lacking nitrogen (-N) as described in section 2.6.6. Plants were grown for 6 days after germination then the plates were scanned using an Epson flatbed scanner before measuring the length of the primary roots using ImageJ (according to section 2.7.3). The mean root lengths of all the mutant, azygote, complemented, and WT Col-0 lines were calculated.

Two separate analyses were undertaken to answer two different questions. The first question: do the factors of plant genotype and nitrogen interact? In other words, knowing whether nitrogen containing media or nitrogen-free media was used is important in understanding the effect of genotype on average root length and *vice versa*. This question was answered using a two-way ANOVA statistical test. The second question was whether growing plants on autophagy inducing media has a disproportionate effect on knockout mutant plants compared to azygote, WT, and the complemented lines on the same media and on the –N or +N media. Two way ANOVA could not be used to answer this question as the factor media only had two levels (+/-N) which does not allow sufficient degrees of freedom to do post-hoc pair wise comparisons such a Tukey tests in SPSS. Splitting the data into its constituent line and media combinations allowed this post-hoc test to be applied. It should be noted that all the data used in the nitrogen starvation assay originates from one single experiment.

Table 6.4: summary data for two-way ANOVA

Source	df (degrees of freedom)	MS (mean square)	F	*P*
Genotype	8	939.2	62.2	<0.001
Media	1	2859.1	189.5	<0.001
Genotype x Media	8	144.3	9.6	<0.001
Error	1107	15.1		

A significant interaction was found between the plant genotype and the type of media the seedlings were sown on (Two-way ANOVA; $F_{8,1107}$=9.6, P<0.001). This strongly indicates that the presence or absence of nitrogen in the media and the genotype of the plants are both important factors for determining primary root length. In general, plants grown on –N media had longer roots than those grown on +N media across all genotypes tested (Figure 6.11). An increase in root length has been observed in seedlings when transferred from +N media plates to –N media plates and this increase in root length was thought to be a nutrient scavenging response in these plants (Zheng *et al.*, 2016).

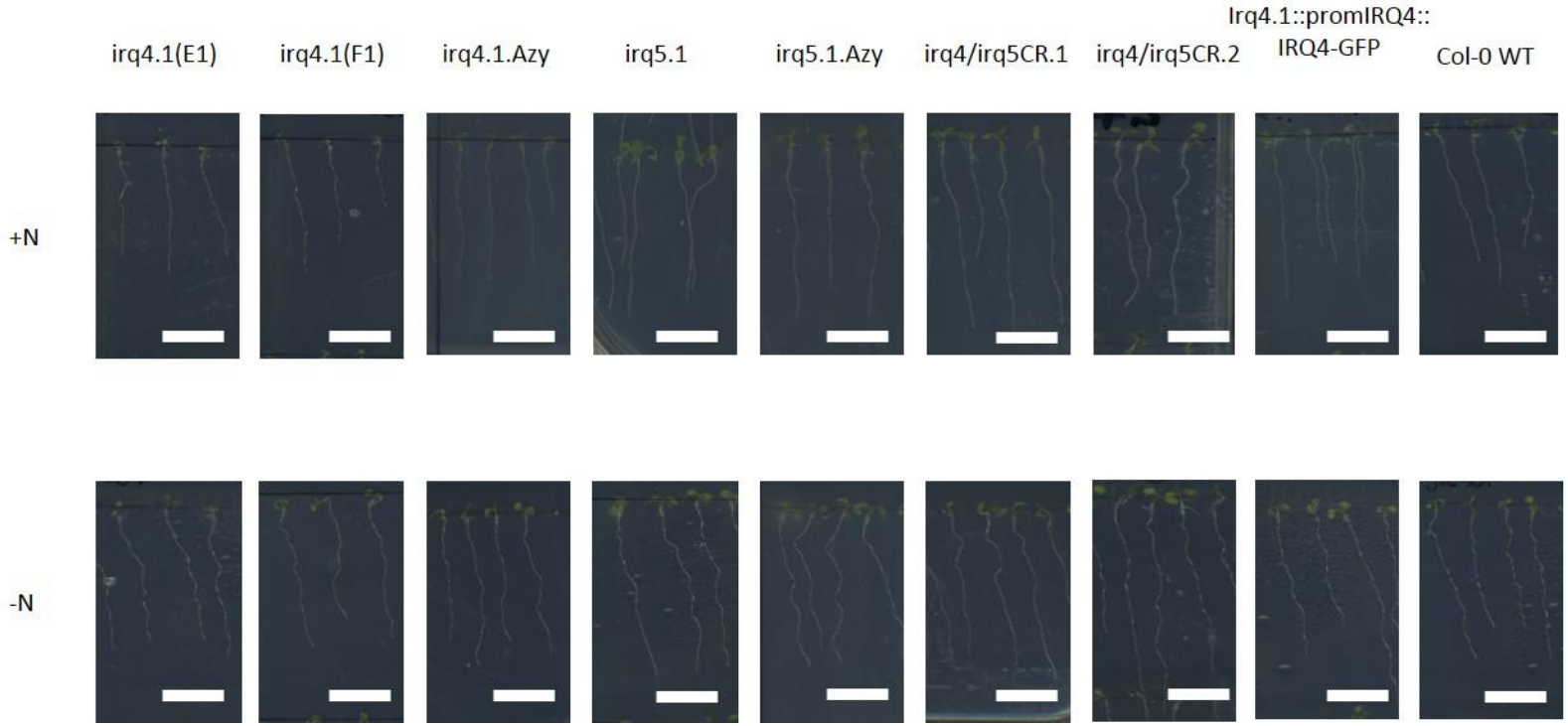


Figure 6.11: Seedling roots of mutant, azygote, and WT Col-0 after six days growth. Seeds of the F3 generation of the mutant, azygote, and WT lines were surface sterilised and sown onto either +N or –N containing plates. Seedlings were left to grow for six days then imaged. Roots were measured and means calculated. Scale bar: 10 mm.

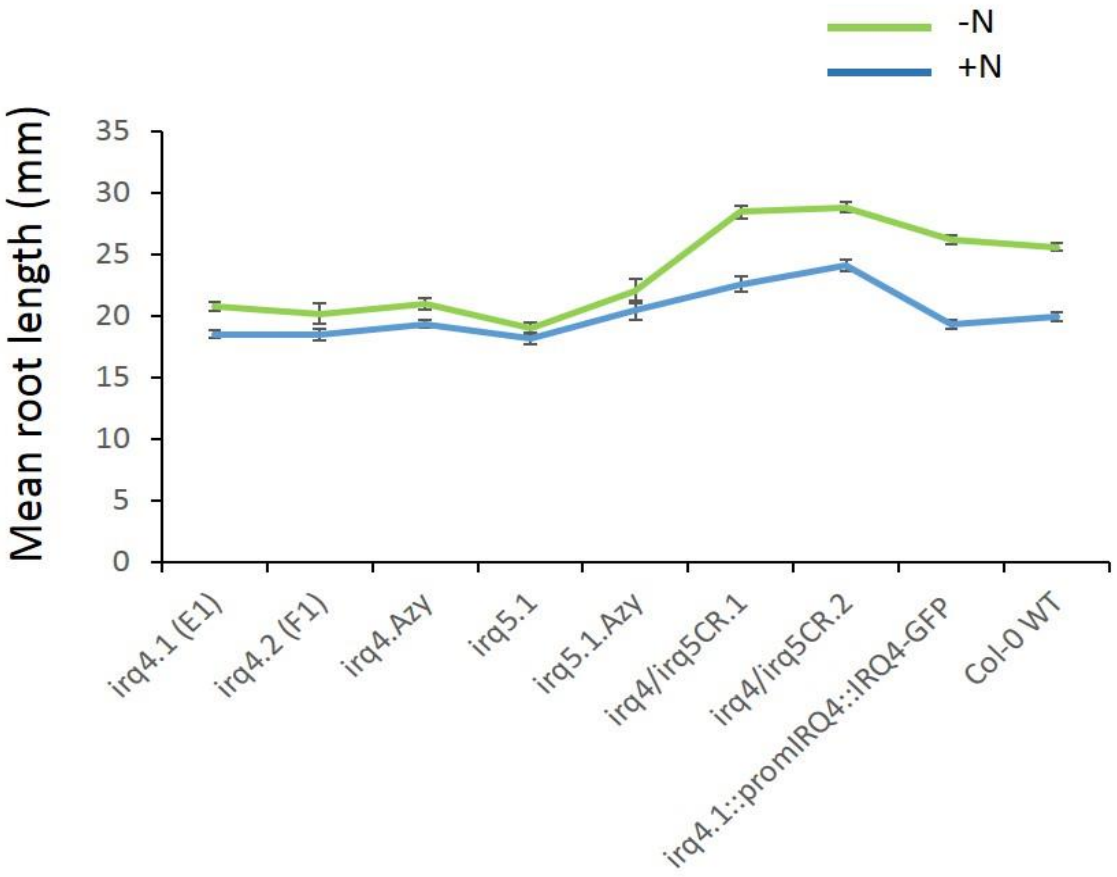


Figure 6.12: Mean root length of mutant, azygote, and WT Col-0 seedlings measured after six days. Seeds of the F3 generation of the mutant, azygote, and WT lines were surface sterilised and sown onto either +N or –N containing plates. Seedlings were left to grow for six days then imaged. Roots were measured and means calculated. Points joined by a blue line indicate plants grown on +N media, points joined by a green line were grown on -N media.

To compare between genotypes and media one-way ANOVA was carried out. Overall there were significant differences in mean primary root length across genotypes and media combinations (ANOVA; $F_{17,1151}$=56.0, P<0.001). The most important comparisons are presented in Table 6.5: number of seedlings measured, mean root lengths (mm) per line, the standard error of the mean (SEM), and 95 % confidence intervals are shown for seedlings grown on +N media (green columns) and –N media (yellow columns). Comparisons between lines were calculated using Tukey tests. Three different overall comparisons were done: 1. Mean root length of mutant lines was compared to their respective azygote lines (in the case of T-DNA insertion lines) or the WT Col-0 line (in the case of CRISPR generated mutants) using only +N grown plants (red column), 2. Mean root length of mutant lines were compared to their respective azygote lines (in the case of T-DNA insertion lines) or the WT Col-0 line (in the case of CRISPR generated mutants) using only -N grown plants (blue column), 3. Mean root length of each line on +N media was compared to itself when grown on –N media (grey column).

The mean root length of the two parent *irq4.1* plants was not significantly different compared to the azygote when grown on the same media. Neither were either of the *irq4.1* parent lines when compared to themselves on the opposite (+N compared to –N) media. Together these results indicate that the *irq4* single line does not have an impaired autophagy response when under starvation conditions.

The mean root length of the *irq5.1* line was not significantly different to the azygote when grown the same media. There was also no significant difference in mean root length of the *irq5.1* when its growth on –N media was compared to its growth on +N media. Together these results indicate that the *irq5* single line does not have an impaired autophagy response to starvation conditions.

The mean root lengths of the *irq4/5CR.1 and irq4/5CR.2* lines were significantly longer when compared to the WT on the same media as shown in the red and blue columns. Similarly both CRISPR lines had significantly longer roots when compared to themselves on +N and –N media. Both CRISPR lines consistently had the longest roots in both +N and –N conditions. These results indicate that the CRISPR lines are more resistant to autophagy inducing conditions compared to the WT and that disrupting both IRQ4 and IRQ5 protein production has an enhancing effect on primary root growth irrespective of media. This enhancement in primary root growth on –N media compared to the WT only occurs when both proteins have been disturbed. The enhancement is not seen in the single *irq4.1* and *irq5* mutant lines when they are compared to their respective azygote lines. Disruption to *IRQ5* transcription leads to insensitivity to abscisic acid (ABA; Ren *et al.*, 2016). ABA is known to cause inhibition of root growth at high concentrations (Luo *et al.*, 2014a; Rowe *et al.*, 2016). This enhancement in root growth was not noted in the current experiment using the single *irq5.1* line compared to the azygote on normal +N media potentially because ABA levels were not experimentally manipulated. However, as IRQ4 and IRQ5 share sequence similarity disturbing both *IRQ4* and *IRQ5* may produce an ABA insensitive response in these lines. Experiments to verify the ABA sensitivity of the single and CRISPR double lines would be required confirm this possibility.

In summary, across the single and double mutant lines none of the lines exhibit a typical autophagy deficiency response to nitrogen starvation. However, this assay has shown that disruption of both *IRQ4* and *IRQ5* transcription leads to significantly longer roots regardless of media type.

Genotype	Number of plants analysed (+N)	Mean primary root length (mm)	Standard error	95 % Confidence Intervals	Number of plants analysed (-N)	Mean primary root length (mm)	Standard error	95% Confidence Intervals	Tukey Test (*P*): +N/+N azygote/WT	Tukey Test (*P*): -N/-N azygote/WT	Tukey Test (*P*): +N/-N
irq4.1.1	57	18.5	0.31	17.87-19.13	63	20.8	0.35	20.06-21.45	0.999	1.000	0.130
irq4.1.2	53	18.5	0.44	17.57-19.37	54	21.0	0.44	18.53-21.82	0.999	1.000	0.697
irq4.1.Azy	65	19.3	0.31	18.73-19.97	62	20.2	0.82	20.07-21.86	1.000	<0.001	0.647
irq5.1	51	18.1	0.48	17.19-19.09	45	19.0	0.45	18.09-19.92	0.691	0.136	1.00
irq5.1.Azy	20	20.5	0.73	18.91-22.05	26	22.0	0.97	19.98-24.05	1.000	0.006	0.998
irq4/5CR.1	77	22.6	0.62	21.32-23.81	70	28.4	0.56	27.31-29.52	0.003	0.001	<0.001
irq4/5CR.2	73	24.1	0.46	23.15-24.98	80	28.8	0.45	27.93-29.72	<0.001	<0.001	<0.001
*irq4::promIRQ4::*IRQ4-GFP	110	19.3	0.36	18.60-20.02	104	26.2	0.40	25.38-26.98	1.000	1.000	<0.001
WT Col-0	76	19.9	-	19.16-20.61	83	25.8	0.33	24.91-26.26	-	-	<0.001

Table 6.5. Statistical analysis of mean root lengths of mutant, azygote and WT plants grown on either +N or –N media

6.4.2.2 Analysis of irq4 and irq5 mutant seedlings under salt stress

Plants deficient in autophagy show reduced tolerance to salt stress (Liu *et al.*, 2009; Luo *et al.*, 2017). In addition salt stress is an osmotic stress (Zhu, 2002; Munns & Tester, 2008). The methodology used for this assay is as described by Wang *et al.* (2016). Seedlings of the F3 generation of *irq4.1*, *irq5.1*, *irq4/5CR.1,* and *irq4/5CR.2* CRISPR, azygous, complemented *irq4.1* and WT Col-0 lines were sterilised and sown on plates containing normal MS media (as described in section 2.6.7). Seedlings were grown until 5 days post germination then either transplanted to fresh plates containing normal media with no salt (-S) or to plates containing media supplemented with 160 mM NaCl (+S). Seedlings were then allowed to grow for a further 6 days after which seedlings were scored for bleaching of hypocotyls. The cotyledons of seedlings which are more sensitive to salt stress bleach faster than those which do not. The number of seedlings which had bleaching symptoms or no bleaching was summed and used to calculate the percentage survival.

All seedlings transferred from –S to –S media showed 100 % survival. Of those transferred to +S media the WT showed the highest survival rate at 70.6 % (Figure 6.12). Survival rates for all other lines varied from 39.4 % to 50 %. Although there is large difference in survival rate between the WT and the *irq4/5CR.1* and *irq4/5CR.2* lines, mutants deficient in autophagy display more dramatic decreases in survival in comparison to their azygote counterparts with survival rates in the order of 2.5 % to 20 % in the case of the *nap1, piri121,* and *arp2/3* mutants screened by Wang *et al.* (2016). The evidence presented here suggests that the mutant plants do not show enhanced susceptibility or resistance to salt stress compared to their azygote counterparts or the WT in the case of the CRISPR mutant lines.

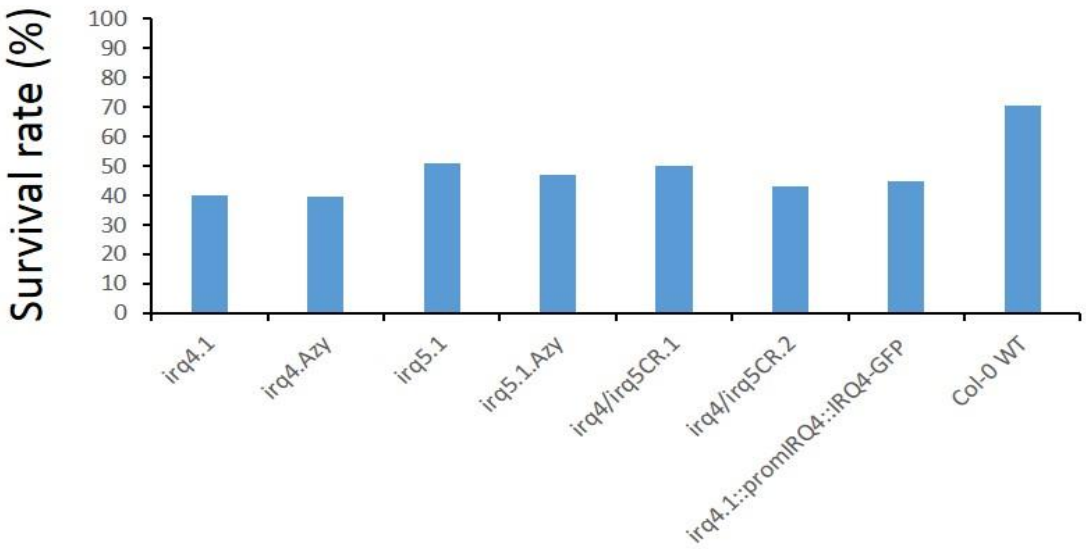


Figure 6.13: Survival rate of mutant, azygote and WT seedlings on high salt media. Seedlings of the F3 generation of mutant, azygote, and WT were sown on normal plates and allowed to grow for five days. At five days plants were transferred to media containing 160 mM NaCl and allowed to grow for a further six days. Plants which showed salt stress symptoms were counted. Survival rate was calculated by dividing the number of seedlings showing salt stress symptoms and dividing by the total number of plants for each respective genotype and multiplying by 100.

6.5 Conclusion

The work in this chapter aimed to determine a function for IRQ4 and a possible genetic interaction with *IRQ5*. Reverse genetics was employed and two different approaches were used to generate mutants in order to test for possible phenotypes. Firstly, two single T-DNA insertional mutants, *irq4.1* and *irq5.1*, were identified and confirmed to have disturbed transcription by RT-PCR, and in the case of *irq4.1*, by western blotting. The second approach involved the use of CRISPR/Cas9 gene editing technology to attempt to generate a double knockout line of *IRQ4* and *IRQ5*. Sequencing confirmed that the bi-allelic mutation in the *irq4/5CR.1* had completely knocked out the vast majority of the protein while the mutation in the *irq4/5CR.2* line could not be amplified by PCR indicating a complete disruption to the gene. The deletion in *IRQ4* had left the remaining sequence in-frame and, through western blotting, was shown to give a truncated protein product. To obtain a full knockout of *IRQ4,* more double *irq4/irq5* mutants are being screened to identify further mutations and the target sites used to target Cas9 are currently being redesigned by the laboratory. The CRISPR lines were still used for the phenotypic analysis as although the *IRQ4* gene was not completely knocked out the truncated protein product may have structural defects which cause mis-localisation or reduce its function within the cell.

Chapters 4 and 5 had established that IRQ4 was expressed in the roots and localised to the vacuole. The morphology of the lytic vacuole in the mutant lines was studied using BCECF-AM

staining. No gross morphological differences in vacuolar shape were observed between any of the T-DNA or CRISPR lines compared to the WT and azygote lines. The localisation of IRQ4 to the PVC/MVB may indicate a non-structural function for IRQ4 at the vacuole but is possible that IRQ4 may have role in younger stages of root development as the embryo develops into a young seedling. Transforming the mutant lines with a vacuolar marker such as VAMP711 will allow the development of the vacuole to be monitored over time using lightsheet microscopy in order to screen for such developmental defects. The imaging technique used here only allowed visualisation of the vacuole at or very close to the epidermal layer of cells. Again, transformation with a vascular marker would allow imaging deeper into the root and across different cell files.

Immunogold labelling using the anti-IRQ4 antibody had suggested that IRQ4 localises to autophagasomes in the roots. Primary root growth was tested using two different autophagy inducing conditions: nutrient starvation, in the form of nitrate starvation, and salt stress. In addition, salt stress is a form of osmotic stress so this test allowed for an initial estimation of osmotic stress tolerance in the mutant lines.

No particular autophagy related phenotype could be detected in the nitrate starvation assay between any of the lines. All lines exhibited increased root growth on –N media, with the CRISPR and *irq4* complemented lines showing significant increases in root length when grown on –N media compared to +N media. Increased root growth on –N media is associated with nutrient foraging behaviour (Zheng *et al.*, 2016). A knockout of *irq5* has been shown to be insensitive to ABA (Ren *et al.*, 2016) showing increased primary root growth when treated with higher concentrations of ABA compared to the WT. The longer root phenotype was not seen in the single *irq5.1* mutant used in this thesis but a longer root phenotype was seen in the CRISPR lines which had disturbed transcription of both *IRQ4* and *IRQ5* regardless of nitrogen content of the media used. The root phenotype may not be apparent without exogenous application of ABA in the *irq5.1* mutant but disturbance of both *IRQ4* and *IRQ5* may impair the response to natural levels of ABA within the root potentially linking IRQ4 to auxin signalling in some way. Further experiments would need to be done to monitor the response of root growth of the single T-DNA insertion lines and the double CRISPR mutants using exogenous application of ABA.

Increases in root length are unusual for plants grown on –N media as generally plants grown this way tend to have shorter roots (Wang *et al.*, 2016a). A possible explanation for this is that these other experiments did not image the roots at the same time point as here so it is difficult to compare results. Also, only a single time point was used here and it is possible that the more 'typical' shorter root response to nutrient starvation may manifest itself at a later time point. The experiment here was only done using one biological replicate so the results of this experiment

would need to be verified by repeating it and perhaps measuring root lengths at multiple time points to see if 1. The shorter root phenotype of roots typical for plants grown on –N media appears and 2. To see if any nutrient stress response is apparent at different stages of root elongation.

An investigation of primary root growth under salt stress was carried out and revealed that no autophagy related response to salt stress was seen in the T-DNA insertion or CRISPR lines. Impaired autophagy responses can be relatively dramatic with survival rates of between 2.5 % to 20 % being seen (Wang *et al.*, 2016a). The lack of any such decrease in survival seen in the mutant lines compared to their azygote or WT counterparts indicates that the mutant lines are not impaired in their autophagy response in response to salt stress. The salt stress assay also functioned as an osmotic stress and may indicate that the mutant lines do not have impaired response to drought stress either. However, in terms of the osmotic stress this assay is only preliminary and more thorough drought-specific stress tests would need to be undertaken using a variety of osmoticants including mannitol and potassium chloride. Protocols for drought tests are different to that used for testing autophagy responses, tending to use lower salt concentrations and monitoring root growth over time in a similar manner to that used for the nitrogen deprivation assay presented here (for examples see (Chen *et al.*, 2014; Rai *et al.*, 2015; Jiang *et al.*, 2016).

No autophagy related phenotype has been detected for knockouts of *IRQ4* and *IRQ5* however this does not rule out that such a phenotype is present. Chapter 5 (section 5.6) demonstrated that localisation of the *promIRQ4*::IRQ4-GFP construct was disturbed when plants were grown on –N media, indicating autophagy inducing conditions do affect IRQ4. However, whether IRQ4 has a role in autophagy is currently unknown. Further experiments could verify a role for IRQ4 in autophagy by repeating the nitrogen starvation assay done here with a greater number of seedlings. The *promIRQ4*::IRQ4-GFP construct containing plants could also be transformed with ATG8, which acts as a marker for autophagy (Contento *et al.*, 2005), and the nitrogen starvation assay repeated to track the increase in autophagasome number under autophagy inducing conditions with the decrease in IRQ4-labelled PVC/MVB to see if there is any overlap in signal.

This chapter has shown that no phenotypic differences were observed in the structure of the lytic vacuole nor in responses to autophagy inducing conditions in the mutant line used here. However, increased primary root growth unlinked to the type of media used and autophagy was noted in the CRISPR lines containing a full knockout of *IRQ5* and a truncated IRQ4 protein. A link between IRQ5 and ABA has already been established (Ren *et al.*, 2016) and further experiments are required to verify if the elongated root phenotype seen here is a result of impaired response

to ABA in the CRISPR lines. In the next chapter potential interaction partners of IRQ4 were investigated using the yeast-2-hybrid (Y2H) system with the aim of understanding the phenotype given here and identify possible further avenues for phenotypic testing of mutants.

Chapter 7 Investigating the Interacting Partners of the IRQ proteins

7.1 Introduction

The previous chapters have demonstrated that IRQ4 is expressed in the roots and is transcribed in this location. Subcellular localisation studies have shown that IRQ4 localises to the tonoplast membrane and the PVC/MVB and potentially autophagasomes. Taken together these results suggest that IRQ4 is involved in trafficking in the late secretory pathway.

Previous work by Dr. D. Mentlak and Dr T.J. Hawkins demonstrated that the IRQ domain plus the C-terminus of NET4B was sufficient and required for the interaction between NET4B and the Rab GTPase RabG3f (Mentlak, 2016). This led to the hypothesis that the IRQ domain is a Rab binding domain and that the IRQ proteins are Rab binding proteins.

This chapter describes experiments undertaken to investigate the potential interacting partners of IRQ4 using the yeast-2-hybrid (Y2H) system. First, an explanation of the yeast-2-hybrid system is given. Second, one-on-one assays between Rab GTPases representative of the various Rabs families and IRQ4 were undertaken. Finally, a yeast-2-hybrid library screen was completed to find other potential interacting proteins for IRQ4. Identifying proteins that interact with IRQ4 may indicate the function of the protein.

7.2 The Yeast-2-Hybrid System

The Y2H system is a well-established technique for investigating protein-protein interactions (Chien *et al.*, 1991) and has been used for proteins derived from a wide range species from mammals to plants. The technique involves the construction of plasmids containing hybrid proteins: one contains the DNA-binding domain (BD) of the GAL4 transcription factor from yeast (*Saccharomyces cerevisiae*) fused to the protein of interest (bait), while the other consists of the GAL4 activation domain (AD) fused to either another known protein (prey; one-on-one assay) or to a protein encoded by a cDNA Library derived from the organism of interest (Y2H library screen). Successful interaction of BD and AD creates a functional GAL4 transcription factor which activates downstream reporter genes which contain binding sites for GAL4.

The Clontech Matchmaker Y2H system was used for all the yeast assays described in this chapter. The system uses two vectors: pGBKT7 and pGADT7. pGBKT7 contains the DNA-binding domain and pGADT7 contains the activation-domain fusions. Both these vectors contain a selectable marker gene to enable selection of positive transformants: pGBKT7 contains a gene for the production of tryptophan while pGADT7 contains a gene for the production of leucine. Yeast

transformed with these vectors are then able to survive on media lacking these amino acids so can be positively selected.

In order to test for interactions both constructs need to be in the same yeast cell and this is achieved by mating the yeast. The two yeast strains used in these experiments were the AH109 and Y187. These strains contain reporter genes under the control of the GAL4 transcription factor and are activated when a protein-protein interaction occurs. Y187 contains the *LacZ* reporter gene which encodes a gene producing β-galactosidase, while AH109 contains *HIS3* encoding a gene producing histidine, *ADE2* encoding a gene producing adenine, and *LacZ*. Yeast are mated and successfully mated yeast are selected on media lacking both tryptophan and leucine to select for diploid colonies.

To test for protein-protein interactions diploid colonies are plated on various different combinations of media lacking tryptophan to select for bait, leucine to select for prey, and adenine and/or histidine to select for reporter genes. Successful protein-protein interactions activate the GAL4 transcription factor allowing the yeast to survive on media lacking amino acids in various combinations to change the stringency of the screen (media lacking all amino acids is more stringent than media just lacking adenine or histidine). β-galactosidase can also be used as an additional marker by growing the mated yeast on media containing the substrate X-α-Gal. Upon successful protein-protein interaction mated colonies turn blue.

7.3 Testing the interaction between full length IRQ4 and Rab GTPases using the Y2H system

As previous work had shown that the IRQ domain was essential for the interaction between NET4B and RabG3f a screen was conducted on a selection of Rab GTPases. As described in chapter 1 Rab GTPases have undergone an adaptive radiation with many Rabs having plant specific functions (Rutherford & Moore, 2002; Woollard & Moore, 2008). The following Rabs were selected for this initial screen: RabF1 (ARA6), RabF2a (RHA1), RabF2b (ARA7), RabG3f, Rab1b (ARA5/RabD2A), and RabA5c (ARA4). As previous chapters have demonstrated IRQ4 localises to compartments associated with the late secretory pathway, so Rabs were chosen which are predominantly involved in these pathways.

RabF1 (ARA6) has been colocalised with the SNARE proteins AtSYP21 and AtSYP22, both of which label the PVC and tonoplast (Robinson *et al.*, 2008). RabF2a (RHA1) and RabF2b (ARA7) have both been associated with trafficking from Golgi to the PVC (Ebine *et al.*, 2011). RabG3f also localises to the PVC and tonoplast (Cui *et al.*, 2016). RabA5c have been localised to Golgi and adjacent vesicles (Rutherford & Moore, 2002). Rab1b is involved in ER to Golgi transport (Barr, 2013).

The full length IRQ4 protein was cloned into the pGBKT7 and pGADT7 expression vectors to create the activation domain and binding domain fusion proteins. Before undertaking the screen an autoactivation test was done. Autoactivation tests ensure that the created construct does not activate the downstream reporter genes itself. IRQ4 in pGBKT7 and pGADT7 were both transformed into yeast strains AH109 and Y187 (two clones in each strain). Yeast were selected in media either lacking tryptophan (-W) or leucine (-L). Successfully transformed strains of yeast containing pGBKT7 (both Y187 and AH109) were resuspended in water (three independent colonies of each) and dropped on the following selective media: -L, -LH (media lacking leucine and histidine), -LH plus 2.5 mM, 5 mM, or 10 mM 3-amino-triazole (3-AT), -LHA (media lacking leucine, histidine, and adenine), -LHA plus 2.5 mM, 5 mM, or 10 mM 3-AT, and -L plus 100 mM X-α-Gal (5-Bromo-4-chloro-3-indoyl-α-D-galactopyranoside). Successfully transformed strains of yeast containing pGADT7 (both Y187 and AH109) were resuspended in water (three independent colonies of each) and dropped on the following selective media: -W (media lacking tryptophan), -WH, -WH plus 2.5 mM, 5 mM, or 10 mM 3-AT, -WHA, -WHA plus 2.5 mM, 5 mM, or 10 mM 3-AT, and -W plus 100 mM X-α-Gal. Colonies were dropped onto -L or -W as a positive control. 3-AT is a competitive inhibitor of the *HIS* gene product and can be used to inhibit the known 'leaky' expression of the *HIS* reporter gene, as basal activity may give rise to false positives (Van Criekinge & Beyaert, 1999; Tian *et al.*, 2012).

Figure 7.1 shows that after 6-7 days growth the control –W and –L plates both allowed colony growth. No colony growth for any construct/strain combination was seen on plates lacking histidine and/or adenine showing that no autoactivation occurs on media lacking these amino acids. A slight blue colour could be seen on the -W/-L plus 100 mM X-α-Gal plates so to remove ambiguity the *LacZ* reporter was not used for any further experiments.

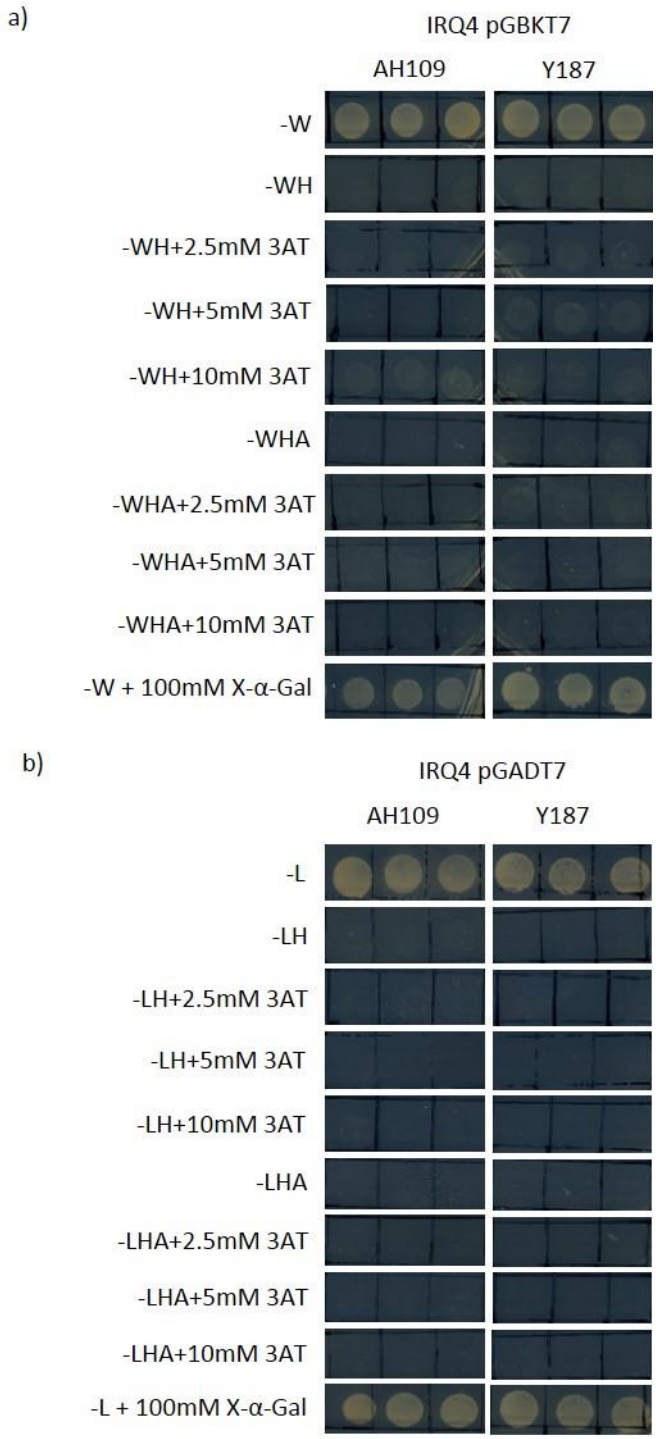


Figure 7.1: Autoactivation tests of IRQ4 in pGBKT7 and pGADT7. The full length IRQ sequence was cloned into the pGBKT7 (a) and pGADT7 (b) vectors and both were transformed into yeast strains AH109 and Y187. Autoactivation of these constructs was tested by plating them on various media lacking tryptophan (to select for pGBKT7), and either histidine (H) or adenine (A). X-α-Gal was used to assess autoactivation of the *LacZ* reporter gene. Three colonies of each construct were grown on the selection plates indicated for a total of 7 days. Both the constructs grew on their respective control –L or –W plates showing they were alive while no growth was recorded on any of the nutrient selection plates demonstrating no autoactivation took place on those plates. There was a slight blue colouration on the X-α-Gal plates so these were not used for selection.

Having confirmed that IRQ4 in either of pGBKT7 or pGADT7 does not cause autoactivation this construct was used to screen the Rabs mentioned earlier. The full length RabA5c, Rab1b, RabG3f, and Rab1b clones were obtained from Dr. D Mentlak while RabF1 and RabF2b were cloned into the entry vector pDONR207 from cDNA using the Gateway system as described in section 2.2.3. The clones were then transformed into either the pGBKT7 or pGADT7 destination vectors for the interaction assay. Note, cloning of Rab1b and RabF1 into the pGBKT7 vector failed and was not used in this assay.

For one-on-one mating the full length Rab clones and IRQ4 were transformed into either yeast strain AH109 (all pGBKT7 containing clones) or Y187 (all pGADT7 containing clones). Mated yeast as negative controls empty pGBKT7 and pGADT7 were mated, IRQ4 and the Rabs in pGBKT7 and pGADT7 were mated with their respective empty controls (either pGADT7 for pGABT7 or pGABKT7 for pGADT7). As a positive control NET4A in pGADT7 was mated with RabG3f in pGBKT7. Two negative controls were also used with NET4A: NET4A in pGADT7 was mated with RabA5c in pGBKT7, an interaction which is known to be negative (Mentlak, 2016), and with the empty pGBKT7. Diploid cells were selected on –WL media and grown on several selection plates: -WL (as a positive control), -WLA, -WLH + 2.5 mM 3AT, and -WLHA + 2.5 mM 3AT.

The results of the screen can be seen in Figure 7.2. All of the mated yeast grew on the –WL plates however none, apart from the positive control reaction, grew on the selective media indicating that IRQ4 does not interact with the Rabs chosen here. Patchy growth was seen in the RabG3f/NET4A interaction on –WLA as had been shown previously but there was strong growth evident on the –WLH+2.5 mM 3AT plates indicating a positive interaction (Mentlak, 2016). Swapping the orientation and putting IRQ4 into pGBKT7 and the Rabs (RabA5c, RabF2a, and RabG3f) did not produce an interaction indicating the orientation of the construct has no effect on the occurrence of the assay for IRQ4.

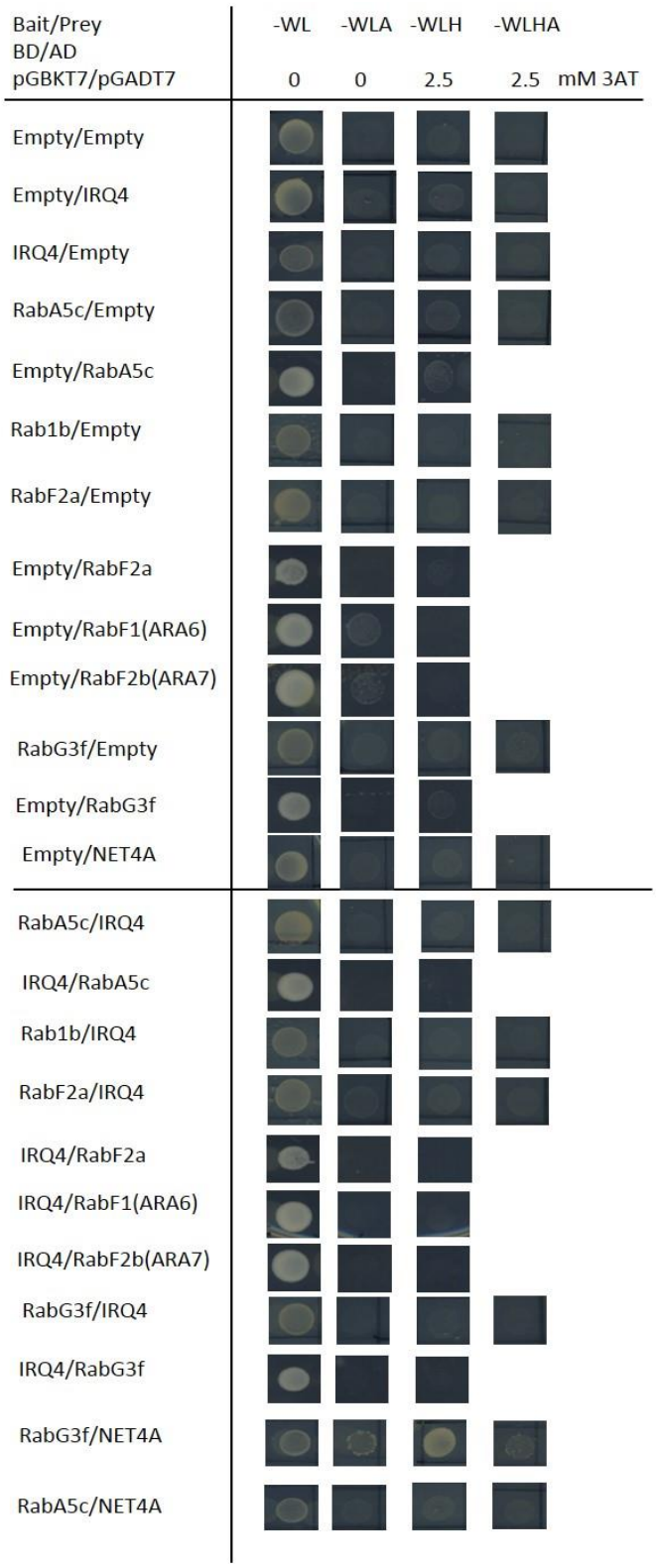


Figure 7.2: One-on-one mating test between representative Rab GTPases and the full length IRQ4 protein. IRQ4 in both the pGBKT7 and pGADT7 vector was used to screen representative Rab from various groups associated with relevant endomembrane trafficking pathways (for details see text). Combinations of IRQ4 in the pGBKT7 and the Rab in the pGADT7 vector and *vice versa* are shown apart from Rab1b, RabF1 (ARA6), and RabF2b (ARA7) which were only used in one orientation. In addition, NET4A was mated with RabG3f (positive control) and RabA5c (negative control). Successfully mated diploids were dropped onto various selective media as shown and allowed to grow for 5 days. One representative colony is shown for each combination. All the –WL plates showed growth while no growth was seen on the other selective plates indicating that no positive interactions took place between IRQ4 and the Rabs. The positive control for NET4A and RabG3f showed colony growth on the –WLH+2.5 mM 3AT indicating that the technique used here was correct.

Truncating proteins can sometimes remove areas of the protein which cause steric hindrance, preventing the interaction from taking place (Rajagopala & Uetz, 2009). The IRQ domain of IRQ4 was cloned into the pDONR entry vector and then transformed into the destination vector pGADT7 using the gateway system as described in section 2.2.3. NET4A, RabG3F, and RabA5C were used as positive and negative controls as described previously. The IRQ4:IRQ construct was transformed into yeast strain Y187 and mated with either the empty pGBKT7 vector control or with the Rabs in the AH109 strain. Mated yeast were selected on –WL plates for diploid selection. Successfully mated colonies were then dropped onto various selective plates: -WL (as a positive control), -WLA, -WLH + 2.5 mM 3AT, and -WLHA + 2.5 mM 3AT.

The results of the assay can be seen in Figure 7.3. The –WL controls all showed colony growth indicating the yeast mated successfully. The positive NET4A control also worked indicating that the assay was performed correctly. No colony growth was seen in the IRQ4:IRQ colonies on the selective plates either with the empty vector control or with the Rabs. The latter result indicates that the truncation has no effect on whether IRQ4 binds to the Rabs shown here.

In summary the Rab binding assay shown here indicates that IRQ4 does not bind to any of the Rabs used in this assay, either in its full length form or when only the IRQ domain of IRQ4 is used.

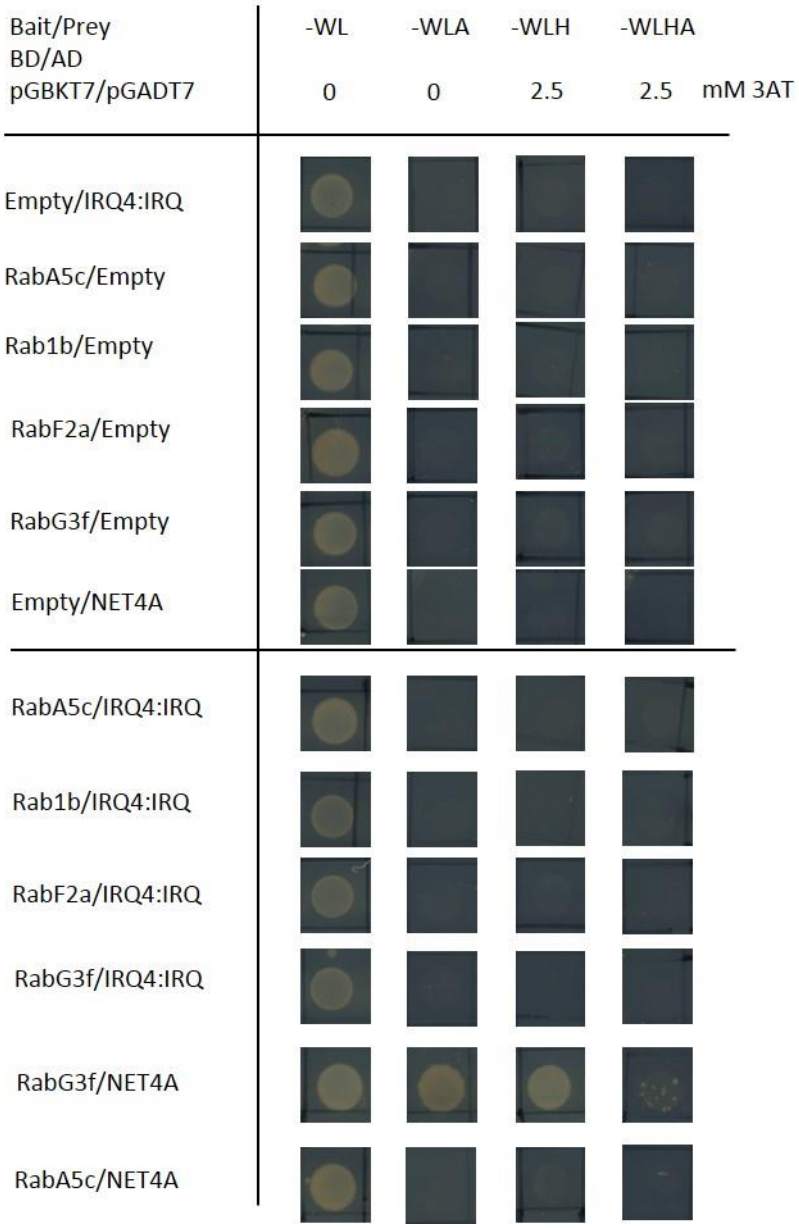


Figure 7.3: One-on-one mating test between the IRQ domain of IRQ4 and Representative Rab GTPases. The IRQ domain of IRQ4 (IRQ4:IRQ) was cloned into the pGADT7 vector and mated with the Rabs in the pGBKT7 vector. IRQ4:IRQ was also mated with the empty pGBKT7 vector as a negative control. In addition, NET4A was mated with RabG3f (positive control) and RabA5c (negative control). Successfully mated diploids were dropped onto various selective media as shown and allowed to grow for 5 days. One representative colony is shown for each combination. All the –WL plates showed growth while no growth was seen on the selective plates indicating that the IRQ4 domain does not interact with selected Rab GTPases. The positive control for NET4A and RabG3f showed colony growth on the –WLH+2.5 mM 3AT indicating that the technique used here was correct.

7.4 Investigating the interacting partners of full length IRQ4 using the Y2H system

7.4.1 Conducting the Y2H library screen using the full length IRQ4 sequence as bait

As has been discussed in the introduction, besides Rab GTPases there are a large number of other proteins involved in the endomembrane system. To investigate other potential interacting proteins for IRQ4, a Y2H library screen using the full length IRQ4 protein as bait was undertaken.

The screen used the full length clone of IRQ4 in pGBKT7 as the cDNA library was cloned into pGADT7. The previous autoactivation tests had shown that media lacking only tryptophan and histidine was sufficient to stop autoactivation of the bait construct. However, prey plasmids within the cDNA library may encode proteins which are able to grow on this media independent of an interaction with IRQ4 and using such a low stringency may produce a large number of false positives. To reduce the probability of such an occurrence a relatively high stringency level was used for the screen using plates of –WLHA+2.5 mM 3AT. The screen was conducted using a mixture of an oligodT and random primed cDNA library from *Arabidopsis* seedlings according to the methods in section 2.3. The mixture of oligodT and random primed cDNA libraries were used to capture both full length transcripts as well as fragments in the screen. Seedling cDNAs were used as IRQ4 is known to be expressed at this stage according to Affymetrix data. Two independent screens were performed with the second screen using a lower amount of library cells and using a larger number of plates in the screen. These changes were made to reduce the level of stress placed on the cells and to increase the efficiency of the screen without reducing the number of potential colonies. The mating efficiency and the approximate number of diploid cells screened are shown in Table 7.1. Mated yeast were then spread on numerous –WLHA+2.5 mM 3AT plates containing 50 $\mu g ml^{-1}$ kanamycin to suppress other fungal growth and left at 30 ^{o}C to grow. Colonies began to appear after 4-5 days and plates were left for 13-14 days. After this period colonies were re-streaked on fresh –WLHA+2.5 mM plates containing 50 $\mu g ml^{-1}$ kanamycin. Plates were left at 30 ^{o}C to grow for 3-4 days. Colonies which grew on these selection plates were presumed to be potential interactors of IRQ4 and the corresponding cDNA clone was rescued according to section 2.3.5. Rescued plasmids were sequenced using the pGADT7 sequencing primers as described in section 2.2.1 (see appendix 1 for primer sequences).

Table 7.1: Mating efficiencies and number of diploids screened in the Y2H library screen

Bait	Approximate mating efficiency (% of diploids)	Approximate number of diploid clones screened
IRQ4 (Screen 1)	32.81	1.14E+08
IRQ4 (Screen 2)	81.7	1.06E+09

7.4.2 Results of the Y2H library screen using the full length IRQ4 sequence as bait

Both screens combined gave a total of 10 rescued clones. A full list of these clones is given in appendix 4. Some of these clones were isolated more than once with some appearing in both screens. The most pertinent are listed in Table 7.2.

Table 7.2: Prey plasmids selected for one-on-one mating

Gene Locus	Gene Annotation	Number of Colonies	Screen (1 or 2)
At3g01090	Arabidopsis thaliana SNF1 kinase homolog 10 (AKIN10)	11	1, 2
At4g29130	Hexokinase 1	3	2

Two of the clones were heat shock proteins, however these are frequently false positives (Serebriiskii *et al.*, 2000) and so were not taken further. Criteria for taking clones through to further validation were: the interactor had been isolated more than once as distinct and independent clones; the transcript is usually expressed *in planta* within the roots as IRQ4 also localises to this tissue; and finally the protein is associated with the late endosomal pathway as IRQ4 is also found to localise to this part of the endomembrane system in plant cells.

Two interactors which met the criteria above were taken further and tested for their interaction in one-on-one Y2H assays: *Arabidopsis thaliana* SNF1 kinase homolog 10 (SnRK1.1; At3g01090) and Hexokinase 1 (HXK1; At4g29130).

SnRK1.1 is part of the *Arabidopsis* SnRK super family. SnRK1 is involved in metabolic responses to sugar deprivation as well as a host of other abiotic and biotic stress responses (Crozet *et al.*, 2014; Lastdrager *et al.*, 2014; Rangarajan *et al.*, 2014; Emanuelle *et al.*, 2015). In the yeast homologue of SnRK, Snf1, there are 3 subunits to the kinase: Snf1 catalytic (α) subunit, the Snf4 stimulatory (γ) subunit, and three isoforms of the β subunit, Sip1, Sip2, and Gal83. The Sip1 β subunit is found in the cytosol when cells are grown in the presence of glucose but, upon switching to glycerol-

ethanol, relocalises with Snf1 to the vacuole membrane (Hedbacker *et al.*, 2004). The vacuole is a central storage hub for carbohydrates and reserve proteins (Marty, 1999) and therefore an interaction with SnRK1.1 may link IRQ4 to the regulation of metabolism in plants. In addition, SnRK1.1 has been shown to be a positive regulator of autophagy (Chen *et al.*, 2017) and as IRQ4 was shown to localise to autophagasomes using immunogold labelling and TEM (see Figure 5.2.3) this is another positive indicator of an interaction and involvement in the autophagasome pathway.

HXK1 is a histidine kinase having a dual role as both a phosphorylating enzyme and acting as a glucose sensor in plants. The glucose sensing and kinase functions are independent however (Harrington & Bush, 2003; Moore *et al.*, 2003). HXK1 is a key enzyme in the regulation of carbohydrate metabolism in plants (Halford & Paul, 2003; Nägele *et al.*, 2014) and was identified in the vacuolar proteome screens (Carter *et al.* 2004; Jaquinod *et al.* 2006). As HXK1 is not regulated directly by SnRK1.1 or *vice versa*, IRQ4 may have a further link with carbohydrate metabolism independent of SnRK1.1.

The remaining proteins were not pursued further in this analysis as they did not meet the validation criteria. Although Dihydrolipoamide succinyltransferase (At5g55070) was returned multiple times each clone returned contained the same sequence. CLK4-associating serine/arginine-rich protein (At4g36980) is a relatively unknown protein which is annotated as being localised to the nucleus and contains a SWAP domain (suppressor-of-white-apricot-protein) which in humans is hypothesised to act as an mRNA splicer in the nucleus. In plants it is annotated as being at the nucleus so is unlikely to interact with IRQ4 at the vacuole. Quinolinate Phosphoribosyltransferase (At2g01350) is involved in NAD synthesis at the chloroplast so is unlikely to interact with IRQ4 at the vacuole. Two of the results from the screen were ambiguous as it was unclear from the sequence returned which of two proteins was actually the true potential interactor: At5g25270 or At2g28950. At5g25270 is a ubiquitin-like super family protein and is annotated on TAIR as being expressed at the nucleus. At2g28950 encodes Expansin 6. Expansins are proteins which are involved in cell expansion (Bayer *et al.*, 2006) and EXPA6 is localised to the cell wall. As neither At5g25270 nor At2g28950 match the all the criteria listed above they were not continued in the analysis.

7.4.3 Confirming interactors of IRQ4 from the screen using on-on-one Y2H mating

To confirm the interaction between IRQ4 and the proteins selected from the screen the fragments were tested in a one-on-one mating test with full length IRQ4. The recovered clones were transformed into *E. coli* DH5α cells, grown, and the plasmids isolated according to sections 2.2.6, 2.2.8. The resulting plasmids were then re-transformed back into the yeast strain AH109.

For the one-on-one mating tests the rescued plasmids in strain AH109 were mated with the full length IRQ4 in yeast strain Y187. As negative control the prey plasmids in pGADT7 were also mated with empty pGBKT7 containing cells while IRQ4 in pGBKT7 was mated with empty pGADT7.

Diploid cells were selected on –WL media and dropped on several selection plates: -WL (as a positive control), -WLA, -WLH + 2.5 mM 3AT, -WLH + 10 mM 3AT, -WLHA + 2.5 mM 3AT, -WLHA + 10 mM 3AT. Different media was used to test for the stringency of the interaction in the one-on-one experimental set-up.

The results of the test are shown in Figure 7.4. The IRQ4 in pGBKT7 mated with the empty pGADT7 only shows colony growth on the –WL mating control plate, showing that IRQ4 does not autoactivate and the empty pGADT7 is truly empty. The SnRK1.1 and HXK1 colonies when mated with their respective empty vectors all grew on the selective media plates potentially, indicating that these represent a false positive result in the screen.

In summary, none of the potential interactors pulled out of the Y2H library screen could be verified to interact with IRQ4 in a one-on-one assay indicating that they may represent false positive results. However, these will revaluated in future work.

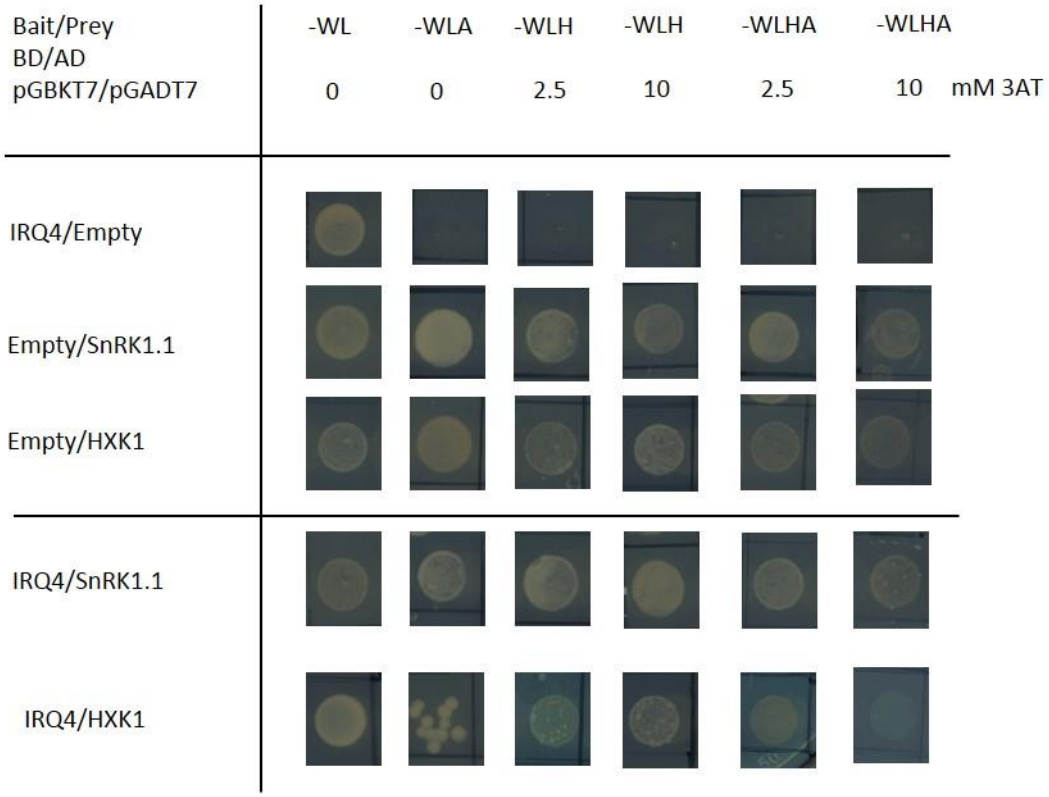


Figure 7.4: One-on-one mating tests between IRQ4 and the SnRK1.1 and HXK1 prey fragments recovered from the Y2H screen. Rescued bait constructs in pGADT7 were mated with the empty pGBKT7 to control for autoactivation and were also mated with full length IRQ4 in pGBKT7. Successfully mated diploids were dropped onto various selective media as shown and allowed to grow for 5 days. One representative colony is shown for each combination. Growth occurred on all the empty vector controls for the potential interacting fragments indicating that these colonies were the result of autoactivation rather than true interaction.

7.5 Conclusion

This chapter aimed to find interacting partners of IRQ4 using the Y2H system through two approaches: one-on-one assays to screen a selection of Rab GTPases for an interaction and conducting a full Y2H library screen on seedling cDNA using the full length IRQ4 protein as bait. None of the Rab GTPases tested against the full length IRQ4 protein or when tested against the IRQ domain alone of IRQ4 was found to interact. This is a surprising results considering the IRQ domain of NET4B was found to interact with the RabG3 subclass of Rab GTPases and it was the IRQ domain that was found to mediate this interaction. The Rabs and full length IRQ4 were tested in both orientations i.e. IRQ4 cloned into the pGBKT7 vector and Rabs in the pGADT7 vector and *vice versa*, therefore the failure of the proteins to interact was not due to a vector compatibility issue.

False negatives can occur in the Y2H system for a number of reasons. The fusion of either or both of the activation domain (AD) and binding domain (BD) to the proteins of interest may be in a spatial configuration such that they can no longer form a functional unit. This lack of ability to forma functional unit could be due to either steric hindrance or their positions may stop the transcriptional activation of the yeast reporter genes. For this reason the IRQ domain alone of IRQ4 was used to screen the Rabs but this also failed (see Figure 7.3). For the activation of the reporter genes both the proteins of interest need to be present in the nucleus. It may be that IRQ4 was not translocated to the nucleus due to protein instability. For the RabG3f failure to be transported to the nucleus is unlikely because of its use a positive control in this thesis in its interaction with NET4A. Also, if the interaction between IRQ4 and the Rabs is transient then this may be missed in the Y2H system.

A common reason for full length proteins to fail to interact in Y2H assays is due to the requirement for post-translational modifications to the protein before interaction can occur (Brückner *et al.*, 2009; Rajagopala & Uetz, 2009). Post translational modifications involve either the transfer of accessory groups to translated proteins to alter or confer additional functionality e.g. phosphates, carbohydrates, lipids, or the proteolytic cleavage of a particular part of the protein (Mann & Jensen, 2003). Some proteins remain inactive until some post-translational modification causes a change in their conformation revealing a specific domain. Within plant cells this occurs due to some stimulus or cellular event, e.g. cold, heat, cell division. Yeast may not possess some or all of the proteins required to modify IRQ4 to enable it to interact with its Rab interacting partner. If the post translational modification required for IRQ4 to interact was known it could be possible to co-express the modifying protein at the same time as IRQ4 and the potential interactor. This has been done successfully with tyrosine-phosphorylation dependent interactions (Osborne *et al.*, 1996).

A further biological reason for the failure of the IRQ4 protein to interact with the Rabs could be that despite the possession of an IRQ domain IRQ4 does not interact with Rabs *in planta*. The Rab binding capability of the IRQ domain has only been shown for two of the more evolutionary ancient NAB/IRQ proteins: NET4A and NET4B. None of the other NAB/IRQ nor IRQ domain alone containing proteins have been tested for Rab binding capability. Possession of a domain known to confer interacting ability with a specific set of proteins or substrates does not imply that all proteins possessing the same domain will interact with the same targets. One example concerns the carbohydrate binding module (CBM) of SnRK1.1 from *Arabidopsis*. CBMs allow binding to carbohydrate such as glycogen. The CBMs in the SnRK1.1 did not bind to starch in any (Emanuelle *et al.*, 2015). This was surprising as the CBMs in SnRK1.1 belong to family 48 CBMs, some members of which are known to bind to starch in plants (Chaen *et al.*, 2012). A further example is

that of Somatic Embryogenesis Receptor-like Kinase 2 (SERK2) and Brassinosteroid Insensitive 1 (BRI1). The SERKs are a group of leucine-rich repeat receptor like kinases (Hecht *et al.*, 2001) involved in several unrelated signalling pathways (Albrecht *et al.*, 2005, 2008; Roux *et al.*, 2011). Residues within the extracellular domain of the protein which are important for its interaction with BRI1 are present in SERK2 but they do not interact *in vivo* (Toorn *et al.*, 2015). A final possibility is that IRQ4 interacts with a very specific Rab which was not used in this Rab screen. Further Rabs will need to be cloned to test this theory.

As a second way to find potential interacting partners of IRQ4, a Y2H library screen was conducted using full length IRQ4 as bait against a cDNA library from 5-7 day old *Arabidopsis* seedlings. The screen was repeated with the differences being fewer library cells were used and the cells were plated on a greater surface area during the second screen. A combination of oligodT and random primed library cells were used each time (see section 2.3.3 for details of screen).

Two of the most interesting potential interacting partners for IRQ4 were SnRK1.1 and Hexokinase 1 (HXK1). Both they and their homologs are well known components of sugar signalling and/or sensing pathways in animals, yeast, and plants. In support for a link between IRQ4 and SnRK1.1, SnRK1.1 has recently been shown to have a role in autophagy (Chen *et al.*, 2017). Using immunogold labelling and TEM, IRQ4 was shown to localise to autophagasomes (see Figure 5.5.5). In addition, in the yeast homolog of SnRK1.1, Snf1, the Sip1 β subunit of the complex relocalises to the vacuole from the cytosol taking Snf1 with it when switched from media containing glucose to media containing glycerol-ethanol (Hedbacker *et al.*, 2004).

Unfortunately none of the potential interactors pulled out from the screen could be verified in a one-on-one assay (see section 7.4) indicating they were potentially false positive results. In addition to the reasons given above for the Rab interaction assays, a potential reason for the failure of the library screen to distinguish true interactors could be because the library used may not have been truly representative and therefore not contain the true interactors. High stringency conditions were used throughout the screen to minimise the potential for autoactivation occurring. The high stringency of the screen may have meant that weaker but true interactors may not have come through the screen and not formed colonies for subsequent isolation. Also, it can be difficult to recover more than 50% of true interactions using a Y2H library screen (Galletta & Rusan, 2015). Further experiments to find interactors may have to use alternative techniques such as proteomics screens and immunoprecipitation.

Chapter 8 Discussion

8.1 Overview

The aim of this thesis was to investigate the IRQ domain containing group of proteins in *Arabidopsis thaliana*, with a focus on IRQ4. The IRQ group of proteins were identified from previous work showing that NET4A and NET4B interact with the RabG class of Rab GTPases via a conserved domain named the IRQ domain (Mentlak, Unpublished thesis). Bioinformatics analysis revealed this domain to be present in the NET3A and NET3C proteins as well as six other proteins possessing no NET Actin Binding (NAB) domain. These six proteins were termed the IRQ proteins due to the characteristic 'IRQ' sequence in the highly conserved part of their IRQ domain.

Building upon the earlier bioinformatics work, in chapter 3 phylogenetic analyses were used across a variety of species from crucial points in the evolutionary history of land plants to show that at first, IRQ domain alone containing proteins evolved from the NET proteins. Second, IRQ1 and IRQ6 likely evolved earliest with IRQ4, IRQ5, IRQ2, and IRQ3 following later. Finally, the IRQ proteins may be a eudicot specific group of proteins.

The localisation of transcription of four of the IRQ genes, *IRQ1, IRQ6, IRQ4,* and *IRQ5*, was analysed in chapter 4 using the GUS reporter system. The main findings were that *IRQ6*, and especially *IRQ1*, localisation strongly indicates that these proteins are involved in the emergence and development of lateral roots. The expression of *IRQ5* generally agreed with previously published data regarding its transcription (Ren *et al.*, 2016). IRQ4 was highly expressed in root tissues.

The localisation of the IRQ4 protein was revealed in more detail in chapter 5 wherein translational fusions to GFP were transiently expressed in *N. benthamiana* leaf epidermal cells and stably expressed in BY2 cell culture and *Arabidopsis* plants. IRQ4 was shown to localise to the PVC/MVB, vacuole, and, using immunogold labelling and TEM imaging, was also likely to localise to autophagasomes in *Arabidopsis* root tips.

To ascertain a function for IRQ4 at these late endocytic compartments, in chapter 6 T-DNA insertional mutants of *IRQ4* and *IRQ5*, along with CRISPR-Cas9 generated double *IRQ4* and *IRQ5* mutant plants, were phenotypically analysed. The CRISPR double mutant plants had an increased root length phenotype.

Finally, in chapter 7 the potential interacting partners of IRQ4 were investigated using the Y2H system. SnRK1.1, a key protein involved in sugar signalling and metabolic regulation, was revealed to be a potential interacting partner for IRQ4.

This chapter will discuss the results obtained in this thesis and put them into a wider context. The results from the Y2H screens will be discussed. The potential contribution of IRQ4 to the elongated root phenotype found within the double CRISPR lines will be assessed and the relationship between IRQ4 and IRQ5 will be analysed. Potential roles for IRQ4 in the late secretory pathway and the links between autophagic processes, endocytosis and PVC/MVB trafficking to the vacuole will be explored with hypotheses on the role of IRQ4. A summary of the roles of IRQ4 is provided in Figure 8.1 Finally, the implications of the IRQ proteins being eudicot specific will be discussed for IRQ1, IRQ6, IRQ2, and IRQ3, looking at what contributions these IRQ proteins may have made to this fundamental split in the angiosperms.

8.3 Potential roles of IRQ4 and IRQ5 based on phenotypic analysis

Overall, no differences in vacuole morphology or differences related to autophagy were found in any of the mutant lines studied in this thesis. The lack of autophagic phenotype at the whole tissue level in relation to nutrient stress and salt stress is in agreement with results from chapter 5 showing a decrease in PVC/MVB number labelled with the *promIRQ4::*IRQ4-GFP construct upon transferring the seedlings to media lacking nitrogen. Typically the number of autophagasomes increases upon transfer to media lacking nitrate as has been shown using the well-established autophagasome marker Atg8 (Bassham, 2015). This does not discount a role in autophagy for IRQ4 and IRQ5 as the autophagic phenotype could be more subtle and not detectable as a gross morphological tissue change. For example there is a relationship between PVC/MVB and autophagasomes (discussed in section 8.4.1), so using specific markers for PVC/MVB and markers specific for autophagasomes in the mutant backgrounds given here may discern a relationship between the two.

The lack of phenotype shown by analysis of the lines transferred to high salt media may suggest that in addition to the lack of whole organ autophagic phenotype there was no phenotype associated with osmotic stress between the lines tested. However, this particular salt assay is typically used only in relation to autophagy and not osmotic stress. More subtle effects on the plant may be missed using a simple survivorship count as no data relating to root length or vacuole morphology were taken. Osmotic stress tests generally use lower salt concentrations in the media and are carried out using different osmoticants such as mannitol and KCl in addition to NaCl.

Although the nutrient stress tests did not reveal an autophagic phenotype a potential phenotype associated with the double CRISPR lines was shown. The CRISPR lines were found to have significantly longer primary roots compared to all other genotypes tested including azygote and WT controls regardless of whether seedlings were grown on media containing nitrate or not (Figure 6.11 and Table 6.5). The single *irq5.1* mutant line used in this study had significantly shorter roots compared to the WT but only when comparing *irq5.1* seedlings grown on +/-N with WT grown on –N. No differences were observed between *irq5.1* seedlings and the azygote *irq5.1Azy* control across either media type. The work shown in this thesis is in agreement with the only other single *irq5* mutant that has been studied to date with no significant differences in primary root length found between the *cip1-1* line and the WT control when grown on media containing nitrate with no additional ABA (Ren *et al.*, 2016).

The significantly elongated roots seen in the CRISPR lines could be associated with the impaired ABA response for the single *irq5* T-DNA insertion mutant (Ren *et al.,* 2016). Auxin inhibits the growth of roots (Rowe *et al.*, 2016). The ABA insensitive phenotype only related to water loss and germination assays, not root elongation. IRQ4 is expressed to a similar level in the root to IRQ5, as shown in chapters 4 and 5. Although no link has been shown in this thesis the similarity between IRQ4 and IRQ5 sequences could lead to functional redundancy. Therefore the elongated phenotype is only seen when both *IRQ4* and *IRQ5* transcription is disrupted. This could reflect a role for IRQ4 in the same capacity as IRQ5, with up-regulation of ABA biosynthesis genes but this needs to be established for IRQ4.

Another related possibility is that disruption to *IRQ4* transcription could lead to an intracellular transport phenotype associated with the movement of substances between the PVC/MVB and vacuole, specifically in relation to ABA. Any phenotype is likely to be subtle as no tissue level phenotype was seen in root length for the single *irq4.1* mutant line compared to the *irq4.1Azy* azygote control. There are known links between size and morphology of the vacuole, ABA signalling, and actin mediated cell growth (Rahman *et al.*, 2007; Lofke *et al.*, 2015; Scheuring *et al.*, 2015). Auxin transporters are typically found on the plasma membrane but are also found on the tonoplast in *Arabidopsis* (Ranocha *et al.*, 2013) and tonoplast localised auxin transporters are also found in other species such as rice (Liu *et al.*, 2016). Disruption to the localisation of these transporters by disruption to IRQ4 could inhibit the release of auxin stored in the vacuole and lead to a lower level of auxin in the nucleus and give rise to an auxin insensitive phenotype in the root. A further alternative is that loss of IRQ4 could lead to auxin insensitivity by affecting the trafficking of the SNARE VTI11. VTI11 is a Qb-SNARE, a vesicle SNARE localised to the to the PVC/MVB and is involved in homotypic fusion of lytic and protein storage vacuoles (Zheng *et al.*, 2014). Loss of VTI11 in the *vti11*/*zig1* mutant causes an auxin insensitive phenotype (Lofke *et al.*,

2015). Mutations of other genes involved in trafficking have been shown to have genetic interactions with *VTI11*. Mutation of *Rab5* (RabF) genes exaggerated the dwarf phenotype of *vti11/zig1* mutants (Kato *et al.*, 2002; Ebine *et al.*, 2014).

It should be noted that as the CRISPR lines contained a truncation for IRQ4 which fell in frame and was translated rather than a full knockout of transcription, the truncated product may be less efficient at its function and therefore a more pronounced phenotype may present itself if a full knockout line for *IRQ4* and *IRQ5* is generated. Further work is currently being done in the lab to identify CRISPR mutants with full knockouts of the *IRQ4* gene in combination with the *IRQ5* gene.

In summary, the elongated root phenotype seen in the CRISPR double mutants could be due to the auxin insensitivity caused by the loss of *IRQ5* which has been enhanced by the loss of IRQ4. Loss of *IRQ4* could contribute to the auxin insensitivity phenotype by altering the trafficking of, for example, proteins involved in auxin transport at the tonoplast membrane or by interfering with proteins known to be the targets of auxin that are involved in changes to vacuole shape and morphology such as VTI11.

8.4 Potential roles of IRQ4 in the late secretory and endocytic pathways

The localisation of IRQ4 was studied by cloning the gene and expressing it either transiently under a 2x35S promoter in epidermal cells (IRQ4-GFP construct) of *N. benthamiana* or stably transforming the IRQ4-GFP into BY2 tissue culture cells. *Arabidopsis thaliana* plants were also stably transformed with IRQ4 expressed under its own promoter fused to GFP (*promIRQ4*::IRQ4-GFP). In addition, anti-IRQ4 polyclonal antibodies were raised against IRQ4 and used in immunogold labelling visualised by TEM. Chapter 5 established that IRQ4 is localised to the late secretory pathway, specifically to the tonoplast, PVC/MVB, and autophagasomes.

8.4.1 IRQ4 as a link between PVC/MVB and autophagy

There has recently been interest in the links between autophagy, PVC/MVB, and endosome trafficking (Lamb *et al.*, 2012; Zhuang *et al.*, 2015; Cui *et al.*, 2016). As IRQ4 localises to both PVC/MVB and autophagasomes this would indicate that IRQ4 may be another such link. NET4B has also been localised to double membrane structures which were potentially autophagasomes (Mentlak, Unpublished thesis). Three examples will be given below to illustrate the variety of links and cross talk between these pathways.

Endocytosis has been directly linked to autophagy by the study of Regulator of G-Protein Signalling 1 (AtRGS1). Presence of D-glucose causes the phosphorylation and endocytosis of

AtRGS1. During this process the expression of five components of the autophagic machinery that are involved in the formation of autophagasomes are upregulated (Yan *et al.*, 2017). When autophagy was interrupted the endocytosis of AtRGS1 was stopped thereby linking autophagy directly with endocytosis.

A link between endosomes, PVC/MVB and autophagy has been exemplified by studies of FREE1. FREE1 is the FYVE domain protein required for endosomal sorting 1 and is a plant unique ESCRT-I component involved in the PVC/MVB and vacuole morphology (Gao *et al.*, 2014; Kolb *et al.*, 2015). FREE1 is linked to autophagy through its interaction with SH3 domain containing protein 2 (SH3P2). The interaction between FREE1 and SH3P2 is mediated by the N-terminal Bin-amphiphysin-Rvs (BAR) domain in SH3P2 and the C-terminal domain of FREE1 (Gao *et al.*, 2015). FREE1 and SH3P2 were also shown to immunoprecipitate with ATG6-YFP indicating that a PI3K complex is formed involving these proteins. FREE1 is therefore an example of a protein with dual roles in the endomembrane system: its C-terminus is responsible for its interaction with SH3P2 and its role in autophagy while the N-terminus of FREE1 interacts with the ESCRT-I component Vps23 via its PTAP-like tetrapeptide motifs (Gao *et al.*, 2014).

In an example of PVC/MVB–based trafficking and autophagy MON1 was shown to be involved in the regulation of tapetal cell death in pollen tubes (Cui *et al.*, 2017). As was discussed in chapter 1 MON1 is part of the MON1-CCZ1 complex that is recruited by Rab5 in order to act as the GEF for Rab7 and activate Rab7 to mediate the maturation of the PVC/MVB before fusion with the tonoplast (Cui *et al.*, 2014). Degradation of the tapetum is vital for microspore development and pollen coat formation and this occurs via programmed cell death (PCD; Varnier *et al.*, 2005; Zhang, Teng and Liang, 2011). Cui *et al.* (2017) monitored the vacuolar processing of cysteine proteases which are involved in the degradation of the tapetum (Zhang *et al.*, 2014b). Rab7-based trafficking of these cysteine proteases was shown to be important for the degradation of the tapetum as mutation of MON1 (causing defects in Rab7 activation) led to delayed tapetum degeneration and tapetal PCD, which caused abnormal pollen coat formation ultimately leading to decreased male fertility (Cui *et al.*, 2017).

The examples of AtRGS1, FREE1, and MON1 given above serve to illustrate how complex the relationship is between the endocytic, secretory and autophagic pathways and the ever increasing number of proteins which have roles in one or more of these pathways is blurring the line between them.

Although IRQ4 was shown to localise to autophagasomes using the polyclonal anti-IRQ4 antibody (Figure 5.2.3), when *Arabidopsis* plants stably expressing the *promIRQ4*::IRQ4-GFP construct were transferred from normal media to media lacking nitrogen no increase in PVC/MVB/autophagsome

number was observed (Figure 5.24). In fact, a significant decrease in the number of punctae was observed in those plants that had been transferred to media lacking nitrogen. This is unexpected as the number of autophagasomes increases in cells under nutrient starvation conditions (Bassham, 2015). This result could be explained by analogy to previous work on proteins involved in regulating autophagy by ubiquitination, the Tumour Necrosis Factor Associated proteins (TRAF). TRAFs localise to the autophagasomes upon starvation and knockouts of TRAF proteins have classic autophagy phenotypes including nutrient starvation sensitivity and increased salicylic acid and reactive oxygen species levels. The expression levels of the TRAF1a, TRAF1b, and an interactor of the TRAFs, ATG6, all increased up to 12 hours after autophagy induction but by 24 hours had decreased indicating these proteins are highly regulated (Qi *et al.*, 2017). The number of PVC/MVB/autophagasomes labelled by IRQ4 was only counted after 24 hours under starvation conditions. If IRQ4 was similarly highly regulated during autophagy, for example if IRQ4 was only involved in the early stages of autophagsome generation or trafficking, then IRQ4 could have been degraded before the images used in the current analysis were taken, giving the reduced number of punctae seen in Figure 5.24. So, rather than a decrease in autophagasomes the result in Figure 5.24 may only represent a decrease in the number of IRQ4 labelled autophagasomes while total autophagasome number actually increased. To verify this theory the experiment would need to be repeated using a much larger sample number, monitor the course of autophagy induction over more time points and earlier than 24 hours e.g. 3, 6, 9, and 12 hours after transfer to nutrient starvation conditions, and also couple the labelling of IRQ4 punctae with a known autophagy marker such as ATG8 to monitor the number of autophagasomes in parallel.

Unrelated to the endocytic or secretory pathway, a potential interactor of IRQ4, SnRK1.1 has also been shown to be directly linked to autophagy. SnRK1.1 shows changes in phosphorylation pattern when grown on media lacking nitrogen (Engelsberger & Schulze, 2012). Over expression (OE) of SnRK1.1 was shown to enhance autophagy, delay leaf senescence, and increase tolerance of plants to starvation and abiotic stress (Chen *et al.*, 2017). In addition OE lines showed enhanced expression of autophagsome forming genes, increased autophagsome formation and degradation, and were shown to enhance the phosphorylation of a known activator of autophagy in animals ATG1a. Although this thesis has not proven a direct interaction between SnRK1.1 and IRQ4, IRQ4 could be involved in localising the SnRK1.1 to the PVC/MVB or autophagasome allowing it to phosphorylate and activate proteins involved in autophagy. Previous work in yeast had shown that when moved from normal glucose containing medium to media containing glycerol-ethanol the Sip1 β subunit relocalises with Snf1 from the cytosol to the vacuole membrane (Hedbacker *et al.*, 2004) so a similar process may operate in plants upon N starvation and autophagy induction.

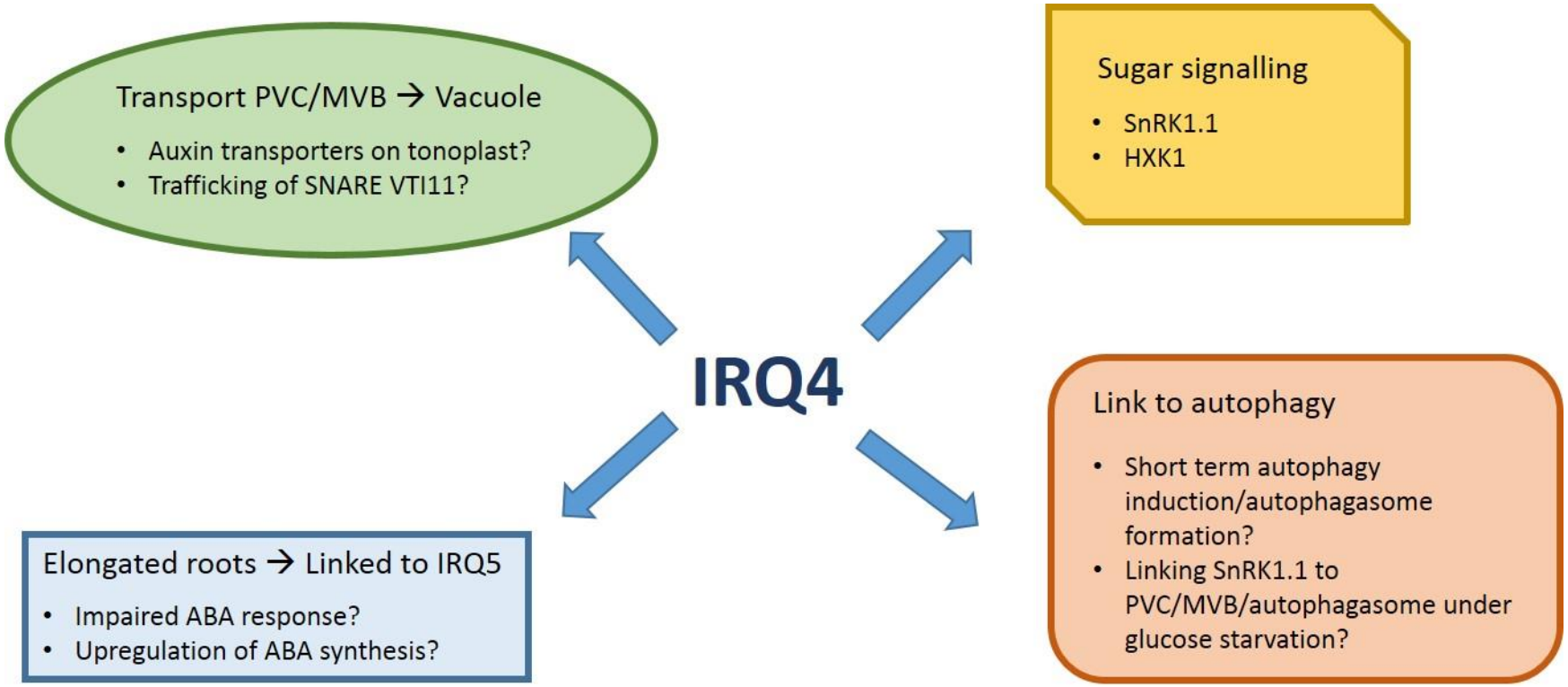


Figure 8.1: Summary of the roles of IRQ4 based on localisation, phenotypic, and interaction data. Based on the research conducted and discussion above the potential roles of IRQ4 are summarised in the boxes.

8.5 IRQ proteins as a eudicot specific group of proteins

As first proposed in 1703 by John Ray (Tudge, 2006) the angiosperms have been divided between monocotyledonous and dicotyledonous plant species based on the number of cotyledons or seed leaves. There are a number of other physiological characteristics which have been used to refine this division (listed in Table 8.1). However, a combination of relatively recent taxonomic and cladistics work has shown that angiosperm phylogeny is more complex. The eudicots and monocots arose from a group of basal angiosperms (Crane *et al.*, 1995) including Nymphaeales (lillies), Piperales (true peppers), Aristolochiaceae, and woody magnoliids. Phylogenetic analysis of 61 plastic genes showed that during angiosperm evolution the five major clades (*Ceratophyllum*, Chloranthaceae, eudicots, magnoliids, and monocots) which form the mesangiosperms have been shown to have undergone very rapid diversification due to the very short branch lengths between these lineages (Moore *et al.*, 2007). The order in which the divisions separating the angiosperms occurred and at what time points major lineages branched off are still an area of debate due to sometimes differing interpretations of fossil data such as pollen palynology data compared with molecular evidence.

Table 8.1: Physiological differences between monocotyledonous and dicotyledonous plants species. Taken from (Willis & McElwain, 2002).

Characteristic	Dicotyledons	Monocotyledons
Flower parts	In fours or fives (usually)	In threes (usually)
Pollen	Usually with three pores or furrows	Usually have one pore or furrow
Cotyledons	Two	One
Leaf venation	Usually netlike	Usually parallel
Primary vascular bundle in stem	In a ring	Complex arrangement
True secondary growth with vascular cambium	Commonly present	Commonly absent

As no direct function for the IRQ domain only containing proteins has yet been found inferences relating to eudicot specificity should be used with caution. However, based on what is known about the localisation of the IRQ proteins some plausible scenarios can be presented.

Genes involved in reproduction often evolve rapidly due to adaptive evolution (Swanson & Vacquier, 2002). Reproductive tissue types, particularly pollen in plants, are crucial in forming a reproductively isolated group of organisms, an essential element of species formation (Howard, 1999). Microarray data from the Affymetrix gene chip has shown that IRQ2 and IRQ3 are male gametophyte specific and both are highly expressed in the mature pollen grain (Figure 3.6c and d). Of the IRQ proteins, IRQ2 and IRQ3 were also potentially the latest to appear, having been

created after a tandem duplication of IRQ4 or IRQ5 (Figure 3.15b). Restriction of IRQ2 and IRQ3 expression to pollen suggests a specialised function for these proteins.

A major difference between the eudicots and other flowering plants is the presence of pollen with three pores set in furrows called colpi, also known as tricolpate pollen. This is in contrast with the rest of the seed plants (gymnosperms, lower angiosperms, monocots) where a single pore is set in a single furrow or sulcus (Edlund *et al.*, 2004). One of the main functions of these pores is to allow the pollen tube carrying the male sperm cells to exit the pollen grain and enter the transmitting tract of the stigma before fertilising the egg and central cells (Edlund *et al.*, 2004). The second main function of the pore is to allow control of water loss, as once released pollen grains are subjected to harsh environmental conditions and would dehydrate. The pores of the pollen grain allow the pollen grain wall to fold in on itself in a precise manner and close the pores, preventing water loss. This folding process is termed harmomegathy (Katifori *et al.*, 2010). As a eudicot and pollen specific set of proteins IRQ2 and IRQ3 may be associated with the structural changes in the tricolpate pollen grain walls of eudicot species, responding to signals associated with germination or desiccation stress.

A further structural role for the IRQ2 and IRQ3 proteins may be related to the remarkable diversity in angiosperm speciation. Over 75 % of extant angiosperm species are eudicots (Doyle *et al.*, 1994). This has in part been attributed to coevolution with pollinators (Darwin, 1862; Gorelick, 2001). The structure of pollen grains act as both an attractant to pollinators and also allows the pollen grain itself to attach to the vector for its dispersal (Piffanelli *et al.*, 1998). IRQ2 and IRQ3 may therefore have a dual role in pollen germination associated with the complex 3D structure of the pollen grain wall in eudicots or a part of the signalling mechanisms involved in pollinator attraction. The numerous coiled coils in the secondary structure of these IRQ proteins may facilitate protein-protein interactions that facilitate the role of these proteins in signalling mechanisms.

IRQ1 and IRQ6 were described by Affymetrix data to localise to meristematic tissues (Figure 3.6a and b). This has been verified in this thesis by promoter GUS fusions to *IRQ1* and *IRQ6* which show IRQ6 was expressed at both the shoot apical meristem and root apical meristems while IRQ1 was mostly found at lateral root primordia as well as two distinct cell files in the lateral roots (see chapter 4).

Significant differences in the physiology of root architecture exist between monocots and eudicot species. In general dicot species possess a ‘taproot’ system while monocots have a more complex fibrous root system. In dicots the primary root is formed embryonically while lateral roots are formed post embryonically. In monocot cereals such as maize both a primary root and a seminal

roots are formed embryonically with crown, brace, and lateral roots (LR) being formed post-embryonically (Smith & Smet, 2012). The architecture of the root apical meristem is also radically different in the monocots compared to the eudicots. *Arabidopsis thaliana* possesses a single cortical cell layer while rice and maize possess 10-15. Even more strikingly *Arabidopsis* has four quiescent centre cells while in rice and maize these cells can number 800-1200 (Hochholdinger & Zimmermann, 2008). With all these differences between monocot and dicot root systems there a large number of possible roles for IRQ1 and IRQ6.

As yet no mutant studies have been completed on IRQ1 or IRQ6, but mutants associated with xylem pole pericycle cells may be illustrative for future work aimed at pinpointing the function of IRQ1 and IRQ6. D-type cyclins associate with their cyclin dependent kinases (CDKs) to phosphorylate the retinoblastoma-related protein which in turn allow the transcription of genes regulated by the E2F transcription factors and leads to commitment to the cell cycle (Kohler *et al.*, 1999; Dewitte & Murray, 2003; Wildwater *et al.*, 2005; Dewitte *et al.*, 2007). The cyclin D mutant, *cycd4;1,* has longer basal meristem cells relative to upper meristem cells giving altered spacing of LRs and hence shows that the basal meristem has a role in determining LR density (Nieuwland *et al.*, 2009). Treatment with auxin led to normal LR initiation showing that LR initiation is not impaired in the *cycd1;4* mutant.

A more extensive structural defect is found in the *lonesome highway* (*lhw*) mutant (Ohashi-Ito & Bergmann, 2007) where the roots only have a single set of xylem and phloem pole pericycle cells, the roots losing the diarch structure and bilateral symmetry of the dicot root. A reduction in the number of cells in the centre of the root indicates that LHW controls the number of cell types by controlling the size of the pool from which they come from. LR density has been shown to be reduced by 47 % in *lhw* mutants (Parizot *et al.*, 2008).

Other mutants associated with the xylem pericycle cells are involved in signalling responses. For example the *fewer roots* (*fwr*) mutant is the result of a mutation in the *GNOM* gene. GNOM is involved in recycling of the PIN auxin efflux carriers. *fwr* mutants have an impaired response to auxin maxima and as a result produce fewer LRs (Okumura *et al.*, 2013). Future work on IRQ1 and IRQ6 should focus on some of the typical lateral root phenotypes described here including lateral root number, distribution along the root and response to exogenous auxin.

8.6 Future work

Although no IRQ domain alone containing proteins were found in the three monocot species studied (*B. distachyon, O. sativa,* and *Z. mays*) this does not preclude an IRQ sequence being found at a later date as the information in the databases is improved, as the example of the additional bases added to the N-terminal of IRQ2 demonstrates. Periodic checks of databases would need to be undertaken in the monocot species studied here. Additional monocot species such as bread wheat (*Triticum aestivum*) should also be screened. It would also be of interest to screen further eudicot species to see the extent of the multiplicity of the IRQ group of proteins in these species.

The localisation data obtained using the IRQ1 and IRQ6 transcriptional fusions at the lateral roots and the shoot apical meristem should be investigated further by creating translational fusions to a fluorophore such as GFP and transforming these in to either BY2 or *Arabidopsis* tissue culture cells. The cell cycle in BY2 and *Arabidopsis* tissue culture cells can be relatively easily synchronised (Samuels *et al.*, 1998; Menges & Murray, 2002). In this way the behaviour of the proteins can be followed at specific stages of the cell cycle. In addition translational fusions encoding IRQ1 and IRQ6 should be transformed into *Arabidopsis* and the localisation of the full length protein investigated using CLSM. This should provide a more precise localisation at the subcellular level in the endogenous host, indicating the organelles IRQ1 and IRQ6 are associated with. These lines could then be crossed or transformed with individual markers known to be associated with particular stages of lateral root development. A marker for the xylem root pericycle cells for example would help confirm the localisation of the *prom*::GUS construct for IRQ1 shown in this thesis. In addition, knockout mutants of *IRQ1* and *IRQ6* should be obtained or created using CRISPR and screened to determine if a lateral root phenotype is shown by these proteins either singly or when both are removed. Mutants involved in cell cycle regulation or if they are associated with the meristematic regions of the plant often generate embryonically lethal phenotypes. In these cases the mutants may have to be generated by RNAi or using an inducible system such as a dexamethasone-based system.

Although IRQ4 was shown to localise to PVC compartments using wortmannin treatment in both the heterologous BY2 system and in the native *Arabidopsis* environment, colocalisation with a known PVC/MVB marker such as the VSR BP-80 or the Rab5 homologs RHA1 or ARA7 could also be used to confirm the PVC/MVB localisation.

The phenotypic analysis has shown that the double CRISPR mutants exhibited an elongated root phenotype regardless of the presence or absence of nitrogen in the media. This experiment will need repeating as although an adequate sample size was used, only one biological replicate was performed in this thesis. In addition, in the double CRISPR mutants the *IRQ4* gene only contained a truncation removing 52 amino acids from the N terminal end of the protein. This deletion fell 'in-frame' resulting in a deletion which produced a truncated protein product. Further CRISPR lines would need to be analysed to try and obtain a full knockout of the *IRQ4* gene. Redesigning the locations of the target sites for Cas9 action may be required to achieve this. If the elongated root phenotype can be confirmed then further work would be needed to investigate the relationship between this phenotype and auxin sensitivity. Application of exogenous auxin could be used to study its effect on the root length of the CRISPR double plants and also the single *IRQ4* and *IRQ5* mutants as auxin sensitivity was not explicitly tested for in this thesis.

This thesis has demonstrated that IRQ4 is likely to be involved in the late endocytic pathway. Investigating the single mutants for phenotypes associated with substance transport could be achieved by transforming them with specific endomembrane markers or substances known to be transported along specific routes and testing using CSLM whether their progression through or localisation in the endomembrane system is altered.

IRQ4 was shown to localise to the tonoplast but no gross morphological differences were observed across any of the single or double CRISPR mutants when the vacuole was imaged using BCECF-AM. However, the vacuole was imaged on plants which under non-stress conditions. Defects in vacuolar morphology may only be noticeable when plants are subjected to some form of stress such as drought or pathogen attack. Defects in vacuolar trafficking or vacuole morphology can make plants more susceptible to osmotic stress (Munns & Tester, 2008). The assay used in this thesis to test drought used a relatively high level of salt (160 mM) as it was predominantly aimed at testing for an autophagy related phenotype (Wang *et al.*, 2016a). If a more subtle phenotype is present in the mutant plants this may be masked by such harsh conditions. A drought assay using a range of salt concentrations (Jiang *et al.*, 2016) in combination with other salts such as KCl or a non-salt such as mannitol (Wathugala *et al.*, 2011) will give a more definitive result relating to drought tolerance of the mutants. Characteristics that could be assessed include visualising overall vacuole morphology using BCECF-AM as used in this thesis or measuring primary root length at a number of different time points.

IRQ4 was shown to localise to autophagasomes using the anti-IRQ4 antibody. This localisation would first need to be verified using a marker for autophagy such as ATG8 attached to a fluorophore and transforming this into the BY2 cells and *Arabidopsis* plants stably expressing IRQ4

constructs linked to fluorophores allowing distinction between IRQ4 and the marker. Subjecting these co-expressing cells and plants to autophagy inducing conditions and analysing the level of colocalisation between them would ascertain whether the localisation to autophagasomes in the immunogold labelling represents a true localisation or some form of artefact. Once the localisation to autophagasomes had been firmly established the nature of the association between IRQ4 and autophagasomes could to be investigated in the manner described earlier, wherein the co-expressing IRQ4 and ATG8 plants are subjected to autophagy inducing conditions e.g. N starvation or salt stress and the progression of autophagy monitored at various time points by counting the number of IRQ4 only labelled compartments, ATG8 only compartments, and colocalised IRQ4 and ATG8 compartments at intervals of 3 hours up to 24 hours. If IRQ4 was involved in early stages of autophagasome formation then co-expressing IRQ4 with proteins known to be involved in the phagaphore assembly site and early stages of phagaphore formation may more firmly establish the connection.

Crucial to finding the function of IRQ4 would be to find true interactors. For the reasons discussed above the Y2H system has not proved successful with IRQ4 but there are other avenues for testing interactions that could be explored. One method could be creating constructs of IRQ4 with affinity purification tags attached to the N- or C-terminal ends to conduct immunoprecipitation experiments and analyse the results using mass spectroscopy. This would have the advantage that the protein complexes extracted would be taken from the native environment in the correct tissues e.g. immunoprecipitation could be performed only on root samples so the proteins would have all the post-translational modifications (if required) to interact with its partner protein(s).

Once potential proteins are found these could be verified by using a technique such as Förster Resonance Energy Transfer-Fluorescent Lifetime Imaging (FRET-FLIM). FRET is the process whereby energy is transferred from a donor fluorophore to an acceptor fluorophore, quenching the excitation and reducing the fluorescence lifetime of the donor fluorophore in the process. The lifetime of the fluorophore can be measured using FLIM. When fluorophores lie in close proximity to each other (within 10 nm, the Förster distance) their fluorescence lifetimes can be altered. Coupling one fluorophore to one potential interactor and another fluorophore to the other potential interactor allows this property of FRET to be exploited to test protein-interactions. The potential interactors are co-expressed as fluorescent fusion proteins and if the proteins interact FRET occurs between them and this is seen as a drop in the fluorescence lifetime of the donor as measured by FLIM (Wallrabe & Periasamy, 2005; Sun *et al.*, 2011, 2013). In addition to confirmation of an interaction between the proteins, FRET-FLIM also gives information related to the localisation of the interaction (Bücherl *et al.*, 2014).

8.7 Conclusion

The endomembrane system is responsible for the movement of membrane and substances around the cell as well as maintenance of organelle identity. This complex system is highly regulated by a plethora of proteins involved in cargo selection, sequestration, packaging into vesicles, targeting, and fusion of these vesicles with the target compartment. In addition to the secretory and endocytic pathways there are numerous anterograde steps between different compartments and new research is shedding light on the cross talk that occurs between processes which were thought to be separate including PVC/MVB trafficking, endocytosis, and autophagy.

This thesis has focussed on the IRQ proteins containing the IRQ domain, a domain that previous work in NET4A and NET4B was shown to be responsible for allowing these NET proteins to bind to the RabG subclass of Rab GTPases. Bioinformatics analysis using the IRQ domain revealed a group of six IRQ domain alone containing proteins as well as two further NETs, NET3A and NET3C, also possessing an IRQ domain.

Different approaches were used in this thesis to gain more information on these six IRQ proteins and attribute a function to them, with a particular focus on IRQ4. Firstly, phylogenetic evolutionary analysis showed that the IRQ proteins evolved from the NETs and that the IRQ proteins are potentially a eudicot specific group of proteins. It is somewhat difficult to ascertain specific areas of eudicot evolution which the IRQ proteins may have had a role in without much more data on the localisation and function of these proteins but some general ideas were proposed based on the information available. During the emergence of the eudicots from the basal angiosperms and monocots, the diversity of flower types along with pollination mechanisms expanded dramatically. IRQ2 and IRQ3 are predicted to be pollen specific so may have a role in the formation of the pollen coat or be associated with pollen germination or pollen tube growth. All of these pollen characteristics are key factors which have contributed to the diversity of eudicot species.

There are a large number of differences between the root system architecture of the monocots and eudicots including a lower number of different forms of post embryonic root development in eudicots and a less complex arrangement of the eudicot root apical meristem. IRQ6 and IRQ1 were shown to be transcriptionally localised to the meristematic regions of the roots. IRQ1 was localised to specific cell files similar to that of xylem pole pericycle cells, cells which are associated with the emergence of lateral roots. The signalling mechanisms involved in lateral root emergence and development may differ between monocots and eudicots and hence IRQ1 and IRQ6 may play a role in LR development processes specific to eudicots.

IRQ4 was shown to be transcriptionally and translationally localised to the root tip. Subcellular analysis has shown that IRQ4 is localised to the tonoplast, PVC/MVB, and autophagasomes. The BFA insensitivity but wortmannin sensitivity of IRQ4 labelled compartments situates IRQ4 in the late secretory pathway. Considerable crosstalk is being discovered between multiple pathways in the endomembrane system and the localisation of IRQ4 to autophagasomes as well as PVC/MVB potentially strengthens these connections. The expression of IRQ4 labelled punctae was shown to decrease after 24 hours under autophagy inducing conditions indicating that IRQ4 may be involved in the early stages of the nutrient starvation response but may be highly regulated and degraded soon after performing its role. Confirmation of the localisation to autophagasomes by using a marker such as ATG8 is vital to understanding more about the relationship between IRQ4, PVC/MVBs, and autophagasomes.

Interaction studies using the Y2H system produced a potential interactor for IRQ4, SnRK1.1, but this was not possible to verify in a one on one assay. Similarly, a screen of representative Rab GTPases from several Rab families did not yield any positive interactions with IRQ4. Finding positive the interactors of IRQ4 is an important part of ascertaining the function of this protein in the late secretory pathway.

Double CRISPR mutants of *IRQ4* and *IRQ5* were shown to have an elongated root phenotype similar to the published auxin insensitivity based phenotype of the single *IRQ5* mutant (Ren *et al.*, 2016). IRQ4 may therefore be involved in auxin signalling by being involved in the trafficking of the auxin efflux carrier PIN proteins. No phenotype related to autophagy was detected in any of the single T-DNA insertion mutants or double CRISPR mutant plants however these assays were completed once, using a large sample size but should be repeated to check the results given here. In addition the mutation of *IRQ4* in the double CRISPR lines only gave a truncated product which was shown to be translated. Double mutants where both *IRQ4* and *IRQ5* are completely knocked using CRISPR technology or crossing the single T-DNA insertion lines are currently being carried out in the laboratory.

In summary, this thesis has shown that the IRQ proteins are a potential eudicot specific group of proteins which evolved from the NET family of proteins. IRQ1 and IRQ6 have been shown to be associated with cells involved in lateral root development while IRQ4 localises to the tonoplast, PVC/MVB, and autophagasomes, potentially adding a further link between PVC/MVB trafficking and autophagy.

Appendices

Appendix 1: Primers

1.1 Primers used for verification of transformation of *Arabidopsis* stable lines

Primer Name	**Sequence**	**Target/Reason for use**
83_GFP_R	CCTTCTTGAGTTTGTAACAGCTGCTG	GFP fluorophore
mRFP_F1	GCTTCAAGGTGCGCATGGAGG	mRFP fluorophore
mRFP_R1	GCTCGTACTGTTCCACGATGG	mRFP fluorophore
MCherry_R1	GGTCTTGACCTCAGCGTCG	mCherry fluorophore
IRQ4_promGUS_F1	CTCTGTGCTTTCTGGTTCCCATTCC	IRQ4 (At1g64330)
IRQ5_promGUS_F1	GAAAGCGTTGGATGATCGGATCGAG	IRQ5 (At5g41790)
IRQ6_promGUS_F1	GTCTATGTTTCTTCACCGTTGC	IRQ6 (At5g05180)
GUS_F1	GTTGGGCAGGCCAGCGTATCGTGC	Check presence of GUS construct
GUS_R2	CGTAATGAGTGACCGCATCGAAACGC	Check presence of GUS construct
IRQ4_F3	GAGGTGGAGAAGAAGCTTGAAGG	Check for presence of *prom*IRQ4::IRQ4-GFP vector

1.2 Primers used to generate IRQ4$^{125-248}$ antigen fragment for polyclonal antibody creation

Primer Name	**Sequence**	**Target/Reason for use**
IRQ4_AB_F1	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCAAACAGATGGAAGATGCAAATC	(At1g64330) IRQ4 fragment used for making antigen
IRQ4_AB_R1	GGGGACCACTTTGTACAAGAAAGCTGGGTCTCTCTCGAGTTCTGCTTCGGTTTCG	(At1g64330) IRQ4 fragment used for making antigen

1.3 Genotyping primers for T-DNA insertion plants

Primer Name	**Sequence**	**Target/Reason for use**
IRQ4284Fw	AGAACGAAACCGAAGCAGAACTCG	*irq4.1* plants
IRQ284Rv	CGTTGTGACCTTCTTTGTCCTCTTGC	*irq4.1* plants
IRQ284_R1	CAGACAAGGATTGGAACCGGTTTAG	*irq4.1* plants
IRQ5_865_F2	GCGAGTTTCAGAGCTGACCAGCGG	*irq5.1* plants
IRQ5_865_R2	CATATCTGCAACCTGTTGATCTGCTG	*irq5.1* plants
LBa1.1	GCGTGGACCGCTTGCTGCAACT	T-DNA border primer for checking presence of insertion
Gabi_8409	ATATTGACCATCATACTCATTGC	T-DNA border primer for checking presence of insertion
Mago Fw	ATTTGGACACGAGTTCCTGGAGTTCG	Control to show genotyping PCR reaction works

Mago Rv	GTAGGAATCCCGCATTGTGAGCCAC	Control to show genotyping PCR reaction works

1.4 Primers used to generate CRISPR mutant plants

Primer Name	**Sequence**	**Target/Reason for use**
VI4_F	CTTGGGAATTGTTGAGAGTGG	Genotyping of CRISPR line specific for IRQ4
VI4_R	GATTCAGTAACGCTGGTTTCTCTTGC	Genotyping of CRISPR line specific for IRQ4
VI4_Fn1	GATATTGAGGAAGATGAATCC	Nested sequencing primer for the IRQ4 gene in the CRISPR lines
VI4_Rn1	GATCAGATTCCGTCTCACCAGC	Nested sequencing primer for the IRQ4 gene in the CRISPR lines
VI5_F	CTTGGGAATGGTTGAGAGTGGAGACG	Genotyping of CRISPR line specific for IRQ5
VI5_R	CTGCTCTGATGCTTCAAGCTGCTGC	Genotyping of CRISPR line specific for IRQ5
VI5_Fn1	GACTTAGTGAAGGAATTCTACAGC	Nested sequencing primer for the IRQ5 gene in the CRISPR lines
VI5_Rn1	GATACCATTGTCCCTCTCAGTTTGC	Nested sequencing primer for the IRQ5 gene in the CRISPR lines

1.5 RT-PCR primers

Primer Name	**Sequence**	**Target/Reason for use**
IRQ4_F1	GAGCCTCATCTTCATCCTGATAATGG	*irq4.1* plants
IRQ4_F3	GAGGTGGAGAAGAAGCTTGAAGG	*irq4.1* plants
IRQ4_R1	CAGACAAGGATTGGAACCGGTTTAG	*irq4.1* plants
IRQ4_R2	CTCAGTTTCAACTCTCAAATTCCCAC	*irq4.1* plants
IRQ5_F1	GAGAAACCTTGAAGTCTTTCTTTGAGC	*irq5.1* plants
IRQ5_F2	GAGAGTAACCGAACAGGTACTAACAG	*irq5.1* plants
IRQ5_R1	GCTGCGACACTCTTCTTTGG	*irq5.1* plants
IRQ5_R2	GACTTTTCGTCCTCTAACTTCTCAG	*irq5.1* plants
N1A_F9	GAGTTCTTGAGTTCAGAGCTGCAGG	NET1A (At3g22790) Control for genomic DNA contamination
Salk_081081R	CTCAAGTTCAATCTCTTCAAACAGC	NET1A (At3g22790) Control for genomic DNA contamination

EF1α_F	CCCATTTGTGCCAATCTCT	(At1g07940) Control for genomic DNA contamination
EF1α_R	CACCGTTCCAATACCACCAA	(At1g07940) Control for genomic DNA contamination

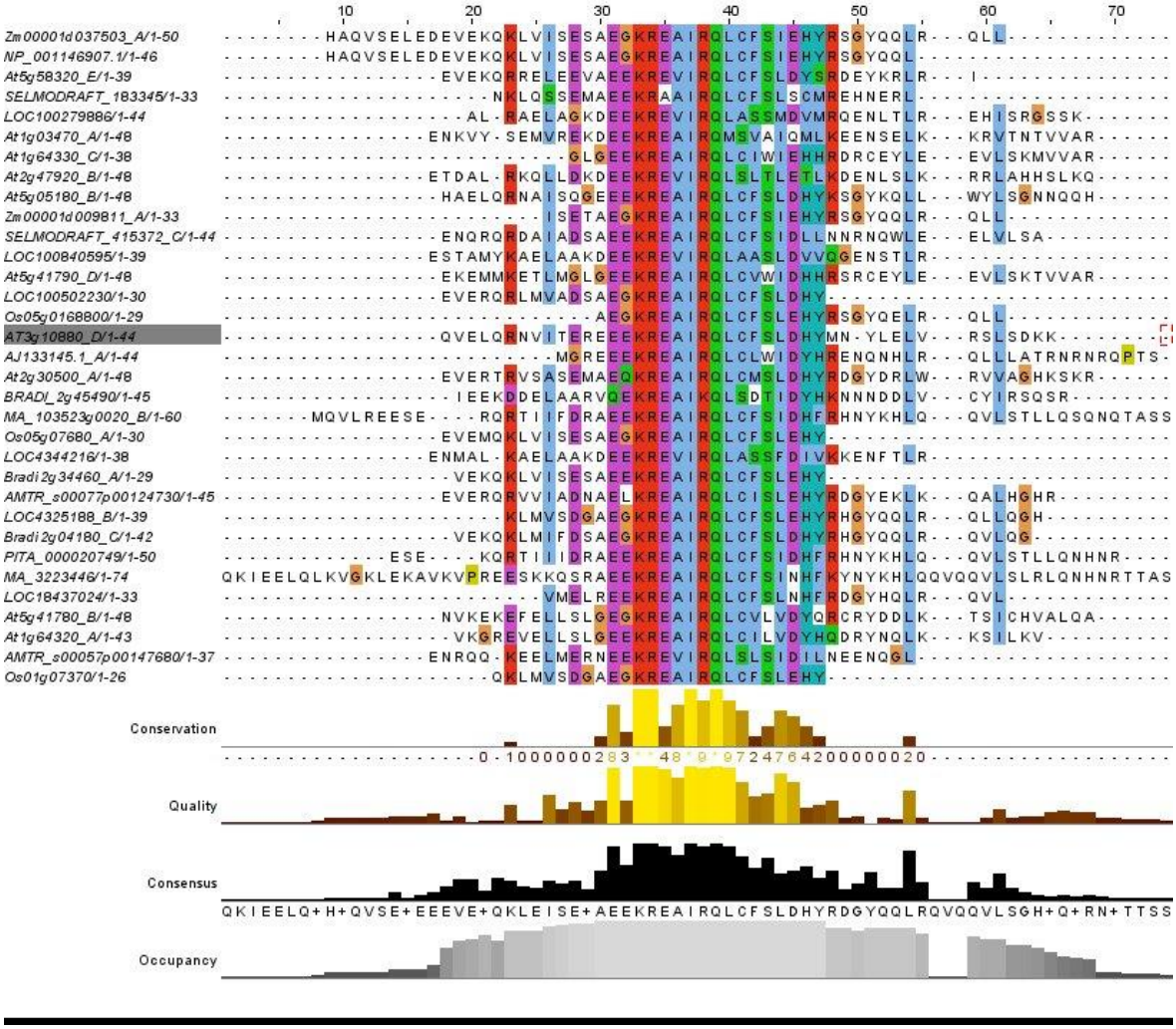
1.6 Primers used to clone full length Rabs for Y2H one-on-one screening

Primer Name	**Sequence**	**Target/Reason for use**
RabF1_F1	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGGGATGTGCTTCTTCTCTTCC	(At3g54840) RabF1 gateway primer
RabF1_F2	GGAACTCTGGAACTTTAAGTGGTC	(At3g54840) RabF1 sequencing primer
RabF1_F3	GCATGTTCTTCATTGAGACGTCAGC	(At3g54840) RabF1 sequencing primer
RabF1_R1	GGGGACCACTTTGTACAAGAAAGCTGGGTCTCATGACGAAGGAGCAGGACGAGGTAGC	(At3g54840) RabF1 gateway primer
RabF1_R2	CCTAATAGAACTAACTTCACACG	(At3g54840) RabF1 sequencing primer
RabF2b_F1	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGGCTGCAGCTGGAAACAAGAGC	(At4g19460) RabF2b gateway primer
RabF2b_F2	GGTGTTGCTTGGAGATGTTGGTGC	(At4g19460) RabF2b sequencing primer
RabF2b_F3	GCAAGAAGGCTACCGCGAGTACAG	(At4g19460) RabF2b sequencing primer
RabF2b_R1	GGGGACCACTTTGTACAAGAAAGCTGGGTCCTAAGCACAACAAGATGAGCTCAC	(At4g19460) RabF2b gateway primer
RabF2b_R2	CCTGGAATTCAACAAACTGATC	(At4g19460) RabF2b sequencing primer

1.7 Primers used in clone IRQ domain alone of IRQ4 for Y2H one-on-one screening

Primer Name	**Sequence**	**Target/Reason for use**
IRQ4_Y2H_T4_F	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCAAGGAGAAGTTGAAGGAGACATTGTTGG	(At1g64330) IRQ4 for IRQ domain alone Gateway primer
IRQ4_Y2H_T2_R:	GGGGACCACTTTGTACAAGAAAGCTGGGTCTTATCCTCTTGCCACGACCATTTTCGAC	(At1g64330) IRQ4 for IRQ domain alone Gateway primer
T7 promoter primer	CGAAATTAATACGACTCACTATAGG	Sequencing primers for pGADT7 and PGBKT7

Appendix 2: Alignment of IRQ domains across various species used for phylogeny calculations



Appendix 3: Expression pattern of *prom*GUS fusions in IRQ1, IRQ6, IRQ4, and IRQ5

IRQ1 (At3g10880)

GUS staining in just germinated seedlings

Line	Root cap	Procambium	Vascular cambium	Cortex-Tip	Cortex-whole root	Epidermis-Tip	Epidermis-Whole root	SAM region	Hypocotyl vasculature	Cotyledon vasculature	Cotyledon tissue
11		x		x					x		
12		x		x							
19		x		x					x		

GUS staining in 5-7 day old seedlings

Line	Root cap	Procambium	Vascular cambium	Cortex-Tip	Cortex-whole root	Epidermis- lateral root initials	Epidermis- Whole root	Root vasculature (specific areas)	Lateral root initials- whole	Specific cell file of lateral roots	Hypocotyl	Stipules	Cotyledon vasculature	Leaf 1+2 vasculature	Leaf 1+2 trichomes
1						x				x					
2						x		x		x					
3						x				x					
9						x		x		x					
11						x		x		x					
12						x		x		x					
19						x		x		x					
20						x				x					
21						x		x		x					

GUS staining in 2 week old plants

Line	Root cap	Procambium	Vascular cambium	Cortex-Tip	Cortex-whole root	Epidermis- lateral root initials	Epidermis- Whole root	Root vasculature (specific areas)	Lateral root initials- whole	Specific cell file of lateral roots	Hypocotyl	Hypocotyl vasculature	Stipules	SAM region	Cotyledon vasculature	Leaf 1+2 vasculature	Leaf 1+2 trichomes
1						x		x		x							
2						x		x		x							
3						x		x		x							
4						x		x		x				x			
5						x		x		x							
6						x		x		x				x			
7						x		x		x				x			
8						x		x		x							
9						x		x		x				x	x	x	
10						x											
11						x				x							
12						x		x		x							
13						x		x		x				x			
14						x		x		x				x			
15						x		x		x				x			
16						x		x		x				x			
17						x				x				x			
18						x		x		x				x			
19						x		x		x					x	x	
20						x		x		x				x			

GUS staining in flowers and siliques of six week old plants

Line	Sepal vasculature	Petal vasculature	Filament	Anthers	Ovary	Style	Stigma	Receptacle at base of flower	Ovules	Silique vasculature	Silique surface	Receptacle at base of silique
2												x
4												x
5								x				x
7								x				x
8												x
9		x	x	x				x				x
16												x
21				x								
22		x	x	x				x				x
23		x		x								x
24								x				
25												x
26												x

IRQ6 (At5g05180)

GUS staining in just germinated seedlings

Line	Root cap	Procambium	Vascular cambium	Cortex-Tip	Cortex-whole root	Epidermis-Tip	Epidermis-Whole root	SAM region	Hypocotyl vasculature	Cotyledon vasculature	Cotyledon tissue
2		x	x	x	x	x	x				
5		x	x	x	x	x	x				
8		x	x	x	x	x	x				

GUS staining in 5-7 day old seedlings

Line	Root cap	Procambium	Vascular cambium	Cortex-Tip	Cortex-whole root	Epidermis- lateral root initials	Epidermis- Whole root (except tip)	Root vasculature	Lateral root whole	Specific cell file of lateral roots	Hypocotyl-vasculature	Stipules	Cotyledon vasculature	Leaf 1+2 vasculature	Leaf 1+2 trichomes
1		x	x		x		x	x	x				x		
2		x	x		x		x	x	x				x		
3		x	x				x	x	x				x		
4		x	x		x			x	x				x		
5		x	x				x	x	x				x		
6									x				x		
9						x		x		x	x		x		

GUS staining in 2 week old plants

Line	Root cap	Procambium	Vascular cambium	Cortex-Tip	Cortex-whole root	Epidermis- lateral root initials	Epidermis- Whole root	Root vasculature (specific areas)	Lateral root-whole	Specific cell file of lateral roots	Hypocotyl	Hypocotyl vasculature	Stipules	SAM region	Cotyledon vasculature	Leaf 1+2 vasculature	Leaf 1+2 trichomes
1									x								
2						x		x	x								
3								x	x						x		
4						x			x							x	
5						x		x	x								
6						x			x							x	
7						x		x	x								
8						x											

GUS staining in flowers and siliques of six week old plants

Line	Sepal vasculature	Petal vasculature	Filament	Anthers	Ovary	Style	Stigma	Receptacle at base of flower	Ovules	Silique vasculature	Silique surface	Receptacle at base of silique
1		x	x			x		x				x
3		x	x			x						x
5		x	x			x						x
6		x	x			x						x
7		x	x			x						x
8		x	x			x						x

IRQ4 (At1g64330)

GUS staining in just germinated seedlings

Line	Root cap	Procambium	Vascular cambium	Cortex-Tip	Cortex-whole root	Epidermis-Tip	Epidermis-Whole root	SAM region	Hypocotyl vasculature	Cotyledon vasculature	Cotyledon tissue
1				x		x					
5				x		x					

GUS staining in 5-7 day old seedlings

Line	Root cap-lateral roots	Cortex-whole root	Epidermis- Whole root	Root vasculature	Lateral root initials- whole	Hypocotyl	Cotyledon vasculature
1	X	X	X				
2	X	X	X				
3		X	X				
4	X	X	X				
5		X	X				
6		X	X				
7		X	X	X			
8		X	X				
9		X	X	X			
10		x	x				

GUS staining in 2 week old plants

Line	Root cap	Procambium	Vascular cambium	Cortex-Tip	Cortex	Epidermis- Whole root	Root vasculature	Hypocotyl	Hypocotyl vasculature	Stipules	SAM region	Cotyledon vasculature	Leaf 1+2 vasculature	Leaf 1+2 trichomes
1				x	X	X	X							
2				x	X	X	X							
4						X	X							
5				x	X	X	X							
6					x	x	X							

GUS staining in mature leaves, flowers, and siliques of six week old plants

Line	Sepal vasculature	Petal vasculature	Filament	Anthers	Ovary	Style	Stigma	Receptacle at base of flower	Ovules	Silique surface	Receptacle at base of silique	Trichomes on mature leaves
1							X				X	X
2							x				x	
3		x					x				x	
5							x				x	
6							X				x	x

IRQ5 (At5g41790)

GUS staining in just germinated seedlings

Line	Root cap	Procambium	Vascular cambium	Cortex-Tip	Cortex-whole root	Epidermis-Tip	Epidermis-Whole root	SAM region	Hypocotyl vasculature	Cotyledon vasculature	Cotyledon tissue
1	x	x	x	x	x	x	x				
5	x	x	x	x	x	x	x				

GUS staining in 5-7 day old seedlings

Line	Root cap	Cortex-whole root	Epidermis-(except meristematic zone)	Root vasculature	Lateral root (except meristematic zone)	Hypocotyl-vasculature	Stipules	Cotyledon non-vasculature	Leaf 1+2 vasculature	Leaf 1+2 trichomes	Leaf 1+2-non-vasculare tissue
1		x	x		x		x			x	x
2		x	x	x							
4		x	x	x	x		x			x	
5		x	x	x	x		x			x	x
6		x	x	x				x			
7		x	x	x							
9		x	x	x				x			
10		x	x	x							

GUS staining in 2 week old plants

Line	Root cap	Procambium	Vascular cambium	Cortex-primary root	Epidermis- lateral root	Epidermis- Whole root	Primary root vasculature	Lateral root-whole	Root hairs	Hypocotyl	Stipules	SAM region	Cotyledon vasculature	Leaf 1+2 vasculature	Leaf 1+2 trichomes
1	x	x	x	x	x	x	x		x	x			x	x	x
2	x	x	x	x	x	x	x								
3	x	x	x	x	x	x									
4		x	x	x		x									
5	x	x	x	x	x	x	x		x	x	x		x	x	x
7	x	x	x	x	x	x									
8	x	x	x	x	x	x									
9	x	x	x	x	x	x	x		x	x	x				

GUS staining in flowers and siliques of six week old plants

Line	Sepal vasculature	Petal vasculature	Filament	Anthers	Ovary	Style	Stigma	Receptacle at base of flower	Ovules	Silique vasculature	Silique surface	Receptacle at base of silique
1		x	x	x			x					x
3												x
4		x	x	x			x	x				x
5		x		x		x	x					x
7				x			x					x
9				x		x	x					x

Appendix 4: Yeast-2-Hybrid library screen results

Gene Locus	Gene Annotation	Number of Colonies	Screen (1 or 2)
At5g55070	Dihydrolipoamide succinyltransferase	14	1, 2
At3g01090	Arabidopsis thaliana SNF1 kinase homolog 10 (AKIN10)	11	1, 2
At4g36980	CLK4-associating serine/arginine-rich protein	5	2
At2g01350	Quinolinate Phosphoribosyltransferase	4	2
At4g29130	Hexokinase 1	3	2
	DNAJ heat shock family protein	2	2
At5g25270 or At2g28950	Ubiquitin-like superfamily protein OR Expansin 6	2	2
At5g08190	Nuclear Factor Y, Subunit 12	1	2
At1g20810	FKBP-like peptidyl-prolyl cis-trans isomerase family protein	1	2
At4g24280	Chloroplast Heat Shock Protein 70-1, CPHSC70-1	1	1

Bibliography

Abiodun MO, Matsuoka K. 2013. Evidence that proliferation of Golgi apparatus depends on both de novo generation from the endoplasmic reticulum and formation from pre-existing stacks during the growth of tobacco BY-2 Cells. *Plant and Cell Physiology* 54: 541–554.

Aebi M, Bernasconi R, Clerc S, Molinari M. 2009. N-glycan structures: recognition and processing in the ER. *Trends in Biochemical Sciences* 35: 74–82.

Akkerman M, Overdijk EJR, Schel JHN, Ketelaar T. 2011. Golgi body motility in the plant cell cortex correlates with actin cytoskeleton organization. *Plant and Cell Physiology* 52: 1844–1855.

Akopian D, Shen K, Zhang X, Shan S. **2013**. Signal recognition particle: an essential protein-targeting machine. *Annual Review of Biochemistry* 82(1): 693-721

Albert VA, Barbazuk WB, dePamphilis CW, Der JP, Leebens-Mack J, Ma H, Palmer JD, Rounsley S, Sankoff D, Schuster SC, 2013. The Amborella Genome and the Evolution of Flowering Plants. *Science* 342: 1241089–1241089.

Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P. 2008a. Intracellular Compartments and Protein Sorting. *Molecular Biology of the Cell*. Garland Science, pp. 695–748.

Albrecht C, Russinova E, Hecht V, Baaijens E, Vries S De. 2005. The *Arabidopsis thaliana* somatic embryogenesis receptor-like kinases1 and 2 control male sporogenesis. *The Plant Cell* 17: 3337–3349.

Albrecht C, Russinova E, Kemmerling B, Kwaaitaal M, Vries SC De. 2008. *Arabidopsis* somatic embryogenesis receptor kinase proteins serve brassinosteroid-dependent and -independent signaling pathways. *Plant Physiology* 148: 611–619.

Alonso M, Stepanova AN, Leisse TJ, Kim CJ, Chen H, Shinn P, Stevenson DK, Zimmerman J, Barajas P, Cheuk R. 2003. Genome-Wide Insertional Mutagenesis of Arabidopsis thaliana. *Science* 301: 653-657.

Asakawa H, Yang H-J, Hiraoka Y, Haraguchi T. 2016. Virtual nuclear envelope breakdown and its regulators in fission yeast meiosis. *Frontiers in Cell and Developmental Biology* 4: 1–8.

Asaoka R, Uemura T, Ito J, Fujimoto M, Ito E, Ueda T, Nakano A. 2013. *Arabidopsis* raba1 gtpases are involved in transport between the trans-golgi network and the plasma membrane, and are required for salinity stress tolerance. *Plant Journal* 73: 240–249.

Atkinson JA, Rasmussen A, Traini R, Voss U, Sturrock C, Mooney SJ, Wells DM, Bennett MJ. 2014. Branching out in roots: uncovering form, function, and regulation. *Plant Physiology* 166: 538–550.

Backues SK, Korasick DA, Heese A, Bednarek SY. 2010. The Arabidopsis dynamin-related protein2 family is essential for gametophyte development. *The Plant Cell* 22: 3218–3231.

Bae S, Park J, Kim JS. 2014. Cas-OFFinder: A fast and versatile algorithm that searches for potential off-target sites of cas9 rna-guided endonucleases. *Bioinformatics* 30: 1473–1475.

Balderhaar HJ k., Ungermann C. 2013. CORVET and HOPS tethering complexes - coordinators of endosome and lysosome fusion. *Journal of Cell Science* 126: 1307–1316.

Barlowe C, Or L, Yeung T, Salama N, Rexach MF. 1994. COPII: A membrane coat formed by set proteins that drive vesicle budding from the endoplasmic reticulum. *Cell* 77.

Barr FA. 2013. Rab gtpases and membrane identity: causal or inconsequential? *Journal of Cell Biology* 202: 191–199.

Le Bars R, Marion J, Le Borgne R, Satiat-Jeunemaitre B, Bianchi MW. 2014. ATG5 defines a phagophore domain connected to the endoplasmic reticulum during autophagosome formation in plants. *Nature Communications* 5: 1–10.

Bassham DC. **2015.** Methods for analysis of autophagy in plants. *Methods* 75: 181–188.

Bassham DC, Brandizzi F, Otegui MS, Sanderfoot A. 2008. The secretory system of *Arabidopsi*s. *The Arabidopsis Book* 65: 1.

Bassham DC, Laporte M, Marty F, Moriyasu Y, Ohsumi Y, Olsen LJ, Yoshimoto K. 2006. Autophagy in development and stress responses of plants. *Autophagy* 2: 2–11.

Batoko H, Zheng HQ, Hawes C, Moore I. 2000. A rab1 GTPase is required for transport between the endoplasmic reticulum and Golgi apparatus and for normal Golgi movement in plants. *The Plant Cell* 12: 2201–2218.

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.

Beeckman T, Smet I De. 2014. Pericycle. *Current Biology 2*4: R378–R379.

Benková E, Michniewicz M, Sauer M, Teichmann T, Seifertová D, Jürgens G, Friml J. 2003. Local, efflux-dependent auxin gradients as a common module for plant organ formation. *Cell* 115: 591–602.

Birnbaum KD. 2016. How many ways are there to make a root? *Current Opinion in Plant Biology* 34: 61–67.

Bliss BJ, Wanke S, Barakat A, Ayyampalayam S, Wickett N, Wall PK, Jiao Y, Landherr L, Ralph PE, Hu Y, et al. 2013. Characterization of the basal angiosperm Aristolochia fimbriata: a potential experimental system for genetic studies. *BMC Plant Biology* 13: 13.

Bloch D, Pleskot R, Pejchar P, Potocký M, Trpkošová P, Cwiklik L, Vukašinović N, Sternberg H, Yalovsky S, Žárský V. 2016. Exocyst SEC3 and phosphoinositides define sites of exocytosis in pollen tube initiation and growth. *Plant Physiology* 172:00690.2016.

Blum T, Briesemeister S, Kohlbacher O. 2009. MultiLoc2 : integrating phylogeny and Gene Ontology terms improves subcellular protein localization prediction. *BMC Bioinformatics* 11: 1–11.

Boens N, Qin W, Basaric N, Orte A, Talavera EM, Alvarez-pez JM. 2006. Photophysics of the fluorescent ph indicator BCECF. *Journal of Physical Chemistry A* 110: 9334–9343.

Boevink P, Oparka K, Cruz SS, Martin B, Betteridge A, Hawes C. 1998. Stacks on tracks: the plant Golgi apparatus traffics on an actin/ER network. *The Plant Journal* 15: 441–447.

Boller T, Kende H. 1979. Hydrolytic enzymes in the central vacuole of plant cells. *Plant Physiology* 63: 1123–1132.

Boruc J, Zhou X, Meier I. 2012. Dynamics of the plant nuclear envelope and nuclear pore. *Plant Physiology* 158: 78–86.

Bouché F. 2017. Arabidopsis Root cell types. Figshare.

Bourge M, Fort C, Soler MN, Satiat-Jeunemaitre B, Brown SC. 2014. A pulse-chase strategy combining click-EdU and photoconvertible fluorescent reporter: tracking Golgi protein dynamics during the cell cycle. *New Phytologist* 205:938-950.

Brandizzi F, Fricker M, Hawes C. 2002. A greener world: the revolution in plant bioimaging. *Nature Reviews Molecular Cell Biology* 3: 520–530.

Briesemeister S, Kohlbacher O. 2010. YLoc-an interpretable web server for predicting subcellular localization. *Nucleic Acids Research* 38: 497–502.

Brückner A, Polge C, Lentze N, Auerbach D, Schlattner U. 2009. Yeast two-hybrid, a powerful tool for systems biology. *International Journal of Molecular Sciences* 10: 2763–2788.

Brüx A, Liu T-Y, Krebs M, Stierhof Y-D, Lohmann JU, Miersch O, Wasternack C, Schumacher K. 2008. Reduced v-atpase activity in the trans-golgi network causes oxylipin-dependent hypocotyl growth inhibition in *Arabidopsis*. *The Plant Cell* 20: 1088–1100.

Bücherl CA, Bader A, Westphal AH, Laptenok SP, Borst JW. 2014. FRET-FLIM applications in plant systems. *Protoplasma* 251: 383–394.

Bush DS. 1995. Calcium regulation in plant cells and its role in signalling. *Annual Review of Plant Physiology* 46: 95–122.

Cai H, Reinisch K, Ferro-Novick S. 2007. Coats, tethers, rabs, and snares work together to mediate the intracellular destination of a transport vesicle. *Developmental Cell* 12: 671–682.

Cai Y, Zhuang X, Wang J, Wang H, Lam SK, Gao C, Wang X, Jiang L. 2012. Vacuolar degradation of two integral plasma membrane proteins, atlrr84a and osscamp1, is cargo ubiquitination-independent and prevacuolar compartment-mediated in plant cells. *Traffic* 13: 1023–1040.

Calcutt JR. 2009. ABP195, a novel plant actin-binding protein. Thesis.

Cao X, Rogers SW, Butler J, Beevers L, Rogers JC. 2000. Structural requirements for ligand binding by a probable plant vacuolar sorting receptor. *The Plant Cell* 12: 493–506.

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.

Carter C, Pan S, Zouhar J, Avila EL, Girke T, Raikhel N V. 2004. The vegetative vacuole proteome of arabidopsis thaliana reveals predicted and unexpected proteins. *The Plant Cell* 16: 3285–3303.

Casimiro I, Beeckman T, Graham N, Bhalerao R, Zhang H, Casero P, Sandberg G, Bennett MJ. 2003. Dissecting *Arabidopsis* lateral root development. *Trends in Plant Science* 8: 165–171.

Casimiro I, Marchant A, Bhalerao RP, Beeckman T, Dhooge S, Swarup R, Graham N, Inze D, Sandberg G, Casero PJ, et al. 2001. Auxin transport promotes arabidopsis lateral root initiation. *The Plant Cell* 13: 843–852.

Chaen K, Noguchi J, Omori T, Kakuta Y, Kimura M. 2012. Crystal structure of the rice branching enzyme I in complex with maltopentaose. *Biochemical and Biophysical Research Communications* 424: 508–511.

Chamala S, Chanderbali AS, Der JP, Lan T, Walts B, Albert VA, dePamphilis CW, Leebens-Mack J, Rounsley S, Schuster SC, et al. 2013. Assembly and validation of the genome of the nonmodel basal angiosperm *Amborella*. *Science 3*42: 1516–1517.

Chanderbali AS, Albert VA, Ashworth VETM, Clegg MT, Litz RE, Soltis DE, Soltis PS. 2008. *Persea* americana (avocado): Bringing ancient flowers to fruit in the genomics era. *BioEssays* 30: 386–396.

Chang J, Lee S, Blackstone C. 2013. Protrudin binds atlastins and endoplasmic reticulum-shaping proteins and regulates network formation. *Proceedings of the National Academy of Science of the United States of America* 110: 14954–14959.

Chanoca A, Kovinich N, Burkel B, Stecha S, Bohorquez-restrepo A, Ueda T, Eliceiri KW, Grotewold E, Otegui MS. 2015. Anthocyanin vacuolar inclusions form by a microautophagy mechanism. *The Plant Cell* 27: 2545–2559.

Chaumont F De, Dallongeville S, Chenouard N, Hervé N, Pop S, Provoost T, Meas-yedid V, Pankajakshan P, Lecomte T, Montagner Y Le. 2012. Icy: an open bioimage informatics platform for extended reproducible research. *Nature* 9(7):690-696.

Chavrier P, Gorvel J-P, Stelzer EHK, Simons K, Gruenberg J, Zerial M. 1991. Hypervariable C-terminal domain of rab proteins acts as a targeting signal. *Nature* 353: 769–772.

Chege NW, Pfeffer SR. 1990. Compartmentation of the Golgi complex: brefeldin-a distinguishes trans-golgi cisternae from the trans-golgi network. *The Journal of Cell Biology* 111: 893–899.

Chen X, Huang Q, Zhang F, Wang B, Wang J. 2014. zmcipk21, a maize cbl-interacting kinase , enhances salt stress tolerance in *Arabidopsis thaliana*. *International Journal of Molecular Sciences*: 14819–14834.

Chen L, Su Z, Huang L, Xia F, Qi H, Xie L, Xiao S, Chen Q. 2017. The amp-activated protein kinase kin10 is involved in the regulation of autophagy in *Arabidopsis*. *Frontiers in Plant Science* 8: 1–11.

Chien CT, Bartel PL, Sternglanz R, Fields S. 1991. The two-hybrid system: a method to identify and clone genes for proteins that interact with a protein of interest. *Proceedings of the National Academy of Sciences of the United States of America* 88: 9578–9582.

Choi Y, Harada E, Wada M, Tsuboi H, Morita Y, Kusano T, Sano H. 2001. Detoxification of cadmium in tobacco plants: formation and active excretion of crystals containing cadmium and calcium through trichomes. *Planta*: 45–50.

Chou K, Shen H. 2010. Plant-mPLoc : a top-down strategy to augment the power for predicting plant protein subcellular localization. *PLoS ONE* 5: e11335.

Chow C-M, Neto H, Foucart C, Moore I. 2008. Rab-a2 and rab-a3 gtpases define a trans-golgi endosomal membrane domain in *Arabidopsis* that contributes substantially to the cell plate. *The Plant Cell* 20: 101–123.

Chung T, Phillips AR, Vierstra RD. 2010. ATG8 lipidation and atg8-mediated autophagy in arabidopsis require atg12 expressed from the differentially controlled atg12a and atg12b loci. *The Plant Journal* 62: 483–493.

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

Cole C, Barber JD, Barton GJ. 2008. The jpred 3 secondary structure prediction server. *Nucleic Acids Research* 36: 197–201.

Contento AL, Bassham DC. 2012. Structure and function of endosomes in plant cells. *Journal of Cell Science* 125: 3511–8.

Contento AL, Xiong Y, Bassham DC. 2005. Visualization of autophagy in *Arabidopsis u*sing the fluorescent dye monodansylcadaverine and a GFP-AtATG8e fusion protein. *The Plant Journal 42*: 598–608.

Crane PR, Friis EM, Pedersen KR. 1995. The origin and early diversification of angiosperms. *Nature* 374: 27–33.

Crepet WL, Friis EM, Nixon KC. 1991. Fossil evidence for the evolution of biotic pollination. *Philosophical Transactions of the Royal Society of London. Series B, Biological sciences* 333: 187–195.

Van Criekinge W, Beyaert R. 1999. Yeast Two-Hybrid: State of the Art. *Biological Procedures Online* 2: 1–38.

Crowley KS, Liao S, Worrell VE, Reinhart GD, Johnson AE. 1994. Secretory proteins move through the endoplasmic reticulum membrane via an aqueous, gated pore. *Cell* 78: 461–471.

Crozet P, Margalha L, Confraria A, Rodrigues A, Martinho C, Adamo M, Elias CA, Baena-GonzÃ¡lez E. 2014. Mechanisms of regulation of SNF1/AMPK/SnRK1 protein kinases. *Frontiers in Plant Science* 5: 1–17.

Cui Y, Shen J, Gao C, Zhuang X, Wang J, Jiang L. 2016. Biogenesis of plant prevacuolar multivesicular bodies. *Molecular Plant* 9: 774–786.

Cui Y, Zhao Q, Gao C, Ding Y, Zeng Y, Ueda T, Nakano A, Jiang L. 2014. Activation of the rab7 gtpase by the mon1-ccz1 complex is essential for pvc-to-vacuole trafficking and plant growth in *Arabidopsis*. *The Plant Cell* 26: 2080–2097.

Cui Y, Zhao Q, Xie H, Wong WS, Wang X, Gao C, Ding Y, Tan Y, Ueda T, Zhang Y, et al. 2017. Monensin sensitivity1 (mon1)/calcium caffeine zinc sensitivity1 (ccz1)mediated rab7 activation regulates tapetal programmed cell. *Plant Physiology* 173: 206–218.

Cvrčková F, Grunt M, Bezvoda R, Hála M, Kulich I, Rawat A, Žárský V. 2012. Evolution of the land plant exocyst complexes. *Frontiers in Plant Science* 3: 1–13.

Czechowski T, Stitt M, Altmann T, Udvardi MK. 2005. Genome-wide identification and testing of superior reference genes for transcript normalization. *Plant Physiology* 139: 5–17.

Danino D, Hinshaw JE. 2001. Dynamin family of mechanoenzymes. *Current Opinion in Cell Biology*: 454–460.

Darwin CR. 1862. On the various contrivances by which British and foreign orchids are fertilised by insects, and on the good effects of intercrossing. London: John Murray.

Day IS, Reddy VS, Shad Ali G, Reddy a SN. 2002. Analysis of EF-hand-containing proteins in *Arabidopsis*. *Genome Biology* 3: 1–24.

Debernardi JM, Mecchia MA, Vercruyssen L, Smaczniak C, Kaufmann K, Inze D, Rodriguez RE, Palatnik JF. 2014. Post-transcriptional control of GRF transcription factors by microRNA miR396 and GIF co-activator affects leaf size and longevity. *The Plant Journal* 413–426.

Debernardi JM, Rodriguez RE, Mecchia MA, Palatnik JF. 2012. Functional specialization of the plant mir396 regulatory network through distinct microrna – target interactions. *PLoS Genetics* 8:1.

Deeks MJ, Calcutt JR, Ingle EKS, Hawkins TJ, Chapman S, Richardson AC, Mentlak DA, Dixon MR, Cartwright F, Smertenko AP. 2012a. A superfamily of actin-binding proteins at the actin-membrane nexus of higher plants. *Current Biology* 22: 1595–1600.

Deeks MJ, Calcutt JR, Ingle EKS, Hawkins TJ, Chapman S, Richardson AC, Mentlak DA, Dixon MR, Cartwright F, Smertenko AP. 2012b. A superfamily of actin-binding proteins at the actin-membrane nexus of higher plants. *Current Biology 22*: 1595–1600.

Deeks MJ, Hussey PJ. 2005. Arp2/3 and scar: plants move to the fore. *Nature Reviews Molecular Cell Biology* 6: 954–964.

Deeks MJ, Karlotti D, Davies B, Mahlo R, Hussey PJ. 2004. Arabidopsis nap1 is essential for arp2/3-dependent trichome morphogenesis. *Current Biology* 14: 1410–1414.

Dempski RE, Imperiali B. 2002. Oligosaccharyl transferase: Gatekeeper to the secretory pathway. *Current Opinion in Chemical Biology* 6: 844–850.

Deng X, Quail PH. 1991. copl : a regulatory locus involved in light-controlled development and gene expression in *Arabidopsis*. *Genes and Development* 5(7): 1172–1182.

Deprez P, Gautschi M, Helenius A. 2005. More than one glycan is needed for er glucosidase ii to allow entry of glycoproteins into the calnexin / calreticulin cycle. *Molecular Cell* 19: 183–195.

Dettmer J, Hong-Hermesdorf A, Stierhof Y-D, Schumacher K. 2006. Vacuolar h+-atpase activity is required for endocytic and secretory trafficking in *Arabidopsis*. *The Plant Cell* 18: 715–730.

Dewitte W, Murray JAH. 2003. The plant cell cycle. *Annual Review of Plant Biology* 54: 235–264.

Dewitte W, Scofield S, Alcasabas AA, Maughan SC, Menges M, Braun N, Collins C, Nieuwland J, Prinsen E, Sundaresan V. 2007. Arabidopsis cycd3 d-type cyclins link cell proliferation and endocycles and are rate-limiting for cytokinin responses. *Proceedings of the National Academy of Sciences of the United States of America* 104: 14537–14542.

Dhonukshe P, Aniento F, Hwang I, Robinson DG, Mravec J. 2007. Clathrin-mediated constitutive endocytosis of pin auxin efflux carriers in *Arabidopsis. Current Biology*: 520–527.

Diekmann Y, Seixas E, Gouw M, Tavares-cadete F, Seabra MC. 2011. Thousands of rab gtpases for the cell biologist. *PLoS Computational Biology* 7: e1002217.

Ding Y, Wang J, Ho J, Lai C, Hoi V, Chan L, Wang X, Cai Y, Tan X, Bao Y. 2013. Exo70E2 is essential for exocyst subunit recruitment and EXPO formation in both plants and animals. *Molecular Biology of the Cell* 25: 412–426.

Dodd ME, Silvertown J, Chase MW. 1999. Phylogenetic analysis of trait evolution and species diversity variation among angiosperm families. *Evolution* 53: 732–744.

Doelling JH, Walker JM, Friedman EM, Thompson AR, Vierstra RD. 2002. The apg8 / 12-activating enzyme apg7 is required for proper nutrient recycling and senescence in *Arabidopsis thaliana*. *The Journal of Biological Chemistry* 277: 33105–33114.

Doudna JA, Batey RT. 2004. Structural insights into the signal recognition particle. *Annual Reviews in Biochemistry* 73: 539–557.

Doyle JA, Donoghue MJ, Zimmer EA. 1994. Integration of morphological and ribosomal rna data on the origin of angiosperms. *Annals of the Missouri Botanical Garden* 81: 419–450.

Drozdetskiy A, Cole C, Procter J, Barton GJ. 2015. JPred4 : a protein secondary structure prediction server. *Nucleic Acids Research* 43: 389–394.

Du W, Tamura K, Stefano G, Brandizzi F. 2013. The integrity of the plant golgi apparatus depends on cell growth-controlled activity of gnl1. *Molecular Plant* 6: 905–915.

Dubrovsky JG, Guttenberger M, Saralegui A, Napsucialy-Mendivil S, Voigt B, Baluška F, Menzel D. 2006. Neutral red as a probe for confocal laser scanning microscopy studies of plant roots. *Annals of Botany* 97: 1127–1138.

Dupree P, Sherrier DJ. 1998. The plant Golgi apparatus. *Biochemica et Biophysica Acta* 1404:259-270.

Ebine K, Fujimoto M, Okatani Y, Nishiyama T, Goh T, Ito E, Dainobu T, Nishitani A, Uemura T, Sato MH.2011. A membrane trafficking pathway regulated by the plant-specific rab gtpase ara6. *Nature Cell Biology* 13: 853–859.

Ebine K, Inoue T, Ito J, Ito E, Uemura T, Goh T, Abe H, Sato K, Nakano A, Ueda T. 2014. Plant vacuolar trafficking occurs through distinctly regulated pathways. *Current Biology* 24: 1375–1382.

Edlund AF, Swanson R, Preuss D. 2004. pollen and stigma structure and function: the role of diversity in pollination. *The Plant Cell* 16: S84–S97.

Edwards K, Johnstone C, Thompson C. 1991. A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. *Nucleic Acids Research* 19: 1349.

Egelie KJ, Graff GD, Strand SP, Johansen B. 2016. The emerging patent landscape of CRISPR–Cas gene editing technology. *Nature* 34(10): 1025–1031.

Eleftheriou EP, Palevitz BA. 1992. The effect of cytochalasin D on preprophase band organization in root tip cells of Allium. *Journal of Cell Science* 998: 989–998.

Elmore JM, Liu J, Smith B, Phinney B, Coaker G. 2012. Quantitative proteomics reveals dynamic changes in the plasma membrane during arabidopsis immune signalling. *Molecular & Cellular Proteomics* 11: M111.014555.

Emanuelle S, Hossain MI, Moller IE, Pedersen HL, Van De Meene AML, Doblin MS, Koay A, Oakhill JS, Scott JW, Willats WGT. 2015. SnRK1 from *Arabidopsis thaliana* is an atypical AMPK. *Plant Journal* 82: 183–192.

Engelsberger WR, Schulze WX. 2012. Nitrate and ammonium lead to distinct global dynamic phosphorylation patterns when resupplied to nitrogen-starved *Arabidopsis* seedlings. *Plant Journal* 69: 978–995.

Engler C, Kandzia R, Marillonnet S. 2008. A one pot, one step, precision cloning method with high throughput capability. *PLoS ONE* 3.

Evans DE, Graumann K. 2014. Dynamics of the plant nuclear envelope during cell division. *Plant Cell Division*. *Humana Press*, 115–126.

Faraco M, Li Y, Li S, Verweij W, Koes R, Quattrocchio FM, Faraco M, Li Y, Li S, Spelt C. 2017. Report a tonoplast p 3b -atpase mediates fusion of two types of vacuoles in petal cells report a tonoplast p 3b -atpase mediates fusion of two types of vacuoles in petal cells. *Cell Reports* 19: 2413–2422.

Fendrych M, Synek L, Pečenková T, Toupalová H, Cole R, Drdová E, Nebesářová J, Šedinová M, Hála M, Fowler JE. 2010. The *Arabidopsis* exocyst complex is involved in cytokinesis and cell plate maturation. *The Plant Cell* 22: 3053–3065.

Feraru E, Paciorek T, Feraru MI, Zwiewka M, De Groodt R, De Rycke R, Kleine-Vehn J, Friml JJ. 2010. The AP-3 β adaptin mediates the biogenesis and function of lytic vacuoles in *Arabidopsis*. *The Plant Cell* 22: 2812–24.

Fiserova J, Kiseleva E, Goldberg MW. 2009. Nuclear envelope and nuclear pore complex structure and organization in tobacco BY-2 cells. *The Plant Journal* 59: 243–255.

Foresti O, DaSilva LLP, Denecke J. 2006. Overexpression of the *Arabidopsis* syntaxin pep12/syp21 inhibits transport from the prevacuolar compartment to the lytic vacuole in vivo. *The Plant Cell* 18: 2275–2293.

Foresti O, Marchis F De, Virgilio M De, Klein EM, Arcioni S. 2008. Protein domains involved in assembly in the endoplasmic reticulum promote vacuolar delivery when fused to secretory GFP, indicating a protein quality control pathway for degradation in the plant vacuole. *Molecular Plant* 1: 1067–1076.

Frigerio L, Hinz G, Robinson DG. 2008. Multiple vacuoles in plant cells: rule or exception? : 1564–1570.

Fukaki H, Tameda S, Masuda H, Tasaka M. 2002. Lateral root formation is blocked by a gain-of-function mutation in the solitary-root/IAA14 gene of *Arabidopsis*. *Plant Journal* 29: 153–168.

Galinha C, Hofhuis H, Luijten M, Willemsen V, Blilou I, Heidstra R, Scheres B. 2007. PLETHORA proteins as dose-dependent master regulators of *Arabidopsis* root development. *Nature* 449: 1053–1057.

Galletta BJ, Rusan NM. 2015. A yeast two-hybrid approach for probing protein-protein interactions at the centrosome. *Methods Cell Biology* 129: 251–277.

Gao C, Luo M, Zhao Q, Yang R, Cui Y, Zeng Y, Xia J, Jiang L. 2014. A unique plant escrt component, free1, regulates multivesicular body protein sorting and plant growth. *Current Biology* 24: 2556–2563.

Gao C, Zhuang X, Cui Y, Fu X, He Y, Zhao Q, Zeng Y, Shen J, Luo M, Jiang L. 2015. Dual roles of an Arabidopsis ESCRT component FREE1 in regulating vacuolar protein transport and autophagic degradation. *Proceedings of the National Academy of Sciences of the United States of America* 112: 1886–1891.

Gardiner J, Overall R, Marc J. 2011. Putative *Arabidopsis* homologues of metazoan coiled-coil cytoskeletal proteins. *Cell Biology International* 35: 767–774.

Gattolin S, Sorieul M, Frigerio L. 2010. Tonoplast intrinsic proteins and vacuolar identity. *Biochemical Society Transactions* 38: 769–773.

Gattolin S, Sorieul M, Frigerio L. 2011. Mapping of tonoplast intrinsic proteins in maturing and germinating arabidopsis seeds reveals dual localization of embryonic tips to the tonoplast and plasma membrane. *Molecular Plant* 4: 180–189.

Geelen DN V, Inze DG. 2001. A bright future for the bright yellow-2 cell culture. *Plant Physiology* 127: 1375–1379.

Geldner N, Dénervaud-Tendon V, Hyman DL, Mayer U, Stierhof YD, Chory J. 2009. Rapid, combinatorial analysis of membrane compartments in intact plants with a multicolor marker set. *Plant Journal* 59: 169–178.

Gelvin SB. 2003. Agrobacterium-mediated plant transformation: the biology behind the 'gene-jockeying' tool. *Microbiology and Molecular Biology Reviews* 67: 16–37.

Gelvin SB. 2010. Plant proteins involved in agrobacterium-mediated genetic transformation. *Annual Review of Phytopathology* 48: 45–68.

Gendre D, McFarlane HE, Johnson E, Mouille G, Sjodin A, Oh J, Levesque-Tremblay G, Watanabe Y, Samuels L, Bhalerao RP. 2013. Trans-Golgi network localized echidna/ypt interacting protein complex is required for the secretion of cell wall polysaccharides in *Arabidopsis*. *The Plant Cell* 25: 2633–2646.

Gendre D, Oh J, Boutte Y, Best JG, Samuels L, Nilsson R, Uemura T, Marchant A, Bennett MJ, Grebe M, 2011. Conserved *Arabidopsis* ECHIDNA protein mediates trans-Golgi-network trafficking and cell elongation. *Proceedings of the National Academy of Sciences of the United States of America* 108: 8048–8053.

Ghoshal K, Theilmann J, Reade R, Sanfacon H, Rochon D. 2014. The Cucumber leaf spot virus p25 auxiliary replicase protein binds and modifies the endoplasmic reticulum via N-terminal transmembrane domains. *Virology* 468–470: 36–46.

Gibson DG, Young L, Chuang R-Y, Venter JC, Hutchison CA, Smith HO. 2009. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods* 6: 343–345.

Gillingham AK, Munro S. 2016. Finding the Golgi: golgin coiled-coil proteins show the way. *Trends in Cell Biology* 26: 399–408.

Goh T, Uchida W, Arakawa S, Ito E, Dainobu T, Ebine K, Takeuchi M, Sato K, Ueda T, Nakano A. 2007. VPS9a, the common activator for two distinct types of rab5 gtpases, is essential for the development of *Arabidopsis thaliana*. *The Plant Cell* 19: 3504–3515.

Gomord V, Denmat L-A, Fitchette-Laine A-C, Satiat-Jeunemaitre B, Hawes C, Faye L. 1997. The C-terminal HDEL sequence is sufficient for retention of secretory proteins in the endoplasmic reticulum (ER) but promotes vacuolar targeting of proteins that escape the ER. *The Plant Journal* 11: 313–325.

Gomord V, Fitchette A, Menu-bouaouiche L, Saint-Jore-Dupas C, Michaud D, Faye L. 2010. Plant-specific glycosylation patterns in the context of therapeutic protein production. *Plant Biotechnology Journal*: 564–587.

Gorelick R. 2001. Did insect pollination cause increased seed plant diversity? *Biological Journal of the Linnean Society* 74: 407–427.

Griffiths G, Simons K. 1986. The trans Golgi network: sorting at the exit site of the Golgi complex. *Science* 234: 438–443.

Grosshans BL, Ortiz D, Novick P. 2006. Rabs and their effectors: achieving specificity in membrane traffic. *Proceedings of the National Academy of Sciences of the United States of America* 103: 11821–7.

Gu Y, Innes RW. 2012. The keep on going protein of *Arabidopsis* regulates intracellular protein traf fi cking and is degraded during fungal infection. *The Plant Cell* 24: 4717–4730.

Guo YL. 2013. Gene family evolution in green plants with emphasis on the origination and evolution of *Arabidopsis thaliana* genes. *Plant Journal* 73: 941–951.

Guo W, Roth D, Walch-Solimena C, Novick P. 1999. The exocyst is an effector for Sec4P, targeting secretory vesicles to sites of exocytosis. *EMBO Journal* 18: 1071–1080.

Güttinger S, Laurell E, Kutay U. 2009. Orchestrating nuclear envelope disassembly and reassembly during mitosis. *Nature Reviews Molecular Cell Biology* 10: 178–191.

Hadlington JL, Denecke J. 2000. Sorting of soluble proteins in the secretory pathway of plants. *Current Opinion in Plant Biology* 3: 461–468.

Hala M, Cole R, Synek L, Drdova E, Pec T, Lamkemeyer T, Madlung J, Hochholdinger F, Fowler JE, Zarsky V. 2008. An exocyst complex functions in plant cell growth in *Arabidopsis* and tobacco. *The Plant Cell* 20: 1330–1345.

Halford NG, Paul MJ. 2003. Carbon metabolite sensing and signalling. *Plant Biotechnology Journal* 1: 381–398.

Hamada T, Nagasaki-Takeuchi N, Kato T, Fujiwara M, Sonobe S, Fukao Y, Hashimoto T. 2013. Purification and characterization of novel microtubule-associated proteins from arabidopsis cell suspension cultures. *Plant Physiology* 163: 1804–1816.

Hanaoka H, Noda T, Shirano Y, Kato T, Hayashi H, Shibata D, Tabata S, Ohsumi Y. 2002a. Leaf senescence and starvation-induced chlorosis are accelerated by the disruption of an Arabidopsis autophagy gene. *Plant Physiology* 129: 1181–93.

Hanaoka H, Noda T, Shirano Y, Kato T, Hyashi H, Shibata D, Tabata S, Ohsumi Y. 2002b. Leaf senescence and starvation-induced chlorosis are accelerated by the disruption of an *Arabidopsis* autophagy gene. *Plant Physiology* 129: 1181–1193.

Hao L, Liu J, Zhong S, Gu H, Qu L. 2016. AtVPS41-mediated endocytic pathway is essential for pollen tube – stigma interaction in *Arabidopsis*. *Proceedings of the National Academy of Sciences of the United States of America* 113: 6307–6312.

Harrington GN, Bush DR. 2003. The bifunctional role of hexokinase in metabolism and glucose signalling. *The Plant Cell* 15: 2493–2496.

Hawes C. 2005. Cell biology of the plant Golgi apparatus. *New Phytologist* 165: 29–44.

Hawes C, Kiviniemi P, Kriechbaumer V. 2015. The endoplasmic reticulum: A dynamic and well-connected organelle. *Journal of Integrative Plant Biology* 57: 50–62.

Hawes C, Osterrieder A, Sparkes IA, Ketelaar T. 2010a. Optical tweezers for the micromanipulation of plant cytoplasm and organelles. *Current Opinion in Plant Biology* 13: 731–735.

Hawes C, Satiat-Jeunemaitre B. 2005. The plant Golgi apparatus — Going with the flow. *Biochimica et Biophysica Acta* 1744: 93–107.

Hawes C, Schoberer J, Hummel E, Osterrieder A. 2010b. Biogenesis of the plant Golgi apparatus. *Biochemical Society Transactions* 38: 761–767.

Hawkins J, Boden M. 2006. Detecting and sorting targeting peptides with neural networks and support vector machines. *Journal of Bioinformatics and Computational Biology* 4: 1–18.

Hawkins J, Davis L, Bode M. 2007. Predicting nuclear localization. *Journal of Proteome Research* 6: 1402–1409.

Hawkins TJ, Deeks MJ, Wang P, Hussey PJ. 2014. The evolution of the actin binding NET superfamily. *Frontiers in Plant Science* 5: 254.

He B, Guo W. 2009. The exocyst complex in polarized exocytosis. *Current Opinion in Cell Biology* 21: 537–542.

He W, Lu Y, Qahwash I, Hu XY, Chang A, Yan R. 2004. Reticulon family members modulate BACE1 activity and amyloid-beta peptide generation. *Nature Methods* 10: 959–965.

Hecht V, Vielle-Calzada J-P, Hartog M V, Schmidt EDL, Boutilier K, Grossniklaus U, Vries SC De. 2001. The *Arabidopsis* SOMATIC EMBRYOGENESIS RECEPTOR KINASE 1 gene is expressed in

developing ovules and embryos and enhances embryogenic competence in culture. *Plant Physiology* 127: 803–816.

Hedbacker K, Townley R, Carlson M. 2004. Cyclic AMP-dependent protein kinase regulates the subcellular localization of snf1-sip1 protein kinase. *Society* 24: 1836–1843.

Heider MR, Gu M, Duffy CM, Mirza AM, Marcotte LL, Walls AC, Farrall N, Hakhverdyan Z, Field MC, Rout MP 2016. Subunit connectivity, assembly determinants and architecture of the yeast exocyst complex. *Nature Structural & Molecular Biology* 23: 59–66.

Helaers R, Milinkovitch MC. 2010. MetaPIGA v2.0: maximum likelihood large phylogeny estimation using the metapopulation genetic algorithm and other stochastic heuristics. *BMC Bioinformatics* 11:379

Helenius A, Aebi M. 2004. Roles of N-linked glycans in the endoplasmic reticulum. *Annual Review of Biochemistry* 73: 1019–49.

Helenius A, Trombetta ES, Herbert DN. 1997. Calnexin, calreticulin and the folding of glycoproteins. *Trends in Cell Biology* 7: 193–200.

Hickey CM, Wickner W. 2010. HOPS initiates vacuole docking by tethering membranes before trans-snare complex assembly. *Molecular Biology of the Cell* 21: 2297–2305.

Hicks GR, Rojo E, Hong S, Carter DG, Raikhel N V. 2004. Geminating pollen has tubular vacuoles, displays highly dynamic vacuole biogenesis, and requires vacuoless1 for proper function. *Plant Physiology* 134: 1227–1239.

Higaki T, Kutsuna N, Okubo E, Sano T, Hasezawa S. 2006. Actin microfilaments regulate vacuolar structures and dynamics: Dual observation of actin microfilaments and vacuolar membrane in living tobacco BY-2 cells. *Plant and Cell Physiology* 47: 839–852.

Hirsch C, Gauss R, Horn SC, Neuber O, Sommer T. 2009. The ubiquitylation machinery of the endoplasmic reticulum. *Nature* 458: 453–460.

Hochholdinger F, Zimmermann R. 2008. Conserved and diverse mechanisms in root development. *Current Opinion in Plant Biology* 11: 70–74.

Hoelz A, Debler EW, Blobel G. 2011. The structure of the nuclear pore complex. *Annual Review of Biochemistry* 80: 613–643.

Hoffmann RD, Palmgren M. 2016. Purifying selection acts on coding and non-coding sequences of paralogous genes in *Arabidopsis thaliana*. *BMC Genomics* 17: 456.

Hofte H, Hubbard L, Reizer J, Ludevid D, Herman EM, Chrispeels MJ. 1992. Vegetative and seed-specific forms of tonoplast intrinsic protein in the vacuolar membrane of *Arabidopsis thaliana*. *Plant Physiology* 15: 561–570.

Horn S, Pab N, Theuß VS, Busch A, Zachgo S. 2014. Analysis of the CYC/TB1 class of TCP transcription factors in basal angiosperms and magnoliids. *The Plant Journal* 81: 559–571.

Horton P, Park K, Obayashi T, Fujita N, Harada H, Nakai K. 2007. WoLF PSORT: protein localization predictor. *Nucleic Acids Research* 35: 585–587.

Houston NL, Fan C, Xiang JQ-Y, Schulze J-M, Jung R, Boston RS, Carolina N, H NCNL. 2005. Phylogenetic analyses identify 10 classes of the protein disulfide isomerase family in plants, including single-domain protein disulfide isomerase-related proteins. *Plant Physiology* 137: 762–778.

Howard DJ. 1999. Conspecific sperm and pollen precedence and speciation. *Annual Review of Ecology and Systematics* 30: 109–132.

Howe GA, Jander G. 2008. Plant Immunity to Insect Herbivores. *Annual Review of Plant Biology* 59: 41–66.

Howell SH. 2013. Endoplasmic reticulum stress responses in plants. *Annual Review of Plant Biology* 64: 477–99.

Hua S, Sun Z. 2001. Support vector machine approach for protein subcellular localization prediction. *Bioinformatics* 17: 721–728.

Hu J, Shibata Y, Voss C, Shemesh T, Li Z, Coughlin M, Kozlov MM, Rapoport TA, Prinz WA. 2008. Membrane proteins of the endoplasmic reticulum induce high-curvature tubules. *Science* 319: 1247–1250.

Hu J, Shibata Y, Zhu PP, Voss C, Rismanchi N, Prinz WA, Rapoport TA, Blackstone C. 2009. A class of dynamin-like gtpases involved in the generation of the tubular er network. *Cell* 138: 549–561.

Hulskamp M, Misera S, Jurgens G. 1994. Genetic dissection of trichome cell development in *Arabidopsis. Cell* 76: 555–566.

Humair D, Felipe DH, Neuhaus J, Paris N, Argand E, Neuchâtel C-. 2007. Demonstration in Yeast of the Function of BP-80, a Putative Plant Vacuolar Sorting Receptor. *The Plant Cell* 13: 781–792.

Hunter PR, Craddock CP, Di Benedetto S, Roberts LM, Frigerio L. 2007. Fluorescent reporter proteins for the tonoplast and the vacuolar lumen identify a single vacuolar compartment in *Arabidopsis* cells. *Plant Physiology* 145: 1371–1382.

Hussey PJ, Allwood EG, Smertenko AP. 2002. Actin-binding proteins in the *Arabidopsis* genome database: properties of functionally distinct plant actin-depolymerizing factors/cofilins. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 357: 791–798.

Hutagalung AH, Novick PJ. 2011. Role of Rab GTPases in membrane traffic and cell physiology. *Physiological Reviews* 91: 119–49.

Isono E, Katsiarimpa A, Mu IK, Anzenberger F, Stierhof Y, Geldner N, Chory J, Schwechheimer C. 2010. The deubiquitinating enzyme amsh3 is required for intracellular trafficking and vacuole biogenesis in *Arabidopsis thaliana*. *The Plant Cell* 22: 1826–1837.

Jackson-Constan D, Keegstra K. 2001. *Arabidopsis* genes encoding components of the chloroplastic protein import apparatus. *Plant Physiology* 125: 1567–1576.

Jaquinod M, Villiers F, Kieffer-Jaquinod S, Hugouvieux V, Bruley C, Garin J, Bourguignon J. 2007. A proteomics dissection of *Arabidopsis thaliana* vacuoles isolated from cell culture. *Molecular & Cellular Proteomics* 6: 394–412.

Jarvis P. 2008. Targeting of nucleus-encoded proteins to chloroplasts in plants. *New Pytologist* 179: 257–285.

Jefferson RA, Kavanagh TA, Bevan MW. 1987. GUS fusions: β-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *The EMBO Journal* 6: 3901–3907.

Jia T, Gao C, Cui Y, Wang J, Ding Y, Cai Y, Ueda T, Nakano A, Jiang L. 2013. ARA7(Q69L) expression in transgenic *Arabidopsis* cells induces the formation of enlarged multivesicular bodies. *Journal of Experimental Botany* 64: 2817–2829.

Jiang K, Moe-lange J, Hennet L, Feldman LJ. 2016. Salt stress affects the redox status of A*rabidopsis* root meristems. *Frontiers in Plant Science* 7: 1–10.

Jin JB, Kim YA, Kim J, Lee H, Kim H, Cheong G, Hwang I. 2001. A new dynamin-like protein, ADL6, is involved in trafficking from the trans-Golgi network to the central vacuole in *Arabidopsis*. *The Plant Cell* 13: 1511–1525.

Jin H, Yan Z, Nam KH, Li J. 2007. Article allele-specific suppression of a defective brassinosteroid receptor reveals a physiological role of uggt in er quality control. *Molecular Cell* 26: 821–830.

Jin Y, Zhuang M, Hendershot LM. 2009. ERdj3, a luminal er dnaj homologue, binds directly to unfolded proteins in the mammalian er : identification of critical residues. *Biohemistry*: 41–49.

Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. 2012. A programmable dual-rna-guided dna endonuclease in adaptive bacterial immunity. *Science* 337: 816–822.

Jinek M, Jiang F, Taylor DW, Sternberg SH, Kaya E, Ma E, Anders C, Hauer M, Zhou K, Lin S, 2014. Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. *Science* 343: 1247997–1247997.

Johannes E, Brosnan JM, Sanders D. 1992. Calcium channles in the vacuolar membrane of plants: multiple pathways for intracellular calcium mobilization. *Philisophical Transactions of the Royal Society B* 338: 105–112.

Johanson U, Karlsson M, Johansson I, Gustavsson S, Sjo S, Fraysse L, Weig AR, Kjellbom P. 2001. The complete set of genes encoding major intrinsic proteins in *Arabidopsis* provides a framework for a new nomenclature for major intrinsic proteins in plants. *Plant Physiology* 126: 1358–1369.

Johnson AE, van Waes MA. 1999. The translocon: a dynamic gateway at the ER membrane. *Annual Review of Cell and Developmental Biology* 15: 799–842.

Juah G-Y, Fischer AM, Grimes HD, Ryan CAR, Rogers JC. 1998. δ-Tonoplast intrinsic protein defines unique plant vacuole functions. *Proceedings of the National Academy of Sciences of the United States of America* 95: 12995–12999.

Juah G, Phillips TE, Rogers JC. 1999. Tonoplast intrinsic protein isoforms as markers for vacuolar functions. *The Plant Cell* 11: 1867–1882.

Jürgens G. 2004. Membrane trafficking in plants. *Annual Review of Cell and Developmental Biology* 20: 481–504.

Kamiya Y, Satoh T, Kato K. 2012. Molecular and structural basis for N -glycan-dependent determination of glycoprotein fates in cells. *Biochemica et Biophysica Acta* 1820: 1327–1337.

Kang JS, Frank J, Kang CH, Kajiura H, Vikram M, Ueda A, Kim S, Bahk JD, Triplett B, Fujiyama K 2008. Salt tolerance of *Arabidopsis thaliana* requires maturation of N-glycosylated proteins in the Golgi apparatus. *Proceedings of the National Academy of Sciences of the United States of America* 105: 5933–5938.

Kang BH, Nielsen E, Preuss ML, Mastronarde D, Staehelin LA. 2011. Electron tomography of raba4b- and pi4k-labeled trans Golgi network compartments in *Arabidopsis*. *Traffic* 12: 313–329.

Katifori E, Alben S, Cerda E, Nelson DR, Dumais J. 2010. Foldable structures and the natural design of pollen grains. *Proceedings of the National Academy of Sciences of the United States of America* 107: 7635–7639.

Kato T, Morita MT, Fukaki H, Yamauchi Y, Uehara M, Niihama M, Tasaka M. 2002. SGR2, a phospholipase-like protein, and ZIG/SGR4, a SNARE, are involved in the shoot gravitropism of *Arabidopsis*. *The Plant Cell* 14: 33–46.

Kim S-J, Bassham DC. 2011. TNO1 is involved in salt tolerance and vacuolar trafficking in *Arabidopsis. Plant Physiology* 156: 514–26.

Kim DI, KC B, Zhu W, Motamedchaboki K, Doye V, Roux KJ. 2014. Probing nuclear pore complex architecture with proximity-dependent biotinylation. *Proceedings of the National Academy of Sciences of the United States of America* 111: E2453-61.

Kim SH, Kwon C, Lee JH, Chung T. 2012. Genes for plant autophagy: Functions and interactions. *Molecules and Cells* 34: 413–423.

Kinchen JM, Ravichandran KS. 2010. Identification of two evolutionarily conserved genes regulating processing of engulfed apoptotic cells. *Nature* 464: 778–782.

Kirsch T, Parist N, Butler JM, Beevers L, Rogerst JC. 1994. Purification and initial characterization of a potential plant vacuolar targeting receptor. . *Proceedings of the National Academy of Sciences of the United States of America* 91: 3403–3407.

Kleffmann T, Russenberger D, Zychlinski A Von, Christopher W, Gruissem W, Baginsky S. 2004. The *Arabidopsis thaliana* chloroplast proteome reveals pathway abundance and novel protein functions. *Current Biology* 14: 354–362.

Klinger CM, Klute MJ, Dacks JB. 2013. Comparative genomic analysis of multi-subunit tethering complexes demonstrates an ancient pan-eukaryotic complement and sculpting in apicomplexa. *PLoS ONE* 8: 1–15.

Klionsky DJ. 2007. Autophagy: from phenomenology to molecular understanding in less than a decade. *Nature Reviews Molecular Cell Biology* 8: 931–937.

Klionsky DJ. 2012. Guidelines for the use and interpretation of assays for monitoring autophagy. *Autophagy*: 445–544.

Knittler MR, Haas IG. 1992. Interaction of BiP with newly synthesized immunoglobulin light chain molecules: cycles of sequential binding and release. *The EMBO Journal* 11: 1573–81.

Kohler S, Mcintosh AR, Moscovitch M, Riou-khamlichi C, Huntley R, Jacqmard A, Murray JAH. 1999. Cytokinin activation of *Arabidopsis* cell division through a d-type cyclin. *Science* 283: 1541–1545.

Kolb C, Nagel M-K, Kalinowska K, Hagmann J, Ichikawa M, Anzenberger F, Alkofer A, Sato MH, Braun P, Isono E. 2015. FYVE1 is essential for vacuole biogenesis and intracellular trafficking in *Arabidopsis*. *Plant Physiology* 167: 1361–1373.

Kondylis V, Rabouille C. 2009. The Golgi apparatus: Lessons from *Drosophila*. *FEBS Letters* 583: 3827–3838.

Kong RPW, Siu SO, Lee SSM, Lo C, Chu IK. 2011. Development of online high-/low-pH reversed-phase – reversed-phase two-dimensional liquid chromatography for shotgun proteomics : A reversed-phase-strong cation exchange-reversed-phase approach. *Journal of Chromatography A* 1218: 3681–3688.

Kramer L, Ungermann C. 2011. HOPS drives vacuole fusion by binding the vacuolar SNARE complex and the Vam7 PX domain via two distinct sites. *Molecular Biology of the Cell* 22: 2601–2611.

Kretsinger RH, Nockolds CE. 1973. Carp muscle calcium-binding protein ii structure determination and general description. *The Journal of Biological Chemistry* 248: 3313–3326.

Kriegel A, Andrés Z, Medzihradszky A, Krüger F, Scholl S, Delang S, Patir-Nebioglu MG, Gute G, Yang H, Murphy AS. 2015. Job sharing in the endomembrane system: vacuolar acidification requires the combined activity of V-ATPase and V-PPase. *The Plant Cell* 27: 3383–3396.

Krysan PJ, Young JC, Sussman MR. 1999. T-DNA as an insertional mutagen in *Arabidopsis*. *The Plant Cell* 11: 2283–2290.

Kulich I, Pečenková T, Sekereš J, Smetana O, Fendrych M, Foissner I, Höftberger M, Žárský V. 2013. *Arabidopsis* exocyst subcomplex containing subunit exo70b1 is involved in autophagy-related transport to the vacuole. *Traffic* 14: 1155–1165.

Lamb CA, Dooley HC, Tooze SA. 2012. Endocytosis and autophagy: Shared machinery for degradation. *Bioessays* 35: 34–45.

Larkin MA, Blackshields G, Brown NP, Chenna R, Mcgettigan PA, Mcwilliam H, Valentin F, Wallace IM, Wilm A, Lopez R. 2007. Clustal W and ClustalX version 2.0. *BMC Bioinformatics* 23: 2947–2948.

Lastdrager J, Hanson J, Smeekens S. 2014. Sugar signals and the control of plant growth and development. *Journal of Experimental Botany* 65: 799–807.

Lau OS, Deng XW. 2012. The photomorphogenic repressors COP1 and DET1 : 20 years later. *Trends in Plant Science* 17: 584–593.

Lee MG, Mishra A, Lambright G. 2009. Structural mechanisms for regulation of membrane traffic by rab GTPases. *Traffic*: 1377–1389.

Lee SY, Yang JS, Hong W, Premont RT, Hsu VW. 2005. ARFGAP1 plays a central role in coupling COPI cargo sorting with vesicle formation. *Journal of Cell Biology* 168: 281–290.

Lemmon AR, Milinkovitch MC. 2002. The metapopulation genetic algorithm : An efficient solution for the problem of large phylogeny estimation. *Proceedings of the National Academy of Sciences of the United States of America* 99(16): 10515-10521.

Letourneur F, Gaynor EC, Demolliere C, Duden R, Emr SD, Riezman H. 1994. Coatomer is essential for retrieval of dilysine-tagged proteins to the endoplasmic reticulum. *Cell* 79: 1199–1207.

Liang H, Ayyampalayam S, Wickett N, Barakat A, Xu Y, Landherr L, Ralph PE, Jiao Y, Xu T, Schlarbaum SE 2011. Generation of a large-scale genomic resource for functional and comparative genomics in *Liriodendron tulipifera* L. *Tree Genetics and Genomes* 7: 941–954.

Lin Y, DING Y, Wang J, Kung C-H, Zhuang X, Yin Z, Xia Y, Robinson DG, Shen J, Lin H-X. 2015. EXPO and autophagosomes are distinct organelles in plants. *Plant Physiology* 169: 1917–1932.

Lister R, Carrie C, Duncan O, Ho LHM, Howell KA, Murcha MW, Whelan J. 2007. Functional definition of outer membrane proteins involved in preprotein import into mitochondria. *The Plant Cell* 19: 3739–3759.

Liu Y, Xiong Y, Bassham DC, Liu Y, Xiong Y, Bassham DC. 2009. Autophagy is required for tolerance of drought and salt stress in plants. *Autophagy* 5: 954–963.

Liu F, Zhang L, Luo Y, Xu M, Fan Y, Wang L. 2016. Interactions of *Oryza sativa* OsCONTINUOUS VASCULAR RING-LIKE 1 (OsCOLE1) and OsCOLE1-INTERACTING PROTEIN reveal a novel intracellular auxin transport mechanism. *New Phytologist* 212: 96–107.

Lofke C, Dunser K, Scheuring D. 2015. Auxin regulates SNARE-dependent vacuolar morphology restricting cell size. *eLIFE* 4: e05868.

Lucas M, Kenobi K, von Wangenheim D, Voβ U, Swarup K, De Smet I, Van Damme D, Lawrence T, Péret B, Moscardi E. 2013. Lateral root morphogenesis is dependent on the mechanical properties of the overlaying tissues. *Proceedings of the National Academy of Sciences of the United States of America* 110: 5229–5234.

Luo X, Chen Z, Gao J, Gong Z. 2014a. Abscisic acid inhibits root growth in *Arabidopsis* through ethylene biosynthesis. *The Plant Journal* 1: 44–55.

Luo F, Fong YH, Zeng Y, Shen J, Jiang L, Wong K. 2014b. How vacuolar sorting receptor proteins interact with their cargo proteins : crystal structures of apo and cargo-bound forms of the protease-associated domain from an arabidopsis vacuolar sorting receptor. *The Plant Cell* 26: 3693–3708.

Luo L, Zhang P, Zhu R, Fu J, Su J, Zheng J, Wang Z. 2017. Autophagy is rapidly induced by salt stress and is required for salt tolerance in *Arabidopsis*. *Frontiers in Plant Science* 8: 1–13.

Ma B, Qian D, Nan Q, Tan C, An L, Xiang Y. 2012. *Arabidopsis* vacuolar h+-atpase (v-atpase) b subunits are involved in actin cytoskeleton remodeling via binding to, bundling, and stabilizing f-actin. *The Journal of Biological Chemistry* 287: 19008–19017.

Mainieri D, Rossi M, Archinti M, Bellucci M, De Marchis F, Vavassori S, Pompa A, Arcioni S, Vitale A. 2004. Zeolin. A new recombinant storage protein constructed using maize gamma-zein and bean phaseolin. *Plant Physiology* 136: 3447–3456.

Maizel A. 2017. Stages of lateral root devlopment. Figshare.

Malamy JE, Benfey PN. 1997. Organization and cell differentiation in lateral roots of *Arabidopsis thaliana*. *Development* 124: 33–44.

Mali P, Esvelt KM, Church GM. 2013. Cas9 as a versatile tool for engineering biology. *Nature Methods* 10: 957–963.

Mann M, Jensen ON. 2003. Proteomic analysis of post-translational modifications. *Nature Biotechnology* 21: 255–261.

Marchesi VT, Steers E. 1968. Selective solubilization of a protein component of the red cell membrane. *Science* 159: 203–204.

Martinez-Garcia JF, Monte E, Quail PH. 1999. A simple, rapid and quantitative method for preparing *Arabidopsis* protein extracts for immunoblot analysis. *Plant Journal* 20: 251–257.

Marty F. 1999. Plant vacuoles. *The Plant Cell* 11: 587–600.

Masterson J. 1994. Stomatal size in fossil plants: evidence for polyploidy oin majority of angiosperms. *Science* 264: 421–424.

Matsuda S, Vert J, Saigo H, Ueda N, Toh H, Akutsu T. 2005. A novel representation of protein sequences for prediction of subcellular location using support vector machines. *Protein Science* 14: 2804–2813.

Matsui M, Deng X. 1995. Cytoskeleton-associated protein. *Proceedings of the National Academy of Sciences of the United States of America* 92: 4239–4243.

Matsuzaki Y, Ogawa-ohnishi M, Mori A, Matsubayashi Y. 2010. Secreted peptide signals required for maintenance of root stem cell niche in *Arabidopsis*. *Science* 329: 1065–1068.

Mazel A, Leshem Y, Tiwari BS, Levine A. 2004. Induction of salt and osmotic stress tolerance by overexpression of an intracellular vesicle trafficking protein AtRab7 (AtRabG3e). *Plant Physiology* 134: 118–128.

McMichael CM, Bednarek SY. 2013. Cytoskeletal and membrane dynamics during higher plant cytokinesis. *New Phytologist* 197: 1039–1057.

Mcnew JA, Parlati F, Fukuda R, Johnston RJ, Paz K, Paumet F, So TH, Rothman JE. 2000. Compartmental specity of cellular membrane fusion encoded in SNARE proteins. *Nature* 407: 153–159.

Menges M, Murray JAH. 2002. Synchronous *Arabidopsis* suspension cultures for analysis of cell-cycle gene activity. *The Plant Journal* 30: 203–212.

Mentlak DA. **2016**. Studies on NET4B and associated proteins. Thesis.

Merkulova EA, Guiboileau A, Naya L, Yoshimoto K, Bourgin IJ, Versailles F-. 2014. Assessment and optimization of autophagy monitoring methods in arabidopsis roots indicate direct fusion of autophagosomes with vacuoles. *Plant Cell Physiology* 55: 715–726.

Meunier L, Usherwood Y, Chung KT, Hendershot LM. 2002. A subset of chaperones and folding enzymes form multiprotein complexes in endoplasmic reticulum to bind nascent proteins. *Molecular Biology of the Cell* 13: 4456–4469.

Michalak M, Groenendyk J, Szabo E, Gold LI, Opas M. 2009. Calretic

ulin, a multi-process calcium-buffering chaperone of the endoplasmic reticulum. *Biochemistry Journal* 666: 651–666.

Mineyuki Y, Palevitz BA. 1990. Relationship between preprophase band organization, F-actin and the division site in *Allium*. *Journal of Cell Science* 97: 283–296.

Mohannath G, Pontvianne F, Pikaard CS. 2016. Selective nucleolus organizer inactivation in *Arabidopsis* is a chromosome position-effect phenomenon. *Proceedings of the National Academy of Sciences of the United States of America* 113: 0–5.

Molinari M, Eriksson KK, Calanca V, Galli C, Cresswell P, Michalak M, Helenius A, Haven N, Tg A. 2004. Contrasting functions of calreticulin and calnexin in glycoprotein folding and er quality control. *Molecular Cell* 13: 125–135.

Molinari M, Galli C, Vanoni O, Arnold SM, Kaufman RJ. 2005. Persistent Glycoprotein Misfolding Activates the Glucosidase II / UGT1-Driven Calnexin Cycle to Delay Aggregation and Loss of Folding Competence. *Molecular Cell* 20: 503–512.

Moore MJ, Bell CD, Soltis PS, Soltis DE. 2007. Using plastid genome-scale data to resolve enigmatic relationships among basal angiosperms. *Proceedings of the National Academy of Sciences of the United States of America* 104: 19363–19368.

Moore I, Schell J, Palme K. 1995. Subclass-specific sequence motifs identified in Rab GTPases. *Trends in Biochemical Sciences* 20: 10–12.

Moore B, Zhou L, Rolland F, Hall Q. 2003. Role of the *Arabidopsis* glucose sensor HXK1 in nutrient, light, and hormonal signaling. *Science* 332: 332–336.

Morita MT, Shimada T. 2014. The plant endomembrane system-A complex network supporting plant development and physiology. *Plant and Cell Physiology* 55: 667–671.

Morré, Mollenhauer. 1974. The Endomembrane Concept: A functional integration of endoplasmic reticulum and Golgi apparatus. ln A.W. Robards (ed.), Dynamic aspects of plant ultrastructure, Maidenhead, pp.84-137.

Morton WM, Ayscough KR, Mclaughlin PJ. 2000. Latrunculin alters the actin-monomer subunit interface to prevent polymerization. *Nature Cell Biology* 2: 376–378.

Muller M, Schmidt O, Angelova M, Faserl K, Weys S, Kremser L, Dalik T, Kraft C, Trajanoski Z, Teis D. 2015. The coordinated action of the MVB pathway and autophagy ensures cell survival during starvation. *eLIFE*: 1–25.

Munns R, Tester M. 2008. Mechanisms of salinity tolerance. *Annual Review of Plant Biology* 59: 651–681.

Murayama KS, Kametani F, Saito S, Kume H, Akiyama H, Araki W. 2006. Reticulons RTN3 and RTN4-B/C interact with BACE1 and inhibit its ability to produce amyloid beta-protein. *European Journal of Neuroscience* 24: 1237–1244.

Nagata T, Sakamoto K, Shimizu T. 2004. Tobacco BY-2 cells: the present and beyond. *In Vitro Cell and Developmental Biology* 40: 163–166.

Nägele T, Weckwerth W, Szymanski JJ, Planck M. 2014. Mathematical modeling reveals that metabolic feedback regulation of SnRK1 and hexokinase is sufficient to control sugar homeostasis from energy depletion to full recovery. *Frontiers in Plant Science* 5: 1–11.

Napier RM, Fowke LC, Hawes C, Lewis M, Hugh RB. 1992. Immunological evidence that plants use both HDEL and KDEL for targeting proteins to the endoplasmic reticulum. *Journal of Cell Science* 271: 261–271.

Nebenfuhr A, Gallagher LA, Dunahay TG, Frohlick JA, Mazurkiewicz AM, Meehl JB, Staehelin LA. 1999. Stop-and-go movements of plant golgi stacks are mediated by the acto-myosin system. *Plant Physiology* 121: 1127–1141.

Niemes S, Langhans M, Viotti C, Scheuring D, San M, Yan W, Jiang L, Hillmer S. 2010. Retromer recycles vacuolar sorting receptors from the trans -Golgi network. *Plant Journal* 61:107–121.

Nieuwland J, Maughan S, Dewitte W, Scofield S, Sanz L, Murray JAH. 2009. The D-type cyclin CYCD4 ; 1 modulates lateral root density in Arabidopsis by affecting the basal meristem region. *Proceedings of the National Academy of Sciences of the United States of America* 106: 22528–22533.

Nikolovski N, Rubtsov D, Segura MP, Miles GP, Stevens TJ, Dunkley TPJ, Munro S, Lilley KS, Dupree P. 2012. Putative glycosyltransferases and other plant golgi apparatus proteins are revealed by lopit proteomics. *Plant Physiology* 160: 1037–1051.

Nishimasu H, Ran FA, Hsu PD, Konermann S, Shehata SI, Dohmae N, Ishitani R, Zhang F, Nureki O. 2014. Crystal structure of Cas9 in complex with guide RNA and target DNA. *Cell* 156: 935–949.

Niu B, Feng YJK, Li WLYCG. 2008. Using AdaBoost for the prediction of subcellular location of prokaryotic and eukaryotic proteins. *Molecular Diversity* 12: 41–45.

Nodzynski T, Feraru MI, Hirsch S, De Rycke R, Niculaes C, Boerjan W, Van Leene J, De Jaeger G, Vanneste S, Friml J. 2013. Retromer subunits VPS35A and VPS29 mediate prevacuolar compartment (PVC) function in arabidopsis. *Molecular Plant* 6: 1849–1862.

Nordmann M, Cabrera M, Perz A, Ostrowicz C, Engelbrecht-vandre S. 2010. The mon1-ccz1 complex is the gef of the late endosomal rab7 homolog ypt7. *Plant Journal* 20: 1654–1659.

Van Norman JM, Xuan W, Beeckman T, Benfey PN. 2013. To branch or not to branch: the role of pre-patterning in lateral root formation. *Development* 140: 4301–4310.

O'Malley RC, Ecker JR. 2010. Linking genotype to phenotype using the *Arabidopsis* unimutant collection. *Plant Journal* 61: 928–940.

Ohashi-Ito K, Bergmann DC. 2007. Regulation of the *Arabidopsis* root vascular initial population by lonesome highway. *Development* 2968: 2959–2968.

Ohta M, Takaiwa F. 2014. Emerging features of ER resident J-proteins in plants. *Plant Signalling and Behaviour* 9 e28194.

Okumura K, Goh T, Toyokura K, Kasahara H, Takebayashi Y. 2013. GNOM/FEWER ROOTS is required for the establishment of an auxin response maximum for arabidopsis lateral root intiation. *Plant and Cell Physiology* 54: 406–417.

Oliviusson P, Heinzerling O, Hillmer S, Hinz G, Tse C, Jiang L, Robinson DG. 2006. Plant retromer, localized to the prevacuolar compartment and microvesicles in arabidopsis , may interact with vacuolar sorting receptors. *The Plant Cell* 18: 1239–1252.

Osborne MA, Zenner G, Lubinus M, Zhang X, Songyang Z, Cantley LC, Majerus P, Burn P, Kochan JP. 1996. The Inositol 5'-Phosphatase SHIP binds to immunoreceptor signaling motifs and responds to high affinity ige receptor aggregation. *The Journal of Biological Chemistry* 271: 29271–29278.

Osterrieder A. 2012. Tales of tethers and tentacles: golgins in plants. *Journal of Microscopy* 247: 68–77.

Osterrieder A, Sparkes IA, Botchway SW, Ward A, Ketelaar T, Ruijter N De, Hawes C. 2017. Stacks off tracks : a role for the golgin AtCASP in plant endoplasmic reticulum-Golgi apparatus tethering. *Journal of Experimental Botany* 68: 3339–3350.

Palenik B, Grimwood J, Aerts A, Rouze P, Salamov A, Putnam N, Dupont C, Jorgensen R, Derelle E, Rombauts S. 2007. The tiny eukaryote *Ostreococcus* provides genomic insights into the paradox of plankton speciation. *Proceedings of the National Academy of Sciences of the United States of America* 104: 7705–7710.

Pantoja O, Gelli A, Blumwald E. 1992. Calcium channels. *Science* 31: 1567–1571.

Paris N, Neuhaus J-M. 2002. BP-80 as a vacuolar sorting receptor. *Plant Molecular Biology* 50: 903–914.

Paris N, Rogers SW, Jiang L, Kirsch T, Beevers L, Phillips TE, Rogers JC. 1997. Molecular cloning and further characterization of a probable plant vacuolar sorting receptor. *Plant Physiology* 115: 29–39.

Paris N, Stanley CM, Jones RL, Rogers JC. 1996. Plant cells contain two functionally distinct vacuolar compartments. *Cell* 85: 563–572.

Parizot B, Laplaze L, Ricaud L, Boucheron-dubuisson E, Bayle V, Bonke M, Smet I De, Poethig SR, Helariutta Y, Haseloff J. 2008. Diarch symmetry of the vascular bundle in *Arabidopsis* root encompasses the pericycle and is reflected in distich lateral root initiation. *Plant Physiology* 146: 140–148.

Park Y, Lee H, Ha J, Kim JY, Park C, Park C. 2017. COP1 conveys warm temperature information to hypocotyl thermomorphogenesis. *New Phytologist* 3: 269–280.

Park M, Song K, Reichardt I, Kim H, Mayer U, Stierhof Y. 2013. *Arabidopsis* μ-adaptin subunit AP1M of adaptor protein complex 1 mediates late secretory and vacuolar traffic and is required for growth. *Proceeedings of the National Academy of Sciences of the United States of America* 110: 1–6.

Parry G. 2015. The plant nuclear envelope and regulation of gene expression. *Journal of Experimental* Botany 66: 1673–1685.

Paul P, Simm S, Mirus O, Scharf K, Fragkostefanakis S, Schleiff E. 2014. The complexity of vesicle transport factors in plants examined by orthology search. *PLoS ONE* 9: e97745.

Pedrazzini E, Giovinazzo G, Bollini R, Ceriotti A, Vitale A. 1994. Binding of BiP to an assembly-defective protein in plant cells. *The Plant Journal* 5: 103–110.

Pereira-Leal JB, Seabra MC. 2000. The mammalian rab family of small gtpases : definition of family and subfamily sequence motifs suggests a mechanism for functional specificity in the ras superfamily. *Journal of Molecular Biology* 301: 1077–1087.

Pereira-Leal JB, Seabra MC. 2001. Evolution of the Rab family of small GTP-binding proteins. *Journal of Molecular Biology* 313: 889–901.

Pereira-Leal B, Strom M, Godfrey RF, Seabra MC. 2003. Structural determinants of Rab and Rab Escort Protein interaction: Rab family motifs define a conserved binding surface. *Biochemical and Biophysical Research Communications* 301: 92–97.

Péret B, De Rybel B, Casimiro I, Benková E, Swarup R, Laplaze L, Beeckman T, Bennett MJ. 2009. *Arabidopsis* lateral root development: an emerging story. *Trends in Plant Science* 14: 399–408.

Pierleoni A, Martelli PL, Fariselli P. 2006. BaCelLo: a balanced subcellular localization predictor. *Bioinformatics* 22: 408–416.

Piffanelli P, Ross JHE, Murphy DJ. 1998. Biogenesis and function of the lipidic structures of pollen grains. *Sexual Plant Reproduction* 11: 65–80.

Plant Illustrations. **2017**. Root Illustrations. Figshare.

Pompa A, Vitale A. 2006. Retention of a bean phaseolin/maize gamma-Zein fusion in the endoplasmic reticulum depends on disulfide bond formation. *The Plant Cell* 18: 2608–21.

Poteryaev D, Datta S, Ackema K, Zerial M, Spang A. 2010. Identification of the switch in early-to-late endosome transition. *Cell* 141: 497–508.

Pourcel L, Irani NG, Lu Y, Riedl K, Schwartz S, Grotewold E. 2010. The formation of anthocyanic vacuolar inclusions in *Arabidopsis thaliana* and implications for the sequestration of anthocyanin pigments. *Molecular Plant* 3(1): 78-90.

Qi H, Xia F-N, Xie L-J, Yu L-J, Chen Q-F, Zhuang X-H, Wang Q, Li F, Jiang L, Xie Q. 2017. TRAF family proteins regulate autophagy dynamics by modulating autophagy protein6 stability in *Arabidopsis*. *The Plant Cell* 29: 890–911.

Rahman A, Bannigan A, Sulaman W, Pechter P, Blancaflor EB, Baskin TI. 2007. Auxin, actin and growth of the *Arabidopsis thaliana* primary root. *Plant Journal* 50: 514–528.

Rai A, Singh R, Shirke PA, Tripathi RD. 2015. Expression of rice CYP450-like gene (Os08g01480) in *Arabidopsis* modulates regulatory network leading to heavy metal and other abiotic tolerance. *PLoS ONE* 10(9): e0138574.

Rajagopala S V., Uetz P. 2009. Analysis of protein-protein interactions using array-based yeast two-hybrid screens. *In: Stagljar I. ed. Yeast Functional Genomics and Proteomics. London: Humana Press,* 223–245.

Rangarajan P, Donahue JL, Hess JE, Gillaspy GE. 2014. Regulation of sucrose non-fermenting related kinase 1 genes in *Arabidopsis thaliana*. *Frontiers in Plant Science* 5: 1–13.

Ranocha P, Martinez Y, Pfrunder S, Jin X, Renou J, Morreel K, Thibaud J, Ljung K, Fischer U, Martinoia E, 2013. *Arabidopsis* WAT1 is a vacuolar auxin transport facilitator required for auxin homoeostasis. *Nature Communications* 4: 1–9.

Ren C, Zhu X, Zhang P, Gong Q. 2016. *Arabidopsis* COP1-interacting protein 1 is a positive regulator of ABA response. *Biochemical and Biophysical Research Communications* 477: 847–853.

Rensing SA, Lang D, Zimmer AD. 2008. Conquest of land by plants. *Science* 319: 64–69.

Richardson LGL, Howard ASM, Khuu N, Gidda SK, Mccartney A, Morphy BJ, Mullen RT. 2011. Protein– protein interaction network and subcellular localization of the *Arabidopsis thaliana* ESCRT machinery. *Frontiers in Plant Science* 2: 1–14.

Richter S, Voß U, Jürgens G. 2009. Post-Golgi traffic in plants. *Traffic* 10: 819–828.

Rink J, Ghigo E, Kalaidzidis Y, Zerial M. 2005. Rab conversion as a mechanism of progression from early to late endosomes. *Cell* 122: 735–749.

Ritzenthaler C, Nebenführ A, Movafeghi A, Stussi-Garaud C, Behnia L, Pimpl P, Staehelin LA, Robinson DG. 2002. Reevaluation of the effects of brefeldin A on plant cells using tobacco Bright Yellow 2 cells expressing Golgi-targeted green fluorescent protein and COPI antisera. *The Plant Cell* 14: 237–261.

Robinson DG, Ding Y, Jiang L. 2015. Unconventional protein secretion in plants: a critical assessment. *Protoplasma* 253(1): 31–43.

Robinson DG, Jiang L, Schumacher K. 2008. The endosomal system of plants: charting new and familiar territories. *Plant Physiology* 147: 1482–1492.

Robinson NJ, Wilson JR, Turner JS. 1996. *Arabidopsis thaliana* in Zn2 + -metallothionein-deficient Synechococcus PCC 7942 : putative role for MT2 in Zn 2 + metabolism. *Plant Molecular Biology* 30: 1169–1179.

Rodriguez RE, Ercoli MF, Debernardi M, Break W, Mecchia MA, Sabatini M, Cools T, Veylder L De, Benfey PN. 2015. MicroRNA miR396 regulates the switch between stem cells and transit-amplifying cells in *Arabidopsis* roots. *The Plant Cell* 27: 3354–3366.

Rodriguez RE, Mecchia MA, Debernardi JM, Schommer C, Weigel D, Palatnik JF. 2010. Control of cell proliferation in *Arabidopsis thaliana* by microRNA miR396. *Devlopment* 137(1): 103–112.

Rojo E, Denecke J. 2008. What is moving in the secretory pathway of plants? *Plant Physiology* 147: 1493–503.

Rojo E, Gillmor CS, Kovaleva V, Somerville CR, Raikhel N V. 2001. VACUOLELESS1 is an essential gene required for vacuole formation and morphogenesis in *Arabidopsis*. *Developmental Cell* 1: 303–310.

Rojo E, Zouhar J, Kovaleva V, Hong S, Raikhel N V. 2003. The AtC-VPS protein complex is localized t the tonoplast and the prevacuolar compartment in *Arabidopsis*. *Molecular Biology of the Cell* 14: 2372–2384.

Rosso MG, Li Y, Strizhov N, Reiss B, Dekker K, Weisshaar B. 2003. An *Arabidopsis thaliana* T-DNA mutagenized population (GABI-Kat) for flanking sequence tag-based reverse genetics. *Plant Molecular Biology* 53: 247–259.

Roth J, Zuber C. 2017. Quality control of glycoprotein folding and ERAD: the role of N - glycan handling, EDEM1 and OS - 9. *Histochemistry and Cell Biology* 147: 269–284.

Rout MP, Aitchison JD, Suprapto A, Hjertaas K, Zhao Y, Chait BT. 2000. The yeast nuclear pore complex. *The Journal of Cell Biology* 148: 635–652.

Roux M, Schwessinger B, Albrecht C, Chinchilla D, Jones A, Holton N, Malinovsky FG, To M, Vries S De, Zipfel C. 2011. The *Arabidopsis* leucine-rich repeat receptor – like kinases BAK1/SERK3 and BKK1/SERK4 are required for innate immunity to hemibiotrophic and biotrophic pathogens. *The Plant Cell* 23: 2440–2455.

Rowe JH, Topping JF, Liu J, Lindsey K. 2016. Abscisic acid regulates root growth under osmotic stress conditions via an interacting hormonal network with cytokinin, ethylene and auxin. *New Phytologist* 211: 225–239.

Ruddock LW, Molinari M. 2006. N-glycan processing in ER quality control. *Journal of Cell Science* 119(21): 4373-4380.

Rutherford S, Moore I. 2002. The *Arabidopsis* rab GTPase family: Another enigma variation. *Current Opinion in Plant Biology* 5: 518–528.

Ryabovol V V, Minibayeva F V. 2016. Molecular mechanisms of autophagy in plants: Role of ATG8 proteins in formation and functioning of autophagosomes. *Biochemistry (Moscow)* 81: 348–363.

Saint-Jore-Dupas C, Nebenfuhr A, Boulaflous A, Follet-Gueye M-L, Plasson C, Hawes C, Driouich A, Faye L, Gomord V. 2006. Plant N-Glycan processing enzymes employ different targeting mechanisms for their spatial arrangement along the secretory pathway. *The Plant Cell* 18: 3182–3200.

Saitou N, Nei M. 1987. The neighbor-joining method : a new method for reconstructing phylogenetic trees. *Molecular Biology Evolution* 4: 406–425.

Samuels AL, Meehl J, Lipe M, Staehelin LA. 1998. Optimizing conditions for tobacco BY-2 cell cycle synchronization *Rapid Communication*. 202(3-4): 232–236.

Sanderfoot AA. 1999. The specificity of vesicle trafficking: coat proteins and SNAREs. *The Plant Cell* 11: 629–642.

Sanderfoot AA, Assaad FF, Raikhel N V. 2000. The *Arabidopsis* genome. An abundance of soluble n-ethylmaleimide-sensitive factor adaptor protein receptors. *Plant Physiology* 124: 1558–1569.

Scheuring D, Löfke C, Krüger F, Kittelmann M, Eisa A, Hughes L, Smith RS, Hawes CR, Schumacher K, Kleine-vehn J. 2015. Actin-dependent vacuolar occupancy of the cell determines auxin-induced growth repression. *Proceedings of the National Academy of Sciences of the United States of America* 113: 452–457.

Scheuring D, Viotti C, Krüger F, Künzl F, Sturm S, Bubeck J, Hillmer S, Frigerio L, Robinson DG, Pimpl P. 2011. Multivesicular bodies mature from the trans-golgi network/early endosome in *Arabidopsis*. *The Plant Cell* 23: 1–20.

Schmid M, Davison TS, Henz SR, Pape UJ, Demar M, Vingron M, Weigel D, Lohmann JU. 2005. A gene expression map of *Arabidopsis thaliana* development. *Nature Genetics* 37: 501–506.

Schumacher K. 2014. pH in the plant endomembrane system-an import and export business. *Current Opinion in Plant Biology* 22: 71–76.

Segev N. 2001. Ypt and rab GTPases: insight into functions through novel interactions. *Current Opinion in Cell Biology* 13: 500–511.

Segui-Simarro JM, Staehelin LA. 2006. Cell cycle-dependent changes in Golgi stacks, vacuoles, clathrin-coated vesicles and multivesicular bodies in meristematic cells of Arabidopsis thaliana: A quantitative and spatial analysis. *Planta* 223: 223–236.

Serebriiskii I, Estojak J, Berman M, Golemis EA. 2000. Approaches to detecting false positives in yeast. *BioTechniques* 28: 328–336.

Shibata Y, Voss C, Rist JM, Hu J, Rapoport TA, Prinz WA, Voeltz GK. 2008. The reticulon and Dp1/Yop1p proteins form immobile oligomers in the tubular endoplasmic reticulum. *Journal of Biological Chemistry* 283: 18892–18904.

Shirakawa M, Ueda H, Koumoto Y, Fuji K, Nishiyama C, Kohchi T, Hara-Nishimura I, Shimada T. 2014. CONTINUOUS VASCULAR RING (COV1) is a trans-Golgi network-localized membrane protein required for Golgi morphology and vacuolar protein sorting. *Plant and Cell Physiology* 55: 764–772.

da Silva Conceição, Marty-Mazars D, Bassham DC, Sanderfoot a a, Marty F, Raikhel N V. 1997. The syntaxin homolog AtPEP12p resides on a late post-Golgi compartment in plants. *The Plant Cell* 9: 571–82.

Simon SM, Blobel G. 1991. A protein-conducting channel in the endoplasmic reticulum. *Cell* 65: 371–380.

Singh J, Freeling M, Lisch D. 2008. A position effect on the heritability of epigenetic silencing. *PLoS Genetics* 4 (10): e1000216.

Singh MK, Kruger F, Beckmann H, Brumm S, Vermeer JEM, Munnik T, Mayer U, Stierhof YD, Grefen C, Schumacher K. 2014. Protein delivery to vacuole requires SAND protein-dependent Rab GTPase conversion for MVB-vacuole fusion. *Current Biology* 24: 1383–1389.

Smertenko AP, Piette B, Hussey PJ. 2011. The origin of phragmoplast asymmetry. *Current Biology* 21: 1924–1930.

Smet I De, Vanneste S, Inze D, Beeckman T. 2006. Lateral root initiation or the birth of a new meristem. *Plant Molecular Biology* 60: 871–887.

Smith S, Smet I De. 2012. Root system architecture: insights from *Arabidopsis* and cereal crops. *Philisophical Transactions of the Royal Society B* 367: 1441–1452.

Sollner T, Whiteheart S, Brunner M, Erdjument-Bromage H, Geromanos S, Tempst P, Rothman JE. 1993. SNAP receptors implicated in vesicle targetign and fusion. *Nature* 362: 318–324.

Soltis DE, Albert VA, Leebens-Mack J, Palmer JD, Wing RA, dePamphilis CW, Ma H, Carlson JE, Altman N, Kim S. 2008. The *Amborella* genome: an evolutionary reference for plant biology. *Genome Biology* 9: 402.

Soltis PS, Brockington SF, Yoo MJ, Piedrahita A, Latvis M, Moore MJ, Chanderbali AS, Soltis DE. 2009. Floral variation and floral genetics in basal angiosperms. *American Journal of Botany* 96: 110–128.

Sparkes IA, Brandizzi F. 2012. Fluorescent protein-based technologies: shedding new light on the plant endomembrane system. *The Plant Journal* 70: 96–107.

Sparkes IA, Graumann K, Martinire A, Schoberer J, Wang P, Osterrieder A. 2011. Bleach it, switch it, bounce it, pull it: Using lasers to reveal plant cell dynamics. *Journal of Experimental Botany* 62: 1–7.

Sparkes IA, Ketelaar T, Norbert CA, Ruijter D, Hawes C, Campus H, Lane G. 2009. Grab a Golgi : laser trapping of Golgi bodies reveals in vivo interactions with the endoplasmic reticulum. *Traffic* 10(5): 567–571.

Sprunck S, Rademacher S, Vogler F, Gheyselinck J, Grossniklaus U, Dresselhaus T. 2012. Egg cell-secreted ec1 triggers sperm cell activation during double fertilization. *Science* 338: 1093–1097.

Staehelin LA, Kang B-H. 2008. Nanoscale architecture of endoplasmic reticulum export sites and of Golgi membranes as determined by electron tomography. *Plant Physiology* 147: 1454–1468.

Stagg SM, LaPointe P, Balch WE. 2007. Structural design of cage and coat scaffolds that direct membrane traffic. *Current Opinion in Structural Biology* 17: 221–228.

Steffen JG, Kang IH, Macfarlane J, Drews GN. 2007. Identification of genes expressed in the *Arabidopsis* female gametophyte. *Plant Journal* 51: 281–292.

Sternberg SH, Redding S, Jinek M, Greene EC, Doudna JA. 2014. DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature* 507: 62–67.

Strachan SD, Hess FD. 1983. The biochemical mechanism of action of the dinitroaniline herbicide oryzalin. *Pesticide Biochemistry and Physiology* 20: 141–150.

Strasser R. 2016. Plant protein glycosylation. *Glycobiology* 26: 926–939.

Suen PK, Shen J, Sun SSM, Jiang L. 2010. Plant science expression and characterization of two functional vacuolar sorting receptor (VSR) proteins, BP-80 and AtVSR4 from culture media of transgenic tobacco BY-2 cells. *Plant Science* 179: 68–76.

Sun Y, Day RN, Periasamy A. 2011. Investigating protein-protein interactions in living cells using fluorescence lifetime imaging microscopy. *Nature Protocols* 6: 1324–1340.

Sun Y, Rombola C, Jyothikumar V, Periasamy A. 2013. Förster resonance energy transfer microscopy and spectroscopy for localizing protein – protein interactions in living cells. *Cytometry Part A Early View*: 1–14.

Sutton RB, Fasshauer D, Jahn R, Brunger AT. 1998. Crystal structure of a SNARE complex involved in synaptic exocytosis at 2.4A˚ resolution. *Nature* 395: 347–353.

Swanson SJ, Jones RL. 1996. Gibberellic acid induces vacuolar acidification in barley aleurone. *The Plant Cell* 8: 2211–2221.

Swanson WJ, Vacquier VD. 2002. The rapid evolution of reproductive proteins. 3: 137-144.

Synek L, Schlager N, Eliáš M, Quentin M, Hauser MT, Žárský V. 2006. AtEXO70A1, a member of a family of putative exocyst subunits specifically expanded in land plants, is important for polar growth and plant development. *Plant Journal* 48: 54–72.

Taiz L. **1992**. The plant vacuole. *Journal of Experimental Botany* 122: 113–122.

Takáč T, Pechan T, Richter H, Müller J, Eck C, Böhm N, Obert B, Ren H, Niehaus K, Šamaj J. 2011. Proteomics on brefeldin A-treated *Arabidopsis* roots reveals profilin 2 as a new protein involved in the cross-talk between vesicular trafficking and the actin cytoskeleton. *Journal of Proteome Research* 10: 488–501.

Tamura T, Akutsu T. 2007. Subcellular location prediction of proteins using support vector machines with alignment of block sequences utilizing amino acid composition. *BMC Bioinformatics* 14: 1–14.

Tan X, Feng Y, Liu Y, Bao Y. 2016. Mutations in exocyst complex subunit SEC6 gene impaired polar auxin transport and PIN protein recycling in *Arabidopsis* primary root. *Plant Science* 250: 97–104.

Tan X, Wei J, Li B, Wang M, Bao Y. 2017. AtVps11 is essential for vacuole biogenesis in embryo and participates in pollen tube growth in *Arabidopsis*. *Biochemical and Biophysical Research Communications* 491: 794–799.

Tanchak MA, Fowke LC. 1987. The morphology of multivesicular bodies in soybean protoplasts and their role in endocytosis. *Protoplasma* 138: 173–182.

Tanchak MA, Griffing LR, Mersey BG, Fowke LC. 1984. Endocytosis of cationized ferritin by coated vesicles of soybean protoplasts. *Planta* 162: 481–486.

Tanchak MA, Rennie PJ, Fowke LC. 1988. Ultrastructure of the partially coated reticulum and dictyosomes during endocytosis by soybean protoplasts. *Planta* 175: 433–441.

Tanz SK, Castleden I, Hooper CM, Vacher M, Small I, Millar HA. 2013. SUBA3: a database for integrating experimentation and prediction to define the SUBcellular location of proteins in *Arabidopsis. Nucleic Acids Research* 41: 1185–1191.

Taylor JS, Raes J. 2004. Duplication and divergence: the evolution of new genes and old ideas. *Annual Review of Genetics* 38: 615–643.

Teh OK, Shimono Y, Shirakawa M, Fukao Y, Tamura K, Shimada T, Hara-Nishimura I. 2013. The AP-1 Mu adaptin is required for KNOLLE localization at the cell plate to mediate cytokinesis in *Arabidopsis*. *Plant and Cell Physiology* 54: 838–847.

The Arabidopsis Genome Initiative. 2000. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana. Nature* 408: 796–815.

Thieme CJ, Rojas-triana M, Stecyk E, Schudoma C, Zhang W, Yang L, Miñambres M, Walther D, Schulze WX, Paz-ares J. 2015. Endogenous *Arabidopsis* messenger RNAs transported to distant tissues. *Nature Plants* 1: 1–8.

Thomas C. 2012. Bundling actin filaments from membranes: some novel players. *Frontiers in Plant Science* 3: 188.

Tian Q, Zhao Y, Liu C. 2012. Modified yeast-two-hybrid system to identify proteins interacting with the growth factor progranulin. *Journal of Visual Experiments* 59: 9–14.

Tominaga M, Ito K. 2015. The molecular mechanism and physiological role of cytoplasmic streaming. *Current Opinion in Plant Biology* 27: 104–110.

Toorn M aan den, Albrecht C, Vries S De. 2015. On the origin of SERKs: Bioinformatics analysis of the somatic embryogenesis receptor kinases. *Molecular Plant* 8: 762–782.

Toyooka K, Goto Y, Asatsuma S, Koizumi M, Mitsui T, Matsuoka K. 2009. A mobile secretory vesicle cluster involved in mass transport from the Golgi to the plant cell exterior. *The Plant Cell* 21: 1212–1229.

Tse YC, Mo B, Hillmer S, Zhao M, Lo SW, Robinson DG, Jiang L. 2004. Identification of multivesicular bodies as prevacuolar compartments in *Nicotiana tabacum* BY-2 cells. *The Plant Cell* 16: 672–693.

Tudge C. 2006. The Variety of Life. Oxford: Oxford University Press.

Twell D, Park SK, Hawkins TJ, Schubert D, Schmidt R, Smertenko A, Hussey PJ. 2002. MOR1/GEM1 has an essential role in the plant-specific cytokinetic phragmoplast. *Nature Cell Biology* 4: 711–4.

Uemura T, Ueda T. 2014. Plant vacuolar trafficking driven by RAB and SNARE proteins. *Current Opinion in Plant Biology* 22: 116–121.

Uemura T, Ueda T, Ohniwa RL, Nakano A, Takeyasu K. 2004. Systematic analysis of SNARE molecules in *Arabidopsis*: Dissection of the post-Golgi network in plant cells. *Cell Structure and Function* 65: 49–65.

Ulengin I, Park JJ, Lee TH. 2015. ER network formation and membrane fusion by atlastin1/SPG3A disease variants. *Molecular Biology of the Cell* 26: 1616–28.

Varnier AL, Mazeyrat-Gourbeyre F, Sangwan RS, Clément C. 2005. Programmed cell death progressively models the development of anther sporophytic tissues from the tapetum and is triggered in pollen grains during maturation. *Journal of Structural Biology* 152: 118–128.

Vernoud V, Horton AC, Yang Z, Nielsen E. 2003. Analysis of the small GTPase gene superfamily of *Arabidopsis*. *Plant Physiology* 131: 1191–1208.

Viotti C, Bubeck J, Stierhof Y-D, Krebs M, Langhans M, van den Berg W, van Dongen W, Richter S, Geldner N, Takano J. 2010. Endocytic and secretory traffic in *Arabidopsis* merge in the trans-Golgi network/early endosome, an independent and highly dynamic organelle. *The Plant Cell* 22: 1344–1357.

Viotti C, Krüger F, Krebs M, Neubert C, Fink F, Lupanga U, Scheuring D, Boutté Y, Frescatada-Rosa M, Wolfenstetter S, 2013. The endoplasmic reticulum is the main membrane source for biogenesis of the lytic vacuole in *Arabidopsis. The Plant Cell* 25: 3434–3449.

Vitale A, Bielli A, Ceriotti A. 1995. The binding protein associates with monomeric phaseolin. *Plant Physiology* 107: 1411–1418.

Vitale A, Boston RS. 2008. Endoplasmic reticulum quality control and the unfolded protein response: Insights from plants. *Traffic* 9: 1581–1588.

Vitale A, Denecke J. 1999. The endoplasmic reticulum-gateway of the secretory pathway. *Plant Cell* 11: 615–628.

Vitale A, Raikhel N V. 1999. What do proteins need to reach different vacuoles? *Trends in Plant Science* 4: 149–155.

Voeltz GK, Prinz WA, Shibata Y, Rist JM, Rapoport TA. 2006. A class of membrane proteins shaping the tubular endoplasmic reticulum. *Cell* 124(3): 573–586.

Voytas DF. 2013. Plant genome engineering with sequence-specific nucleases. *Annual Review of Plant Biology* 64: 327–350.

Vukasinovic N, Žárský V. 2016. Tethering complexes in the *Arabidopsis* endomembrane system. *Frontiers in Cell and Developmental Biology* 4: 1–13.

Wagner GJ, Wang E, Shepherd R. 2004. New approaches for studying and exploiting an old protuberance, the plant trichome. *Annals of Botany* 93: 3–11.

Wahid A, Gelani S, Ashraf M, Foolad MR. 2007. Heat tolerance in plants: An overview. *Environmental and Experimental Botany* 61: 199–223.

Wallrabe H, Periasamy A. 2005. Imaging protein molecules using FRET and FLIM microscopy. *Current Opinion in Biotechnology* 16: 19–27.

Wang J, Cai Y, Miao Y, Lam SK, Jiang L. 2009. Wortmannin induces homotypic fusion of plant prevacuolar compartments. *Journal of Experimental Botany* 60: 3075–3083.

Wang J, Ding Y, Wang J, Hillmer S, Miao Y, Lo SW, Wang X, Robinson DG, Jiang L. 2010. EXPO, an exocyst-positive organelle distinct from multivesicular endosomes and autophagosomes, mediates cytosol to cell wall exocytosis in Arabidopsis and tobacco cells. *The Plant Cell* 22: 4009–30.

Wang P, Hawkins TJ, Richardson C, Cummins I, Deeks MJ, Sparkes I, Hawes C, Hussey PJ. 2014b. The plant cytoskeleton, NET3C, and VAP27 mediate the link between the plasma membrane and endoplasmic reticulum. *Current Biology* 24: 1397–1405.

Wang H, Hsu Y, Guo C, Jane W, Wang H, Jiang L, Jauh G. 2017. VPS36-Dependent multivesicular bodies are critical for plasmamembrane protein turnover and vacuolar biogenesis. *Plant Physiology* 173: 566–581.

Wang P, Hussey PJ. 2015. Interactions between plant endomembrane systems and the actin cytoskeleton. *Frontiers in Plant Science* 6: 422.

Wang P, Hussey PJ. 2017. NETWORKED 3B: A novel protein in the actin cytoskeleton-endoplasmic reticulum interaction. *Journal of Experimental Botany* 68: 1441–1450.

Wang J, Li Y, Lo SW, Hillmer S, Sun SSM, Robinson DG, Jiang L. 2007. Protein mobilization in germinating mung bean seeds involves vacuolar sorting receptors and multivesicular bodies. *Plant Physiology* 143: 1628–1639.

Wang P, Richardson C, Hawes C, Hussey PJ. 2016a. *Arabidopsis* NAP1 regulates the formation of autophagosomes. *Current Biology* 26: 2060–2069.

Wang P, Richardson C, Hawkins TJ, Sparkes I, Hawes C, Hussey PJ. 2016b. Plant VAP27 proteins: Domain characterization, intracellular localization and role in plant development. *New Phytologist* 210: 1311–1326.

Wang YH, Wang YH. 2008. How effective is T-DNA insertional mutagenesis in *Arabidopsis*? *Journal of Biochemical Techniques* 1: 11–20.

Wang Z-P, Xing H-L, Dong L, Zhang H-Y, Han C-Y, Wang X-C, Chen Q-J. 2015. Egg cell-specific promoter-controlled CRISPR/Cas9 efficiently generates homozygous mutants for multiple target genes in *Arabidopsis* in a single generation. *Genome Biology* 16: 144.

Wang Y, Zhang W-Z, Song L-F, Zou J-J, Su Z, Wu W-H. 2008. Transcriptome analyses show changes in gene expression to accompany pollen germination. *Plant Physiology* 148: 1201–1211.

Von Wangenheim D, Fangerau J, Schmitz A, Smith RS, Leitte H, Stelzer EHK, Maizel A. 2016. Rules and self-organizing properties of post-embryonic plant organ cell division patterns. *Current Biology* 26: 439–449.

Watanabe E, Shimada T, Kuroyanagi M, Nishimura M, Hara-nishimura I. 2002. Calcium-mediated association of a putative vacuolar sorting receptor PV72 with a propeptide of 2S albumin. *The Journal of Biological Chemistry* 277: 8708–8715.

Waterhouse AM, Procter JB, Martin DMA, Clamp M, Barton GJ. 2009. Jalview Version 2 — a multiple sequence alignment editor and analysis workbench. *Bioinformatics* 25: 1189–1191.

Waters GM, Serafini T, Rothman JE. 1991. 'Coatomer': a cytosolic protein complex containing subunits of non-clathrin-coated golgi transport vesicles. *Nature* 349: 248–251.

Wathugala DL, Richards SA, Knight H, Knight MR. 2011. OsSFR6 is a functional rice orthologue of SENSITIVE TO FREEZING-6 and can act as a regulator of COR gene expression, osmotic stress and freezing tolerance in *Arabidopsis*. *New Phytologist* 191(4): 984–995.

Welply JK, Shenbagamurthi P, Lennarz WJ, Naider F. 1983. Substrate recognition by oligosaccharyltransferase. *Journal of Biological Chemistry* 258: 11856–11863.

Werner T, Schmulling T. 2009. Cytokinin action in plant development. *Current Opinion in Plant Biology* 12: 527–538.

Wildwater M, Campilho A, Perez-perez JM, Heidstra R, Blilou I, Korthout H, Chatterjee J, Mariconti L, Gruissem W, Scheres B. 2005. The RETINOBLASTOMA-RELATED gene regulates stem cell maintenance in *Arabidopsis* roots. *Cell* 123: 1337–1349.

Willis KJ, McElwain JC. 2002. The Evolution of Plants. Oxford: Oxford University Press.

Wilson SM, Bacic A. 2012. Preparation of plant cells for transmission electron microscopy to optimize immunogold labeling of carbohydrate and protein epitopes. *Nature Protocols* (9):1716-27.

Winter V, Hauser M. 2006. Exploring the ESCRTing machinery in eukaryotes. *Trends in Plant Science* 11: 115–123.

Winter D, Vinegar B, Nahal H, Ammar R, Wilson G V, Provart NJ. 2007. An electronic fluorescent pictograph browser for exploring and analyzing large-scale biological data sets. *PLoS ONE* 2(8): e718.

Witkos TM, Lowe M. 2016. The golgin family of coiled-coil tethering proteins. *Frontiers in Cell and Developmental Biology* 3: 1–9.

Witkos TM, Lowe M. 2017. Recognition and tethering of transport vesicles at the Golgi apparatus. *Current Opinion in Cell Biology* 47: 16–23.

Woodhouse FG, Goldstein RE. 2013. Cytoplasmic streaming in plant cells emerges naturally by microfilament self-organization. *Proceedings of the National Academy of Sciences of the United States of America* 110: 14132–7.

Woollard AA, Moore I. 2008. The functions of Rab GTPases in plant membrane traffic. *Current Opinion in Plant Biology* 11: 610–619.

Wright A V, Nun JK, Doudna JA. 2016. Biology and applications of CRISPR systems: Harnessing nature's toolbox for genome engineering. *Cell* 164: 29–44.

Xing H-L, Dong L, Wang Z-P, Zhang H-Y, Han C-Y, Liu B, Wang X-C, Chen Q-J. 2014. A CRISPR/Cas9 toolkit for multiplex genome editing in plants. *BMC Plant Biology* 14: 327.

Xu H, Jun Y, Thompson J, Yates J, Wickner W. 2010. HOPS prevents the disassembly of trans - SNARE complexes by Sec17p / Sec18p during membrane fusion. *The EMBO Journal* 29: 1948–1960.

Yamazaki M, Shimada T, Takahashi H, Tamura K, Kondo M, Nishimura M, Hara-nishimura I. 2017. *Arabidopsis* VPS35, a retromer component , is required for vacuolar protein sorting and involved in plant growth and leaf senescence. *Plant Cell and Physiology* 49: 142–156.

Yan A, Lennarz WJ. 2005. Unraveling the mechanism of protein N-glycosylation. *Journal of Biological Chemistry* 280: 3121–3124.

Yan Q, Wang J, Fu ZQ, Chen W. 2017. Endocytosis of AtRGS1 is regulated by the autophagy pathway after d-glucose stimulation. *Frontiers in Plant Science* 8: 1–11.

Yang JS, Lee SY, Gao M, Bourgoin S, Randazzo PA, Premont RT, Hsu VW. 2002. ARFGAP1 promotes the formation of COPI vesicles, suggesting function as a component of the coat. *Journal of Cell Biology* 159: 69–78.

Yi C, Deng XW. 2005. COP1 – from plant photomorphogenesis to mammalian tumorigenesis. *Trends in Cell Biology* 15: 9–16.

Yoo MJ, Chanderbali AS, Altman NS, Soltis PS, Soltis DE. 2010. Evolutionary trends in the floral transcriptome: Insights from one of the basalmost angiosperms, the water lily *Nuphar advena* (Nymphaeaceae). *Plant Journal* 64: 687–698.

Yoshida K, Ohnishi M, Fukao Y, Okazaki Y, Fujiwara M, Song C, Nakanishi Y, Saito K, Shimmen T, Suzaki T. 2013. Studies on vacuolar membrane microdomains isolated from *Arabidopsis* suspension-cultured cells: local distribution of vacuolar membrane proteins. *Plant and Cell Physiology* 54: 1571–1584.

Yoshimoto K. 2012. Beginning to understand autophagy, an intracellular self-degradation system in plants. *Plant and Cell Physiology* 53: 1355–1365.

Žárský V, Kulich I, Fendrych M, Pečenková T. 2013. Exocyst complexes multiple functions in plant cells secretory pathways. *Current Opinion in Plant Biology* 16: 726–733.

Zhang C, Hicks GR, Raikhel N V. 2014a. Plant vacuole morphology and vacuolar trafficking. *Frontiers in Plant Science* 5: 476.

Zhang D, Liu D, Lv X, Wang Y, Xun Z, Liu Z, Li F, Lu H. 2014b. The cysteine protease cep1, a key executor involved in tapetal programmed cell death, regulates pollen development in *Arabidopsis*. *The Plant Cell* 26: 2939–2961.

Zhang J, Teng C, Liang Y. 2011. Programmed cell death may act as a surveillance mechanism to safeguard male gametophyte development in *Arabidopsis*. *Protein and Cell* 2: 837–844.

Zhang X, Yazaki J, Sundaresan A, Cokus S, Chan SW, Chen H, Henderson IR, Shinn P, Pellegrini M, Jacobsen SE. 2006. Resource mapping and functional analysis of DNA methylation in *Arabidopsis*. *Cell* 126: 1189–1201.

Zhao Y, Sato Y, Isaji T, Fukuda T, Matsumoto A, Miyoshi E. 2008. Branched N-glycans regulate the biological functions of integrins and cadherins. *FEBS Journal* 275: 1939–1948.

Zheng J, Han SW, Rodriguez-welsh MF, Rojas-pierce M. 2014. Homotypic vacuole fusion requires VTI11 and is regulated by phosphoinositides. *Molecular Plant* 7: 1026–1040.

Zheng H, Pan X, Yuxia D, Wu H, Liu P, Li X. 2016. AtOPR3 specifically inhibits primary root growth in *Arabidopsis* under phosphate deficiency. *Scientific Reports* 6: 24778.

Zheng H, Staehelin LA. 2011. Protein storage vacuoles are transformed into lytic vacuoles in root meristematic cells of germinating seedlings by multiple, cell type-specific mechanisms. *Plant Physiology* 155: 2023–2035.

Zhou X, Graumann K, Meier I. 2015. The plant nuclear envelope as a multifunctional platform LINCed by SUN and KASH. *Journal of Experimental Botany* 66: 1649–1659.

Zhu J. 2002. Salt and drought stress signal transduction in plants. *Annual Review of Plant Biology* 53: 247–273.

Zhu J, Gong Z, Zhang C, Song C-P, Damsz B, Inan G, Koiwa H, Zhu J-K, Hasegawa PM, Bressan R a. 2002. OSM1/SYP61: a syntaxin protein in *Arabidopsis* controls abscisic acid-mediated and non-abscisic acid-mediated responses to abiotic stress. *The Plant Cell* 14: 3009–28.

Zhuang X, Cui Y, Gao C, Jiang L. 2015. Endocytic and autophagic pathways crosstalk in plants. *Current Opinion in Plant Biology* 28: 39–47.

Zhuang X, Pan K, Cui Y, Lin W, Gao C, Kang B. 2017. ATG9 regulates autophagosome progression from the endoplasmic reticulum in *Arabidopsis*. *Proceedings of the National Academy of Sciences of the United States of America*: E426–E435.

Zuccolo A, Bowers JE, Estill JC, Xiong Z, Luo M, Sebastian A, Goicoechea J, Collura K, Yu Y, Jiao Y. 2011. A physical map for the *Amborella trichopoda* genome sheds light on the evolution of angiosperm genome structure. *Genome Biology* 12: R48.

Zwiewka M, Feraru E, Möller B, Hwang I, Feraru MI, Kleine-vehn J, Weijers D, Friml J. 2011. The AP-3 adaptor complex is required for vacuolar function in *Arabidopsis*. *Cell Research* 21: 1711–1722.