Functional Analysis of *Arabidopsis* DOA10-like E3 Ligases

by

ROSS DAVID ETHERINGTON

A thesis submitted to
The University of Birmingham
for the degree of

DOCTOR OF PHILOSOPHY

School of Biosciences
College of Life and Environmental Sciences
University of Birmingham
January 2022

Abstract

The ubiquitin proteasome system (UPS) performs the bulk of targeted protein degradation within eukaryotic cells. As such it is crucial for regulating protein abundance, destroying harmful misfolded proteins and for facilitating responses to stimuli. UPS substrates are identified by E3 ligase enzymes, which catalyse their ubiquitination through recognition of degrons (sequences of amino acids encoding degradation signals).

ScDoa10 is an E3 ligase of the yeast endoplasmic reticulum-associated degradation (ERAD) system that has been reported to recognise acetylated N-termini as degrons, as part of the Ac/N-degron pathway. It is shown here that one *Arabidopsis thaliana* homolog of ScDoa10 (AtDOA10A) can functionally complement the yeast E3 both in the response to hygromycin and in the degradation of a substrate (AtSQE1), demonstrating conservation of function. A second homolog, AtDOA10B, which is specific to *Brassicaceae*, was unable to complement ScDoa10 despite sequence and structural similarities.

Despite evidence of functional conservation for AtDOA10A, multiomic analyses, including transcriptomics, quantitative proteomics and N-terminal (Nt-)acetylomics, of seedlings with reduced levels of AtDOA10s revealed no evidence of an accumulation of Nt-acetylated substrates, suggesting that an AtDOA10-mediated Ac/N-degron pathway is not a significant contributor to global protein degradation in *Arabidopsis*. Similarly, targeting of AtSQE1, the first identified likely substrate of AtDOA10A, was also independent of its Nt-acetylation status, although, interestingly, disruption of Nt-acetylation machinery did indirectly affect AtSQE1's stability. This supports other recent investigations which suggest that Nt-acetylation does not act as a universal degradation signal but also highlights the significant indirect impacts of this post-translational modification on proteostasis.

Acknowledgements

Table of Contents

List of Figures

List of Tables

List of Abbreviations

aa = Amino acid

ABA = Abscisic acid

Ac = Acetylation

AZC = Azetidine-2-carboxylic acid

CBB = Coomassie brilliant blue

CHX = Cycloheximide

CRISPR = Clustered regularly interspaced short palindromic repeats

DEG = Differentially expressed gene

DEP = Differentially expressed protein

DO = Synthetic drop-out

DSB = Double strand break

EBR = 24-epibrassinolide

(E)GFP / (E)YFP = (Enhanced) green / yellow fluorescent protein

ER = Endoplasmic reticulum

ERAD = Endoplasmic reticulum-associated degradation

GA = Gibberellic acid (gibberellin)

GUS = β-glucuronidase

HR = Homologous recombination non-homologous end joining (NHEJ)

HRM = High resolution melt

IP = Immunoprecipitation

Lad = DNA ladder (molecular weight size marker)

LB = Lysogeny broth

NAA = 1-Naphthaleneacetic acid

NAT = N-acetyltransferase

NHEJ = Non-homologous end joining

Nt = Amino-terminal

ORF = Open reading frame

PAC = Paclobutrazol

PAM = Protospacer adjacent motif

PCR = Polymerase chain reaction

qRT-PCR = Quantitative RT-PCR

RIPA = Radioimmunoprecipitation assay

RT-PCR = Reverse transcription PCR

SOB = Super optimal broth

SOC = Super optimal broth with catabolite repression

UPR = Unfolded protein response

Yeast = *Saccharomyces cerevisiae* (unless stated otherwise)

YPD = Yeast-extract peptone dextrose

Chapter I -

Background Literature Review

1.1 Protein turnover and the ubiquitin proteasome system (UPS)

Proteins are crucial components of cells which are responsible for the majority of cellular functions, including maintaining structure, transmitting signals and catalysing metabolic reactions. They are important for maintaining plant health and in the development of seeds and fruit and therefore control crop yields (Weber et al., 2005; Ruan et al., 2012). Tight regulation of the abundance of each individual protein within a cell is therefore essential. Improving our understanding of this regulation has the potential to aid the crop improvement efforts required to ensure global food security.

1.1.1 Protein degradation and proteostasis

Protein homeostasis (proteostasis) is regulated by the constant production, modification and destruction of cellular proteins. The abundance of individual proteins is dependent on a balance between their specific rates of translation and turnover. As such, protein turnover is a key mechanism of response to external stimuli and genetic developmental programs that can produce a more rapid response than alterations in gene expression (Hershko and Ciechanover, 1998). In eukaryotes, the ubiquitin proteasome system (UPS) is responsible for the bulk of intracellular selective protein degradation, with a huge number of diverse proteins targeted based on short amino acid sequences known as degrons (Ravid and Hochstrasser, 2008). Degrons are important determinants of the rate of a protein's degradation and can be encoded anywhere within the amino acid sequence. The nature and number of degrons within a protein sequence can significantly influence protein stability, with half-lives varying from minutes to years in some cases (Glickman and Ciechanover, 2002). For example, in humans the enzyme ornithine decarboxylase (HsODC) is turned over within a matter of minutes (Pegg, 2006), whereas structural proteins, such as nucleoporin 205 (HsNUP205), of the nuclear pore complex and crystallin proteins from the eye lens can remain stable for a number of years (Toyama and Hetzer, 2013).

1.1.2 Ubiquitination of proteins

Degradation of eukaryotic proteins via the UPS occurs through the addition of the small ubiquitous protein ubiquitin, in a process known as protein ubiquitination. Ubiquitinated proteins can be recognised and destroyed by the 26S proteasome (Ciechanover, 1994). Ubiquitin is a 76 amino acid regulatory protein that is highly conserved throughout eukaryotes both in terms of residue sequence and overall structure (Vijay-Kumar et al., 1987). Ubiquitin is encoded by twelve genes in the model plant *Arabidopsis thaliana* (hereafter *Arabidopsis*). Unusually, none of these genes code for monomeric ubiquitin proteins but rather they encode homo- or hetero-meric ubiquitin fusions (Callis, 2014). Five ubiquitin genes (*AtUBQ3*, *AtUBQ4*, *AtUBQ10*, *AtUBQ11*, *AtUBQ14*) encode homomeric ubiquitin fusions, where a polyubiquitin chain of identical monomers is translated, with the addition of 1 or 2 amino acids at the C-terminus. A further seven ubiquitin genes encode heteromeric ubiquitin fusions. Five such genes lead to the production of ubiquitin fused to one of two ribosomal proteins (*AtUBQ1*, *AtUBQ2*, *AtUBQ5*, *AtUBQ6*, *AtUBQ17*) and an additional two genes code for ubiquitin fused to ubiquitin-like RUB (Related to ubiquitin) proteins (*AtUBQ7*, *AtUBQ15*). The ubiquitin proteins translated are non-functional in their fused-state and therefore require cleavage by proteases known as deubiquitinases (DUBs) in order to function (Callis, 2014).

Ubiquitination of substrates predominantly occurs at lysine residues. The presence of multiple lysine residues within the amino acid sequence of substrates, and of ubiquitin itself, introduces further complexity by allowing the formation of a variety of ubiquitination patterns. For substrates with multiple exposed lysine residues, multi-monoubiquitination can occur, where ubiquitin monomers are covalently bonded to several sites. Additionally, the presence of seven lysine residues within the amino acid sequence of a ubiquitin monomer, along with its N-terminus, provides multiple additional sites for potential subsequent ubiquitination and the development of a ubiquitin chain (Akutsu et al., 2016). This is known as polyubiquitination (Figure 1). In *Arabidopsis*, six of these seven lysine residues have been observed to act as ubiquitination sites, allowing polyubiquitination chains to develop in a range of branching

topologies (Callis, 2014). Historically, a chain of four ubiquitin monomers or more, linked by lysine-48, was believed to be required in order to generate a degradation signal, but the 26S proteasome is now known to target a range of polyubiquitin chains (Clague and Urbe, 2010). Additionally, due to the complexity of ubiquitination patterns, plus the further complication of post-translational modifications of ubiquitin, ubiquitin tagging can also provide alternative (non-degradation) signals (Akutsu et al., 2016). Most commonly, non-degradative signals are produced by lysine-63 linked polyubiquitin chains. For example, the activation of mammalian nuclear factor κB (NF-κB) transcription factors is dependent on lysine-63 linked ubiquitination (Komander and Rape, 2012).

The targeted addition of ubiquitin to substrate proteins occurs via an enzymatic cascade where a ubiquitin monomer is first activated, through the addition of adenosine monophosphate (AMP) to its C-terminus. AMP is yielded from the energy-releasing hydrolysis of adenosine triphosphate (ATP). The adenylation of ubiquitin is catalysed by an E1 (ubiquitin-activating) enzyme, which subsequently binds to ubiquitin in place of the AMP (Schulman and Harper, 2009). Following this, the activated ubiquitin protein is transferred from an E1 to an E2 (ubiquitin-conjugating) enzyme. The E2 enzyme then combines with an appropriate E3 (ubiquitin-ligase) enzyme to form an E2-E3 complex which also binds to a protein substrate using a recognition domain of the E3 enzyme. By binding to both proteins simultaneously, the E3 catalyses the transfer of ubiquitin from the E2 enzyme to the substrate (Figure 2) (Ciechanover, 1994; Vierstra, 2009; Callis, 2014). The targeted aspect of the cascade, whereby only specific proteins are ubiquitinated, is dependent on the action of the E3 enzyme (Hershko and Ciechanover, 1998; Vierstra, 2009). Individual E3s target specific protein degrons, meaning that only certain proteins are recognised by a particular E3 and are therefore ubiquitinated.
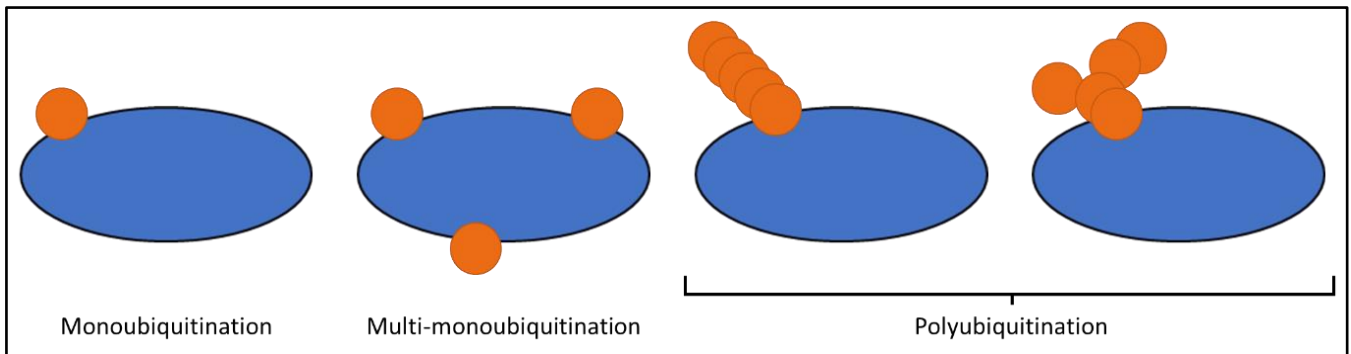
Figure 1 – Patterns of ubiquitination

Ubiquitin (orange) can be added to target proteins (dark blue) at a single site (monoubiquitination) or at multiple sites (multi-monoubiquitination). Ubiquitin itself can also be ubiquitinated, allowing the generation of polyubiquitin chains, which can be linear or branched depending on linkage type, e.g. lysine-48 and lysine-63 linkages. Based on Callis (2014).
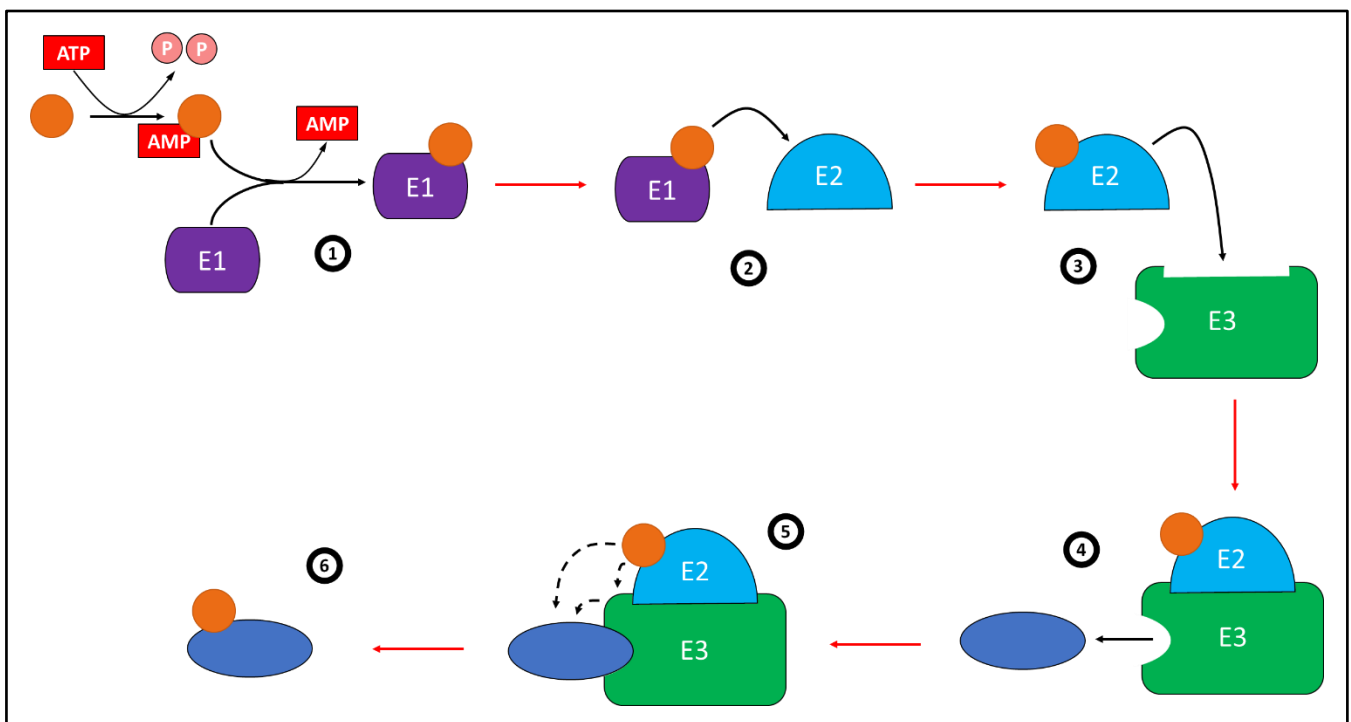
Figure 2 – The ubiquitination cascade

(1) Ubiquitin (orange) is activated through adenylation following ATP hydrolysis and is then bound by an E1 ubiquitin-activating enzyme (purple) which substitutes for the AMP (2) Activated ubiquitin in then transferred from the E1 to an E2 ubiquitin-conjugating enzyme (sky blue). (3) The E2 binds to an E3 ubiquitin-ligase enzyme (green) to form an E2-E3 complex. (4) Target proteins (dark blue) are recognised and bound by the E3 ligase. (5) Ubiquitin is added to the target protein either directly from the E2 or via the E3. (6) The ubiquitinated target protein can be potentially subject to further ubiquitination and degradation. Based on information from Callis (2014), Ciechanover (1994) and Vierstra (2009).

1.1.3 Ubiquitin E3 ligases

Due to the specific binding to only certain protein targets, E3s are by far the most diverse class of enzyme within the ubiquitin cascade. Whilst only 2 E1 and 37 E2 enzymes exist in *Arabidopsis*, more than 1500 E3s have been identified (Hua and Vierstra, 2011; Chen and Hellmann, 2013). This number greatly exceeds the number of E3 ligases present in other eukaryotes such as humans and yeast, showing heightened importance of the UPS in plants, relative to other species (Vierstra, 2009). The range of E3 ligases can be categorised into three main classes: RING, U box and HECT-type E3s, according to the E2-binding domain present within the E3. The structure of RING and U box domains are highly similar, with the only significant differences between the domains arising from how the structure is established. The binding of two zinc ions is used within RING domains to bring about this structure, whereas the same structure is formed in U box domains through the formation of non-covalent interactions between proximal amino acids (Ohi et al., 2003; Morreale and Walden, 2016). The concurrent binding of a RING or U box E3 to the E2 and substrate protein catalyses the direct transfer of ubiquitin from the E2 to the substrate (Pickart and Eddins, 2004; Ardley and Robinson, 2005; Berndsen and Wolberger, 2014). HECT domains not only contain an E2 binding site but also a catalytic cysteine residue, which is able to form a thioester bond with ubiquitin. Ubiquitination of substrates by HECT-type E3 ligases therefore occurs through the transfer of ubiquitin from the E2 to the E3, and then to the substrate (Pickart and Eddins, 2004; Ardley and Robinson, 2005; Berndsen and Wolberger, 2014). More recently an additional hybrid class of E3 ligases, known as RBR (RING-Between-RING) type, has also been identified. E3s of this type possess two RING domains, however, only one of these domains assumes the expected RING domain structure. Despite the presence of RING domains, substrate ubiquitination is achieved via a HECT-type E3-like mechanism, whereby an E2 binds to the canonical RING domain and ubiquitin is transferred to the substrate via a thioester bond with a cysteine residue within the non-canonical RING domain (Spratt et al., 2014; Morreale and Walden, 2016).

1.1.3.1 RING E3 ligases

The increased diversity of E3 ligases in plants in comparison to other eukaryotes is predominantly due to the expanded number of RING-type E3s (Chen and Hellmann, 2013). RING motifs were first identified by Freemont et al. (1991), as sequences that closely resembled those of DNA-binding zinc-finger motifs. Whilst RING domains do share a structural similarity to zinc-fingers, RING domains allow the binding of other proteins, rather than DNA. The RING structure is dependent on the specific spacing of eight conserved cysteine and histidine residues required for coordination to two Zn^{2+} ions. These amino acids are arranged in a cross-brace formation to hold the metal ions within the core of the domain, maintaining the overall structure (Deshaies and Joazeiro, 2009; Metzger et al., 2014). E3 ligases containing RING domains can exist both as monomers, where both the E2 and substrate bind to two domains within a single protein, and as multi-subunit complexes. These multi-subunit complexes are often referred to as CULLIN-RING E3s, as a shared feature of these complexes is the presence of a ~90 kDa CULLIN protein, which functions as the scaffold upon which the rest of the complex is built. The C-terminal region of the CULLIN protein allows the binding of a RING domain-containing protein, and the N-terminal region can bind to a range of receptor proteins (sometimes via an adaptor protein), which in turn recognise specific ubiquitination targets (Petroski and Deshaies, 2005; Hua and Vierstra, 2011). Ubiquitination is therefore carried out when an E2 binds to the RING domain and its associated ubiquitin is transferred to a receptor-bound substrate (Figure 3).

Arabidopsis possesses six CULLIN-like proteins, which all bind the RING-containing protein AtRBX1 but differ in terms of the adaptor or receptor proteins they bind to at their Nt-region. Due to the variability of the adaptor or receptor proteins within CULLIN-RING complexes, these E3s outnumber any other class of E3 ligase in *Arabidopsis* and have known roles in many key processes including hormone signalling, response to light, cell cycle regulation and stress responses (Hua and Vierstra, 2011; Choi et al., 2014). For example, the AtSCF[TIR1] complex functions as an auxin receptor by ubiquitinating AtAUX/IAA proteins in the

presence of auxin (Dharmasiri et al., 2005; Kepinski and Leyser, 2005) and AtSCF$^{ZTL/FKF1/LKP2}$ can target circadian rhythm-related proteins in response to the detection of blue light (Han et al., 2004; Takase et al., 2011). AtTIR1 and AtZTL1/FKF1/LKP2 are examples of F-box proteins that, via an adaptor protein (AtSKP1), can both bind to AtCULLIN1 (forming a SKP1-CULLIN-F-box (SCF) complex) highlighting how a diverse range of substrates can be targeted by AtCULLIN1 complexes alone.

Monomeric RING E3s also play important roles in many key processes in *Arabidopsis*. Ranging from nodule formation to gravitropism (Chen and Hellmann, 2013). A significant number of monomeric E3s have also been shown to be involved in stress responses. The AtRGLG1 and AtRGLG2 RING-E3s, for example, are known to negatively regulate drought and salinity responses by targeting the AtERF53 transcription factor (Cheng et al., 2012).
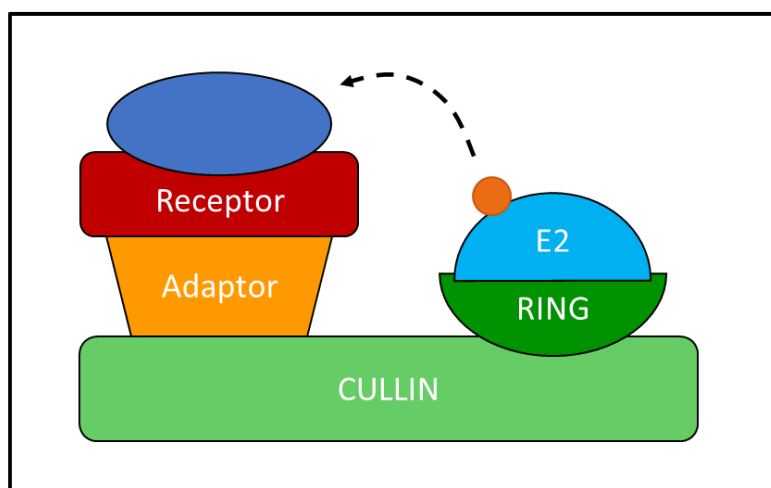
Figure 3 – Ubiquitination by multi-subunit E3 ligases (CULLIN-RING E3s)
The CULLIN protein scaffold allows binding of RING domain-containing proteins and substrate receptors often via interactions with adaptor proteins. Binding of an E2 and substrate protein (dark blue) to the RING and receptor proteins respectively catalyses the substrates ubiquitination. Based on Hua and Vierstra, (2011).

1.2 Endoplasmic Reticulum-Associated Degradation (ERAD) and the UPS

1.2.1 The role of ERAD and the Unfolded Protein Response (UPR)

A key function of the UPS is the degradation of aberrant proteins at the endoplasmic reticulum (ER). As the gateway to the secretory pathway, a third of all cellular proteins are folded and modified at the ER (Vembar and Brodsky, 2008; Strasser, 2018). Protein folding is a complex process involving the action of a range of ER-associated enzymes and, as such, routinely results in errors and the formation of misfolded proteins that require degradation in order to maintain homeostasis.

Nascent polypeptides are translocated into the ER via translocon complexes containing the Sec61 channel (Matlack et al., 1998). Upon entry to the ER, polypeptides are met by a suite of enzymes and chaperones such as BiP proteins, which maintain the polypeptide's foldable state, in order to facilitate subsequent folding (Liu and Li, 2014; Strasser, 2018). Correct folding is often also dependent on the appropriate attachment of oligosaccharides to amine groups within asparagine residues (N-linked glycosylation) (Liu and Li, 2014; Strasser, 2018). Precise quality control mechanisms exist within the ER to monitor the N-linked glycosylation and folding of proteins. Proteins recognised as terminally misfolded, as opposed to intermediately folded, are removed from the ER and degraded in the cytosol in a process known as endoplasmic reticulum-associated degradation (ERAD). Additionally, some correctly folded proteins can also undergo ERAD under certain conditions (Meusser et al., 2005). One such protein is 3-hydroxy-3-methylglutaryl-coenzyme A Reductase (HMGR), a rate limiting enzyme of the mevalonate pathway, which undergoes rapid ERAD in response to high levels of sterols (Jo and DeBose-Boyd, 2010).

ERAD occurs via the overarching mechanism of the UPS, but with the added condition that proteins need to be retrotranslocated from the ER into the cytoplasm for their ubiquitination and degradation by the 26S proteasome (Figure 4). The retrotranslocation process is only beginning to be understood (Wu et al., 2020), but is known to conclude with

the extraction of ubiquitinated proteins from the surface of the ER-membrane by the ATPase complex CDC48 (Liu and Li, 2014). The key mechanism by which misfolded proteins are recognised in order to be translocated also remains unclear, although the exposure of hydrophobic regions is thought to have some influence (Meusser et al., 2005; Vembar and Brodsky, 2008). Such a mechanism is required to be finely tuned, so that excess stringency (degradation of correctly folded proteins) and leniency (export of misfolded proteins) are avoided.

In *Arabidopsis*, environmental stresses such as heat, salinity and certain biotic stresses can significantly increase the rate of generation of misfolded proteins at the ER (Howell, 2013; Chen et al., 2020). If this increase exceeds the capacity for ERAD, the accumulation of misfolded proteins can trigger a cellular response known as the unfolded protein response (UPR). The UPR involves increasing the level of ERAD, inducing bulk degradation of proteins by autophagic or vacuolar degradation and, in a worst-case scenario, initiating cell death. The UPR has also been shown to slow the rate of translation in animals, but this feature of the UPR is yet to be fully confirmed in plants (Howell, 2013). These UPR responses are enacted via two signalling mechanisms involving members of the bZIP family of transcription factors. One mechanism involves the release of bZIP membrane proteins from the ER membrane and their transfer to the Golgi apparatus where they are proteolytically cleaved, allowing their cytoplasmic regions to move into the nucleus and alter gene expression (e.g. of ERAD genes). The alternative signalling mechanism also involves changes in gene regulation following the movement of a normally membrane-bound bZIP transcription factor into the nucleus, however this is brought about through the production of an alternatively spliced isoform rather than proteolytic cleavage (Howell, 2013; Liu and Howell, 2016).
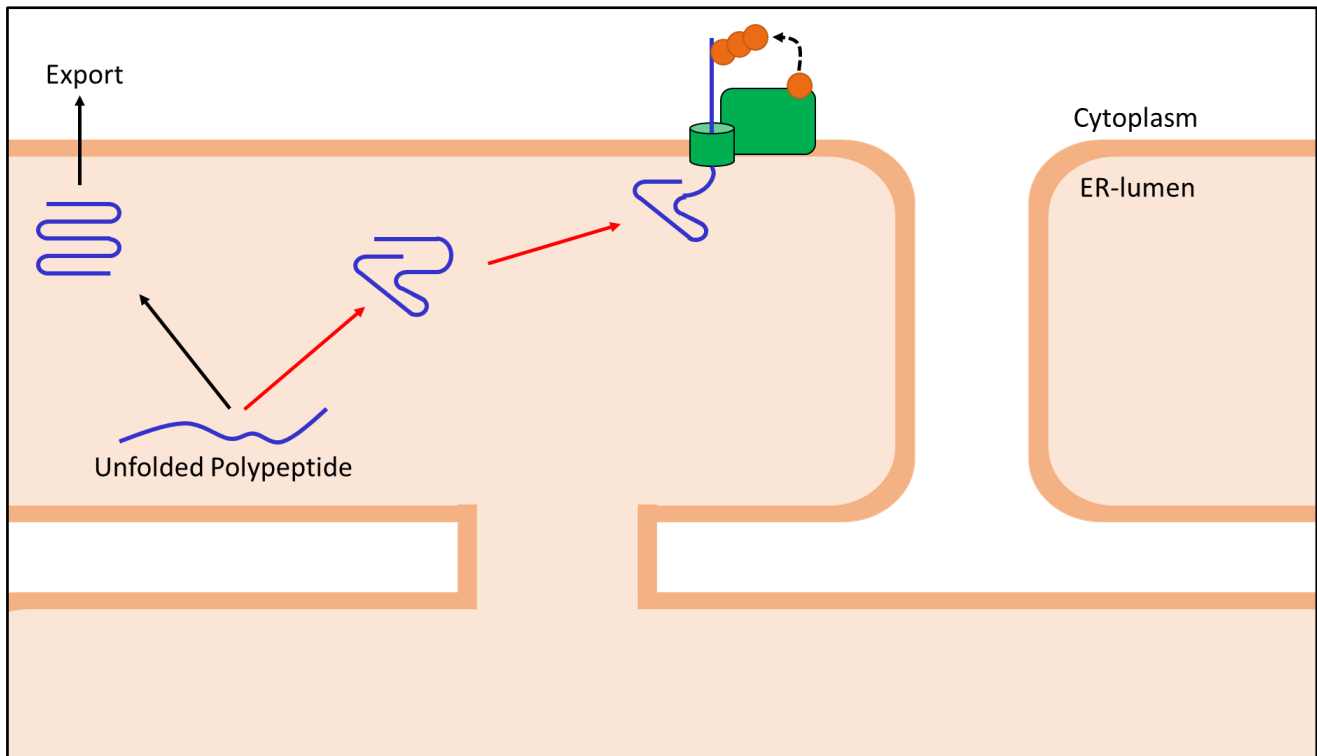
Figure 4 – Degradation of misfolded proteins by the ERAD system
Polypeptides are folded in the ER using a range of enzymes and chaperones. Correctly folded proteins are exported from the ER (solid black arrows). Misfolded proteins (solid red arrows) are retrotranslocated from within the ER-lumen, across the plasma membrane and into the cytoplasm, where they are ubiquitinated by membrane-bound E3 ligases (green). Based on Strasser (2018).

1.2.2 ERAD defects

Inefficiencies and errors within the ERAD process can have a significant impact on the health of a cell or entire organism. In general, ERAD defects are a result of either over-leniency or over-stringency, leading to the lack of degradation of aberrant proteins, or the premature degradation of functional proteins, respectively. The considerable effect of ERAD defects can be understood by considering the many human diseases that have been attributed to errors within the protein folding and quality control systems. ER stress, the generation of misfolded proteins, and formation of large protein aggregates, has been linked to both Parkinson's and Alzheimer's diseases, for example (Ogen-Shtern et al., 2016). Alternatively, cystic fibrosis is a disease arising from the degradation of functional HsCFTR chloride channel proteins. Mutations in the *HsCFTR* gene can lead to the generation of an

incorrectly folded, but still functional, channel protein. However, recognition of a folding error by the ER quality control systems results in ERAD of mutated HsCFTR proteins before they can be exported to epithelial cell membranes (Rowe et al., 2005). The lack of HsCFTR chloride channels results in membrane transports problems and the characteristic development of thickened mucus in the lungs.

1.2.3 ERAD E3 ligases

The specific E3 ligases involved in the ubiquitination of ERAD substrates have been most thoroughly studied in *Saccharomyces cerevisiae*, where two ER-localised E3s, HMG-coA Reductase (HMGR) Degradation 1 (ScHrd1) and Degradation of α2 10 (ScDoa10), have been identified (Liu and Li, 2014; Wu and Rapoport, 2018). Both E3s are membrane-bound but differ in regard to the substrates they target. ERAD can be divided into three pathways depending on whether the degron or lesion in the misfolded protein is located in the ER-lumen (ERAD-L), the ER-membrane (ERAD-M) or in the cytoplasm (ERAD-C). Substrates of the ERAD-L pathway are ubiquitinated by ScHrd1, whereas ScDoa10 is in involved in ERAD-C (Hirsch et al., 2009; Strasser, 2018; Wu and Rapoport, 2018). ERAD-M ubiquitination was previously thought to also be performed by ScHrd1, but a recent study has shown that ScDoa10 can target an ERAD-M substrate (Habeck et al., 2015), suggesting that this pathway may include some overlap.

ScHrd1 and ScDoa10 are both RING E3 ligases that work with the E2s ScUbc1/7 and ScUbc6/7 respectively (Hirsch et al., 2009). The RING-domains of both E3s are in the cytoplasm, and therefore, the ubiquitination of substrates can only occur on the cytosolic side of the ER-membrane. This means that proteins to be degraded by either ERAD-L or ERAD-M are required to be transported through the ER-membrane prior to ubiquitination. The identity of the proteins facilitating this retrotranslocation has been a subject of much recent debate (Roemisch, 2017) but recent evidence has identified ERAD E3 ligases themselves as possessing retrotranslocase activity (Schmidt et al., 2020; Wu et al., 2020). By forming a complex with several other proteins (ScHrd3, ScDer1, ScUsa1, and ScYos9), ScHrd1 is able

to form a membrane channel to permit the movement of its substrates across the ER membrane as well as ubiquitinating them (Strasser, 2018; Wu and Rapoport, 2018; Wu et al., 2020). The more recently discovered retrotranslocase activity of ScDoa10 has been shown to function, in combination with ScCdc48, in the extraction of the ScDoa10-associated E2, ScUbc6 (Schmidt et al., 2020). The ubiquitination of substrates is proposed to sterically block the return movement of the protein back into the ER-lumen prior to extraction from the ER by ScCdc48 (Wu and Rapoport, 2018). Ubiquitination of ERAD-C proteins by ScDoa10 does not require the transportation of substrates.

The *Arabidopsis* genome contains two known homologues of both ScHrd1 (AtHRD1A and AtHRD1B) and ScDoa10 (AtDOA10A and AtDOA10B) (Liu et al., 2011). In comparison to *S. cerevisiae*, the functions of these enzymes in *Arabidopsis* are poorly understood, although they are thought to be conserved (Liu and Li, 2014). *Athrd1* and *Atdoa10* null mutants only exhibit significant phenotypes in non-physiological conditions, i.e. when protein misfolding is likely to be increased. *Athrd1a Athrd1b* double mutants, for example, have been shown to have an increased sensitivity to salt and the toxic sulfate analogue selenate and also present with an improved tolerance to heat stress (Strasser, 2018). Additionally, the ER-localised brassinosteroid receptor AtBRI1 has been shown to be degraded by AtHRD1A/B when bearing C69Y or S622F mutations, further supporting their perceived role in ERAD (Su et al., 2011). AtBRI1 stabilisation is not observed in single *Athrd1* mutants, which also do not demonstrate the same phenotypes as *Athrd1a Athrd1b*, suggesting that AtHRD1A and AtHRD1B may function redundantly. Unlike the ScHrd1 homologues, there is little evidence of AtDOA10A and AtDOA10B working redundantly, with all relevant phenotypes being attributed solely to mutations in AtDOA10A (see 1.5.4.2). These phenotypes include grater thermotolerance and enhanced tolerance to drought stress (Strasser, 2018). AtDOA10A has also been shown to have roles in the synthesis of cuticle lipids and in ABA responses (Lu et al., 2012; Zhao et al., 2014).

1.3 N-terminal acetylation

The acetylation of a protein's N-terminus is a highly prevalent protein modification that can be performed, on 80% of proteins in both humans and plants (Arnesen et al., 2009; Bienvenut et al., 2012; Aksnes et al., 2016). The occurrence of N-terminal (Nt-) acetylation is dependent on the identity of the Nt-amino acids. In humans for example, acetylation occurs on 95% of Nt-alanines but only 50% of Nt-cysteines and 20% Nt-valines are acetylated (Aksnes et al., 2016).

1.3.1 Nt-acetyl transferases (NATs)

Nt-acetylation is carried out at by a group of enzyme complexes known as N-acetyltransferases (NATs), which bind to acetyl-CoA and catalyse the transfer of the acetyl moiety to the alpha amino group of a polypeptide (Varland et al., 2015). This reaction is presumed to be irreversible, since, as yet, no Nt-deacetylases have been discovered (Starheim et al., 2012; Aksnes et al., 2019). Several members of the NAT family have been identified across a range of organisms, including seven in plants (NatA-NatG) (Aksnes et al., 2016; Linster and Wirtz, 2018). Homologues of cytosolic NatA–NatE can be found in all eukaryotes; membrane-associated NatF exists in multicellular eukaryotes and chloroplast-localised NatGs are specific to plants (Giglione and Meinnel, 2021). The specific acetylation targets of each NAT complex are governed by the amino acids present at the N-terminus of a protein and are conserved in humans and yeast (Table 1). NatA is recognised as the complex required to acetylate Nt-Ala, -Val, -Ser, -Thr, -Cys and -Gly following Met cleavage, whereas NatB is known to N-terminally acetylate uncleaved Met, when followed by -Asn, -Gln, Asp or Glu residues. NatC, NatE and NatF are all thought to catalyse the acetylation of Met followed by hydrophobic or amphipathic residues such as -Leu, -Ile, -Phe, -Met and -Tyr, and NatG has been observed to predominantly acetylate Nt-Met, -Ala, - Ser and -Thr in the chloroplast. Acetylation by NatD appears to be much more specific, with its only known targets being the histone proteins H2A and H4 (Aksnes et al., 2016; Linster and Wirtz, 2018; Aksnes et al., 2019). Recently an additional cytosolic NAT (NatH) has been identified in animals that is

dedicated to the Nt-acetylation of actin (Drazic et al., 2018; Goris et al., 2018; Wiame et al., 2018). Homologues of NatH are not present in plants or yeast. The presence of a proline residue at either position one or two has been shown to prevent Nt-acetylation by any of the NAT complexes (Goetze et al., 2009). Similarly, in *S. cerevisiae*, proteins containing a lysine residue at position two are very rarely acetylated (Polevoda and Sherman, 2003; Arnesen et al., 2009). Several NATs (NatD-NatGs) are also known to possess some lysine acetylation capability (Bienvenut et al., 2020; Giglione and Meinnel, 2021).

Most Nt-acetylation is performed co-translationally by ribosome-associated NATs (NatA-NatE) (Aksnes et al., 2016; Linster and Wirtz, 2018). These NATs generally consist of a catalytic subunit and one or more auxiliary subunits which anchor the complex to the ribosome (Figure 5) (Polevoda et al., 2008; Aksnes et al., 2016). Two such NATs, NatA and NatB, are responsible for the Nt-acetylation of around 75 % of all proteins in humans and plants (Linster and Wirtz, 2018). NatA is made up of the catalytic subunit Naa10 and the ribosome-binding auxiliary subunit Naa15, to which the Huntingtin Interacting Protein K (HYPK) has also been shown to be associated in humans (Aksnes et al., 2019). Naa50, the catalytic subunit of NatE, shares these auxiliary subunits and has a regulatory influence on Naa10 as well as carrying out NatE-catalysed Nt-acetylation. In yeast, ScNaa50 no longer displays catalytic activity and instead solely regulates NatA activity (Deng et al., 2019). The NatB complex consists of the catalytic subunit Naa20 and the auxiliary subunit Naa25. NatC contains two auxiliary subunits Naa35 and Naa38, in addition to the catalytic subunit Naa30. To date, no auxiliary subunit of Naa40 (NatD) has been identified to facilitate its ribosome binding (Aksnes et al., 2019). Auxiliary subunits can also influence substrate specificity. For example, HsNaa10 has been reported to carry out post-translational Nt-acetylation in a monomeric, non-ribosomal, form where it exhibits altered substrate specificity (Van Damme et al., 2011). Other recent studies have shown that several additional NATs also function away from the ribosome, suggesting that post-translational Nt-acetylation occurs more regularly than previously assumed (Aksnes et al., 2015; Dinh et al., 2015; Drazic et al., 2018; Bienvenut

et al., 2020; Linster et al., 2020). The catalytic subunit of NatF, Naa60, performs post-translational Nt-acetylation as a membrane bound monomer. In humans, a HsNAA60 isoform has been shown to be bound to the Golgi apparatus, whereas an *Arabidopsis* isoform localises to the plasma membrane (Aksnes et al., 2015; Aksnes et al., 2017; Linster et al., 2020). AT2G39000 was formerly referred to as AtNAA70 (now GCN5-related N-acetyltransferase 4 (AtGNAT4)), the plant-specific catalytic subunit of chloroplastic NatG (Dinh et al., 2015). However, recent analysis has revealed at least five further plastidic enzymes with NAT capability known as AtGNAT2, 5, 6, 7 and 10 (Bienvenut et al., 2020). The term "NatG" now generally refers to the entire family of chloroplastic NATs. Other AtGNATs have also been discovered that function away from the chloroplast in plants (Bienvenut et al., 2020). The metazoan cytosolic Naa80, specific to the post-translational Nt-acetylation of actin, exists as a monomer (NatH) (Aksnes et al., 2019).

Table 1 – NAT complex substrate specificities

NAT complex	Organism	N-termini of substrates
NatA	All eukaryotes	A-, V-, S-, T-, C-, G-
NatB	All eukaryotes	MN-, ME-, MD-, MQ-
NatC	All eukaryotes	MΦ-
NatD	All eukaryotes	Histones H2A and H4
NatE	All eukaryotes	MΦ-
NatF	Multicellular eukaryotes	MΦ-
NatG	Plants	M-, A-, S-, T-
NatH	Metazoans	Actin

Φ = hydrophobic/amphipathic residue
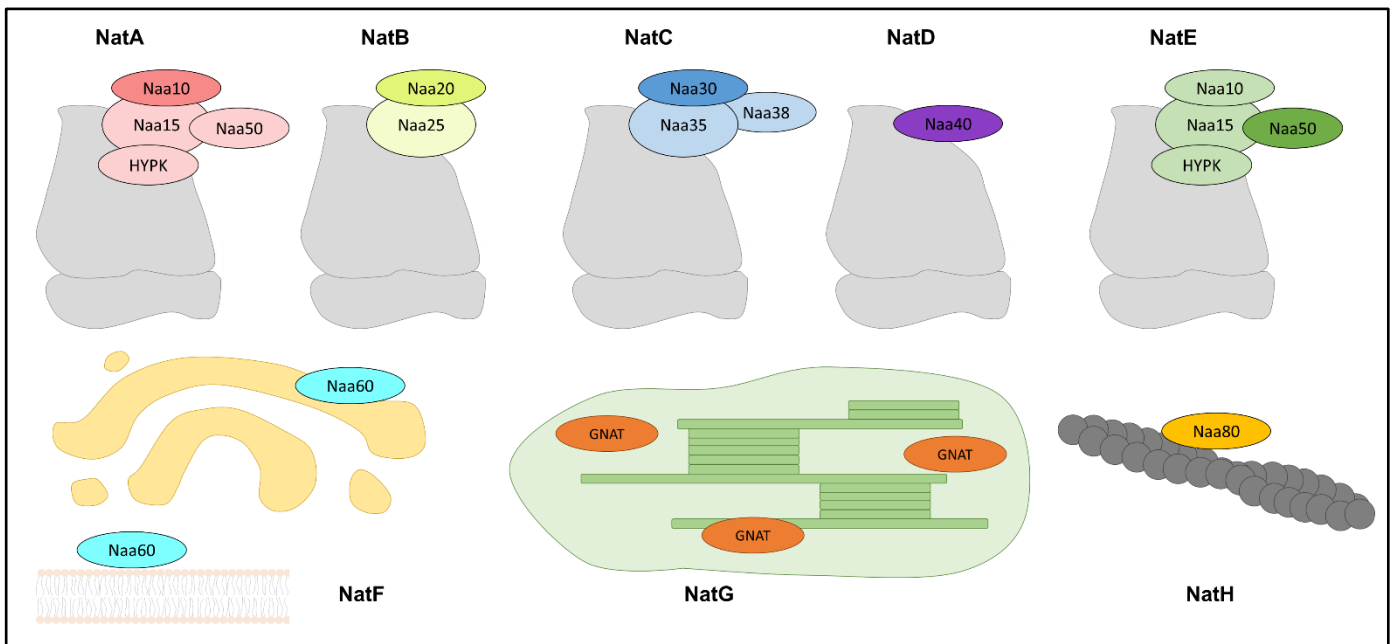
Figure 5 – Subunit composition of eukaryotic NAT complexes
Catalytic subunits of each NAT complex (e.g. Naa10 of NatA) are shown in dark shading and known auxiliary subunits (e.g. Naa15 of NatA) in pale shading. NatA-E complexes are ribosome associated, NatF is found at the Golgi in humans and the plasma membrane in plants. NatG refers to a family of chloroplast GCN5-related N-acetyltransferase (GNATs). Cytosolic NatH, acetylates actin and is specific to animals. Based on Aksnes et al. (2016).

1.3.2 Functions of Nt-acetylation

The addition of an acetyl group has the effect of increasing the size of N-termini, as well as removing their positive charge and increasing hydrophobicity. This modification can alter the fate of a protein in a number of ways, including affecting its interactions, localisation, folding and stability.

1.3.2.1 The effect of Nt-acetylation on protein interactions

At its simplest, the increased hydrophobicity brought about by Nt-acetylation can drive interactions between hydrophobic regions of proteins and lead to complex formation (Nguyen et al., 2018). An example of this is the Nt-acetylation of the human E2 enzymes HsUBC12 and HsUBC2F by NatC which allows the E2s' N-termini to dock into a hydrophobic binding pocket within the HsDCNL family of E3 ligases (Scott et al., 2011; Monda et al., 2013). Alternatively, Nt-acetylation can also block interactions with other enzymes that target N-termini for modification and therefore significantly influence the substrate protein's activity.

1.3.2.2 The effect of Nt-acetylation on protein localisation

In addition to interactions with other proteins, acetylation of N-termini can also change a protein's affinity for plasma membranes. The Parkinson's disease associated protein α-synuclein's membrane affinity is increased by Nt-acetylation which stabilises its Nt-α-helix (Dikiy and Eliezer, 2014). It has also been observed that Nt-acetylation of proteins can oppose their translocation into the ER (Forte et al., 2011). This suggests that either Nt-acetylation can prevent translocation into the ER or that binding of the signal recognition particle (SRP) can block Nt-acetylation.

1.3.2.3 The effect of Nt-acetylation on protein folding

Nt-acetylation is also implicated in the regulation of protein folding, with the loss of positive charge at the N-terminus thought to affect internal interactions within a polypeptide (Aksnes et al., 2016). Deletions of either NatA or NatB in yeast have been shown to lead to the accumulation of misfolded protein aggregates and the upregulation of chaperones

(Holmes et al., 2014; Friedrich et al., 2021). However, the reason for the observed increase in global protein aggregation is not fully understood, since proteins within aggregates are not restricted to those that are targeted by the deleted NAT (Friedrich et al., 2021).

1.3.2.4 The effect of Nt-acetylation on protein stability

In addition to influencing protein interactions, localisation, folding, and therefore activity, Nt-acetylation can also alter protein abundance by affecting their stability. Whilst initially considered a modification that increased protein stability, Nt-acetylation is now recognised as something that can both extend and reduce a protein's half-life (Aksnes et al., 2019). The *Drosophila melanogaster* protein DmHyx and human HsTHOC7 are two examples of proteins that show an increased longevity when Nt-acetylation is blocked by changing their second amino acid residue to an unacetylatable proline (Goetze et al., 2009; Myklebust et al., 2015). Nt-acetylation is also known to prevent the targeting of some proteins by degradative pathways that target the N-terminus. For example, in humans, the degradation of a regulator of G protein signalling 2 (HsRGS2) variant by the Arginylation (Arg)/N-degron pathway can be blocked by Nt-acetylation (Park et al., 2015). Conversely, acetylation of the N-terminus has also been shown to act as a signal for degradation by a separate branch of the N-degron pathway: the acetylation (Ac)/N degron pathway (Lee et al., 2016; Nguyen et al., 2018) (see 1.4). Indeed, whilst the unacetylated HsRGS2 isoform can be targeted for degradation by the Arg/N-degron pathway, Nt-acetylated HsRGS2 variant is susceptible to degradation via the Ac/N degron pathway (Park et al., 2015).

1.3.3 Physiological importance of Nt-acetylation

Given the impact of Nt-acetylation on a range of protein functions, it is unsurprising that perturbation of the NAT machinery can result in significant physiological effects including numerous consequences for human health. As previously mentioned, Nt-acetylation by NatB increases the affinity of α-synuclein for the plasma membrane (Dikiy and Eliezer, 2014). Nt-unacetylated α-synuclein is more prone to aggregate and form Lewy bodies, which are characteristic of Parkinson's disease (Bartels et al., 2014).

Mutations in HsNAA10 are also now known to underly numerous human diseases and developmental defects (Aksnes 2019). A Ser37Pro mutation is responsible for the most significant of these, Ogden syndrome, which results in severe delays in development and ultimately infant mortality (Rope et al., 2011). Silencing of *HsNAA30* has been shown to reduce the expression levels of a range of proteins, particularly those required for mitochondrial functioning (Van Damme et al., 2016). Nt-acetylation by NATs is also thought to have a significant role in the growth and regulation of cancer cells (Kalvik and Arnesen, 2013).

Disruption of Nt-acetylation also leads to a diverse range of defects in *Arabidopsis* (Gibbs, 2015). The impacts of NAT mutations are thought to be the result of large-scale proteome changes, as phenotypes are generally not-attributable to the lack of acetylation of individual NAT substrates (Linster and Wirtz, 2018). Null mutations (also referred to as gene knockouts: "KO") of NatA are embryo lethal in *Arabidopsis* due to disruptions in auxin signalling (Feng et al., 2016) and a decrease in the expression of either *AtNAA10* or *AtNAA15* leads to a significant reduction in growth (Linster et al., 2015). Further characterisation by Linster et al. (2015) revealed a role for NatA in the ABA-induced drought response, whereby ABA treatment reduced the expression level and abundance of NatA, causing a decrease in global protein acetylation and leading to characteristic drought responses, such as stomatal closure and changes to root morphology. Additionally, the NatA complex has been implicated in plant immune responses, as *Atnaa15* mutant plants have been shown to accumulate the Nod-like immunity receptor Suppressor of NPR1, Constitutive 1 (AtSNC1) and have an improved resistance to pathogens (Xu et al., 2015). Degradation of AtSNC1 therefore appears dependent on Nt-acetylation, suggesting that AtSNC1 may be degraded via the Ac/N-degron pathway, although this is yet to be confirmed. AtSNC1 bears a Met-Met-Asp N-terminus, and translation of AtSNC1 is able to initiate with either Met residue, yielding the alternative N-termini of Met-Met and Met-Asp (Xu et al., 2015). Due to the variation in the identity of the second residue, Nt-acetylation by the NatA complex only occurs on the Met-Met-AtSNC1 isoform. Met-Asp-AtSNC1 is Nt-acetylated by the NatB complex, which was unexpectedly

observed to lead to stabilisation of the protein, opposite to the effect of Nt-acetylation by NatA (Xu et al., 2015). This shows that Nt-acetylation can both increase and decrease protein half-lives depending on the context.

The substrate specificity of NatA in yeast and humans has been shown to be relatively conserved in *Arabidopsis* (Linster et al., 2015). NatA is therefore perceived to only target Met-cleaved N-termini (Aksnes et al., 2019) but the unexpected NatA-catalysed Nt-acetylation of Met-Met-AtSNC1 demonstrates that reported Nt-specificities of the NAT complexes are not definitive. Like NatA, NatE has also recently been shown to be involved in negatively regulating defence responses such as those triggered by pathogen attack (Neubauer and Innes, 2020). This similarity is perhaps not surprising, considering they share the same auxiliary subunits (see 1.3.1), although the loss of AtNAA50 has been shown not to affect NatA activity (Armbruster et al., 2020). *Atnaa50* KO plants are also not embryo lethal like *Atnaa10* but they are severely dwarfed and sterile (Armbruster et al., 2020; Feng, J. et al., 2020; Neubauer and Innes, 2020). By reducing (knocking down) *AtNAA50* expression using artificial micro RNAs, Neubauer and Innes (2020) were also able to identify a range of other processes AtNAA50 is involved in, including root hair development and vacuole maturation. Additionally, reduced expression of *AtNAA50* was observed to cause stress response proteins to accumulate and constitutively induce environmental and ER stress responses (Armbruster et al., 2020; Feng, J. et al., 2020; Neubauer and Innes, 2020).

Null mutations of the NatB complex are also thought to be lethal in *Arabidopsis*, since depletions of either AtNAA20 or AtNAA25 have a significant impact on normal growth and development (Huber et al., 2020). A range of pleiotropic phenotypes has been observed as a result of NatB depletion including defects in embryo, leaf and inflorescence development; earlier flowering time, a reduced tolerance of salt and osmotic stress and hypersensitivity to DTT due to constitutive over-reduction of the cytosol (Ferrandez-Ayela et al., 2013; Huber et al., 2020; Huber et al., 2022). Interestingly, NatB has also been linked to gene silencing, due to the generation of synergistic phenotypes when AtNAA25 and ARGONAUTE10 (AtAGO10)

of the microRNA pathway were mutated (Ferrandez-Ayela et al., 2013). Given the Met-Pro N-terminus of AtAGO10, it is not predicted that this link is a result of AtAGO10 being Nt-acetylation substrate of NatB. As is the case for NatA, substrate specificity of NatB also appears to be conserved in yeast, humans and plants (Huber et al., 2020). This is supported by recent reports showing that *Arabidopsis* NatB contributed to the Nt-acetylation of AtSIB1, a regulator of salicylic acid-induced cell death, and three AtACO enzymes, involved in ethylene biosynthesis, which all bear NatB-specific ME- N-termini (Huber et al., 2020; Li et al., 2020; Liu et al., 2021).

Other *Arabidopsis* NATs are less well studied. NatC null mutants are viable but display a reduced photosynthetic efficiency due to a decrease in the levels of the D1 and CP47 subunits of photosystem II (Pesaresi et al., 2003). As for the post-translational NATs, membrane-bound NatF is required for salt stress responses (Linster et al., 2020) and mutations of chloroplastic AtGNAT2, result in subtle photosynthetic defects (Koskela et al., 2018), although this defect may be caused by changes in Lys-acetylation activity, which AtGNATs also possess.

1.4 The N-degron pathways of targeted proteolysis

1.4.1 General principles of the N-degron pathways

A significant division of the ubiquitin proteasome system in eukaryotes is the N-degron pathways (formerly N-end rule pathways), in which proteins are targeted for ubiquitination and subsequent degradation according to a degron located at their N-terminus; an N-degron (Tasaki et al., 2012; Gibbs et al., 2014a). The specific class of monomeric RING E3 ligases responsible for the recognition and ubiquitination of N-degron-bearing proteins are known as N-recognins. The N-degron pathways are sub-divided into two major branches based on two different classes of N-degron: the so-called arginylation (Arg/N) and acetylation (Ac/N) branches of N-degron pathways (Bachmair et al., 1986; Hwang et al., 2010; Gibbs et al., 2014a; Lee et al., 2016). More recently, additional classes of N-degron have been observed leading to the identification of small additional subsets of the eukaryotic N-degron pathways: the proline (Pro/) N-degron pathway (Chen et al., 2017; Dougan and Varshavsky, 2018), the glycine (Gly/) N-degron pathway (Timms et al., 2019) and the formylmethionine (fMet/) N-degron pathway of plastid proteins (Kim et al., 2018; Kim, 2019), that also occurs in prokaryotes (Piatkov et al., 2015). Analogous pathways of degradation depending on the amino acid sequence of a protein's C-terminus has also been recently reported and termed C-degron pathways, emphasizing the importance of both protein termini in regulating their stability (Chatr-Aryamontri et al., 2018; Koren et al., 2018; Lin et al., 2018; Varshavsky, 2019; Timms and Koren, 2020).

This study focusses on the N-degron pathways in plants but significant aspects of the N-degron pathways in other eukaryotes will also be mentioned in brief. The general principles of the pathways are conserved within all eukaryotic organisms.

All proteins are synthesised with a methionine (Met) at the N-terminus (fomylmethionine (fMet) in prokaryotes and plastids). The majority of these Nt-Met residues are co- or post-translationally cleaved by peptidases, thereby producing variation amongst

cellular proteins according to their newly exposed Nt-residue (Varshavsky, 2011; Giglione et al., 2015). Following Met cleavage, the nature of the amino acids at the N-terminus determine the susceptibility of a protein molecule to degradation via the N-degron pathways (Tasaki et al., 2012; Gibbs et al., 2014a). Within an N-degron, it is thought to be the Nt-residue (i.e. very first residue) specifically that is the single most significant determinant of stability, with an additional key role of the second residue (Varshavsky, 2011; Kats et al., 2018). According to their chemical properties, Nt-amino acids can be categorised as either stabilising (maintaining a long protein half-life) or destabilising (permitting degradation via N-degron pathways).

In addition to the presence of an N-degron, to be a substrate of the N-degron pathways, a protein's N-terminus must also be accessible to an N-recognin, and additionally contain a suitable downstream ubiquitination site (lysine residue) (Tasaki et al., 2012; Gibbs et al., 2016). Proteins fitting these criteria can conditionally be subjected to degradation via the N-degron pathways.

1.4.2 The arginylation (Arg/N) branch of the N-degron pathways

The Arg/N-degron pathway involves N-recognins which target either basic or hydrophobic Nt-amino acids. Such amino acids are classed as type I (basic) or type II (hydrophobic) destabilising residues (Varshavsky, 2011; Tasaki et al., 2012). The arginylation branch of the N-degron pathways was first identified in the model eukaryote *S. cerevisiae* by Bachmair et al. (1986) following the observation that β-galactosidase proteins, expressed in the yeast were degraded at differential rates depending on the amino acid present at the N-terminus. Variations in Nt-residue were generated through the expression of ubiquitin-β-galactosidase fusion proteins, which exposed different Nt-amino acids following their constitutive deubiquitination (Figure 6). Exposure of an Nt-Arg was demonstrated to reduce the *in vivo* half-life of β-galactosidase more than any other amino acid (Figure 6) and hence targeted protein degradation of this type was initially referred to as the Arg/N-end rule (Bachmair et al., 1986). More recently, the pathway has been renamed the Arg/N-degron pathway to reflect the subsequent observations that any of the 20 amino acids can act as a

destabilising Nt-residue depending on the identity of the other amino acids close to the N-terminus (Varshavsky, 2019). Therefore, it is the nature of an Nt-degron, rather than the Nt-residue alone, that can govern a protein's stability, as is the case with internal and C-terminal degrons. However, it should be remembered that the identity of the Nt-residue is still believed to contribute the largest influence on the overall nature of the degron (Kats et al., 2018).

The presence of a destabilising Nt-residue as part of an N-degron, following Met excision, allows a protein to be recognised by an N-recognin. Unlike in yeast and animals, where type I and type II destabilising residues are recognised by the same N-recognins (Ubr1-type) at different binding sites (Kwon et al., 2001; Xia et al., 2008), in plants individual N-recognins target only one class of destabilising residue (Sriram et al., 2011). At present, two such Arg/N-recognins have been identified in plants: the E3 ligases PROTEOLYSIS1 and PROTEOLYSIS6 (AtPRT1 and AtPRT6), which target type II and type I destabilising residues respectively (Potuschak et al., 1998; Stary et al., 2003; Garzon et al., 2007). The binding of an appropriate E3 ligase to a substrate protein catalyses the ubiquitination of the substrate, hence tagging the protein for degradation via the 26S proteasome. Given the differential targeting of type II and type I substrates compared to other eukaryotes, it has recently been proposed that the overarching Arg/N-degron pathway nomenclature should be altered in plants, with degradation pathways instead referred to by the appropriate N-recognin (i.e. the AtPRT1/N-degron pathway and AtPRT6/N-degron pathway) (Holdsworth et al., 2019; Millar et al., 2019).

As well as differing in basicity and hydrophobicity, destabilising Nt-amino acids also vary in terms of how they are targeted by N-recognins. Some residues, such as histidine and tyrosine, are able to be recognised directly by N-recognins and are therefore known as primary destabilising residues (Dougan et al., 2012). Alternatively, other amino acids, can require additional chemical alterations in order to facilitate N-recognin binding. Residues of this type are known as secondary or tertiary destabilising residues, based on the fact that they require one or two further chemical modifications, respectively, before they will be targeted (Figure 7)

(Varshavsky, 2011; Gibbs et al., 2014a). The amide-containing amino acids asparagine (Asn) and glutamine (Gln) are both tertiary destabilising residues. These amino acids firstly undergo deamidation by Nt-amidases, which yields the secondary destabilising aspartic acid (Asp) and glutamic acid (Glu) respectively. A second modification then follows in which Asp or Glu N-termini are arginylated by arginyl-tRNA transferase (ATE) enzymes (Varshavsky, 2011; Tasaki et al., 2012; Gibbs et al., 2014a). The addition of arginine, a primary destabilising residue, to a protein's N-terminus allows recognition by an N-recognin, therefore leading to ubiquitination and degradation following two modifications to the original N-terminus.

An additional and unique tertiary destabilising amino acid in most eukaryotes is cysteine (Cys), which is exemplified in the degradation of the first identified substrates of the Arg/N-degron pathway in plants: the *Arabidopsis* group VII Ethylene Response Factors (AtERFVIIs) (Gibbs et al., 2011; Licausi et al., 2011). Following cleavage of Met by methionine amino peptidases (AtMetAPs), the AtERFVII transcription factors possess an Nt-Cys; which requires two additional modifications prior to recognition by the N-recognin AtPRT6. Initially, the exposed Nt-Cys is oxidised through the addition of two oxygen atoms by plant cysteine oxidases (AtPCOs) (Weits et al., 2014; White et al., 2017). This results in the formation of Cys sulfinic acid, a residue which bears a highly similar structure to the secondary destabilising residue Asp (White et al., 2017). Following oxidation, the N-terminus is then able to be arginylated by AtATEs (Gibbs et al., 2011; Licausi et al., 2011). These consecutive N-terminal alterations of the AtERFVIIs permit the binding of AtPRT6, leading to their ubiquitination and degradation (Figure 8). This degradation pathway of the AtERFVIIs has been shown to be the key mechanism of oxygen sensing used by plants (Gibbs et al., 2011; Licausi et al., 2011). The initial oxidation of the AtERFVIIs' tertiary destabilising Nt-Cys is dependent on atmospheric oxygen availability (Hu et al., 2005; Gibbs et al., 2014b; White et al., 2017). Hypoxic conditions, such as submergence, therefore prevent cysteine oxidation and hence stabilise the ERFVIIs. ERFVII stabilisation thereby signals a lack of available oxygen. In addition to oxygen, the oxidation of Nt-Cys also requires nitric oxide (NO). The specific role of

NO in the oxidation process remains unknown, but the Nt-Cys oxidation of the AtERFVIIs and of the best-studied Nt-Cys substrates in humans (the regulator of G protein signalling proteins HsRGS4, HsRGS5 and HsRGS16), cannot occur without both O_2 and NO (Hu et al., 2005; Gibbs et al., 2014b).

With many proteins possessing Met-Cys-N-termini, there is the potential that numerous other proteins may be subject to degradation via this oxygen and NO dependent-mechanism. Indeed, recently VERNALIZATION2 (AtVRN2), a subunit of the Polycomb Repressive Complex 2 (AtPRC2), and the LITTLE ZIPPER2 (AtZPR2) transcription factor have been shown to be a substrates of the PCO branch of the Arg/N-degron pathway in *Arabidopsis* (Gibbs et al., 2018; Weits et al., 2019). In addition to RGS proteins, in humans the cytokine interleukin-32 (HsIL-32) has also been recognised as a substrate of this degradation pathway, with 2-aminoethanethiol dioxygenase (HsADO) identified as the human enzyme required for Nt-Cys oxidation (Masson et al., 2019). The lack of a PCO homologue prevents the oxidation and hence degradation of Met-Cys-proteins in fungi and, as such, Cys is not considered a destabilising residue in yeast (Masson et al., 2019).

No obvious correlation exists between the recognised Arg/N-degron pathway substrates. Substrates vary significantly both in terms of their structures and functions, demonstrating the potential for a wide range of protein classes to be linked to the Arg/N-degron pathway (Gibbs et al., 2014a). The pathway has been shown to have roles in numerous cellular processes including the cell cycle, apoptosis, environmental sensing and signalling mechanisms and defects in the Arg/N-degron pathway can result in several human diseases (Varshavsky, 2011). Johanson-Blizzard syndrome is a human disease which is the result of mutations of the *HsUBR1* gene (a homologue of the *AtPRT6* E3 ligase) and causes several severe abnormalities such as congenital exocrine pancreatic insufficiency, physical deformities and intellectual disability (Zenker et al., 2005). Additionally, alterations in the Arg/N-degron pathway mediated degradation of specific proteins has been linked to breast cancer and Parkinson's disease (Xu et al., 2012; Yamano and Youle, 2013).

X	Degradation Time
Arginine	2 minutes
Phenylalanine	
Leucine	3 minutes
Aspartic acid	
Lysine	
Glutamine	10 minutes
Tryptophan	
Isoleucine	30 minutes
Glutamic acid	
Methionine	
Serine	
Alanine	>20 hours
Threonine	
Valine	
Glycine	

Met — Ubiquitin — Gly — X — β-galactosidase — Lys

Figure 6 – Discovery of the Arg/N-degron pathway

The degradation time of β-galactosidase was found to vary depending on the Nt-amino acid residue, with an Nt-Arg leading to the fastest degradation. Different N-termini were generated following the constitutive deubiquitination of ubiquitin-β-galactosidase fusion proteins (dashed red line) when expressed in yeast, leading to the discovery of the Arg/N-degron pathway. Based on Bachmair et al. (1986).
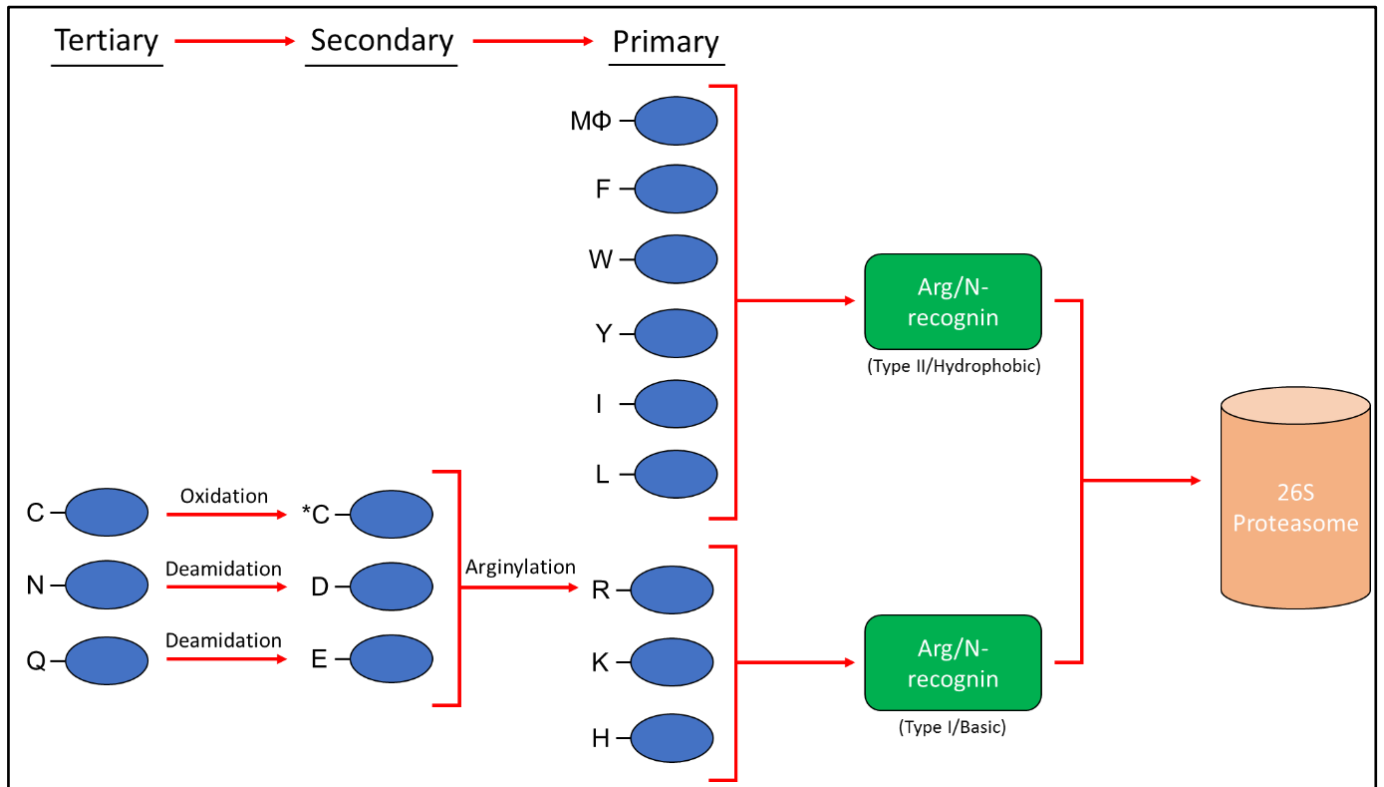
Figure 7 – The eukaryotic Arg/N-degron pathway
Following cleavage of methionine, destabilising Nt-residues follow different routes through the Arg/N-degron pathway depending on whether they are primary, secondary or tertiary destabilising residues and whether they are basic or hydrophobic. Basic and hydrophobic primary destabilising residues are recognised by different Arg/N-recognins in plants or by alternative domains of the same N-recognins in yeast and mammals. Methionine itself can also be targeted by Arg/N-recognins when followed by a hydrophobic residue (Φ) (see 1.4.4). Based on Gibbs et al. (2014a).
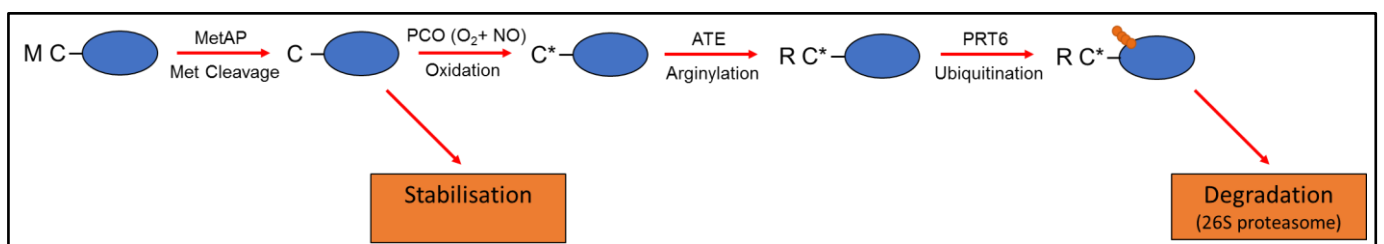
Figure 8 – ERFVII degradation via the Arg/N-degron pathway regulates oxygen sensing in plants
Following Met cleavage, the Nt-cysteine of AtERFVIIs can only be oxidised in the presence of oxygen and nitric oxide. When oxygen and nitric oxide are available, Nt-cysteine oxidation and arginylation allow ubiquitination, catalysed by the N-recognin AtPRT6, and therefore degradation via the 26S proteasome. In hypoxic conditions, Nt-cysteine oxidation does not take place, hence the AtERFVIIs are stabilised and hypoxia responses can be induced. Based on Gibbs et al. (2016).

1.4.3 The acetylation (Ac/N) branch of the N-degron pathways

Rather than targeting basic or hydrophobic Nt-amino acids like the Arg/N-degron pathway, substrates of the more recently discovered Ac/N-degron pathway are targeted because they bear acetylated N-termini (Varshavsky, 2011; Lee et al., 2016; Nguyen et al., 2018). In a similar manner to the Arg/N-degron pathway, proteins are targeted for acetylation and degradation via this pathway according to the identity of the amino acids at their N-terminus.

Proteins with appropriate Nt-amino acids can be Nt-acetylated by NATs either co- or post-translationally (see 1.3). The resulting steric and chemical alterations of the N-terminus, such as increased size and hydrophobicity, can allow binding of the appropriate N-recognin, resulting in a protein's ubiquitination and ultimately degradation via the Ac/N-degron pathway (Figure 9) (Aksnes et al., 2015). Indeed, high throughput screening of yeast peptides has recently suggested that proteins may be targeted for degradation according to the hydrophobicity of their N-terminus rather than as a direct result of Nt-acetylation (Kats et al., 2018). This recent work, combined with the historical theory that Nt-acetylation protected proteins from degradation by preventing ubiquitination of the N-terminus (Hershko et al., 1984), has cast doubt on the universality of Nt-acetylation acting as a broad degradation signal; degradation via the Ac/N-degron pathway is now considered more likely to occur only under specific circumstances (Eldeeb et al., 2019).

The Ac/N-degron pathway was initially discovered in *S. cerevisiae* during an investigation carried out by Hwang et al. (2010) originally intended to determine whether downstream residues could enable Nt-Cys to act as a destabilising residue in yeast, as had been observed in studies of the mammalian Arg/N-degron pathway (Tasaki and Kwon, 2007). During the investigation it was observed that protein constructs bearing Cys-Leu-N-termini were short-lived whereas those with Cys-Lys-N-termini were long-lived. Further analysis revealed that comparable half-lives were observed when the primary Cys residue was substituted for a range of other amino acids. Similar to Cys-Lys-, Met-Lys-proteins were stable

and, like Cys-Leu, Met-, Ala-, Val-, Ser- and Thr-Leu-proteins remained unstable. This instability was shown to be proteasome dependent and require the presence of the ScDoa10 E3 ligase (Hwang et al., 2010). Using liquid chromatography-tandem mass spectrometry, the unstable Met-Leu-protein was shown to be Nt-acetylated, as was ScMATα2, the canonical substrate of ScDoa10 (see 1.5). By studying a derivative of ScMATα2 (Met-Asn-α2), Hwang et al. (2010) demonstrated that ScDoa10 specifically targets the acetylated N-terminus of Met-Asn-α2. Unacetylated variants of Met-Asn-α2 accumulated even when a functioning ScDoa10 was present. Similar Nt-acetylation-dependent degradation patterns were also observed for several other proteins, including the telomere regulator ScTbf1, the chromosomal segregation regulator ScSlk19, the uncharacterised protein ScYmr090w and the histidine biosynthesis enzyme ScHis3 (Hwang et al., 2010).

Following the identification of ScDoa10 as the first N-recognin of the Ac/N-degron pathway, the cytosolic E3 ligase ScNot4 was also recognised as being able to fulfil this role in *S. cerevisiae* (Shemorry et al., 2013). ScNot4 was implicated in the UPS-mediated degradation of ScCog1, a subunit of the conserved oligomeric Golgi (COG) complex, following its Nt-acetylation by NatB. Additionally, the targeting of ScCog1 was shown to be stoichiometry-dependent, whereby degradation of ScCog1 was inhibited by the concurrent expression of other subunits of the COG complex, suggesting that interactions with binding partners helped to protect the protein from degradation by shielding its N-terminus from ScNot4 (Shemorry et al., 2013). Similarly, SpHcn1, a subunit of the *Schizosaccharomyces pombe* E3 ligase APC/C, also appears to be degraded by the Ac/N-degron pathway in the absence of its binding partner SpCut9 (Zhang et al., 2010; Shemorry et al., 2013). The control of subunit stoichiometries is now considered a potentially key role of the Ac/N-degron pathway, as is the regulation of protein quality control, with surplus subunits and misfolded proteins theoretically exposing their N-termini, hence making them susceptible to degradation (Figure 10) (Lee et al., 2016; Nguyen et al., 2018; Eldeeb et al., 2019). This hypothesis also correlates with the observation that 80% of proteins in higher eukaryotes are Nt-acetylated (Arnesen et

al., 2009; Bienvenut et al., 2012; Aksnes et al., 2016), but only a small proportion of these are degraded.

The Ac/N-degron pathway has now also been characterised in mammals, following the observation that the human HsRGS2 protein is a substrate of the Ac/N-degron pathway (as previously mentioned in 1.3.2.4). HsRGS2 is ubiquitinated by the human homologue of ScDoa10 HsMARCH6 (also known as HsTEB4) but this only occurs following the Nt-acetylation of HsRGS2 by NatB (Park et al., 2015). Interestingly, a Gln2Leu mutation of HsRGS2 permits degradation of the protein by both the Ac/N- and Arg/N-degron pathways (see 1.4.4). This mutation, and the similar Gln2Arg mutation, leads to a reduction in the half-life of HsRGS2 which is associated with hypertension (Yang et al., 2005). A further substrate of the Ac/N-end rule in humans is the lipid droplet binding protein perilipin 2 (HsPLIN2). Nt-acetylation of HsPLIN2 is performed by NatA, but shielding of the N-terminus through interactions with the phospholipid monolayer of lipid droplets is believed to stabilise the protein (Takahashi et al., 2016; Bersuker and Olzmann, 2017). However, the acetylated N-terminus of unbound HsPLIN2 can be recognised and ubiquitinated by HsMARCH6 and therefore degraded by the 26S proteasome (Nguyen et al., 2019). The absence of HsMARCH6 leads to an increased abundance of HsPLIN2 and has been shown to cause a significant increase in the number of lipid droplets within cells. This is associated with many human health conditions such as type 2 diabetes and obesity (Conte et al., 2016; Nguyen et al., 2019).

Despite the presence of NAT complexes and putative homologues of both yeast Ac/N-recognins (ScDoa10 and ScNot4), the lack of a proven substrate means that existence of the Ac/N-degron pathway is yet to be confirmed in plants (Gibbs et al., 2016). The stabilisation of AtSNC1 induced by NatB-catalysed acetylation and the recent discovery that AtSIB1 is also stabilised in response to Nt-acetylation demonstrate that the modification does not act as a definitive degradation signal in plants (Xu et al., 2015; Copeland, 2020; Li et al., 2020). However, the observation that Nt-acetylation by NatA induces the degradation of AtSNC1 in *Arabidopsis* (Xu et al., 2015) may suggest the functioning of a plant Ac/N-degron pathway

(see 1.3.3). Additionally, null mutations of AtDOA10A result in ABA hypersensitive and drought tolerant plants (Lu et al., 2012; Zhao et al., 2014), as is seen NatA depleted plants (Linster et al., 2015, see 1.3.3), which may indicate a role for the Ac/N-degron pathway in ABA signalling. Despite these hypothesised links, no confirmed role of the Ac/N-degron pathway in plants has been identified.
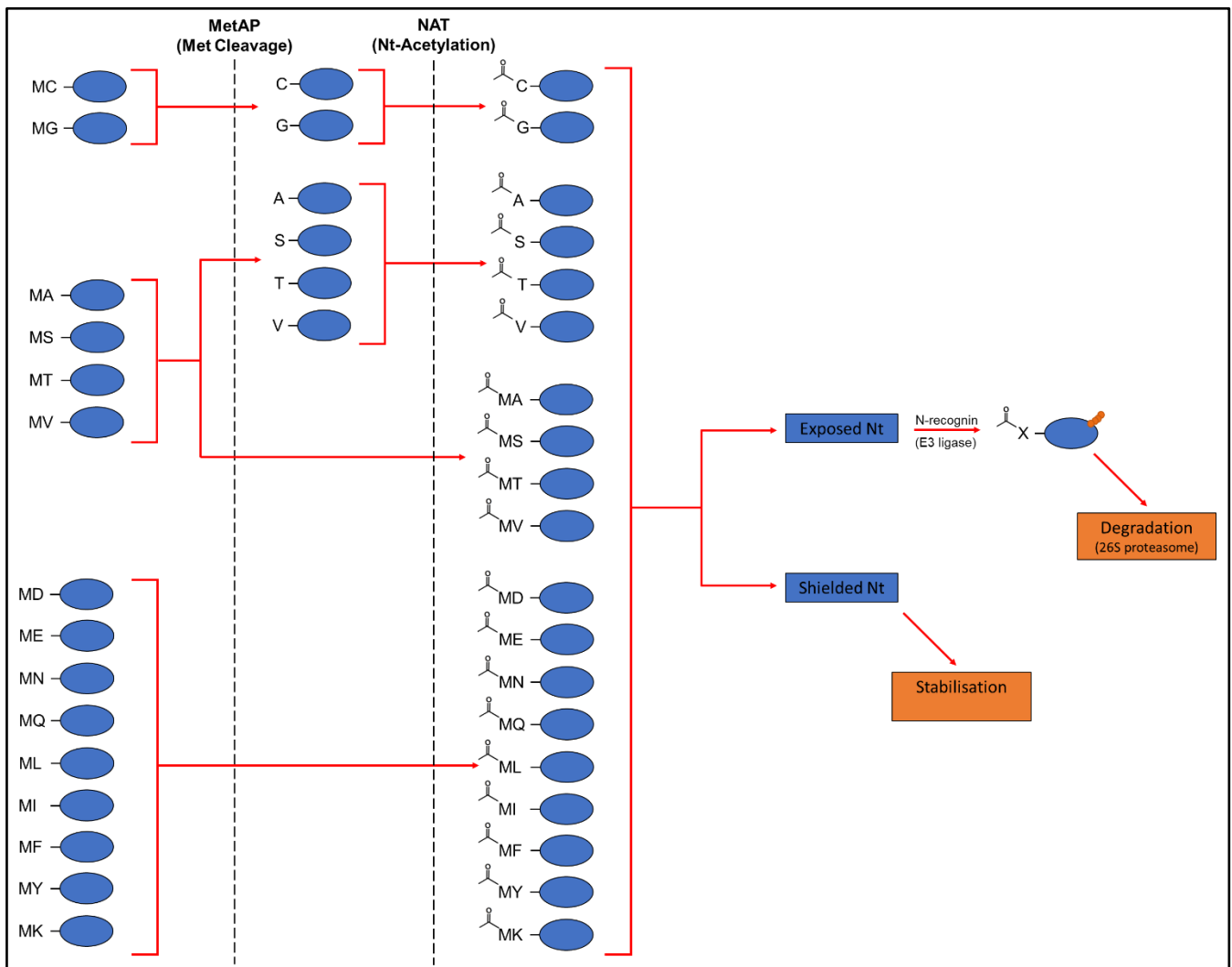
Figure 9 –The Eukaryotic Ac/N-degron pathway
NAT catalysed Nt-acetylation can occur on Nt-Met or following Met cleavage depending on the identity of the subsequent amino acid. The majority of Nt-amino acids can be acetylated. Proteins with acetylated N-termini can be targeted for ubiquitination by N-recognins only if the N-terminus is relatively unstructured and therefore accessible. Ubiquitinated proteins can then be degraded by the 26S proteasome. Shielded N-termini cannot be bound by N-recognins and such proteins are therefore stabilised. Based on Nt-acetylation and processing described in Aksnes et al. (2019).
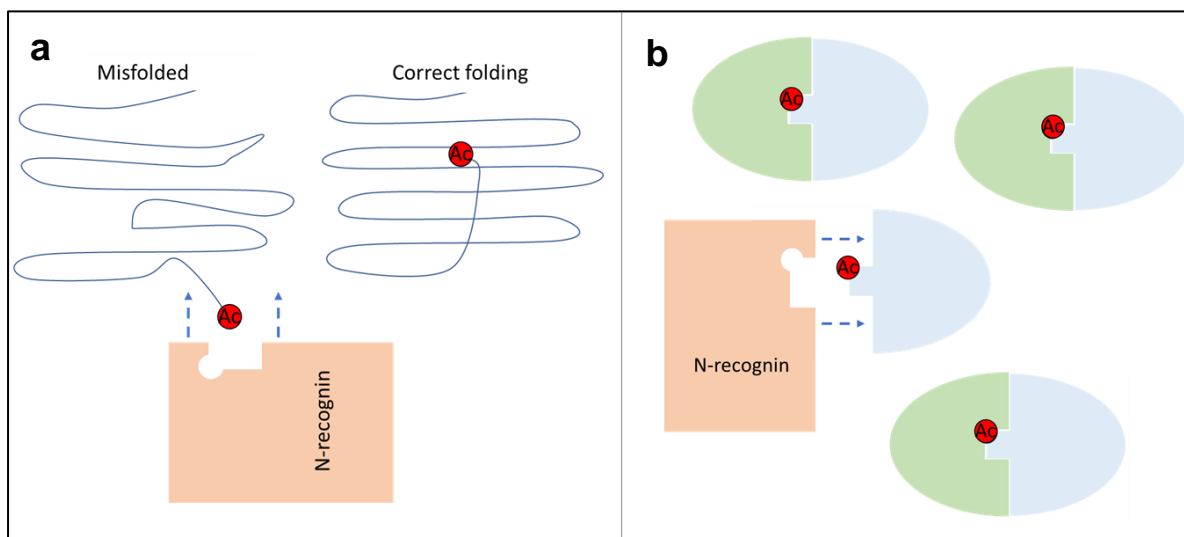
Figure 10 – The Ac/N-degron pathway as a regulator of protein quality and stoichiometry
(a) Misfolding of a protein may expose an acetylated N-terminus (red) making it accessible for an N-recognin. (b) The lack of a binding partner (green) may also expose the acetylated N-terminus of a protein subunit (blue), allowing the binding of an N-recognin.

1.4.4 Crosstalk between the Arg/N- and Ac/N-degron pathways

Since many proteins are only partially acetylated, the potential degradation pathways of acetylatable proteins that avoid acetylation was investigated by Kim et al. (2014). From this study, the existence of crosstalk between the Ac/N- and Arg/N-degron pathways was identified, whereby certain proteins are able to be degraded via either branch of the pathways, depending on their acetylation status (Kim and Hwang, 2014). It was found that in addition to basic (type I) and hydrophobic (type II) residues, N-recognins of the Arg/N-degron pathway were also capable of ubiquitinating proteins with uncleaved Nt-Met residues, provided that they were followed by a hydrophobic amino acid (Φ) (Kim et al., 2014). This targeting could be blocked through the NAT-catalysed acetylation of the Nt-Met. However, whilst Nt-acetylation of Met-Φ-N-termini can protect proteins from the Arg/N-degron pathway, it also potentially enables their degradation by the Ac/N-degron pathway (Kim et al., 2014). Therefore, the degradation of certain proteins is controlled by the combined action of both the Ac/N- and Arg/N-degron pathways.

Several Met-Φ-substrates have been shown to be degraded by this combined mechanism, such as the Gln2Leu mutant of human HsRGS2; rat RnAANAT proteins; and the yeast proteins ScMsn4 (a transcriptional activator), ScSry1 (a 3-hydroxyaspartate dehydratase enzyme), ScArl3 (a GTPase) and ScPre5 (a subunit of the 20S proteasome) (Kim et al., 2014; Park et al., 2015; Wadas et al., 2016). Additionally, the Arg/N-degron pathway was shown to degrade misfolded synthetic peptides bearing Met-Φ-N-termini, demonstrating that this branch of the pathways is also able to target misfolded proteins for degradation (Kim et al., 2014).

Similarly, Nt-acetylation has also been shown to protect artificial peptide reporters that contain specific Nt-sequences, from the Arg/N-degron pathway in yeast. Acetylation of Nt-Glu by NatA can prevent arginylation and acetylation of Met-Asp N-termini appears to preclude Met cleavage, blocking the Arg/N-degron pathway in both cases (Kats et al., 2018).

1.5 DOA10 E3 ligases

ScDoa10 is a yeast RING-E3 ligase, which has been shown to function as part of both the ERAD system and the Ac/N-degron pathway (see 1.2.3 and 1.4.3). It was first identified as an E3 ligase involved in the degradation of the yeast mating type protein ScMATα2 by targeting a degron contained within the first 67 amino acids of the protein (Swanson et al., 2001). The addition of these 67 amino acids to otherwise stable proteins was also able to induce their degradation by ScDoa10. ScDoa10's cognate E2 enzymes ScUbc6 and ScUbc7 were also observed to affect the degradation time of these proteins (Swanson et al., 2001).

1.5.1 The localisation and structure of ScDoa10

Subsequent characterisation of ScDoa10 revealed that it is localised to the endoplasmic reticulum and both the inner and outer membranes of the nuclear envelope (Swanson et al., 2001; Deng and Hochstrasser, 2006; Zattas and Hochstrasser, 2015). It contains 1319 amino acids (≈151 kDa) which make up 14 transmembrane domains, with both the N- and C-termini being cytosolic (Figure 11) (Kreft et al., 2006). The E2-binding RING domain of ScDoa10 resides close the N-terminus, before the first transmembrane domain, and therefore on the cytosolic side of the ER-membrane (Kreft et al., 2006). Another feature of ScDoa10 is the conserved ≈130 aa TEB4-Doa10 (TD) domain, which contains the 5th, 6th and 7th transmembrane helices and is essential for the functioning of ScDoa10. The TD domain is the perceived site of association of the E2 enzyme ScUbc6 via its C-terminal membrane anchor, since mutations in this region significantly affect the stability of the E2 (Kreft and Hochstrasser, 2011). A further highly conserved 16 residue motif at the C-terminus has also been identified as being important for the recognition of some ScDoa10 substrates (Zattas et al., 2016).

Unlike most E3s, ubiquitination of the majority of ScDoa10 substrates appears to require the formation of a complex including two separate E2s, ScUbc6 and ScUbc7, as well as the membrane binding ScUbc7 co-factor, ScCue1 (Swanson et al., 2001; Ravid et al.,

2006). The reason why the combined action of two E2 enzymes is required is poorly understood. ScUbc6 and ScUbc7 are thought to generate alternative types of ubiquitin chains, with ScUbc6 favouring K11-linked chains, and ScUbc7 producing K48-linked chains (Xu et al., 2009). This has led to the proposal that ubiquitination by ScDoa10 may occur via alternate ubiquitin transfer from the two E2s. Specifically, primary ubiquitination by ScUbc6, followed by the bulk transfer of a pre-assembled ubiquitin chain by ScUbc7 to extend the ubiquitination is one possibility (Zattas and Hochstrasser, 2015).
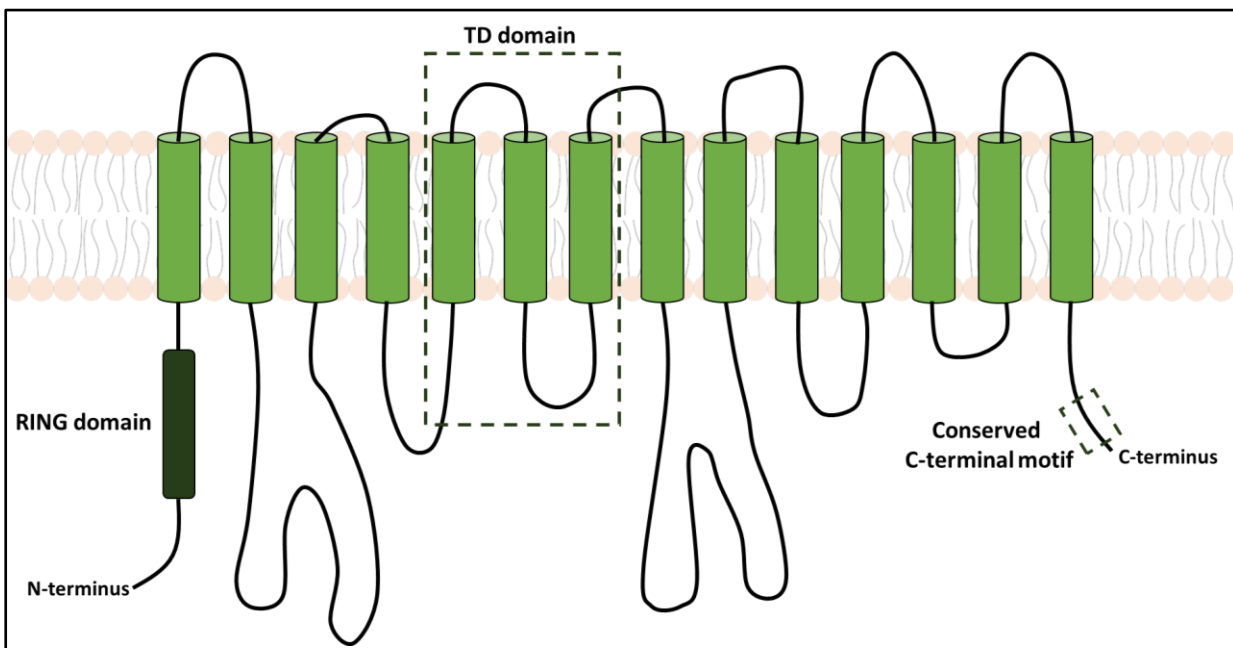
Figure 11 – Structure of ScDoa10
The predicted structure of ScDoa10 determined both experimentally and through modelling by Kreft et al. (2006). The 151 kDa protein contains 14 transmembrane domains, with the catalytic RING domain located close to the N-terminus. The overall conformation of ScDoa10 is believed to be similar to human and plants homologues. The highly conserved TEB4-Doa10 (TD) domain and conserved C-terminal motif are highlighted (dashed boxes). Based on Kreft et al. (2006).

1.5.2 Functions of ScDoa10

The ScDoa10 complex is involved in protein quality control, whereby it targets certain proteins when they contain folding errors, however, it is also able to target some correctly folded proteins as part of their standard regulation (Ravid et al., 2006). Substrates can be cytosolic, like ScMATα2, or membrane-bound proteins that contain cytosolic degrons (ERAD-C, see 1.2.3), such as the spindle tethering ScMps2 (Zattas and Hochstrasser, 2015). Whilst

ScDoa10-dependent degradation of cytosolic proteins only requires the core subunits of the ScDoa10 complex (ScDoa10, ScUbc6, ScUbc7 and ScCue1), the degradation of membrane bound substrates additionally requires the action of the ScCdc48 ATPase and the ubiquitination shuttle factors ScRad23 and ScDsk, presumably for membrane extraction (Ravid et al., 2006). The Sec61 β-subunit homologue 2 (ScSbh2) is now also recognised as a substrate of ScDoa10 that is targeted based on a transmembrane degron (Habeck et al., 2015), suggesting that ScDoa10 can participate in ERAD-M in addition to ERAD-C.

As a reported N-recognin of the Ac/N-degron pathway, some substrates of ScDoa10 are recognised according to their acetylated N-termini. However, many reported substrates of ScDoa10 are not Nt-acetylated and therefore it is possible that the E3 may contain multiple recognition domains (Aksnes et al., 2016). Indeed, contradictory reports have suggested that degradation of ScMATα2 is both dependent and independent of Nt-acetylation by NatB (Hwang et al., 2010; Zattas et al., 2013). ScDoa10 is also known to target substrates bearing degrons away from the N-terminus, such as the artificial C-terminal degron CL1 (Metzger et al., 2008), meaning that ScDoa10 does not function exclusively as an E3 ligase of the Ac/N-degron pathway.

1.5.3 Homologues of ScDoa10

1.5.3.1 The human homologue of ScDoa10 – HsMARCH6

ScDoa10 homologues have been identified in a broad range of eukaryotes (Zattas and Hochstrasser, 2015). The human homologue HsMARCH6 (also known as HsTEB4) is similar to ScDoa10 both in terms of structure and localisation and, like many E3s, HsMARCH6 was shown to self-ubiquitinate, hence regulating its own abundance (Hassink et al., 2005; Kreft et al., 2006). A notable difference between HsMARCH6 and ScDoa10, is that HsMARCH6 appears to work solely in conjunction with the human homologue of ScUbc7, potentially in a dimerised configuration, rather than also associating with ScUbc6 homologues (Kostova et al., 2007; Zattas and Hochstrasser, 2015). HsMARCH6 and ScDoa10 share many substrates, notably the enzyme squalene epoxidase (ScERG1 in yeast, HsSM in humans), a key enzyme

of the sterol biosynthesis pathway (Foresti et al., 2013; Zelcer et al., 2014). In humans, HsMARCH6 has also been shown to influence the stability of HsHMGR and HsLDM, another two significant enzymes of this pathway (Zelcer et al., 2014; Scott et al., 2020). Despite this, HsMARCH6 was unable to complement mutant *ScDoa10Δ* yeast phenotypes when tested (Kreft et al., 2006).

1.5.3.2 The *Arabidopsis* homologues of ScDoa10 – AtDOA10A and AtDOA10B

The *Arabidopsis* genome has two presumed homologues of ScDoa10: AtDOA10A (also known as AtCER9 and AtSUD1) and AtDOA10B (as discussed in 1.2.3). Like ScDoa10 and HsMARCH6, AtDOA10A is predicted to consist of a 14 transmembrane domain structure and is known to affect the sterol biosynthesis pathway, however, as with HsMARCH6, complementation of *ScDoa10Δ* yeast with AtDOA10A proved unsuccessful (Doblas et al., 2013). Mutations of AtDOA10A result in upregulation of both UPS and UPR-related genes, supporting its perceived ER-associated E3 ligase activity (Lu et al., 2012). Given ScDoa10's roles in ERAD and the Ac/N-degron pathway, mutations in *AtDOA10A* and *AtDOA10B* would be anticipated to affect a range of processes within plants such as stress responses, ABA signalling, immune responses, flowering time and photosynthesis, based on the plethora of known impacts of disruption of ERAD and the NAT machinery (see 1.2 and 1.3.3) (Pesaresi et al., 2003; Ferrandez-Ayela et al., 2013; Linster et al., 2015; Xu et al., 2015; Koskela et al., 2018; Strasser, 2018; Armbruster et al., 2020; Feng J. et al., 2020; Huber et al., 2020; Linster et al., 2020; Neubauer and Innes, 2020).

Whilst little work has been done to characterise AtDOA10B, the impact of mutations of AtDO10A have been studied more closely, following its initial identification as a so-called eceriferum (cer) mutant due to the generation of semi-glossy shoots (Koornneef et al., 1989). This phenotype was seemingly a result of a significant elevation in the amount of cuticular wax in *Atdoa10a*, as well as changes in wax composition, with a large increase in very long chain fatty acids and a decrease in the levels of other organic compounds such as alkanes, alcohols and aldehydes (Jenks et al., 1995; Lu et al., 2012; Reisberg et al., 2013). Cutin, the other

component of plant cuticles, was also substantially more abundant in *Atdoa10a*, resulting in a thicker but more irregularly structured membrane overall (Lu et al., 2012). Suberin content was also increased in the roots of *Atdoa10a* plants (Lu et al., 2012). It is likely that AtDOA10A regulates the enzymes involved in the biosynthesis of cuticular waxes and cutin, which are also predominantly ER-localised (Kunst and Samuels, 2009). Given their increased cuticle membrane thickness, it was unsurprising that that *Atdoa10a* mutants were shown to exhibit lower rates of transpiration and, as a result, a tolerance of drought stress (Lu et al., 2012). However, this tolerance was suggested to be a consequence of decreased stomatal pore aperture, due to the formation of cuticular ledges, rather than a direct effect of the changes in cuticle properties, since the shift in wax composition towards fatty acids would be expected to increase the permeability of the barrier (Lu et al., 2012). The changes in cuticle composition have also been reported to alter the microbial communities present on *Atdoa10a* plants (Reisberg et al., 2013).

Consistent with the drought tolerant phenotype, mutations of AtDOA10A have been shown to repress the drought hypersensitivity of *Atsqe1* mutants (Doblas et al., 2013). AtSQE1 is an *Arabidopsis* homologue of ScErg1 and encodes a squalene epoxidase enzyme of the sterol synthesis pathway. Both ScDoa10 and HsMARCH6 are known to ubiquitinate yeast and human squalene epoxidases (see 1.5.3.1) (Foresti et al., 2013; Zelcer et al., 2014). In addition to drought hypersensitivity, the loss of AtSQE1 alters the composition of major sterols in the roots and causes severe defects that include stunted and abnormal root and shoot growth, sterility and ABA insensitivity (Rasbery et al., 2007; Pose et al., 2009; Doblas et al., 2013). *Atdoa10a* was identified by a suppressor screen as a mutation that fully suppressed almost all the *Atsqe1* mutant phenotypes (Doblas et al., 2013). Surprisingly however, the phenotype suppression was not a result of the sterol composition being restored. Instead, the recovery was associated with the prevention of a root-specific accumulation of squalene and other sterol intermediates, involving reducing the activity of the upstream AtHMGR enzyme. AtHMGR is known to undergo ERAD under certain conditions (see 1.2.1), suggesting that, as

in yeast and humans, AtDOA10A is likely part of the ERAD system. Studies of the *Arabidopsis* homologues of ScUbc6 and ScUbc7, ScDoa10's cognate E2s, are also consistent with a perceived role ERAD (Cui et al., 2012a; Cui et al., 2012b; Chen et al., 2016; Chen et al., 2017a; Feng, H. et al., 2020), although they are also known to interact with other *Arabidopsis* ER-associated E3s (Ahn et al., 2018). Furthermore, it has been shown that mutations of AtHRD1A, AtHRD1B and AtDOA10A additively improve resistance to ERAD-inducing heat stress (Li et al., 2017).

As well as its links to AtSQE1, AtDOA10A has also been directly implicated in ABA signalling, as *Atdoa10a* mutants are hypersensitive to ABA during the transition from embryo germination to developing seedling. As a result, embryos remain dormant (Zhao et al., 2014). This hypersensitivity correlates with ABA phenotypes observed for *Atsqe1* mutants (Pose et al., 2009; Doblas et al., 2013), and the ABA responsiveness of the NatA complex, of the Ac/N-degron pathway (Linster et al., 2015). However, whilst exogenous applications of ABA reduce expression of NatA, AtDOA10A expression is upregulated by ABA treatment. Mutations of AtPRT6, an *Arabidopsis* Arg/N-recognin (see 1.4.2), have also been shown to result in ABA hypersensitivity (Holman et al., 2009). The ABA-hypersensitivity phenotype can be at least partially explained by an increase in ABA abundance in *Atdoa10a* mutants (Zhao et al., 2014).

1.6 Focus of this study

Although, currently published results are not inconsistent with AtDOA10A's perceived role as an E3 ligase involved in ERAD and the so far undescribed plant Ac/N-degron pathway, further work is required to confirm this. The aim of this doctoral work was to clarify the function of AtDOA10A in *Arabidopsis*, as well as investigating the predicted similar role of the lesser-studied AtDOA10B. Significant efforts were made to identify substrates of these E3s to provide insight into their modes of action. Results generated were also used to consider the overall contribution of the Ac/N-degron pathway towards global proteolysis in *Arabidopsis*.

Chapter II -

Materials and Methods

2.1 Bacterial strains

2.1.1 *Escherichia coli*

2.1.1.1 DH5α

Laboratory strain containing *lacZ∆M15*, *recA1* and *endA* mutations to improve transformation and screening efficiency which was used for plasmid propagation.

2.1.1.2 DB3.1

Competent cells possessing the *gyrA462* mutation which provides resistance to the toxic *ccdB* gene. DB3.1 cells were used for propagation of *ccdB*-containing empty vectors.

2.1.1.3 XL10-Gold® (Agilent)

Ultracompetent cells (TetrD(mcrA)183 D(mcrCB-hsdSMR-mrr)173 endA1 supE44 thi-1 recA1 gyrA96 relA1 lac Hte [F´ proAB lacIqZDM15 Tn10 (Tetr) Amy Camr]) used for transformations of mutagenized plasmids (included with QuikChange Lightning Site-Directed Mutagenesis Kit).

2.1.2 *Agrobacterium tumefaciens*

2.1.2.1 GV3101

A. tumefaciens strain used for transformations of *Arabidopsis thaliana*, which contains the disarmed Ti plasmid pMP90 providing gentamycin resistance in addition to chromosome-encoded rifampicin resistance.

2.2 Yeast Strains

2.2.1 *Saccharomyces cerevisiae*

2.2.1.1 BY4743

WT and mutant yeast cells used in all recombinant protein analysis were homozygous diploid BY4743 cells derived from the non-flocculent S288C strain (Dharmacon yeast KO collection, Horizon Discovery). BY4743 cells carry mutations in selectable marker genes (*his3*, *leu2*, *met15* and *ura3*) and insertional mutants also contain the *KanMX4* gene which provides resistance to the antibiotic G418.

2.3 List of vectors

2.3.1 Gateway® entry vectors

Vector	Size (bp)	Promoter	Epitope Tag	Selection	
				E. coli	Target
pENTR™/D-TOPO® (Invitrogen)	2,820	N/A	N/A	Kan	N/A
pDONR™221 (Invitrogen)	4,762	N/A	N/A	Kan	N/A
pEn-Sa-Chimera (Steinert et al., 2015)	3,054	N/A	N/A	Amp	N/A

2.3.2 Gateway® destination vectors

Vector	Size (bp)	Promoter	Epitope Tag	Selection	
				E. coli	Target
pGWB533 (Nakagawa et al., 2007)	12,977	N/A	GUS (C-term)	Spec	Hyg (*Arabidopsis*)
pGWB540 (Nakagawa et al., 2007)	11,882	N/A	EYFP (C-term)	Spec	Hyg (*Arabidopsis*)
pGWB17 (Nakagawa et al., 2007)	17,407	CaMV 35S	4x MYC (C-term)	Kan	Hyg/Kan/Neo (Arabidopsis)
pEarleyGate103 (Earley et al., 2006)	12,411	CaMV 35S	mGFP5 (C-term)	Kan	Basta (*Arabidopsis*)
pDe-Sa-Cas9 (Steinert et al., 2015)	15,522	PcUbi4-2	N/A (C-term)	Spec	Kan (*Arabidopsis*)
pAG413GPD-ccdB-HA (Susan Lindquist, Addgene plasmid #14238)	7,668	GPD	3x HA (C-term)	Amp	HIS (Yeast)
pAG416GPD-ccdB-EGFP (Susan Lindquist, Addgene plasmid #14196)	8,203	GPD	EGFP (C-term)	Amp	URA (Yeast)

2.4 List of antibodies

2.4.1 Primary Antibodies

Antibody	Host Species
α – β-Actin (Abcam ab184220)	Mouse
α – β-Tubulin (Sigma-Aldrich T8328)	Mouse
α – AtCNX1/2 (Agrisera AS12 2365)	Rabbit
α – AtUGPase (Agrisera AS05 086)	Rabbit
α – GUS (Sigma-Aldrich G5420)	Rabbit
α – GFP/YFP (ROCHE 11814460001)	Mouse
α – HA (Sigma-Aldrich H3663)	Mouse
α – MYC (Antibodies.com A85281)	Mouse

2.4.2 Secondary Antibodies

Antibody	Host Species
α – Mouse (Sigma-Aldrich A5278)	Goat
α – Rabbit (Cell Signalling Technology 7074)	Goat

2.5 List of selective agents (antibiotics/herbicides)

Selective Agent	Working Concentration
Kanamycin	50 µg ml^{-1}
Ampicillin	100 µg ml^{-1}
Spectinomycin	50 µg ml^{-1}
Rifampicin	50 µg ml^{-1}
Gentamycin	25 µg ml^{-1}
G418 (Geneticin)	200 µg ml^{-1}
Hygromycin	30 µg ml^{-1}
Basta (glufosinate ammonium)	20 µg ml^{-1}

2.6 Growth conditions, propagation and treatments

2.6.1 *Arabidopsis thaliana*

2.6.1.1 Growth on soil

Arabidopsis seeds were sown on media composed of Levington M3 compost, vermiculite and perlite (in a 4:2:1 ratio) and grown in 16/8 hours light/dark in a glasshouse or growth chamber (90–100 µmol m^{-2} s^{-1} white fluorescent light) with environmental temperatures restricted to 20-22 °C.

2.6.1.2 Sterile Growth

Seeds were sterilised through 10 minutes of agitation in 10% bleach solution and then rinsed three times in sterile deionised water. Seeds were then plated onto half strength Murashige and Skoog (½ MS) media (Table 2) with selective agents (see 2.5) or supplements added where required before being stratified at 4 °C for a minimum of 48 hours. Following stratification, plates were moved into a constant temperature growth room (20-22 °C) with 16/8 hours light/dark (90–100 µmol m^{-2} s^{-1} white fluorescent light).

Table 2 – *Arabidopsis* sterile growth media recipe

½ MS media
2.15 g l^{-1} Murashige and Skoog Basal Salt Mixture (Sigma-Aldrich)
1 % Agar
KOH added dropwise to pH 5.8

2.6.1.3 Cross fertilisation of *Arabidopsis thaliana*

Unopened buds (petals not visible) were selected on female parent plants and were isolated by the removal of neighbouring buds with scissors. Sepals, petals and stamens were then removed using fine forceps (pre-sterilised with 70% ethanol) to leave only the pistil. From male parent plants mature flowers with fully splayed petals were removed. Removed flowers were squeezed at the base of the bud with sterile forceps to expose stamens. Pollination was then performed by touching the pollen-containing anthers against the exposed pistil of the

female plant. Pollinated pistils were wrapped in clingfilm to prevent contact with additional pollen and allowed to develop into mature seed-bearing siliques before being harvested.

2.6.1.4 Histochemical staining

Histochemical staining was performed to assess the abundance of β-glucuronidase (GUS)-tagged recombinant proteins through the accumulation of the coloured product of the GUS enzyme. The staining solution was prepared by combining 1M monosodium phosphate and 1M disodium phosphate in a 39:61 ratio to generate a pH 7.0 phosphate buffer. The phosphate buffer was then diluted 1 in 10 to produce a 0.1M, pH 7.0 solution. X-gluc was then added to a concentration of 2 mM followed by the addition of Triton-X-100 (0.1% v/v).

Immediately prior to staining, 1/100x the total volume of GUS stain required of both potassium ferricyanide and potassium ferrocyanide (both 100 mM) were added to the GUS stain solution. Seedlings or adult tissues were then submerged in up to 1 ml of GUS stain solution and were incubated at 37 °C in the dark for 24 hours.

After 24 hours, GUS stain solution was removed from samples by pipetting and was replaced by up to 1 ml of fixative (3:1 ethanol:acetic acid plus 1 % (v/v) Tween). Samples were incubated at room temperature with gentle shaking and periodic removal and replacement of fixative to maximise chlorophyll removal until tissues appeared cleared. Cleared samples were then mounted onto microscope slides in 50 % glycerol in order to be studied under the microscope.

2.6.1.5 Chemical treatments

Arabidopsis seedlings were subjected to a range of chemical treatments such as phytohormones, ERAD stressors and sterol synthesis inhibitors. Treatments were applied either through direct addition of the chemical to the ½ MS growth media just before plates were poured or via a transient liquid treatment. For liquid treatments, seedlings were placed in wells containing water or liquid ½ MS and the specified concentration of a given chemical. Seedlings were removed from the transient treatments at the reported time points.

2.6.2 *Saccharomyces cerevisiae*

In general, *S. cerevisiae* was grown on stationary yeast-extract peptone dextrose (YPD) or synthetic drop-out (DO) plates (Table 3) incubated at 30 °C. Liquid cultures were prepared through the inoculation of colonies into liquid YPD or DO media, which were incubated at 30 °C with 200 rpm shaking. G418 antibiotic was added to media used for the growth of mutant strains.

Table 3 – Yeast growth media recipes

YPD media	DO Media (single)	DO Media (Double)	DO Media (Quadruple)
10.0 g l^{-1} Yeast Extract	Manufacturer specified (~1.9) g l^{-1} DO base (Formedium)	Manufacturer specified (~1.8) g l^{-1} DO base (Formedium)	Manufacturer specified (~1.5) g l^{-1} DO base (Formedium)
20.0 g l^{-1} Peptone	1.7 g l^{-1} Yeast nitrogen base without amino acids (Formedium)	1.7 g l^{-1} Yeast nitrogen base without amino acids (Formedium)	1.7 g l^{-1} Yeast nitrogen base without amino acids (Formedium)
20.0 g l^{-1} Glucose	1.0 g l^{-1} L-glutamic acid	1.0 g l^{-1} L-glutamic acid	1.0 g l^{-1} L-glutamic acid
(20.0 g l^{-1} Agar)	10.0 g l^{-1} Sucrose	10.0 g l^{-1} Sucrose	10.0 g l^{-1} Sucrose
	(20 g l^{-1} Agar)	(20 g l^{-1} Agar)	(20 g l^{-1} Agar)

2.6.3 *Escherichia coli*

Lysogeny broth (LB) containing the appropriate antibiotics was used to grow colonies on agar plates and liquid cultures of *E. coli.* Colonies and cultures were grown at room temperature to 37 °C, with 200 rpm shaking for liquid cultures.

2.6.4 *Agrobacterium tumefaciens*

A. tumefaciens was grown on LB agar plates and in liquid LB containing rifampicin, gentamycin and additional vector-dependent antibiotics. Colonies and cultures were grown at 28 °C, with 200 rpm shaking for liquid cultures.

2.7 Nucleic acid isolations and manipulations

2.7.1 Genomic DNA Extraction from *Arabidopsis thaliana*

DNA was extracted from *Arabidopsis* tissue (most commonly rosette leaf discs) through the addition of 40 µl of extraction buffer (Table 4) followed by grinding of the leaf tissue using a 200 µl pipette tip. Samples were then heated to 95 °C for 10 minutes using a heat block before being placed on ice for 5 minutes and then diluted with 40 µl of dilution buffer (Table 4). Samples were then centrifuged at 16300 x g for 30 seconds in a benchtop centrifuge and 1-3 µl of the resulting supernatant used for subsequent analysis.

Table 4 – *Arabidopsis* DNA extraction buffers

Extraction Buffer	Dilution Buffer
0.25 M KCl	3% BSA (w/v)
10 mM EDTA	
100 mM Tris-HCl (pH 9.5)	

2.7.2 DNA Extraction from *Saccharomyces cerevisiae*

A 1 µl inoculation loop was used to suspend a small portion of a *S. cerevisiae* pellet or colony in 100 µl of extraction buffer (Table 5). The suspension was then briefly vortexed and heated to 70 °C for 5 minutes. DNA was then precipitated through the addition of 300 µl of 100% ethanol, followed by vortexing for 5 seconds. Precipitated DNA was then pelleted by centrifugation at 15000 x g for 3 minutes and the supernatant disposed. To wash the pellet, 500 µl of 70% ethanol was then added and briefly vortexed. Samples were then centrifuged

at 15000 x g for 30 seconds and supernatant removed by pipetting, with any remaining supernatant then allowed to evaporate by air-drying for 2 minutes. The dry pellet was then dissolved in 100 µl of SDW before being centrifuged again at 15000 x g for 15 seconds. 60 µl of the resulting supernatant was then removed with 1-3 µl of this supernatant then used for subsequent analysis.

Table 5 – Yeast DNA extraction buffer

Extraction Buffer
0.2 M Lithium acetate
1% SDS

2.7.3 Plasmid DNA Extraction from *Escherichia coli* (miniprep)

Plasmid DNA was extracted from liquid cultures of *E. coli* grown for a minimum of 16 hours at room temperature to 37 °C. Bacterial cells were pelleted through centrifugation at 3000 rpm in a large benchtop centrifuge (MSE Mistral 3000i) at 4 °C for 15 minutes. Supernatant was removed and QIAprep Spin Miniprep Kit (QIAGEN) was used to isolate plasmid DNA from the resulting pellet, according to manufacturer's instructions. The quantity and quality of extracted plasmid DNA was checked using a NanoDrop™ 1000 spectrophotometer (ThermoFisher Scientific) followed by PCR or a restriction digest to confirm the vector appeared correct.

2.7.4 Restriction Digests

Restriction digests were performed using a range of restriction endonucleases (New England Biolabs). Reactions were carried out with 0.5 µg of plasmid DNA combined with 0.5 µl of restriction enzyme(s) and 2 µl of reaction buffer in a 20 µl reaction volume. Reactions were incubated at 37 °C for 15 minutes to 2 hours and were inactivated by heat or the addition of gel loading dye. In cases where distinct reaction buffers were required for the different endonucleases being used in a single reaction, sequential digests were performed using enzyme-specific protocols devised using NEBcloner® (New England Biolabs).

2.7.5 Total RNA Extraction from *Arabidopsis thaliana*

Arabidopsis tissue was snap frozen in liquid nitrogen. RNA was extracted from frozen samples using an RNeasy minikit (QIAGEN) according to the manufacturer's plant tissue-specific instructions. The extracted RNA was then analysed using a NanoDrop™ 1000 spectrophotometer (ThermoFisher Scientific).

2.7.6 Total RNA Extraction from *Saccharomyces cerevisiae*

RNA was extracted from a *S. cerevisiae* pellet equivalent to a minimum of 2 OD_{600} units (e.g. 2 ml of culture with a 600 nm absorbance value of 1.0) using a phenol-chloroform extraction method. 0.5 ml of TRI reagent® (Sigma-Aldrich) was added to the pellet and triturated for 1 minute to lyse cells. 200 µl of chloroform was then added, with the solution then vortexed for 15 seconds and incubated at room temperature for 5 minutes. Centrifugation for 10 minutes at 16300 x g and 4 °C was used to separate aqueous and organic phases. 300 µl of the aqueous phase only was transferred to a new microcentrifuge tube and incubated on ice. RNA was precipitated through the addition of 300 µl of isopropanol, which was mixed by inversion and held on ice for a minimum of 15 minutes before being pelleted by centrifuging for 10 minutes at 18900 x g and 4 °C. Supernatant was disposed and the pellet washed with 70% ethanol, which was removed by pipetting following a further 18900 x g centrifugation for 1 minute at 4 °C. Any remaining ethanol was allowed to evaporate by air-drying for 10 minutes. The RNA pellet was then dissolved in 50 µl of RNase-free water and analysed using a NanoDrop™ 1000 spectrophotometer (ThermoFisher Scientific).

2.7.7 RNA Sequencing (RNA-seq)

RNA was extracted from 10-day old *Arabidopsis thaliana* seedlings (as 2.7.5). The quality of RNA samples obtained was checked using a NanoDrop™ 1000 spectrophotometer (ThermoFisher Scientific), and by gel electrophoresis with 1 µg of RNA on an agarose gel containing 1% (v/v) bleach, the addition of which denatures RNA secondary structure and prevents degradation by RNases (Aranda et al., 2012). The presence of clear bands

corresponding to 28S and 18S ribosomal RNAs, as well as the absence of genomic DNA bands confirmed the purity of the samples.

Samples were shipped to Novogene (HK) Company Limited (Hong Kong) where RNA integrity was checked using a 2100 bioanalyzer and RNA 6000 Nano kit (Agilent) and sample purity and quantity was again assessed by gel electrophoresis and NanoPhotometer® spectrophotometer (IMPLEN). A cDNA library was generated from the RNA samples using a NEBNext® UltraTM RNA Library Prep Kit for Illumina® (NEB) to enrich mRNA using oligo dT-tagged magnetic beads, randomly fragment the resultant mRNA and synthesise cDNA using random hexamer primers. Subsequent processing of the termini and addition of sequencing adaptors followed by PCR amplification, produced the final cDNA library, which was re-assessed for quality using a 2100 bioanalyzer (Agilent), and quantified using a Qubit 2.0 fluorometer (Life Technologies), and by qPCR. The library was then sequenced using an Illumina Hiseq platform.

The resulting raw sequencing reads were filtered to isolate clean reads by removing low quality reads and those containing adapter sequences. The remaining reads were mapped to the *A. thaliana* genome using TopHat v2.0.12 software and the number of reads corresponding to each gene were counted using the same software. From this, Fragments Per Kilobase of transcript per Million mapped reads (FPKM) values for each gene were calculated by normalising for the total number of fragments sequenced and the length of the gene. Differences in gene expression between Col-0 and *Atdoa10a*/b (RNAi line 4-2(2)) were determined using the DESeq R package (1.18.0) with resulting p-values adjusted for false discoveries to give q-values using the Benjamini and Hochberg's approach (Benjamini and Hochberg 1995).

2.7.8 DNase Treatment of RNA

DNase treatments were performed on a maximum of 1.5 μg of total RNA extract using RQ1 RNase-free DNase (Promega). Up to 8 μl of RNA extract was added to 1 μl of RQ1

DNase and 1 µl of 10X reaction buffer. The reaction was then topped up to a final volume of 10 µl with RNase-free water and incubated at 37 °C for 30 minutes. 1 µl of RQ1 stop solution was then added and the reaction incubated at 65 °C for 10 minutes to terminate the reaction.

2.7.9 cDNA synthesis

cDNA was synthesised from DNase-treated total RNA extracts using oligo(dT) primers and SuperScript™ II Reverse Transcriptase (ThermoFisher Scientific) according to the manufacturer's protocol. The resultant cDNA samples were then diluted to a final volume of 100 µl with SDW.

2.7.10 Sanger DNA sequencing

Sanger sequencing of DNA samples was performed by the University of Birmingham Functional Genomics Laboratory. Submitted samples contained either 50 – 100 ng of purified PCR product or 300 ng of plasmid DNA isolated by miniprep in addition to 3.2 pmol of the appropriate primer in a total reaction volume of 10 µl. Sequencing was carried out using an ABI 3730 (Applied Biosystems) capillary sequencer.

2.7.11 Site directed mutagenesis

Entry vectors were altered by changing the identity of a single nucleotide using a QuikChange Lightning Site-Directed Mutagenesis Kit (Agilent) according to the manufacturer's instructions. The primers used were designed to match the DNA sequence required except for the inclusion of the mutation.

2.8 PCR and associated methods

2.8.1 General

MyTaq™ Red Mix (Meridian Bioscience) or Phusion® High-Fidelity DNA Polymerase (New England Biolabs) was used to amplify specific fragments of DNA according to the manufacturer's instructions and using the primers listed in Appendix Table 1. Optimum

annealing temperatures were estimated based on the T_m of primers or, where necessary, determined by gradient PCR.

2.8.2 Genotyping PCR

PCR-based genotyping was conducted using MyTaq™ Red Mix (Meridian Bioscience) according to manufacturer's instructions. Separate reactions were carried out using either two flanking primers (forward and reverse primers either side of the T-DNA insertion point) or using a T-DNA specific border primer with the appropriate flanking primer.

2.8.3 Colony PCR

Colonies of *E. coli*, *A. tumefaciens* or *S. cerevisiae* were screened by PCR to determine the presence of transformed vectors including the gene of interest. Colonies were suspended in 20 to 30 µl of SDW. PCR was then performed using 2 µl of the suspended cells and MyTaq™ Red Mix (Meridian Bioscience). Primer pairs contained one gene-specific primer and one vector-specific primer in order to identify colonies containing recombinant plasmids.

2.8.4 qPCR

qPCR reaction mixes were prepared using Brilliant III Ultra-Fast SYBR® Green QPCR Master Mix with Low ROX (Agilent) according to the manufacturer's instructions. Reactions contained 45 ng of template cDNA (assuming 100% conversion from RNA). qPCR was performed using an AriaMx Real-time PCR System (Agilent). qPCR primers were designed so that at least one oligonucleotide spanned an exon-exon junction in order to prevent amplification from any contaminating genomic DNA. In addition to this prior to quantification, melt curves were checked to confirm they contained a single peak only and NTC samples were used to show that no primer dimers were formed. Relative levels of gene expression were quantified using the delta delta C_t method (Livak and Schmittgen 2001).

2.8.5 High Resolution melt (HRM) analysis

HRM analysis was performed in order to detect small indels in amplified regions of DNA from plants transformed with CRISPR-Cas9 constructs. Primers were designed to

amplify a short (≈100 bp) DNA sequence from the *Arabidopsis* genome which included the anticipated Cas9 cut site. Using the AriaMx Real-time PCR System (Agilent) PCR was carried out for 40 amplification cycles followed by HRM analysis. HRM analysis was performed through the denaturation of PCR products at 95 °C for 30 seconds followed by cooling to 65 °C for 30 seconds to allow DNA strands to re-anneal. The temperature was then raised in 0.2 °C increments in 10 second time intervals until a final temperature of 95 °C was reached.

Melt curves were then analysed to identify the precise temperature at which annealed DNA strands were denatured. Melt curves that deviated from the melt curve of control plants were recognised as those potentially containing indels. Deviations took the form of shifts in melting temperature of the primary product and the presence of secondary, lower temperature melt peaks, corresponding to alterations in the length of the PCR amplicon and the existence of mismatches between DNA strands formed during the 50 °C annealing.

2.8.6 Purification of PCR products

PCR purifications were carried out using a QIAquick® PCR Purification Kit (QIAGEN) according to the manufacturer's instructions.

2.8.7 Gel Electrophoresis

Gels were prepared by adding agarose to TBE buffer (0.66 to 2.00 % w/v) before heating solutions in a microwave until the agarose had completely dissolved. The solutions were then allowed to cool to approximately 50 °C, at which point ethidium bromide (0.05 µl ml^{-1}) was added and swirled to mix. The gel solutions were then poured into an appropriately sized gel cast with comb and allowed to set for around 40 minutes. Once set, gels were submerged in TBE running buffer within electrophoresis tanks and 6 to 24 µl of DNA samples (including loading dye) were loaded into wells. A 60 to 120 V potential difference was then applied to drive DNA migration. After sufficient migration, gels were removed from running buffer and visualised using a Gel Doc™ XR+ Gel Documentation System (BioRad).

2.8.8 Purification of digested DNA from agarose gels (gel extraction)

Fragments of DNA produced from digestion with restriction endonucleases (see 2.7.4) were separated by gel electrophoresis. Sections of gel containing single DNA fragments of interest were excised from the agarose gel with a scalpel, under UV light. DNA fragments were then extracted from the agarose using a QIAquick Gel Extraction Kit (QIAGEN), according to the manufacturer's instructions.

2.9 Protein extractions and analysis

2.9.1 Total protein extraction from *Arabidopsis thaliana*

Arabidopsis tissue was frozen and ground up in liquid nitrogen. A volume of lysis buffer was added to samples following grinding. The volume and composition of lysis buffer were individually determined based on the quantity of tissue and the nature of the protein of interest (commonly 200 µl of RIPA buffer – Table 6). Samples were then centrifuged at 21700 x g for 10 minutes to pellet cell debris and 80 – 120 µl of supernatant removed to a new tube. An appropriate volume of 5x sample buffer (Table 6) was then added before samples were boiled at 98 °C for 5 minutes. Samples were allowed to cool to room temperature prior to proceeding to SDS-PAGE or long-term storage.

Table 6 – *Arabidopsis* protein extraction and sample buffers

RIPA Lysis Buffer	5x Sample Buffer
10 mM Tris-HCl (pH 8.0)	300 mM Tris-HCl (pH 6.8)
150 mM NaCl	50% Glycerol
0.5 mM EDTA	25% β-mercaptoethanol
0.1 % SDS	10% SDS
1 % Triton-X-100	0.05% Bromophenol blue
1 % Sodium deoxycholate	
cOmplete™, Mini, EDTA-free Protease Inhibitor Cocktail (Roche)	

2.9.2 Total protein extraction from *Saccharomyces cerevisiae*

Protein was extracted from a *S. cerevisiae* pellet equivalent to a minimum of 2 OD_{600} units. The pellet was resuspended in 50 µl of extraction buffer (Table 7) and heated to 90 °C for 10 minutes. 0.67 µl of 3M acetic acid was added and mixed by vortexing for 30 seconds. Samples were then re-heated at 90 °C for 10 minutes. 12.5 µl of sample buffer (Table 7) was then added and mixed by vortexing for 2 seconds. Centrifugation at 12000 x g for 5 minutes was used to pellet cell debris and 55 µl of the resulting supernatant was transferred to a new tube. Protein samples were then boiled at 98 °C for 1 minute before being allowed to cool to room temperature prior to proceeding to SDS-PAGE or long-term storage.

Table 7 – Yeast protein extraction and sample buffers

Extraction Buffer	Sample Buffer
0.1 M NaOH	250 mM Tris-HCl (pH 6.8)
50 mM EDTA	50% Glycerol
2% β-mercaptoethanol	0.05% Bromophenol blue
2% SDS	

2.9.3 Quantification of protein concentration (Bradford Assay)

Concentrations of protein extracts were estimated using the Bradford Assay (Bradford 1976). Prior to the addition of sample buffer, 1 µl samples of protein extract were added to 19 µl of SDW. 200 µl of Bradford Reagent (AppliChem) was then added to the diluted protein samples and the absorbance at 595 nm was then measured using a spectrophotometer. Measured absorbances were compared to a BSA standard curve to provide an estimation of protein concentration.

2.9.4 Immunoprecipitations (IPs) using GFP-Trap® beads

Immunoprecipitation of (E)GFP/YFP-tagged recombinant proteins was performed using GFP-Trap® magnetic agarose beads (ChromoTek) and 60 seven-day old *Arabidopsis* seedlings. Cells were disrupted in 400 µl of lysis buffer (Table 8) as described in 2.9.1. Following vortexing and centrifugation at 21700 x g for 10 minutes, the lysate was transferred to a new, pre-cooled tube and diluted with 600 µl of dilution buffer (Table 8). The diluted lysate was then added to a 25 µl slurry of GFP-Trap® beads previously equilibrated by washing with 500 µl of dilution buffer. The lysate was then rotated at 20 rpm at 4 °C for 3 hours to allow the binding of (E)GFP/YFP-tagged proteins to the beads. Following binding, beads were separated by magnet and the supernatant was removed. Beads were then washed three times with wash buffer (Table 8) and transferred to a new tube. Bound proteins were then eluted through the addition of 80 µl of 2x SDS sample buffer and boiling at 95 °C for 5 mins. Beads were again separated by magnet and the supernatant used for subsequent SDS-PAGE and Western blotting.

Table 8 – Immunoprecipitation buffers

Lysis/Wash Buffer	Dilution Buffer	2X SDS Sample Buffer (Elution)
10 mM Tris/Cl (pH 8.0)	10 mM Tris/Cl (pH 8.0)	120 mM Tris-HCl (pH 6.8)
150 mM NaCl	150 mM NaCl	20 % Glycerol
0.5 mM EDTA	0.5 mM EDTA	10 % β-mercaptoethanol
0.5 % Nonidet P-40		4 % SDS
cOmplete™, Mini, EDTA-free Protease Inhibitor Cocktail (Roche)	cOmplete™, Mini, EDTA-free Protease Inhibitor Cocktail (Roche)	0.02 % Bromophenol blue

2.9.5 SDS-PAGE and Western blotting

SDS-PAGE and Western blotting were performed using the BIO-RAD Mini-PROTEAN system. Proteins were separated according to size by running on a 6-10% polyacrylamide gel (Table 9). Gels were submerged in running buffer (Table 9) and a potential difference of 40-100 V applied. Separated proteins were transferred to a PVDF membrane by setting up a blotting cassette submerged in transfer buffer (Table 9) and applying a 25 V potential difference for approximately 16 hours.

Following protein transfer, membranes were blocked by being placed into 5% non-fat milk in TBST (Table 9) with gentle rocking for a minimum of 1 hour. Blocking solution was then replaced by the appropriate primary antibody (diluted in either TBST or 5 % milk-TBST) and rocked for a further 3 to 24 hours. The primary antibody was then removed, the corresponding secondary antibody applied and the membrane was rocked for an additional 50 minutes or more before being removed.

Protein bands were visualised though the addition of Pierce™ ECL Western Blotting Substrate (ThermoFisher Scientific), with the resulting luminescence used to expose X-ray film (Amersham), developed using a Curix 60 (Agfa) film processor.

Table 9 – SDS-PAGE and Western blotting recipes

5% Polyacrylamide Gel (Stacking)	10% Polyacrylamide Gel (Resolving)	Running Buffer	Transfer Buffer	TBST
5 % Acrylamide/bis-acrylamide (30%)	10 % Acrylamide/bis-acrylamide (30%)	0.1 % SDS	10 % methanol	150 mM NaCl
125 mM Tris-HCl (pH 6.8)	375 mM Tris-HCl (pH 8.8)	25 mM Tris	25 mM Tris	50 mM Tris (pH adjusted to 7.5 with HCl)
0.1 % (w/v) SDS	0.1 % (w/v) SDS	192 mM glycine	192 mM glycine	0.1 % Tween 20
0.1 % (w/v) Ammonium persulfate	0.1 % (w/v) Ammonium persulfate			
0.01 % (v/v) TEMED	0.004 % (v/v) TEMED			

2.9.6 Coomassie staining

As a loading control for Western blot analysis, PVDF membranes were submerged and agitated in Coomassie stain (Table 10). After 20 minutes the stain was removed and replaced with destain (Table 10) for a further 15 minutes. Destain was then removed and stained membranes were left to air-dry.

Table 10 – Coomassie staining solution

Coomassie Stain solution	Destain solution
50 % (v/v) Acetic acid	10 % (v/v) Acetic acid
20 % (v/v) Methanol	45 % Methanol
0.5 g R250 Coomassie Brilliant Blue (Thermo Fisher Scientific)	

2.9.7 Cycloheximide chase assays

2.9.7.1 *Saccharomyces cerevisiae*

Cycloheximide (CHX) chases were performed based on the protocol published by Buchanan et al. (2016) for yeast cells. Colonies of yeast to be analysed by CHX chase were inoculated into 10 ml of the appropriate double DO media and grown at 30 °C (with shaking)

for a minimum of 16 hours. The OD_{600} of each culture was measured and the required volume of the primary culture was subcultured into 30 ml of double DO media to generate an OD_{600} of 0.2. The secondary cultures were then incubated at 30 °C (with shaking) and their growth monitored regularly until an OD_{600} of 1.0 was reached, at which point 8 ml of culture was transferred to a 15 ml Falcon® tube held on ice (volume adjusted for small deviations away from ideal OD_{600} of 1.0). Cultures were centrifuged in a large benchtop (MSE Mistral 3000i) at 3000 rpm for 20 minutes with the supernatant then removed. The pellet was resuspended in 3.2 ml of double DO media prewarmed to 30 °C and the Falcon® tube then incubated at 30 °C in a water bath for 5 minutes. CHX (AppliChem) (or DMSO solvent-only control) was then added up to a concentration of 250 µg ml^{-1} and briefly vortexed. A 950 µl sample of each suspension was immediately removed and added to 50 µl of ice-cold Stop Mix (1M NaN_3, 100 µg ml^{-1}BSA) and held on ice. Further 950 µl samples were collected at 30-minute and 60-minute time points. Following collection of all samples, centrifugation for 30 seconds at 6500 x g was used to generate cell pellets. Supernatants were discarded and pellets were snap-frozen in preparation for protein extraction.

2.9.7.2 *Arabidopsis thaliana*

Seedlings were grown for 7 days on vertical ½ MS plates before being added to 0.3 g l^{-1} CHX (or DMSO solvent-only control) in water. 20 to 30 seedlings were removed from the treatment at specified time points (commonly 0-, 3-, and 6-hours), blotted dry on tissue paper and snap frozen in liquid nitrogen prior to protein extraction and SDS-PAGE (see 2.9.1 and 2.9.5).

2.9.8 Cell Fractionation

Cell fractionation was performed based on the protocol described by Abas and Luschnig (2010) designed for use with small amounts of tissue and without the need for ultracentrifugation. Tissue was lysed by grinding in a detergent-free extraction buffer (Table 11). Preclearance of the lysate was then performed by centrifugation at 630 x g for 10 minutes to pellet cell debris and large organelles such as nuclei. High speed centrifugation (21000 x

g) was then carried out for 2 hours at 4 °C in order to pellet the microsomal fraction. The supernatant (the soluble fraction) was removed for analysis and the pellet washed with 150 µl of wash buffer (Table 11). The microsomal fraction was then resuspended in resuspension buffer (Table 11). A volume of 5x sample buffer (Table 6) was added to each of the collected fractions to produce 1x solutions, which were then boiled for 5 minutes at 98 °C and analysed by SDS-PAGE and Western blotting.

Table 11 – Cell fractionation buffers

Detergent-free Extraction Buffer	Wash Buffer	Resuspension Buffer
100 mM Tris (pH 8)	100 mM Tris (pH 8)	10mM Tris (pH 8),
5 % Glycerol	5 mM EDTA	0.5 mM EDTA
10 mM EDTA	150 mM NaCl	150 mM NaCl
10 mM EGTA		
5 mM KCl		
1 mM DTT		
cOmplete™, Mini, EDTA-free Protease Inhibitor Cocktail (Roche)	cOmplete™, Mini, EDTA-free Protease Inhibitor Cocktail (Roche)	cOmplete™, Mini, EDTA-free Protease Inhibitor Cocktail (Roche)

2.9.9 Quantification of the proteome and protein Nt-acetylation

Analysis of the Nt-acetylome was carried out by the Protein maturation, cell fate and therapeutics lab (Institute of Integrative Biology of the Cell, CNRS, Université Paris-Sud, France) according to the method described by Bienvenut et al. (2017). Briefly, total protein was extracted from 10-day old *Arabidopsis* seedlings through grinding with 3mm and 5mm iron beads at 30 Hz in a Mixer Mill (Retsch). Ground tissue was then resuspended in 1 ml of protein extraction buffer (50 mM HEPES at pH = 7.2 with NaOH, 1.5 mM MgCl2, 1 mM EGTA, 10% Glycerol, 1% Triton X-100, 2 mM PMSF, 150 mM NaCl and one antiprotease tablet (for 50 mL of buffer)), before being centrifuged at 12,000 x g for 30 minutes, with the supernatant then taken for analysis.

Proteins were then purified via acetone precipitation at -20 °C and then dissolved in denaturation buffer (6 M guanidine hydrochloride, 50 mM Tris buffered with HCl at pH 8, 4 mM dithiothreitol). Iodoacetamide was added in order to alkylate cysteine residues and prevent the action of cysteine peptidases. A further acetone precipitation was then performed, with the resulting pellet resuspended in phosphate buffer. A 50-fold molar excess of N-acetoxy-$[^2H_3]$ succinimide was then added to induce trideuterium-labelled acetylation of free N-termini. Unwanted acetylation of oxygen containing side chains was then removed by hydrolysis through the addition of hydroxylamine.

Peptides were then transferred to digestion buffer (50 mM ammonium bicarbonate in water) using a PD-10 column and a trypsin digestion performed. The solution was then desalted using a Sep-Pak tC18 cartridge (Waters), dried and resuspended in Strong Cation Exchange (SCX) buffer (5 mM KH_2PO_4, 30% acetonitrile, and 0.05% formic acid; adjusted to pH = 3 with H_3PO_4). Peptides were then separated based on charge by SCX using an Alliance HPLC system and an increasing concentration of KCl.

The resulting fractions were then loaded into an LTQ-Orbitrap™ Velos (Thermo Scientific) mass spectrometer via a pre-column and Easy NanoLC-II system (Thermo Scientific) for mass determination. Detected protein fragments were identified using Mascot Distiller (Matrix Science) and levels of Nt-acetylation were quantified based on the ratio of trideuterium-labelled and unlabelled Nt-acetyl groups using the EnCOUNTer tool (Bienvenut et al., 2017).

Analysis of the proteome was performed in the same manner but without the trideuterium-labelled acetylation of N-termini and subsequent hydrolysis of acetyl groups added to side chains.

2.10 Gateway® recombination cloning

2.10.1 Generation of entry clones

Amplified genes of interest were incorporated into Gateway® entry vectors by either BP recombination reaction or TOPO cloning (see 2.10.1.1 and 2.10.1.2). Entry clones were then transformed into DH5α competent *E. coli* cells (see 1.11.2). Antibiotic resistant colonies were then screened by colony PCR (see 2.8.3), with positive colonies then being cultured overnight in LB broth (with addition of appropriate antibiotic). After a minimum of 16 hours of growth, entry clones were extracted from cultures by miniprep (see 2.7.3). Gene sequences were fully sequenced by Sanger sequencing to confirm the gene contained no errors before entry vectors were used in LR ligation reactions.

2.10.1.1 BP recombination reactions

Primers were designed specific to the full length of genes of interest but with an additional 31 bp (forward) or 30 bp (reverse) overhang (Table 12) including the Gateway® attB sites (underlined). PCR was performed with the proofreading Phusion® High-Fidelity DNA Polymerase (NEB) in order to generate a blunt end double stranded PCR product flanked by attB forward and reverse sequences. The products were cloned into donor vectors to generate entry clones using Gateway® BP Clonase® II enzyme mix (Thermo Fisher Scientific). BP reaction mixes containing 50 fmol of attB-flanked DNA, 150 ng of donor vector (pDONR™221) and 1 µl of Gateway® BP Clonase® II enzyme mix were incubated at room temperature for a minimum of one hour. Following incubation, 0.5 µl of Proteinase K (Invitrogen) was added, and the reaction was then incubated at 37 °C for 10 minutes.

Table 12 – BP recombination reaction primer design

Primer	Overhang sequence
Forward	GGGG<u>ACAAGTTTGTACAAAAAAGCAGGCT</u>NN(Gene of interest 5' end)
Reverse	GGGG<u>ACCACTTTGTACAAGAAAGCTGGGT</u>N(Gene of interest 3' end)

2.10.1.2 TOPO Cloning

Primers were designed specific to the full length of genes of interest but with the addition of CACC at the 5' end of the forward primer. Phusion® High-Fidelity DNA Polymerase (NEB) was to amplify the gene, preceded by CACC. The blunt ended product was incorporated into the pENTR™/D-TOPO® entry vector through the action of the associated topoisomerase enzyme using pENTR™/D-TOPO™ Cloning Kit (Thermo Fisher Scientific), according to the manufacturer's instructions.

2.10.2 Generation of expression clones by ligation reactions

Genes of interest were transferred from entry to destination vectors by LR ligation reactions containing 75 ng of entry clone, 75 ng of destination vector and 0.5 µl of Gateway® LR Clonase® II Enzyme Mix (Thermo Fisher Scientific), made up to a total volume of 4 µl with TE Buffer (10 mM Tris (pH 8.0), 1 mM EDTA). Reactions were then incubated at room temperature for a minimum of 90 minutes before termination through the addition of 0.5 µl of Proteinase K and incubation at 37 °C for 10 minutes.

The entire LR reaction mix was then used the transform competent DH5α cells (as described in 2.11.2). Resulting colonies were screened by colony PCR, with positive colonies then grown in 10 ml of LB broth (plus appropriate antibiotic) for at least 16 hours. Expression clones were then extracted by miniprep (see2.7.3) in order to be used for transformations.

2.11 Transformation Methods

2.11.1 Generation of DH5α *Escherichia coli* competent cells

A glycerol stock of DH5α cells were streaked onto an antibiotic-free LB agar plate and grown at 37 °C for 16 hours. A single colony was then inoculated into 10 ml of liquid LB and incubated at 37 °C, with 200 rpm shaking, until an OD_{600} of 0.6 to 1.0 was reached. 1 ml was then subcultured into 200 ml of SOB media (Table 13) in a 1 litre conical flask and grown at

18 °C with 200 rpm shaking until the OD_{600} had increased to ≈0.6. The conical flask was then cooled in an ice bucket for 10 mins before the culture was centrifuged at 2000 x g and 4 °C for 12 minutes, in a large benchtop centrifuge (MSE Mistral 3000i), to pellet the cells. The supernatant was discarded and the pellet resuspended in 80 ml of 4 °C TB buffer (Table 13) and then returned to the ice bucket for an additional 10 minutes. Centrifugation was repeated, as before, with the supernatant removed and the pellet resuspended in a further 12 ml of 4 °C TB buffer. DMSO was then added up to a concentration of 7.5% and gently mixed. Aliquots were then prepared by adding 100 µl to pre-chilled microcentrifuge tubes and snap-freezing in liquid nitrogen. Aliquots were stored at -80 °C to be used for transformations when required.

Table 13 – Buffers for generation of competent DH5α cells

SOB Media	TB Buffer
20 g l^{-1} Tryptone	100 mM MOPS
5 g l^{-1} Yeast extract	55 mM $MnCl_2$
10 mM NaCl	15 mM $CaCl_2$
2.5 mM KCl	250 mM KCl
10 mM $MgCl_2$	
10 mM $MgSO_4$	

2.11.2 Transformation and growth of *Escherichia coli*

Competent DH5α cells stored at -80 °C were thawed on ice. The entire contents of the BP, topoisomerase or LR reaction was added to the cells and gently mixed. Cells were then incubated on ice for 30 minutes. A 42 °C water bath was then used to heat shock the cells for exactly 30 seconds before they were returned to ice for 5 minutes. 950 µl of SOC media (SOB media, Table 13, with added 20 mM glucose) was then added to the heat shocked cells and the mixture was incubated at 37 °C with 200 rpm shaking for 1 hour. Cells were then centrifuged at 2000 x g in a microcentrifuge to produce a pellet. All but 50 µl of supernatant was then removed by pipetting and the pellet resuspended. Suspensions were then spread

across LB agar plates containing the appropriate antibiotic (see 2.5) with plates then incubated at 37 °C for approximately 16 hours.

After 16 hours, antibiotic resistant colonies were suspended in 20 µl of sterile deionized water, 2 µl of which was used in colony PCRs using MyTaq™ Red Mix (Meridian Bioscience) to screen for the gene of interest. For positive colonies, the remaining 18 µl of suspension was added to 10 ml of LB broth (containing the appropriate antibiotic) and incubated at 37 °C with 200 rpm shaking for approximately 16 hours.

2.11.3 Transformation and growth of *Saccharomyces cerevisiae*

S. cerevisiae cells were transformed via a lithium acetate mediated method. 2 µg of expression clone was added to 100 µl of transformation buffer in a microcentrifuge tube (Table 14). A 1 µl inoculation loop of yeast (freshly grown on YPD) was suspended in the buffer and briefly vortexed. Tubes were then mounted at a 45 °C angle and incubated at 37 °C for 45-60 minutes with 200 rpm shaking. Suspensions were then spread on the appropriate DO media plates which were then incubated at 30 °C for a minimum of 48 hours to allow the growth of transgenic colonies. Colonies were screened by extracting DNA (see 2.7.2) followed by PCR.

Table 14 – Yeast transformation buffer

Transformation Buffer
33% PEG_{3350}
0.33 M Lithium Acetate
0.66% β-mercaptoethanol

2.11.4 Generation of GV3101 *Agrobacterium tumefaciens* competent cells

A glycerol stock of the GV3101 *A. tumefaciens* strain, which contains the helper plasmid pMP90, was streaked onto a LB agar plate with added rifampicin and gentamycin and incubated at 28 °C for at least 48 hours. 10 ml of liquid LB (plus rifampicin and gentamycin) was inoculated with one of the resulting colonies and grown at 28 °C with 200 rpm shaking. After 24 hours of growth, 1 ml was subcultured into a further 200 ml of LB (plus rifampicin and

gentamycin) and allowed to continue growing at 28 °C with 200 rpm shaking for 18 hours. The culture was then transferred to chilled 50 ml Falcon® tubes and the cells pelleted by centrifugation at 4000 rpm and 4 °C for 15 minutes (MSE Mistral 3000i). Following removal of the supernatant, pellets were resuspended in a total volume of 10 ml of 4 °C TE buffer (10 mM Tris (pH 8.0), 1 mM EDTA) before centrifugation was repeated (as previous). The supernatant was again discarded and the pellet resuspended in 20 ml of 4 °C LB liquid media and aliquoted (200 µl) into pre-chilled microcentrifuge tubes, which were then snap-frozen in liquid nitrogen and stored and -80 °C.

2.11.5 Transformation and growth of *Agrobacterium tumefaciens*

The freeze-thaw method of transformation was used to introduce plant transformation vectors into *A. tumefaciens* (Weigel and Glazebrook, 2006). 200 µl of competent *A. tumefaciens* cells were thawed on ice and 1 µg of expression clone added and mixed by gentle agitation. Cells were then incubated on ice for 5 minutes before being placed into liquid nitrogen for 5 minutes followed by a 37 °C water bath for a further 5 minutes. Subsequently, 1 ml of LB was added and the mixture incubated at 28 °C with 200 rpm shaking for 2 hours. Cells were pelleted by 1 minute of centrifugation at 2000 x g in a microcentrifuge and 1 ml of the supernatant discarded. The pellet was resuspended in the remaining 200 µl of LB by gentle pipetting and then added to LB agar plates with appropriate antibiotics added. Plates were incubated at 28 °C for 48-72 hours and the resulting colonies were screened by colony PCR (see 2.8.3).

2.11.6 Transformation and growth of *Arabidopsis thaliana*

Arabidopsis plants were transformed by floral dip (Clough and Bent, 1998) based on the method described by Zhang et al. (2006).

Successfully transformed *A. tumefaciens* colonies were inoculated into 10 ml of LB liquid media, containing rifampicin, gentamycin and vector-specific antibiotics, and placed at 28 °C with 200 rpm shaking until a high level of growth was achieved (16 to 48 hours). 2 ml of

this culture was then added to a further 200 ml of LB liquid media containing the same antibiotics. This was allowed to grow for 16 to 24 hours in the same conditions until an OD_{600} of 1.5 to 2.0 was reached, at which point cultures were centrifuged at 4000 x g for 10 minutes in a large, free-standing centrifuge (Sorvall RC-26 Plus). The supernatant was removed and the pellet resuspended in 400 ml of 5% sucrose solution. 80 µl of Silwett L77 was added and mixed to produce the final dipping solution.

Arabidopsis plants were clipped shortly after bolting to induce the formation of multiple secondary bolts. Plants with secondary bolts of approximately 5 to 10 cm in height, were used for floral dipping. Suitable plants were inverted and aerial tissues were submerged into the dipping solution with gentle agitation for around 10 seconds. Plants were then removed from the dipping solution and laid horizontally inside plastic bags to maintain high humidity and kept in darkened conditions for around 48 hours. After his period, plants were unwrapped and returned to standard growth conditions to produce seeds (see 2.6.1.1).

T1 seeds were screened on ½ MS media with appropriate antibiotics or herbicides added. In cases where the inserted T-DNA provided Basta resistance, T1 selection was carried out on soil soaked in a BASTA-based herbicide: Harvest® (Bayer Crop Science). Concentrated Harvest® was diluted 1:1000 to produce a glufosinate-ammonium concentration of 150 mg l^{-1} prior to soil soaking.

Antibiotic-resistant individuals were transplanted and grown in standard growth conditions. The next generation of seeds (T2) were then screened on ½ MS plates in order to identify seed-sets with a 3:1 ratio of resistance:susceptibility to the selective agent. Such seed sets can be identified as those produced by a parent plant containing only a single T-DNA insertion. A selection of resistant individuals were then transplanted to soil and T3 seeds were screened to identify homozygous lines, where all T3 plants demonstrated resistance to the selective agent.

Chapter III -

Evaluation of AT4G34100 and AT4G32670 as homologues

of the yeast ERAD component ScDOA10

3.1 Introduction

In yeast, the ER-associated RING E3 ligase, ScDoa10, has been implicated in both ERAD and the Ac/N-degron pathway of targeted proteolysis (see 1.2.3 and 1.4.3) (Swanson et al., 2001; Hwang et al., 2010). Basic Local Alignment Search Tool (BLAST) searches performed by Liu et al. (2011) identified AT4G34100 (*AtDOA10A*) and AT4G32670 (*AtDOA10B*) as putative homologues of ScDoa10 in *Arabidopsis*. Located only 0.57 Mb apart on chromosome four, the genomic DNA sequences of *AtDOA10A* and *AtDOA10B* are 4275 bp and 3321 bp in length respectively (start codon to stop codon). *AtDOA10A* and *AtDOA10B* both contain 9 annotated exons and produce proteins 1108 aa (AtDOA10A) and 860 aa (AtDOA10B) in length. By comparison, ScDoa10 contains 1319 aa's. Mutations of *AtDOA10A* are reported to induce upregulation of some UPR-related genes (Lu et al., 2012; Li et al., 2017) and enhance thermotolerance, when combined with mutations of the other ERAD E3 ligases *AtHRD1A* and *AtHRD1B* (Li et al., 2017), thereby supporting its perceived role in ERAD. Further impacts of the *Atdoa10a* mutation include alterations in cuticular wax composition, ABA hypersensitivity and the suppression of the drought-hypersensitive phenotype of the *Atsqe1* sterol-synthesis mutant (Jenks et al., 1995; Lu et al., 2012; Doblas et al., 2013; Zhao et al., 2014). The sterol-synthesis pathway is known to be regulated by ERAD in humans and yeast (Loertscher et al., 2006; Jo and DeBose-Boyd, 2010; Foresti et al., 2013; Zelcer et al., 2014). Whilst this limited evidence is consistent with a role in ERAD for AtDOA10A, links between AtDOA10B and ERAD have not yet been explored, although *AtDOA10B* has been identified as strongly upregulated in response to treatment with the protein misfolding agent azetidine-2-carboxylic acid (AZC) (Kim et al., 2017). The lack of an identified substrate of either AtDOA10A or AtDOA10B means that their participation in a hypothesised *Arabidopsis* Ac/N-degron pathway is yet to be confirmed.

This chapter aims to analyse AtDOA10A and AtDOA10B and evaluate their similarity to ScDoa10 using both bioinformatics and an experimental approach to assess their properties and functions.

3.2 AtDOA10A and AtDOA10B share high protein sequence homology with ScDoa10

The key indicative feature of an E3-ligase is the presence of an E2 binding domain within its protein sequence (see 1.1.3). ScDoa10 contains a RING domain between its N-terminus and first transmembrane domain (Kreft et al., 2006), although this domain does not fully conform to the recognised RING consensus sequence. The $C-X_2-C-X_{10-45}-C-X-C-X_7-H-X_2-C-X_{11-25}-C-X_2-C$ sequence of ScDoa10's E2 binding domain is known as a RING-CH variant domain, due to the positional changes of the zinc-coordinating Cys (C) and His (H) amino acids relative to the classical RING-HC domain ($C-X_2-C-X_{9-39}-C-X_{1-3}-H-X_{2-3}-C-X_2-C-X_{4-48}-C-X_2-C$). Analysis of the amino acid sequences of AtDOA10A and AtDOA10B was performed using The SMART protein domain annotation resource (http://smart.embl-heidelberg.de/) (Letunic and Bork, 2018) to confirm the presence of such a domain within the *Arabidopsis* proteins. Both E3s were predicted to contain RING-CH domains in the same location as ScDoa10; between the N-terminus and first transmembrane domain. Protein sequence alignments, studied using EMBL-EBI Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/) (Sievers et al., 2011), also showed a high degree of similarity between these RING-CH domains and that of ScDoa10. The RING-CH domains of both AtDOA10A and AtDOA10B were shown to be 48.98% identical to ScDoa10, including complete conservation of all zinc-binding Cys and His amino acids (Figure 12a). In comparison, this is a greater level of conservation than the 44.90% identity ScDoa10 shares with its well-studied human homologue HsMARCH6 at the RING-CH domain (Swanson et al., 2001). The RING-CH domains of AtDOA10A and AtDOA10B share 57.14% identity (Figure 12a).

Another significant region within the protein sequence of ScDoa10 is the so-called TD domain, which has been shown to aid the degradation of the ScUbc6 which is a substrate of ScDoa10 as well as being an associated E2 enzyme (see 1.5.1) (Kreft and Hochstrasser 2011). The TD domain has also been shown to be well conserved across several species (e.g. 26.72% identity between ScDoa10 and HsMARCH6) (Swanson et al., 2001). Alignments of

this domain with AtDOA10A and AtDOA10B revealed that they are 32.06% and 21.31% identical to ScDoa10 respectively and 35.66% identical to each other (Figure 12b).

More recently, a 16-residue motif close to the C-terminus of ScDoa10, thought to be important in the recognition of certain of substrates, has also been identified as being well conserved across multiple orthologs (Zattas et al., 2016). Five of these 16 residues (31.25%) are conserved between ScDoa10 and HsMARCH6 with a further 4 highly similar residues (Zattas et al., 2016). When aligned, AtDOA10A shows a 37.50% identity compared to ScDoa10 and AtDOA10B a 25.00% identity. AtDOA10A and AtDOA10B share 43.75% of the same amino acids in this region (Figure 12c).

The full-length amino acid sequences for AtDOA10A and AtDOA10B are 22.87% and 19.23% identical to ScDoa10 respectively (Appendix 2).

These alignments suggest that AtDOA10A and AtDOA10B are closely related to ScDoa10, with particularly high levels of conservation in key protein domains, such as the RING-CH domain. At the RING-CH domain, the TD domain and the 16 amino acid C-terminal motif, ScDoa10 shares a greater similarity with AtDOA10A than it does with HsMARCH6, a well-characterised ScDoa10 homologue that is able to target the same substrates (Foresti et al., 2013; Zelcer et al., 2014). This therefore suggests that AtDOA10A is very likely to be a true homologue of the yeast E3 ligase. The similarity to ScDoa10 of AtDOA10B also exceeds that of HsMARCH6 at the RING-CH domain, however, AtDOA10B is significantly shorter than ScDoa10 and conservation is weaker within the TD domain and C-terminal motif. The increased divergence of these regions, but conservation of the RING-CH domain may suggest that AtDOA10B can function similarly as a RING-CH E3 ligase but has evolved to target a different range of substrates.

a

i

```
ScDoa10    38  TCRICRGEATEDNPLFHPCKCRGSIKYMHESCLLEWVASKNIDISKPGADVKCDICH
AtDOA10A   67  VCRICRNPGDADNPLRYPCACSGSIKFVHQDCLLQWLNHSNA-------RQCEVCK
AtDOA10B   32  ICRICQSPEEPDNPLRHPCACRGSLKYIHSDCLFLWLNRRKR-------NHCEICK
```

ii

		ScDoa10
	AtDOA10A	48.98%
AtDOA10B	57.14%	48.98%

b

i

```
ScDoa10   625  KVFTLFFIELAGFPILAGVMLDFSLFCPILASNSRMLWVPSICAIWPPFSLFVYWTIGTL
AtDOA10A  524  KVAFLLVIELGVFPLVCGWWLDVCTVRMF---GKTMSHRVQFLSISPLASSLVHWVVGIM
AtDOA10B  349  KDTFILSFKLGVLPWLLGCWLHFCTFPIL---GKTASHTVEVLSDYPLMA-DKHWLMGTL

ScDoa10   685  YMYWFAKYIGMIRKNIIRPGVLFFIRSPEDPNIKILHDSLIHPMSIQLSRLCLSMFIYAI
AtDOA10A  581  YMLQISIFVSLLR-GVLRPGVLYFLRDPADPNYNPFRDLIDDPVHKHARRVLLSVAVYCS
AtDOA10B  405  YLVSALSCMELIQ-KIVQKRALWYILDVAFPNYKVTKL-------HLGPILLAFALHGT

ScDoa10   745  FIVLGFGFH-------TRIFFP
AtDOA10A  640  LIVMLVFLPVKLAIRMAPSIFP
AtDOA10B  456  MVVIVLHLPIKTISLISQSFFP
```

ii

		ScDoa10
	AtDOA10A	32.06%
AtDOA10B	35.66%	21.31%

c

i

```
ScDoa10   1301  VKDEVYTKGRALENLP
AtDOA10A  1051  IRDDRYLIGRRLHNFG
AtDOA10B   841  EFDNRYKVGERLVDFT
```

ii

		ScDoa10
	AtDOA10A	37.50%
AtDOA10B	43.75%	25.00%

Figure 12 – Amino acid sequence alignments of functional domains of ScDoa10, AtDOA10A and AtDOA10B
(i) Sequence alignments of amino acids within (a) RING-CH domain, (b) TD-domain and (c) 16 aa C-terminal motif. Black shading signifies identical amino acids and grey shading shows similar amino acids. (ii) Percentage identity values between proteins for each domain.
Alignments were generated using EMBL-EBI Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/) (Sievers et al., 2011) with shading added using BoxShade 3.21 (https://embnet.vital-it.ch/software/BOX_form.html).

3.3 AtDOA10A and AtDOA10B are structurally similar to ScDoa10

In addition to similarly positioned RING-CH domains, the *Arabidopsis* homologues are also predicted to share a similar overall structure to ScDoa10. ScDoa10 is located at the endoplasmic reticulum and has been shown experimentally to contain 14 transmembrane domains, with both termini and several large unstructured loops situated on the cytosolic side of the membrane (see 1.5.1, Figure 11) (Kreft et al., 2006). The number of transmembrane domains in ScDoa10 can be correctly predicted using membrane topology prediction programs such as TMHMM 2.0 (http://www.cbs.dtu.dk/services/TMHMM/) (Krogh et al., 2001). The distribution of the 14 transmembrane domains predicted by this program, however, does not fully match the experimentally derived topology.

To assess conservation of structure between ScDoa10, AtDOA10A and AtDOA10B, TMHMM 2.0 was used to predict the membrane topologies of the *Arabidopsis* proteins. Like ScDOA10 a high density of transmembrane domains were also predicted for AtDOA10A and AtDOA10B, with 14 and 13 transmembrane domains predicted respectively (Figure 13 and Appendix 3). The predicted 14 transmembrane domains and cytosolic RING-CH domain of AtDOA10A, matches the experimentally determined number of transmembrane domains of ScDoa10, thereby suggesting that they are likely to share a similar structure. With one fewer transmembrane domain predicted for AtDOA10B, the TMHMM 2.0 model predicts a luminal N-terminus; a potentially significant divergence from the structure of ScDoa10. However, due to the high chance of prediction errors, especially with proteins containing such a high density of transmembrane domains, the likelihood that both *Arabidopsis* homologues contain 14 transmembrane domains remains high. Alternative membrane topology models, such as TMpred (https://embnet.vital-it.ch/software/TMPRED_form.html) (Hofmann and Stoffel, 1993) and Phobius (http://phobius.sbc.su.se/index.html) (Kall et al., 2004), predicted the presence of the expected 14 transmembrane domains in AtDOA10B. It should be noted, though, that these models also predicted increased numbers of transmembrane domains for ScDoa10 and AtDOA10A (Table 15).

Therefore, as with comparisons of amino acid sequences (see 3.2), membrane topology predictions also suggested a greater similarity between ScDoa10 and AtDOA10A rather than AtDOA10B.

Table 15 – DOA10 homologue membrane topology predictions

Prediction Model	Predicted number of Transmembrane domains		
	ScDoa10	AtDOA10A	AtDOA10B
TMHMM 2.0	14	14	13
TMpred	14	16	14
Phobius	16	16	14
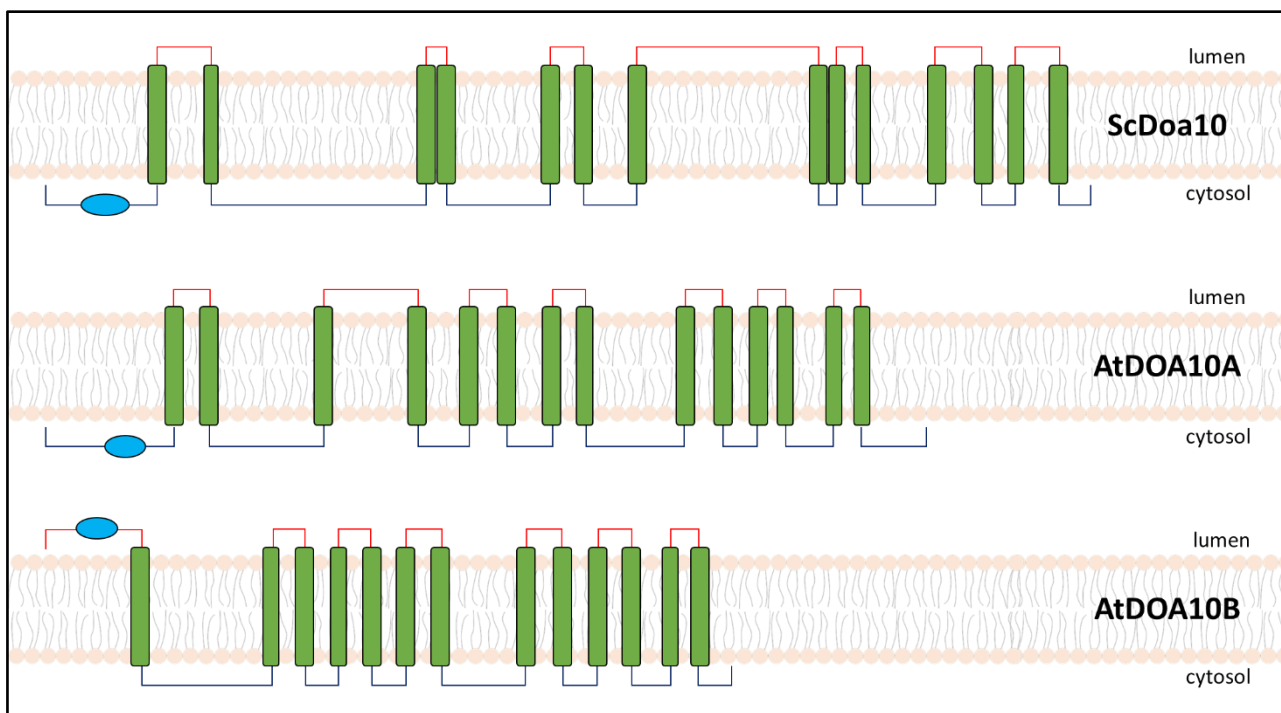
Figure 13 – Predicted distribution of DOA10 homologue transmembrane domains
To scale representations of membrane topologies of ScDoa10, AtDOA10A and AtDOA10B as predicted by TMHMM 2.0 (http://www.cbs.dtu.dk/services/TMHMM/) (Krogh et al., 2001) (for outputs see appendix 3). Transmembrane domains are represented by green boxes and cytosolic and luminal loops by blue and red links respectively. Blue ovals represent the predicted location of the RING-CH domain within each protein.

3.4 AtDOA10B may be specific to the *Brassicaceae* family

During the analysis of the *Arabidopsis* homologues, it was noted that whilst many plant species possessed well conserved homologues of AtDOA10A, conservation of AtDOA10B was more variable. BLAST searches were performed using the amino acid sequences of AtDOA10A and AtDOA10B submitted to NCBI BLAST (https://blast.ncbi.nlm.nih.gov/Blast.cgi). From this, the most similar protein, according to the percentage of identical amino acids, was identified for a selected 8 plant species (Tables 16 and 17). The strong conservation of AtDOA10A-like proteins is demonstrated by the 68-98% of amino acids that they share with AtDOA10A. For AtDOA10B however, the most similar proteins in some species are only 30% identical. In fact, in the non-*Brassicaceae* species analysed, the closest protein relatives to AtDOA10B were all much more similar to AtDOA10A (Table 17), despite the presence of multiple DOA10-like proteins in some cases. In *Glycine max*, for example, two ScDOA10 homologues have been identified (Glyma02g11570.1 (XP_006574896) and Glyma07g33840.1 (XP_006583895)) but both are more closely related to AtDOA10A (Silva et al., 2015).

This raises the possibility that AtDOA10B-like proteins may be specific to the *Brassicaceae* family. However, even within this family conservation of the AtDOA10B sequence is much weaker than for AtDOA10A. In *Capsella rubella* for example, the closest relative of AtDOA10B only shares 48 % of its amino acids with AtDOA10B. In contrast, its AtDOA10A homologue is 96 % identical to AtDOA10A. This may indicate a higher level of divergence in DOA10B substrates between species or show that DOA10As are subject to selection pressure ensuring their strict conservation, potentially implying that they fulfil crucial functions.

Table 16 – AtDOA10A-like (closest relative) proteins from a range of plant species		% identical (AtDOA10A)	% identical (AtDOA10B)
Brassicaceae	**Arabidopsis lyrata** XP_002869176 (probable E3 ubiquitin ligase SUD1)	**98**	31
	Capsella rubella XP_006283036 (probable E3 ubiquitin ligase SUD1)	**96**	31
	Brassica rapa XP_009117224 (probable E3 ubiquitin ligase SUD1)	**93**	31
	Brassica napus XP_013657810 (probable E3 ubiquitin ligase SUD1)	**93**	31
Non-Brassicaceae	**Gossypium hirsutum** XP_016723032 (probable E3 ubiquitin ligase SUD1)	**79**	31
	Glycine max XP_006574896 (probable E3 ubiquitin ligase SUD1)	**78**	30
	Medicago truncatula XP_003607923 (probable E3 ubiquitin ligase SUD1)	**77**	31
	Hordeum vulgare KAE8776302 (putative E3 ubiquitin ligase SUD1)	**68**	30

Table 17 – AtDOA10B-like (closest relative) proteins from a range of plant species		% identical (AtDOA10A)	% identical (AtDOA10B)
Brassicaceae	**Arabidopsis lyrata** XP_020875327 (probable E3 ubiquitin ligase SUD1)	30	**89**
	Capsella rubella EOA22078 (hypoth. protein CARUB_v10002618mg)	30	**48**
	Brassica rapa XP_033135275 (probable E3 ubiquitin ligase SUD1)	29	**66**
	Brassica napus XP_013666703 (probable E3 ubiquitin ligase SUD1)	29	**66**
Non-Brassicaceae	**Gossypium hirsutum** XP_016723032 (probable E3 ubiquitin ligase SUD1)	**79**	31
	Glycine max XP_006574896 (probable E3 ubiquitin ligase SUD1)	**78**	30
	Medicago truncatula XP_003607923 (probable E3 ubiquitin ligase SUD1)	**77**	31
	Hordeum vulgare KAE8776302 (putative E3 ubiquitin ligase SUD1)	**68**	30

3.5 AtDOA10A and AtDOA10B are widely expressed in *Arabidopsis* tissues

Due to the importance of protein quality control within all cells, genes coding for ERAD components are broadly expressed in all plant tissues throughout development. Such widespread expression has been observed by tissue-specific RT-PCR analysis of the ERAD genes *AtOS9*, a putative subunit of the *Arabidopsis* HRD1 complex, and *AtUBC32*, a homologue of yeast *ScUbc6* (Cui et al., 2012a; Hüttner et al., 2012) and is backed up by freely-available microarray data for other ERAD components via the *Arabidopsis* eFP Browser 2.0 database (http://bar.utoronto.ca/efp2/Arabidopsis/Arabidopsis_eFPBrowser2.html) (Winter et al., 2007). To establish whether this was also the case for *AtDOA10A* and *AtDOA10B*, their level of expression in a range of different tissues was assessed by quantitative RT-PCR (qRT-PCR). RNA was extracted from samples of 7-day old shoots and roots, as well as from the rosette leaves, cauline leaves, stem, inflorescences and siliques of mature plants. cDNA was synthesised from the extracted RNA and qRT-PCR was then performed to determine the abundance of each transcript (primers 101+102, 1+2 and 78+79, see Appendix 1).

Transcripts of both genes were detected in a relatively ubiquitous manner, although there was some variation between different tissues (Figure 14). These slight differences were likely influenced by variations in the expression level of the *AtACTIN7 (AtACT7)* gene used as a housekeeping control and therefore may be less significant than they appear. For example, *AtACT7* has been reported to have lower expression in leaves than in the stem (McDowell et al., 1996). The universal expression patterns observed were consistent with those reported in the TraVA RNA-seq database, available at http://travadb.org/ (Klepikova et al., 2016), where both genes are shown to be relatively uniformly expressed, although significant elevations of the *AtDOA10A* and *AtDOA10B* transcripts are reported in some tissues that were not tested here, such as dry seeds. The distribution of *AtDOA10A* expression observed by qRT-PCR also matched that reported by *Arabidopsis* eFP Browser 2.0 (Winter et al., 2007). *AtDOA10B* is not listed within the *Arabidopsis* eFP Browser 2.0 database.

The TraVA database was also used to compare the gene expression levels of *AtDOA10A* and *AtDOA10B*. Read counts showed that *AtDOA10B* expression was much lower than that of *AtDOA10A* (17.4%). Using template serial dilutions, qRT-PCR reaction efficiencies were calculated to allow the comparison of *AtDOA10A* and *AtDOA10B* transcript abundance in the seven tissues analysed here by qRT-PCR. The average relative expression level of *AtDOA10B* was found to be even lower than that suggested by the TraVA database at just 3.03 ± 0.30 % of the level of *AtDOA10A* expression. The lower relative expression level of *AtDOA10B* calculated by qRT-PCR was likely to be a result of the limited number of tissue types selected for analysis. For example, seeds, reported to have elevated AtDOA10B expression, were not analysed by qRT-PCR.

To analyse the distribution of these E3s at the protein level, β-glucuronidase (GUS) reporter lines were generated through the transformation of *Atdoa10a* and *Atdoa10b* KO plants with the *AtDOA10A/B* genomic DNA sequence followed by the GUS gene (see 4.2.2 for generation of lines). A ≈2kb sequence of DNA from immediately upstream of *AtDOA10A/B* was also included to include the gene's native promoters. The generation of these complementation lines allowed histochemical staining to be performed in order to visualise the location of the GUS tagged proteins in *Arabidopsis* whole seedlings and individual organs (see 2.6.1.4). Incubation of these complementation lines in X-gluc solution for 24 hours produced widespread staining, especially in regions of rapid growth and high levels of protein production.

Both AtDOA10A-GUS and AtDOA10B-GUS lines exhibited a similar pattern of staining, with apical and root meristems in particular showing a strong signal (Figures 15 and 16). As expected, based on the lower transcript levels of *AtDOA10B* (Figure 14), observed staining was generally both more intense and more widespread in AtDOA10A-GUS lines compared to AtDOA10B-GUS lines. Staining of both proteins appeared stronger in younger tissues, with weaker staining observed in 14-day old seedlings in comparison to 7-day old seedlings. Young rosette leaves also showed more staining than older rosette leaves, although AtDOA10B-GUS staining was very weak even in young rosette leaves (Figures 15g

and 16g). AtDOA10B appeared completely absent from older rosette leaves, whilst AtDOA10A produced some small patches of staining in older leaf tissue (Figures 15h and 16h). In contrast to qRT-PCR results, cauline leaves were largely free from staining for both AtDOA10A and AtDOA10B (Figures 15i and 16i), which may indicate a significant role of posttranslational regulation. Staining of both AtDOA10A and AtDOA10B was observed in inflorescences, with developing seeds in particular being strongly stained (Figures 15k/j and 16k/j). The contribution of free-GUS to overall staining should also be considered, as free-GUS, resulting from the cleavage of AtDOA10A/B-GUS, was detected in each of these lines (Figure 22).
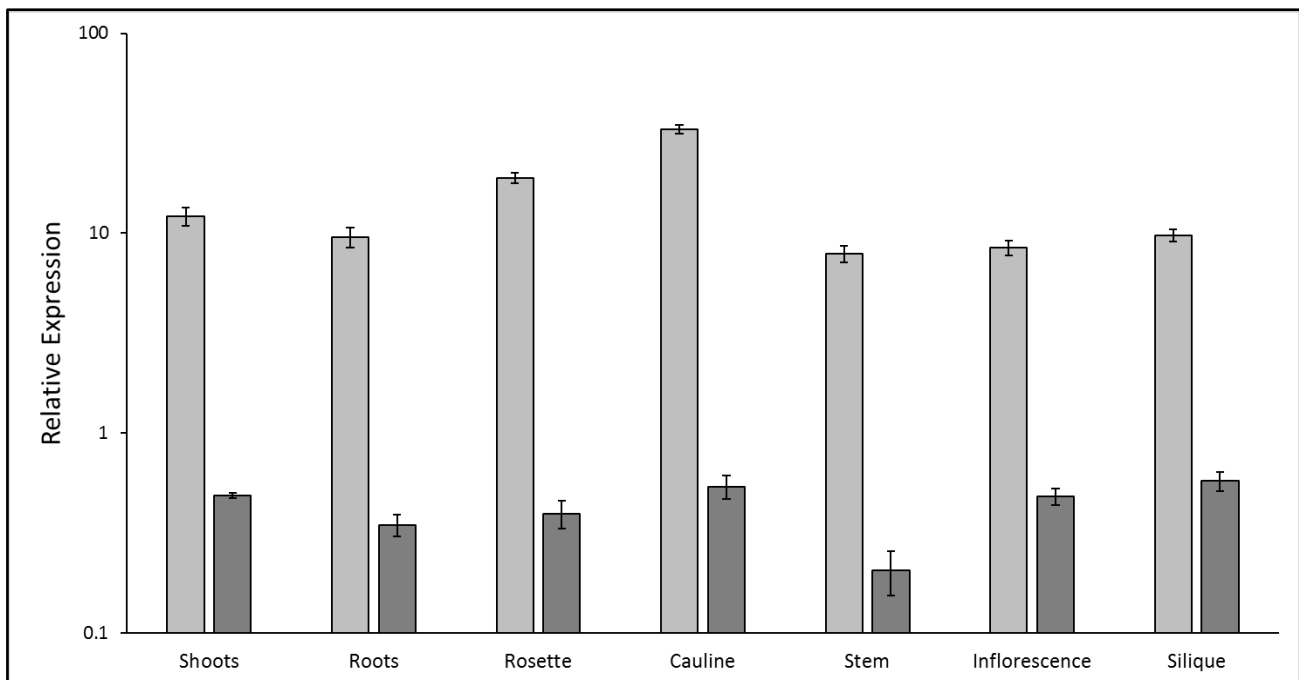
Figure 14 – *AtDOA10A* and *AtDOA10B* expression atlas
AtDOA10A (light grey) and *AtDOA10B* (dark grey) are broadly expressed throughout *Arabidopsis* tissues. *AtDOA10B* expression is only 3.03 ± 0.3 % that of *AtDOA10A*. Shoots = green leaves of 7-day old seedlings, Roots = roots of 7-day old seedlings, Rosette = rosette leaf harvested at bolting, Cauline = second cauline leaf harvested after the growth of the third cauline leaf, Stem = stem tissue taken from 3cm below inflorescence, Inflorescence = entire inflorescence, Silique = mature (full length and filled but still green) siliques, stage 17 (Smyth et al., 1990). Relative expression levels were calculated through normalisation to *AtACT7* (AT5G09810) and are the average of three biological repeats. Values of relative expression for each tissue are given as a relative percentage (±SE) of the total expression of *AtDOA10A* from all tissues. Differences in expression level between *AtDOA10A* and *AtDOA10B* were calculated via the delta delta C_t method using C_t values adjusted for differences in reaction efficiency (C_{tadj}) (*AtDOA10A* reactions = 96.70%, *AtDOA10B* reactions = 97.70%) according to the formula: $C_{tadj} = C_t \times \log_2(1+\text{efficiency})$. Values of relative expression are plotted on a $\log_{10}$ scale.
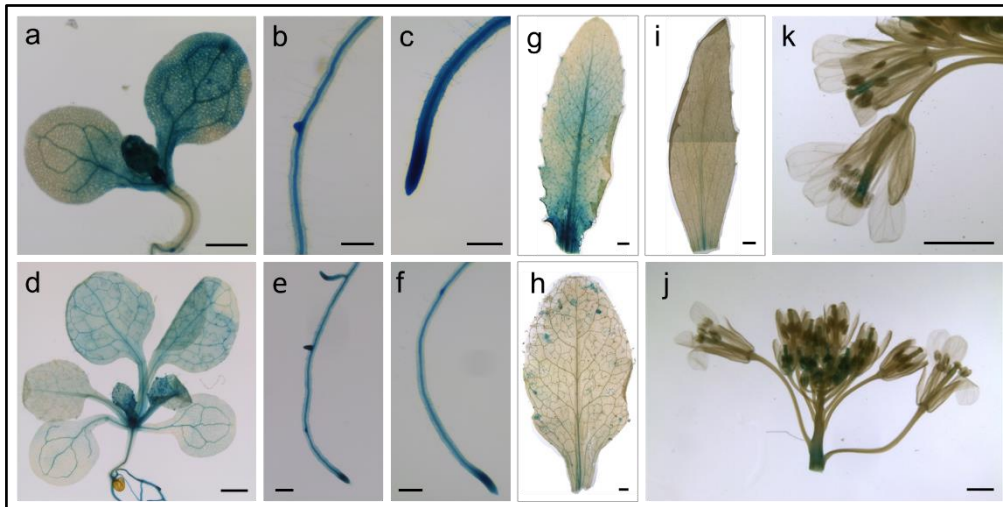
Figure 15 – Histochemical staining of AtDOA10A-GUS reporter lines
Histochemical staining pattern of transgenic lines expressing *pAtDOA10A::AtDOA10A-GUS*
produced after 24 hours of exposure to GUS stain solution. (a) cotyledons of 7-day old seedling. (b
and c) roots of 7-day-old seedlings. (d) leaves of 14-day old seedling. (e and f) roots of 14-day old
seedlings. (g) young rosette leaf. (h) old rosette leaf. (i) cauline leaf. (j and k) inflorescence.
Representative images from eight seedlings or six mature plants from two independent
AtDOA10A-GUS reporter lines. Scale bars = 1mm.
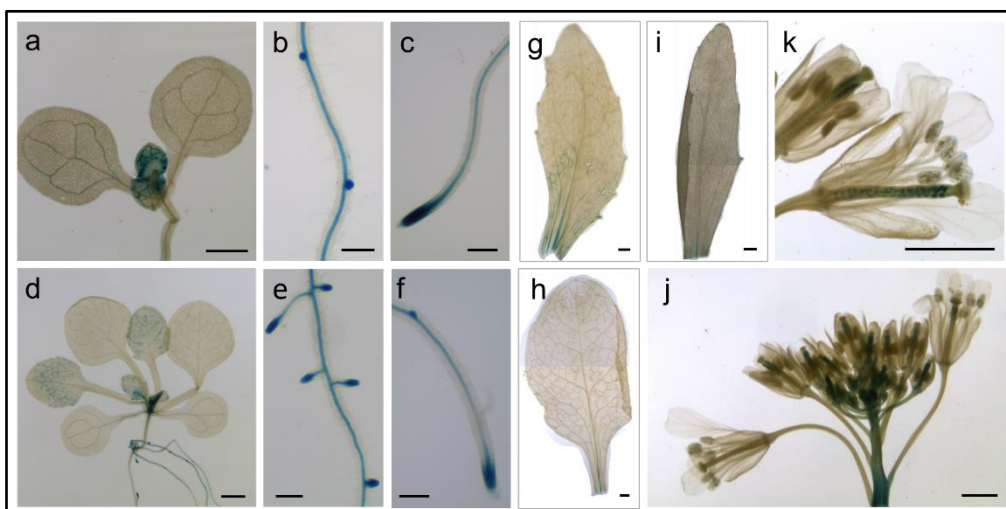
Figure 16 – Histochemical staining of AtDOA10B-GUS reporter lines
Histochemical staining pattern of transgenic lines expressing *pAtDOA10B::AtDOA10B-GUS*
produced after 24 hours of exposure to GUS stain solution. (a) cotyledons of 7-day old seedling. (b
and c) roots of 7-day-old seedlings. (d) leaves of 14-day old seedling. (e and f) roots of 14-day old
seedlings. (g) young rosette leaf. (h) old rosette leaf. (i) cauline leaf. (j and k) inflorescence.
Representative images from eight seedlings or six mature plants from two independent
AtDOA10B-GUS reporter lines. Scale bars = 1mm.

3.6 AtDOA10A and AtDOA10B are localised to the endoplasmic reticulum

To accompany expression analyses, it was also important to determine the subcellular localisation of AtDOA10A and AtDOA10B to provide insight into their cellular functions. The subcellular localisation of the GUS-fusion proteins was determined using cell fractionation (see 2.9.8). Protein extracts were precleared by centrifugation at 630 x g for 10 minutes to remove cell debris and large organelles such as nuclei. The resulting supernatant was centrifuged at 21000 x g for 2 hours yielding microsomal (pellet) and soluble (supernatant) fractions according to the method described by Abas and Luschnig (2010). To fulfil their perceived role as part of the ERAD system, AtDOA10A and AtDOA10B were expected to be found within the microsomal fraction, as has been shown previously for ScDOA10 and HsMARCH6 (Swanson et al., 2001; Hassink et al., 2005).

As expected, Western blot analysis showed that both AtDOA10 homologues were present in microsomal fractions (Figure 17), colocalising with the ER marker proteins Calnexin1 and Calnexin2 (AtCNX1 and AtCNX2). The cytosolic marker AtUGPase was detectable only in soluble fractions. The co-localisation of AtDOA10A/B-GUS with AtCNX1/2 was consistent with the anticipated ER-localisation and therefore a role in ERAD. This is further supported by confocal microscopy work carried out by the Markus Wirtz Laboratory, Universität Heidelberg, Germany using EYFP-DOA10A fusion proteins expressed in *Nicotiana benthamiana* (Huber et al., 2022).

By comparing the strength of the bands relative to the strength of the AtCNX1/2 band produced by the same sample, clear differences in protein abundance were observed. AtDOA10A-GUS produced a significantly stronger band than AtDOA10B-GUS (Figure 17), re-affirming qRT-PCR and GUS staining results from 3.5 showing that *AtDOA10A* is much more strongly expressed that *AtDOA10B*.
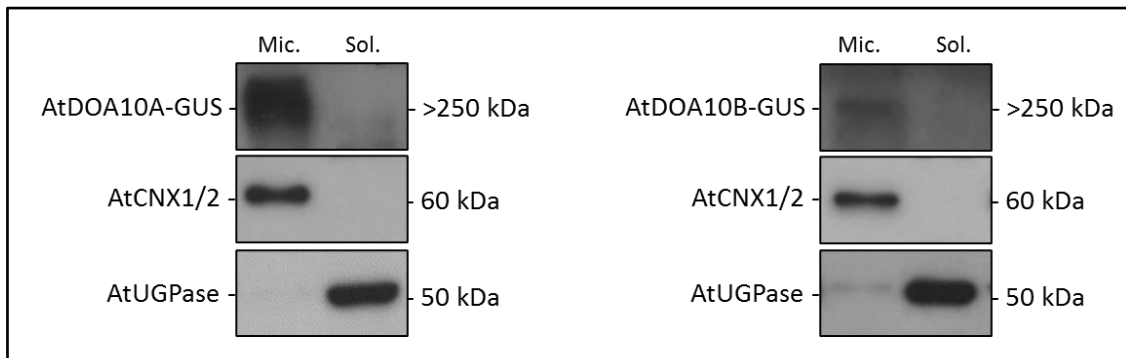
Figure 17 – AtDOA10A/B-GUS cell fractionation
Following cell fraction AtDOA10A-GUS and AtDOA10B-GUS fusion proteins were detected in the microsomal fraction, along with the ER-marker AtCNX1/2 and separate from AtUGPase which was found in the soluble fraction. Proteins were detected using anti-GUS, anti-AtCNX1/2 and anti-AtUGPase antibodies (see 2.4.1).

3.7 Expression levels of *AtDOA10A* and *AtDOA10B* are responsive to protein misfolding stress

The production of misfolded proteins due to the application of an external stress is known to alter the gene expression of a range of *Arabidopsis* E3 ligases to facilitate their destruction and avoid the formation of misfolded protein aggregates (Kim et al., 2017). With the degradation of misfolded proteins being a key function of the ERAD process, a strong upregulation of ERAD E3s in response to protein misfolding stress is to be expected. Indeed, RNA-seq analysis following treatment with the proline analog azetidine-2-carboxylic acid (AZC), which is known to induce protein misfolding (Fowden and Richmond, 1963; Trotter et al., 2002), identified the ERAD ubiquitin ligase genes *AtHRD1B* and *AtDOA10B* as two such examples, with their expression increasing 9.0- and 11.0-fold respectively (Kim et al., 2017). Surprisingly, however, the expression levels of the other AtHRD1 and AtDOA10 homologues were relatively unaffected by the treatment (*AtHRD1A* up 1.1-fold, *AtDOA10A* up 1.3-fold).

To validate these observations, we monitored the responses of *AtDOA10A* and *AtDOA10B* transcripts during the activation of the UPR, using the alternative ERAD-stressor tunicamycin (TM). TM is an inhibitor of N-linked glycosylation that is used routinely to induce protein misfolding and the UPR in *Arabidopsis* (Chen and Brandizzi, 2013). The misfolding stress was applied by subjecting 6-day old seedlings to a 4-hour treatment with 5 μg ml^{-1} of

TM, or an equivalent volume of DMSO solvent as a control. Following the treatment, RNA was extracted from the seedlings and qRT-PCR was performed to assess the abundance of the *AtDOA10A* and *B* transcripts (primers 101+102 and 1+2).

The abundance of AtDOA10B mRNA was significantly elevated (4.99-fold) in response to TM, relative to the DMSO control treatment, mirroring the upregulation observed by Kim et al. (2017) in response to AZC. Conversely, levels of the *AtDOA10A* transcript were significantly reduced (0.64-fold), providing an indication that AtDOA10B may play a more important role than AtDOA10A in the degradation of misfolded proteins.
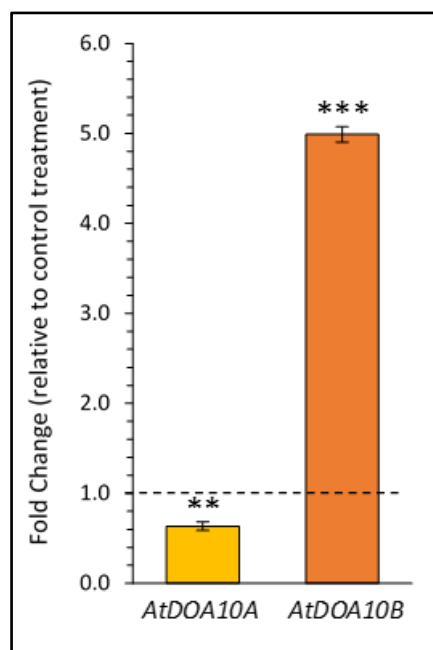
Figure 18 – Gene regulation of *AtDOA10A* and *AtDOA10B* during the UPR
Fold changes of *AtDOA10A* and *AtDOA10B* expression in 10x 6-day old seedlings following 4-hour treatment with tunicamycin, as determined by qRT-PCR, relative to DMSO-treated seedlings. Fold changes were calculated according to the delta delta C_t method and data presented are the averages of 3 biological repeats (±SE). Statistical significance was determined using a student's t-test.

3.8 UPR genes are not upregulated in *Atdoa10* mutants

Whilst the expression levels of *AtDOA10A* and *AtDOA10B* were shown to be altered during the UPR, confirming a relationship between these genes and protein misfolding, this did not provide any indication of their relative importance in the ERAD process. Impeding ERAD function can cause misfolded proteins to accumulate, leading to activation of the UPR (Howell, 2013). Mutations of several components of the *Arabidopsis* HRD1 complex, and hence disruption of ERAD, have been shown to increase the magnitude of the UPR in response to protein misfolding stress and, in some cases, induce its constitutive activation (Liu et al., 2011; Su et al., 2011; Hüttner et al., 2012; Liu et al., 2015; Li et al., 2017; Lin et al., 2019). Little analysis of the UPR in *Arabidopsis doa10* mutants has been performed, although microarray analysis has previously identified a small number of upregulated UPR genes in *Atdoa10a* (Lu et al., 2012). Similarly, qRT-PCR has been used to show a slight increase in expression of some UPR genes in an *Atdoa10a* mutant both at basal levels and following a heat stress (Li et al., 2017). However, this increase was only significant when combined with *Athrd1a* and *Athrd1b* mutations.

To assess the impact of *Atdoa10a* and *Atdoa10b* mutations on the UPR, qRT-PCR was performed on TM-treated seedlings, as described in 3.7, to determine the expression levels of three known UPR marker genes: *AtBIP1/2*, *AtCNX1* and *AtbZIP60* (primers 154+155, 152+153 and 156+158). *AtBIP and AtCNX* genes encode the molecular chaperones Binding Protein and Calnexin, whereas *AtbZIP60* codes for a transcription factor that is active only when spliced in response to ERAD stress (Howell 2013; Liu and Howell 2016). The *AtBIP* primers used amplified a 54 bp fragment of both *AtBIP1* and *AtBIP2*, as these proteins are 99% identical, and the *AtbZIP60* primers were designed to be specific to the spliced sequence.

The TM treatment significantly upregulated the expression of each of the genes tested, confirming the efficacy of the TM treatment. In Col-0 plants, on average *AtBIP1/2* expression increased 6.07-fold, *AtCNX1* 5.80-fold and *AtbZIP60* 72.16-fold, when compared with the control treatment. However, combinations of mutations and suppressions of *AtDOA10A* and

AtDOA10B (see Chapter 4 for details of the lines used) had no significant impact on the expression level of any of the UPR genes following either the control or TM treatment (Figure 19 a, c and e). The scale of upregulation relative to Col-0 seedlings treated concurrently was also analysed but differences again proved insignificant (Figure 19 b, d and f). The lack of perturbation of the UPR in plants lacking *AtDOA10A* and *AtDOA10B* suggests that their role within the ERAD system may be relatively minor.

Previous reports have shown some subtle increases in UPR-genes in *Atdoa10a* but these increases were also not significant. *AtbZIP60*, for example, was reported to be insignificantly upregulated at basal levels in an *Atdoa10a* mutant (Li et al., 2017), whereas here a slight decrease in expression was observed. The primers used in this study were specific to the UPR-responsive spliced variant of the *AtbZIP60* transcript. Those used by Li et al. (2017) were specific to the inactive unspliced transcript and therefore were an inferior indicator of the UPR. In other studies, a heightened UPR in ERAD-mutants has been confirmed by Western blots showing an increased abundance of UPR-associated proteins rather than qRT-PCR analysis of transcriptional changes (Liu et al., 2011; Su et al., 2011; Hüttner et al., 2012; Liu et al., 2015; Li et al., 2017; Lin et al., 2019). During this investigation the levels of the UPR-associated AtCNX1 and AtCNX2 proteins were also studied in the different *Atdoa10* mutants treated with and without TM, to accompany the AtCNX1 qRT-PCR analysis. Correlating with the qRT-PCR data, the TM treatment caused an increase in the abundance of AtCNX1/2 proteins across all samples, although not to the same magnitude as that observed for the AtCNX1 transcript, and no clear differences between WT and mutant lines were observed (Appendix 4). The qRT-PCR approach was pursued in preference to Western blot analysis due to the stronger induction of the transcript and therefore greater opportunity to identify variation between mutant lines.
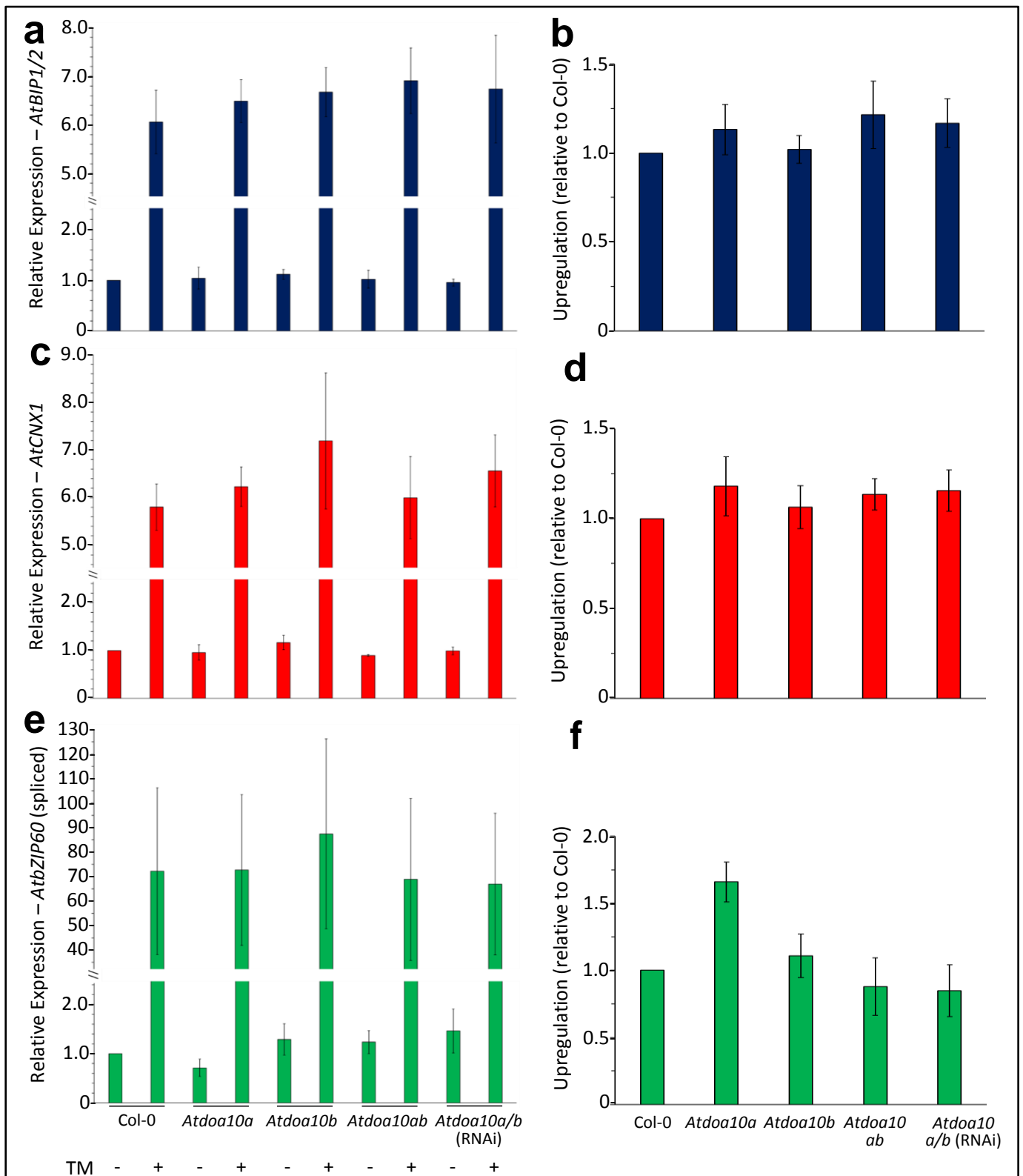
Figure 19 – The impact of *Atdoa10* mutations on UPR gene expression
Expression levels of (a, blue) *AtBIP1/2*, (c, red) *AtCNX1* and (e, green) *AtbZIP60* in 10x 6-day old seedlings following 4-hour treatment with TM or with DMSO (control), as determined by qRT-PCR, relative to DMSO-treated Col-0 seedlings. Data corresponding to the TM-induced upregulation of (b) *AtBIP1/2*, (d) *AtCNX1* and (f) *AtbZIP60* in each of the mutant backgrounds, relative to the upregulation of the same genes in a Col-0 sample treated concurrently are also presented. *Atdoa10ab* is CRISPR line 3-11(2), *Atdoa10a/b* (RNAi) is line 4-2(2). Expression levels and upregulation values were calculated according to the delta delta C_t method and data presented are the averages of 3 biological repeats (±SE). No differences were determined to be statistically significant using a one-way ANOVA.

3.9 AtDOA10A can functionally complement for the loss of ScDoa10 in yeast

To provide evidence that AtDOA10A and AtDOA10B are functional homologues of ScDoa10 as well as sequence homologues, the *Arabidopsis* E3s were expressed in mutant *ScDoa10Δ* yeast to assess conservation of function. *ScDoa10Δ* yeast was obtained from the Dharmacon yeast KO collection (Horizon Discovery) (see 2.2.1.1) and was genotyped to confirm the presence of a gene cassette including the kanamycin resistance gene *KanMX4* in place of *ScDoa10* (Figure 20). Genotyping was performed by PCR on DNA extracted from both WT and *ScDoa10Δ* yeast, using primers: 37, 38, 39, 40, 45 and 46. *ScDoa10*-specific amplification products were only detected in WT yeast and *KanMX4*-specific products only in *ScDoa10Δ* samples, confirming the absence of *ScDoa10* in the mutant (Figure 20).

ScDoa10Δ has been previously shown to vary from WT in its resistance to chemical stresses (Bockhorn et al., 2008; Theodoraki et al., 2012; Cuesta-Marban et al., 2013; Crowder et al., 2015). Using the transformation vector pAG416GPD-ccdB-EGFP, *ScDoa10Δ* was complemented with the CDS of *AtDOA10A* or *AtDOA10B* in order to establish whether either *Arabidopsis* homologue could rescue WT phenotypes. Successful transformants were selected on synthetic drop out media lacking uracil (DO-URA) and were then grown overnight in liquid DO-URA media. A portion of this culture was then diluted to an OD_{600} of 1.0 and spotting assays performed with serial dilutions. Contrary to previous reports of hypersensitivity (Crowder et al., 2015), in our growth conditions, *ScDoa10Δ* was found to have increased resistance to hygromycin. This disparity in phenotype may be explained by differences in the growth medium used. Due to the use of transgenes, here yeast were grown on synthetic drop out media, whereas previous characterisations were performed on YPD which is likely to have a major effect on survival and growth.

ScDoa10Δ exhibited significantly faster growth than WT on DO-URA media containing either 50 or 75 µg ml^{-1} Hygromycin B, when both were transformed with empty vector controls (Figure 21a). However, complementation of *ScDoa10Δ* with *AtDOA10A-EGFP* was able to

return hygromycin susceptibility to a level similar to WT (Figure 21a). This indicated that AtDOA10A was able to compensate for the loss of ScDoa10 function within the response to hygromycin and demonstrated conservation between kingdoms. Alternatively, *AtDOA10B-EGFP* complementation appeared to have little effect on growth levels (Figure 21a), suggesting divergence away from the role of ScDoa10. RNA was extracted from pellets of the yeast cultures used in the spotting assays to confirm the expression of the appropriate transgene by RT-PCR with primers 18/19+12, 9+100, 10+4 and 110+111 (Figure 21b).

Previous attempts to complement *ScDoa10Δ* with *AtDOA10A* have been unsuccessful, thought to be because AtDOA10A is very unstable in yeast (Doblas et al., 2013). The successful rescue observed here may therefore indicate that the presence of AtDOA10A, even at very low levels, is enough to rescue the hygromycin phenotype.
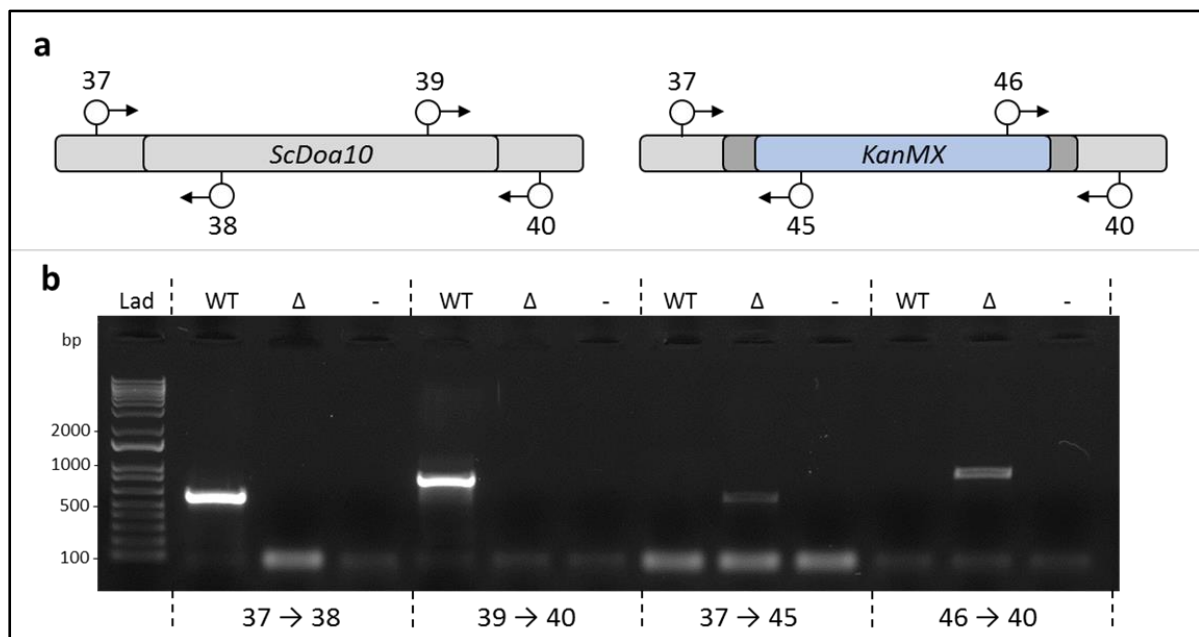
Figure 20 – PCR genotyping of *ScDoa10Δ*
(a) Schematic diagram of *ScDoa10* locus in WT (left) and *ScDoa10Δ* (right) yeast including genotyping primer positions. *ScDoa10* is replaced by a gene cassette including the *KanMX4* gene in *ScDoa10Δ*. (b) Gel electrophoresis of PCR products resulting from amplification of DNA from WT and *ScDoa10Δ* (Δ) with specified primers. No template control PCR products (-) and molecular weight size marker (Lad) also shown.
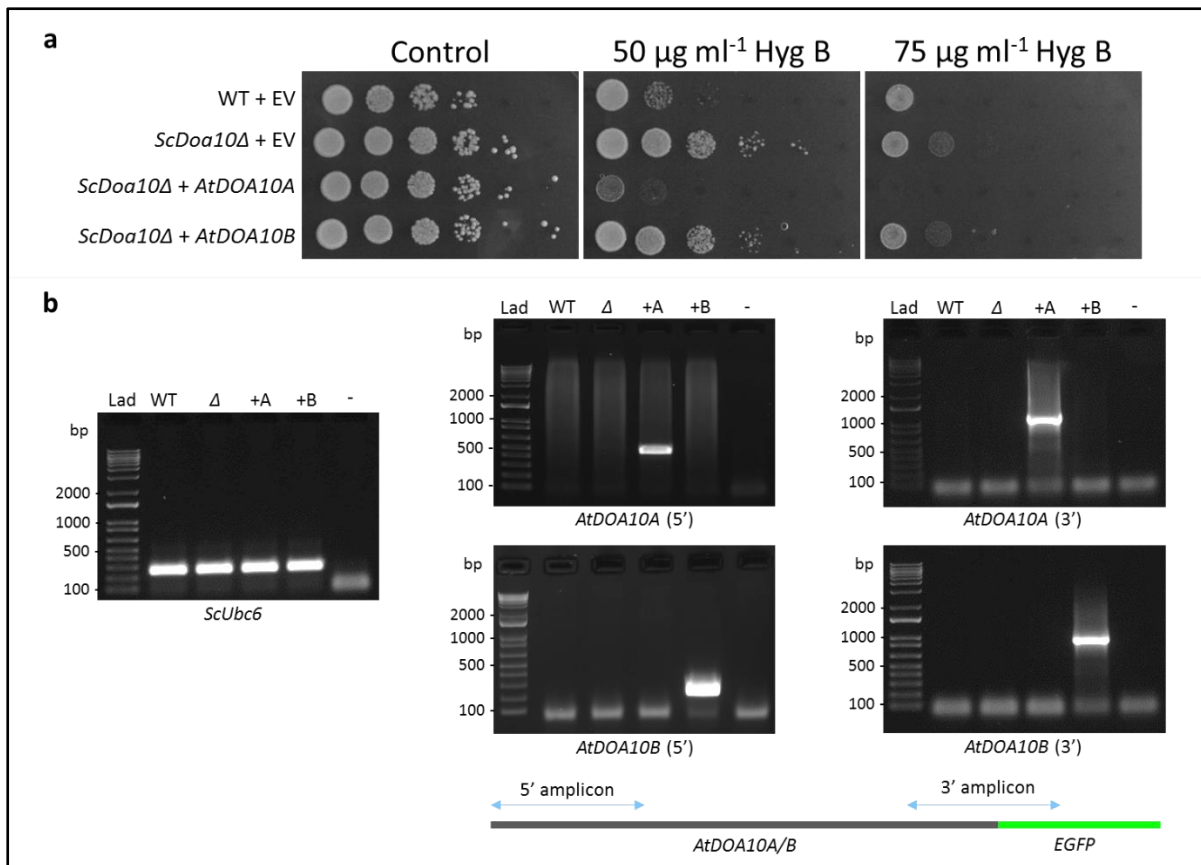
Figure 21 – The effect of AtDOA10A/B-EGFP complementation on *ScDoa10*Δ's hygromycin resistance phenotype

(a) Spotting assays after 2 days of growth following addition of 5 µl spots of diluted yeast culture with absorbances of OD_{600} = 1.0×10^0, 1.0×10^{-1}, 1.0×10^{-2}, 1.0×10^{-3}, 1.0×10^{-4} and 1.0×10^{-5} (left to right). *ScDoa10*Δ showed increased growth on DO-URA media relative to WT (both transformed with empty vector (EV) pAG416GPD-ccdB-EGFP). Complementation with *AtDOA10A-EGFP* but not *AtDOA10B-EGFP* restored the WT phenotype. Representative images from three replicates. (b) RT-PCR of amplicons from 5' and 3' ends confirmed that both *AtDOA10A-EGFP* and *AtDOA10B-EGFP* were being successfully transcribed. 3' amplicons were generated from primers specific to both the gene and C-terminal EGFP tag. No template control PCR products (-) also shown. Expression of the constitutive *ScUbc6* was consistent.

3.10 Discussion

AtDOA10A and B were identified as homologues of ScDOA10 due to the similarity of their amino acid sequence (Liu et al., 2011). Here, analysis has shown that similarity is particularly high within functionally significant regions of their sequences; namely the RING-CH domain, TD domain and C-terminal motif (Figure 12). Protein topology models and cell fractionation experiments have also shown that this similarity extends to their overall structure and subcellular localisation, with AtDOA10A and B both containing a high density of transmembrane domains and colocalising with the ER-marker AtCNX1/2 (Figures 13 and 17). These analyses also suggested that AtDOA10A is more closely related than AtDOA10B to ScDoa10.

Consistent with ScDoa10's known role in ERAD (Swanson et al., 2001), *AtDOA10A* and *AtDOA10B* are broadly expressed throughout the plant (Figure 14) in a similar manner to other known components of the *Arabidopsis* ERAD system (Cui et al., 2012a; Hüttner et al., 2012), although the expression level of *AtDOA10B* was observed to be much lower than that of *AtDOA10A*. This widespread expression was mirrored by GUS-staining used to highlight the distribution of AtDOA10A and AtDOA10B proteins (Figures 15 and 16). Staining intensity also indicated a reduced abundance of AtDOA10B relative to AtDOA10A. For both proteins, greater levels of staining were also observed in young and fast-growing regions, where a higher rate of protein production and therefore increased ERAD demand is to be expected. It should be noted, however, that similar GUS staining patterns have been observed for several other E3 ligases and their promoters, including those not associated with ERAD (Weber and Hellmann, 2009; Deng et al., 2017; Zhang et al., 2017). The absence of GUS staining from tissues shown to contain *AtDOA10A* and *AtDOA10B* transcripts, may indicate a significant role of posttranslational regulation in controlling the abundance and localisation of the enzymes at the protein level. This could potentially occur by autoubiquitination, another common feature of E3 ligases (de Bie and Ciechanover, 2011). HsMARCH6 is known to autoubiquitinate (Hassink et al., 2005).

Further evidence of a role in ERAD was provided by the observation that an applied protein misfolding stress (TM), induced significant changes in the gene expression of both *AtDOA10A* and *AtDOA10B* (Figure 18). Surprisingly, given their apparent similarities, the expression levels of the two genes responded differently to the ERAD stress, with *AtDOA10A* expression decreasing by 36%, whilst *AtDOA10B* transcript abundance increased 5-fold. These results correlate with differing responses of *AtDOA10A* and *AtDOA10B* to the alternative ERAD-stressor AZC (Kim et al., 2017). Intriguingly, Kim et al. (2017) also observed a similar pattern for the *Arabidopsis* HRD1 homologues, with only *AtHRD1B* being strongly upregulated in response to AZC. This may suggest that, in plants, both B homologues may have diverged away from their roles in yeast, since activation of the UPR in yeast using TM and the reducing agent dithiothreitol (DTT) has been shown to induce only very minor changes in the expression of *ScDoa10* (Gasch et al., 2000; Travers et al., 2000; Leber et al., 2004). However, some reports do show a UPR-associated upregulation of ScHrd1 (Travers et al., 2000; Leber et al., 2004). Even so, the results presented here, in combination with those of Kim et al. (2017) may suggest that in *Arabidopsis* AtDOA10A is involved in the basal turnover of ERAD substrates, with AtDOA10B being specifically expressed to degrade misfolded proteins during periods of high demand. The lack of AtDOA10B-like proteins in non-*Brassicaceae* plant species (Table 17) raises the question of how their strategies for dealing with the accumulation of ERAD-substrates may differ.

In comparison to AtDOA10A, AtDOA10B has an increased divergence from the amino acid sequence of ScDoa10 in the TD-domain and C-terminal motif (Figure 12b and c). Since both of these regions have been implicated in substrate recognition (Kreft and Hochstrasser 2011; Zattas et al., 2016), this variation may enable AtDOA10B to target an alternative and potentially broader range of substrates. This could facilitate the targeting of abnormal misfolded proteins at the expense of losing specificity for basal ERAD substrates recognised by AtDOA10A. Such a role would also be consistent with the lower levels of AtDOA10B in unstressed conditions (Figures 14-16) and might explain why AtDOA10B was unable to

complement the hygromycin sensitivity phenotype of *ScDoa10Δ*, unlike AtDOA10A (Figure 21). Alternatively, AtDOA10B may be highly unstable in yeast (as has been reported for AtDOA10A (Doblas et al., 2013)) and unable to rescue the phenotype because of this.

Despite AtDOA10A and AtDOA10B being homologues of one of only two yeast ERAD E3 ligases, the loss of both enzymes had no significant impact on the magnitude of the TM-induced UPR (Figure 19). This indicated that these enzymes are unlikely to play a major role in the total degradation of all misfolded proteins at the ER. The degradation of a specific subset of ERAD substrates in their folded and misfolded states may be more likely. Alternatively, the lack of effect on the UPR may be a result of functional overlap with other ERAD E3s, such as the HRD1 homologues. In yeast, large increases in the magnitude of the UPR are only induced by simultaneous deletion of *ScDoa10* and *ScHrd1* (Swanson et al., 2001). Similar observations have been made in *Arabidopsis,* with *Atdoa10a Athrd1a Athrd1b* triple mutants exhibiting an elevated UPR and a thermotolerance phenotype not observed in *Atdoa10a* and *Athrd1a Athrd1b* single and double mutants respectively (Li et al., 2017). Recent discoveries of plant-specific components of the *Arabidopsis* HRD1 complex (Liu et al., 2015; Lin et al., 2019) may indicate that the roles of AtHRD1A and AtHRD1B have been expanded, marginalising the position of AtDOA10A and AtDOA10B within the ERAD process. Mutations of these and other components of the *Arabidopsis* HRD1 complex have been reported to significantly exacerbate the UPR (Liu et al., 2011; Su et al., 2011; Hüttner et al., 2012; Liu et al., 2015; Li et al., 2017; Lin et al., 2019). Surprisingly however, mutations of the AtHRD1A and AtHRD1B E3 ligases themselves have little impact on the UPR (Li et al., 2017). The existence of undiscovered plant-specific ERAD components that may alleviate the impact of *Atdoa10 and Athrd* mutations, remains possible.

Overall, the consideration of AtDOA10A and AtDOA10B as homologues of ScDoa10 is justified based on the high levels of sequence homology, structural similarities and conserved subcellular localisation. This is particularly true for AtDOA10A, which was also able to successfully complement ScDoa10 showing conservation of function. The increased

divergence of AtDOA10B may indicate a slightly altered role of this E3, although this is still likely to be as part of the ERAD system given its localisation and expression increase in response to TM. However, their contribution to the global degradation of all misfolded proteins within the *Arabidopsis* ERAD system appears to be relatively minor since their mutations had no significant impact on the magnitude of the TM-induced UPR. A degree of redundancy between DOA10 and HRD1 homologues, along with as yet undiscovered *Arabidopsis* ERAD components may also explain this.

Chapter IV -

Generating genetic tools for characterising AtDOA10 function

4.1 Introduction

The confirmation of the similarities between ScDoa10, AtDOA10A and AtDOA10B, outlined in Chapter III supported the classification by Liu et al. (2011) that the *Arabidopsis* proteins are ScDoa10 homologues. The observed functional complementation of ScDoa10 by AtDOA10A in particular, demonstrated some conservation of function as well as sequence homology (Figure 21). The failure of AtDOA10B to rescue defects associated with the loss of ScDOA10 in yeast, as well as its slightly reduced similarity in terms of amino acid sequence and structure (Figures 12 and 13), may indicate a diversification of its function or substrates. Nevertheless, analysis of expression pattern (Figure 14), subcellular localisation (Figure 17) and gene expression changes in response to TM (Figure 18) suggested that both E3s are part of the ERAD system in *Arabidopsis*, like ScDoa10 in yeast. However, the lack of impact on the UPR caused by the loss of AtDOA10A and AtDOA10B (Figure 19) indicated that their contribution to total ERAD may be small or overlap with the function of other E3s. Analyses in Chapter III did not assess whether AtDOA10A or AtDOA10B may participate in an *Arabidopsis* Ac/N-degron pathway beyond showing their similarities to ScDoa10, which acts as an N-recognin in yeast (Hwang et al., 2010).

To further investigate the *in planta* roles of AtDOA10A and AtDOA10B as ERAD components and potentially Ac/N-recognins, a range of *Arabidopsis* lines with altered expression levels of the E3s and other associated enzymes were developed. This chapter describes the development of these genetic tools that were used in phenotypic, transcriptomic and proteomic comparisons in Chapter V.

4.2 Characterisation of T-DNA insertional mutants

To assess the importance of AtDOA10A, AtDOA10B, a postulated *Arabidopsis* Ac/N-degron pathway and the processes they influence during plant growth and development, a selection of T-DNA insertional mutants were identified. In addition to *Atdoa10a* and *Atdoa10b* mutants, an *Atnaa20* T-DNA mutant was also acquired, as this was needed for subsequent studies outlined in Chapter VI. *AtNAA20* encodes the catalytic subunit of NatB (see 1.3.1). Seeds of all T-DNA insertional mutants used were obtained from the Nottingham Arabidopsis Stock Centre (NASC).

4.2.1 *Atdoa10a* (AT4G34100)

The T-DNA insertion mutant GABI_588A06 investigated in the present study was first characterised by Lu et al. (2012), who reported that the T-DNA insert lies within exon 6. However, amplification and sequencing of gDNA (primers 22+23) confirmed that the T-DNA is actually present within exon 5 (2670 bp from start codon, Appendix 5). This sequencing data matches that reported by GABI-KAT (http://signal.salk.edu/cgi-bin/tdnaexpress?JOB=TEXT&TYPE=DATA&QUERY=GABI_588A06). Despite this discrepancy, GABI_588A06 has been reported to produce a truncated *AtDOA10A* transcript (Lu et al., 2012).

PCR-based genotyping using primers 21+22 confirmed the absence of the WT allele in *Atdoa10a* plants as well as confirming the presence of a T-DNA insertion within the open reading frame (ORF) (primers 22+23) (Figure 22a and b). RT-PCR was also performed but the full-length 3.3 kb *AtDOA10A* transcript could not be amplified from either Col-0 or *Atdoa10a* cDNA (not shown). Instead, primers designed to amplify a smaller fragment of complementary DNA (cDNA) (1294 bp) including the T-DNA insertion site were used (primers 210+102). RT-PCR with these primers generated a product for Col-0 only (Figure 22c). This RT-PCR result confirmed that transcription of *AtDOA10A* was being disrupted at the point of the T-DNA insertion. However, as was observed by Lu et al. (2012), expression of truncated transcripts could still be detected away from the insertion site (not shown). The presence of

similar starting amounts of genomic or complementary DNA within each PCR or RT-PCR reaction was confirmed by amplification of the housekeeping gene *AtACT7* using primers 14+15 (Figure 22d and e). The null mutation of *AtDOA10A* was further confirmed by screening seedlings on ½ MS media containing ABA. The *Atdoa10a* seedlings exhibited the published ABA hypersensitivity phenotype (Figures 25c and 42) associated with a loss of functioning AtDOA10A (Zhao et al., 2014).
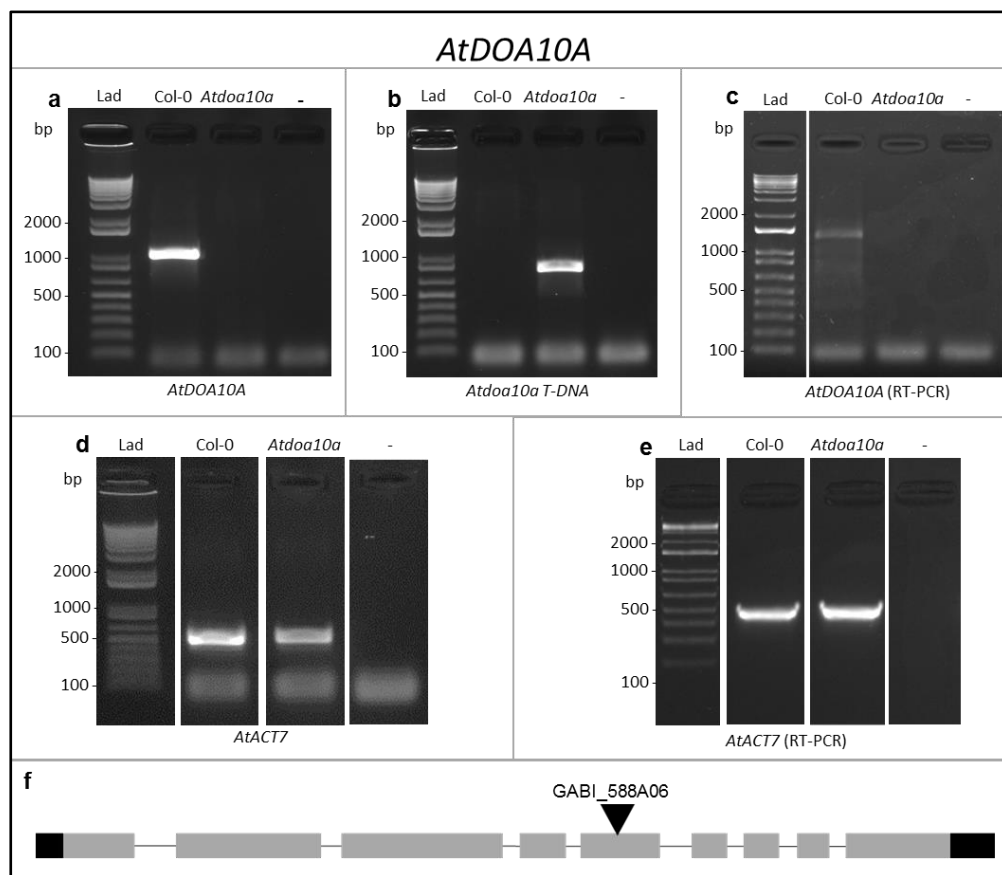
Figure 22 – PCR/RT-PCR genotyping of *Atdoa10a Arabidopsis* mutant
Gel electrophoresis of PCR amplification products and no template control reactions (-). (a) WT-gene product from amplification with left and right flanking primers. (b) T-DNA specific product from amplification with T-DNA border primer and flanking primer. (c) RT-PCR of region including T-DNA insertion site. (d) *AtACT7* PCR product demonstrating successful DNA extraction. (e) RT-PCR amplification of *AtACT7.* (f) Schematic diagram of T-DNA insertion position (black triangle). Black boxes = UTR, grey boxes = exons, lines = introns.

4.2.2 *Atdoa10b* (AT4G32670)

The T-DNA line SAIL_1303_E09 contains a T-DNA insertion within the 5' untranslated region of *AtDOA10B.*

PCR genotyping using primers 24+25 and 25+26 confirmed that the T-DNA successfully disrupts the DNA sequence in this region in the *Atdoa10b* mutant plants (Figure 23a and b). A full-length *AtDOA10B* transcript was detected by RT-PCR in Col-0 cDNA but was absent from *Atdoa10b* cDNA using primers 10+11 (Figure 23c).
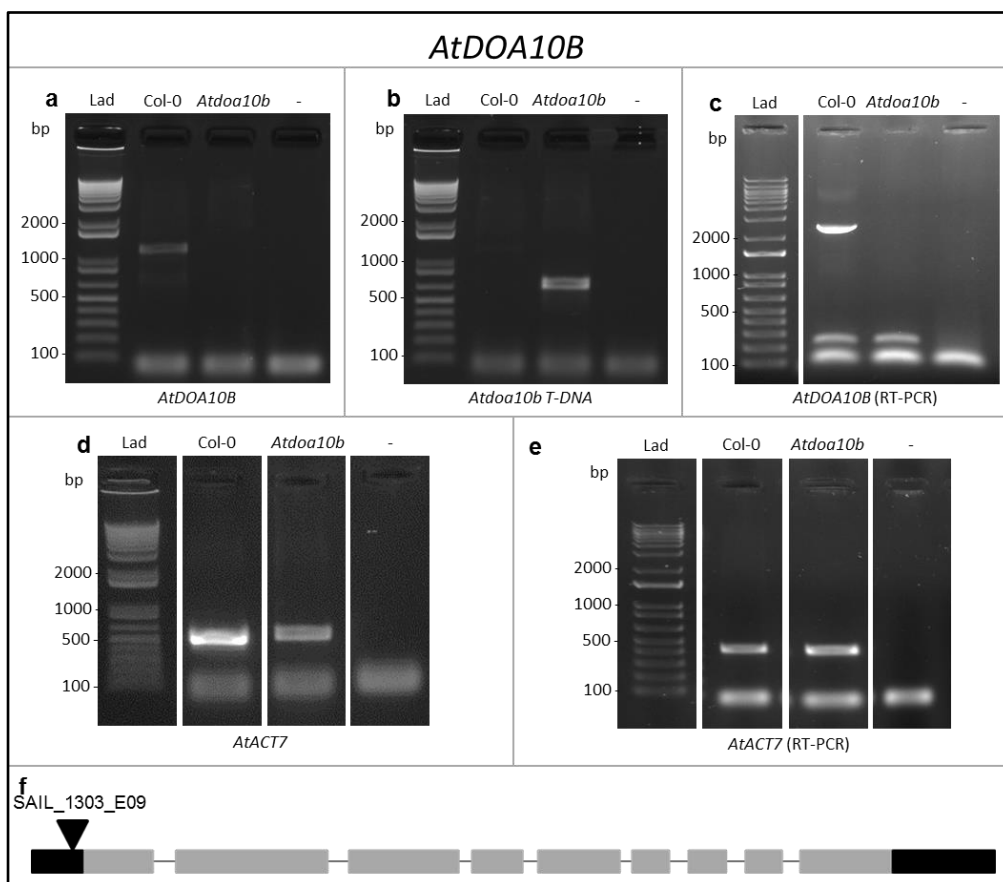
Figure 23 – PCR/RT-PCR genotyping of *Atdoa10b Arabidopsis* mutant
Gel electrophoresis of PCR amplification products and no template control reactions (-). (a) WT-gene product from amplification with left and right flanking primers. (b) T-DNA specific product from amplification with T-DNA border primer and flanking primer. (c) RT-PCR of full-length gene transcript. (d) *AtACT7* PCR product demonstrating successful DNA extraction. (e) RT-PCR amplification of *AtACT7.* (f) Schematic diagram of T-DNA insertion position (black triangle). Black boxes = UTR, grey boxes = exons, lines = introns.

4.2.3 *Atnaa20* (AT1G03150)

AT1G03150 codes for *AtNAA20*, the catalytic subunit of the NatB complex in *Arabidopsis* (alternatively named *AtNBC1* by Ferrandez-Ayela *et al.* (2013)). The T-DNA insertion mutant line SAIL_323_B05, which contains a T-DNA insertion 85 bp upstream of the AT1G03150 start codon has been characterised previously (Ferrandez-Ayela et al., 2013; Xu et al., 2015). Here, the alternative SALK_027687 *Atnaa20* mutant (Linster et al., 2015) was studied. This mutant contains a T-DNA insertion within the third intron of the gene and has been reported to cause a ≈60% depletion of *AtNAA20* transcripts (Huber et al., 2020).

Interruption of the ORF was confirmed by PCR using primers 27+28 and 28+29 (Figure 24a and b). In contrast to the reported depletion, RT-PCR of the full-length transcript with primers 92+93 failed to generate any product, as if a null mutant (Figure 24c). Huber et al. (2020) were able to detect *AtNAA20* transcripts downstream of the T-DNA insertion site. Despite discrepancies, both results indicate a significant reduction in *AtNAA20* gene expression.
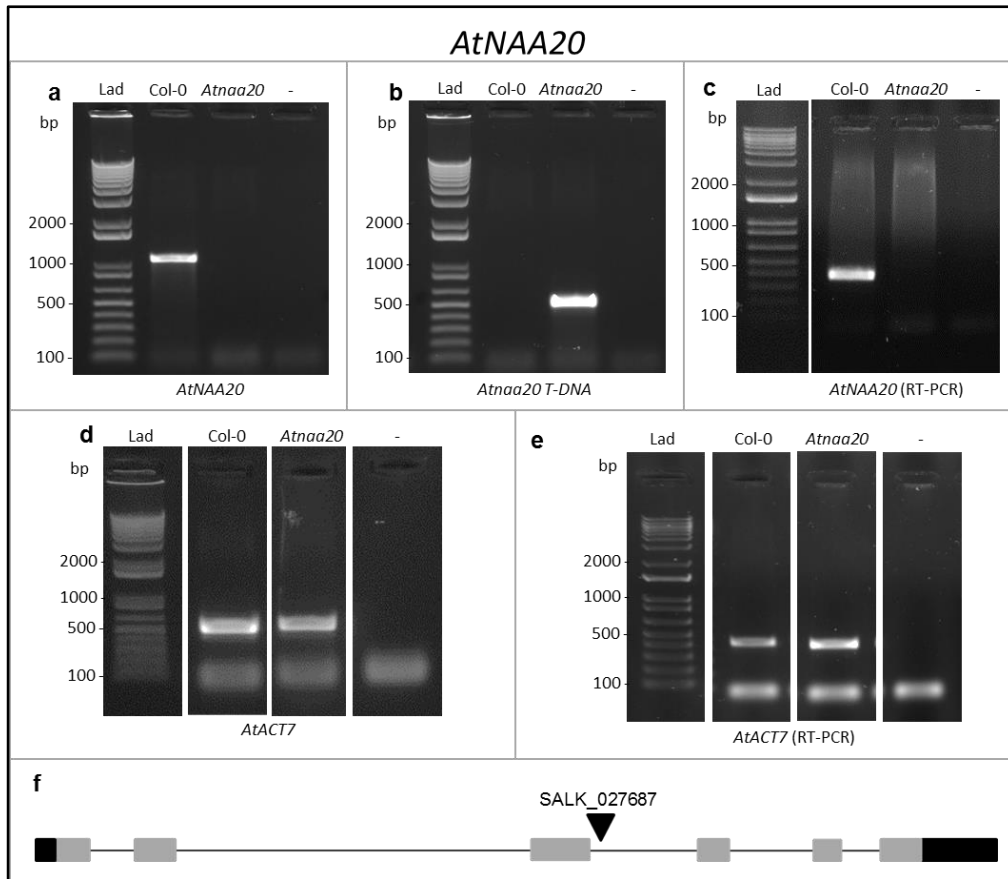
Figures 24 – PCR/RT-PCR genotyping of *Atnaa20 Arabidopsis* mutant
Gel electrophoresis of PCR amplification products and no template control reactions (-). (a) WT-gene product from amplification with left and right flanking primers. (b) T-DNA specific product from amplification with T-DNA border primer and flanking primer. (c) RT-PCR of full-length gene transcript. (d) *AtACT7* PCR product demonstrating successful DNA extraction. (e) RT-PCR amplification of *AtACT7.* (f) Schematic diagram of T-DNA insertion position (black triangle). Black boxes = UTR, grey boxes = exons, lines = introns.

4.3 Generation of AtDOA10A and AtDOA10B GUS and EYFP reporter lines

To further dissect the localisation and function of AtDOA10A and AtDOA10B in *Arabidopsis*, plants expressing C-terminal β-glucuronidase (GUS) and enhanced yellow fluorescent protein (EYFP)-tagged variants, under the control of approximately 2 kb of their native promoters, were developed. Genomic WT alleles of At*DOA10A* and At*DOA10B* were amplified from *Arabidopsis* DNA extracts using the Gateway® cloning compatible primers 16+200 and 201+11 respectively. These primers were designed to produce products from 2006 bp upstream of *AtDOA10A* and 1997 bp upstream of *AtDOA10B*. The resulting amplicons were cloned into either pDONR™221 (*AtDOA10A*) or pENTR™/D-TOPO™ (*AtDOA10B*) entry vectors (Invitrogen) and then subcloned into the destination binary vectors pGWB533 and pGWB540, which include GUS and EYFP tags respectively (Nakagawa et al., 2007). *Agrobacterium tumefaciens* was transformed with these expression clones and a floral dip was then performed. The resultant T1 seeds were harvested and successfully transformed plants were identified by their hygromycin resistance. All steps from cloning to T1 selection were carried out by Dr. Mark Bailey.

Subsequent generations were screened for a 3:1 ratio of hygromycin resistance at T2 and 100% resistance at T3. Following selection, two AtDOA10A-GUS lines, two AtDOA10A-EYFP lines, three AtDOA10B-GUS lines and three AtDOA10B-EYFP lines were identified as homozygous. RT-PCR was performed on these lines to monitor the expression of the fusion constructs (primers 18/19+13/12). Appropriate cDNA bands were present in the following lines: AtDOA10A-EYFP-2, AtDOA10A-EYFP-4, AtDOA10A-GUS-1, AtDOA10A-GUS-3, AtDOA10B-EYFP-4, AtDOA10B-GUS-2, AtDOA10B-GUS-4 and AtDOA10B-GUS-6 (Figure 25a).

These lines were then assessed for expression of the fusion proteins via a Western blot. Following protein extraction with RIPA buffer, a protein band was detected, using an anti-GUS antibody (see 2.4.1), in each of the AtDOA10A-GUS and AtDOA10B-GUS lines that was

not present in Col-0 protein extracts (Figure 25b). However, the observed bands did not correspond to the expected molecular weights of the fusion proteins: ≈198 kDa and ≈174 kDa for AtDOA10A- and AtDOA10B-GUS fusions respectively. Instead, all bands were detected at a molecular weight in excess of 250 kDa, potentially as a result of post translational modifications. In addition to the high molecular weight bands at over 250 kDa, a second non-WT band at around 70 kDa was also seen for each GUS line (Figure 25b). Given its size, this band was thought to correspond to free GUS protein (68 kDa), i.e., GUS protein molecules that were not fused to AtDOA10A or AtDOA10B, potentially due to proteolytic cleavage. Bands corresponding to both the fusion proteins and free GUS were much stronger in AtDOA10A-GUS lines compared to AtDOA10B-GUS lines, consistent with RT-PCR results (see 3.5).

AtDOA10A-EYFP and AtDOA10B-EYFP fusion proteins could not be directly detected by Western blot, following RIPA buffer protein extraction, perhaps due to low expression levels. To circumvent this, an additional immunoprecipitation step using ChromoTek GFP-Trap® Magnetic Agarose was performed (see 2.9.4). SDS-PAGE and Western blotting of immunoprecipitation products resulted in the detection of protein bands at the expected molecular weights (≈151 kDa and ≈128 kDa) using an anti-GFP/YFP antibody (see 2.4.1) (Figure 25b).

To test the functionality of the AtDOA10A-GUS/EYFP complementation lines, their response to ABA was assessed, given that *Atdoa10a* plants have been previously shown to be ABA hypersensitive (Zhao et al., 2014). Two AtDOA10A-EYFP and two AtDOA10A-GUS lines were tested for ABA sensitivity by assessing germination and growth on ½ MS media containing 0.5 or 1.0 µM ABA. All four lines exhibited at least a partial rescue of the mutant phenotype (Figures 25c and d), with the EYFP lines especially appearing visually similar to WT. The weaker rescue by AtDOA10A-GUS fusions may be the result of impairment of protein function due to the larger size of the GUS-tag (68 kDa) relative to EYFP (27 kDa).
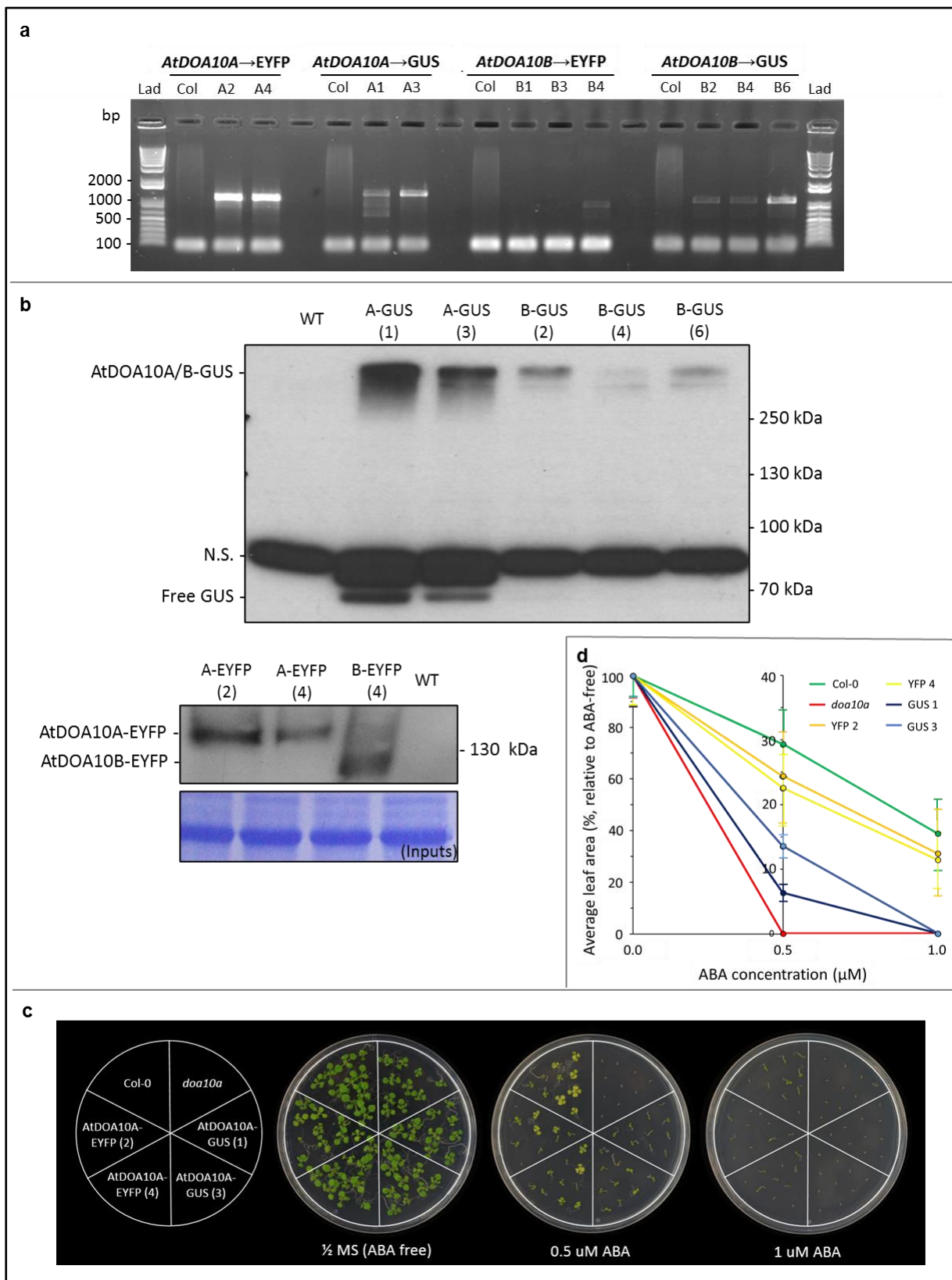
Figure 25 – Characterisation of *AtDOA10A/B* complementation lines
(a) RT-PCR identification of lines successfully expressing complementation constructs using a gene-specific forward primer and tag-specific reverse primer. (b) anti-GUS/anti-YFP (see 2.4.1) Western blot analysis of GUS and EYFP tagged proteins. GUS lines appear to contain free GUS protein bands at ≈68 kDa. Non-specific (N.S.) band also detected at ≈80 kDa. (c) Germination and growth of AtDOA10A complementation lines in response to ABA treatment. (d) Quantification of average leaf area of lines in (c) relative to ABA-free plate, as measured using FIJI (Schindelin et al., 2012).

4.4 Generation of an *Atdoa10ab* double mutant

Given the possibility of functional redundancy between AtDOA10A and AtDOA10B, an *Atdoa10a Atdoa10b* (*Atdoa10ab*) double mutant was required in order to assess the physiological functions of the two homologous E3 ligases.

4.4.1 Crossing of *Atdoa10a* and *Atdoa10b* single mutants

A cross of the *Atdoa10a* and *Atdoa10b* single mutants was performed in an attempt to generate an *Atdoa10ab* double mutant. Resulting heterozygous F1 individual plants were identified by PCR. F2 seeds were harvested from heterozygous plants but were not progressed further due to the very low probability of producing a double mutant by this method. As both genes are located on chromosome four (Figure 26), the generation of a double mutant would require a homologous recombination event to occur between both loci during meiosis. Due to their close proximity (0.57 Mb apart) and the therefore very low probability of the required meiotic recombination occurring, alternative approaches for the generation of a double mutant were pursued.
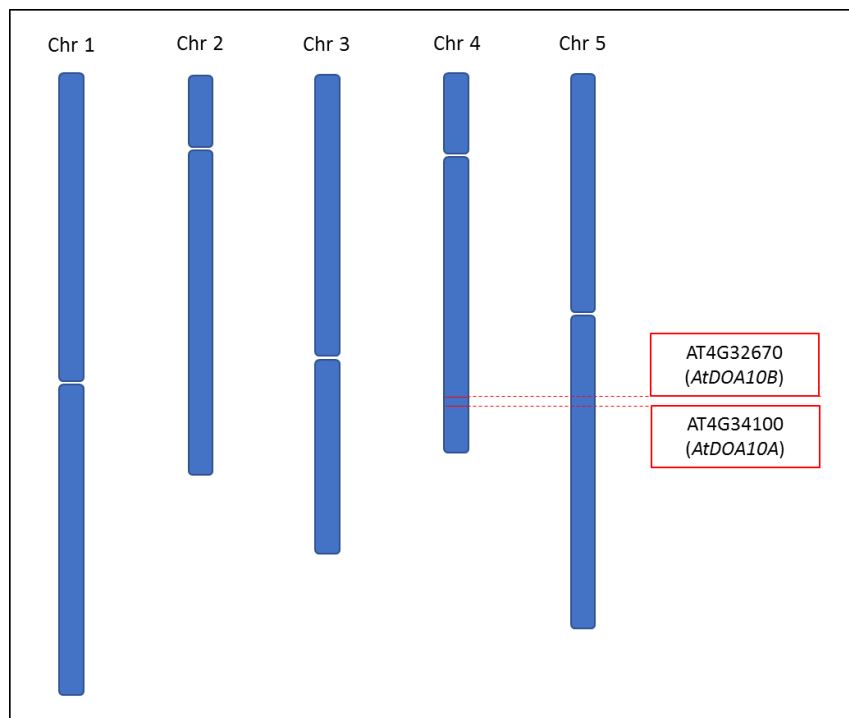
Figure 26 – Proximity of *AtDOA10A* and *AtDOA10B* within *Arabidopsis* genome
Arabidopsis chromosome diagram showing *AtDOA10A* and At*DOA10B* loci are only 0.57Mb apart on chromosome four. Based on The Arabidopsis Information Resource (TAIR), [https://www.arabidopsis.org/jsp/ChromosomeMap/tool.jsp], on www.arabidopsis.org, [18/06/19].

4.4.2 Generation of an *AtDOA10B* knockdown in the *Atdoa10a* KO background by RNAi

As well as the predicted crossing difficulties, the possibility remained that an *Atdoa10ab* double mutant could be embryonically lethal. For this reason, RNA interference (RNAi) constructs were developed, designed to reduce, but not eliminate, *AtDOA10B* expression in the *Atdoa10a* mutant background. RNAi is able to decrease the levels of gene transcripts through post-transcriptional gene silencing whereby double stranded RNA (dsRNA) can induce cleavage of mRNA with the corresponding sequence. Long sections of dsRNA are split into short pieces (≈22 nucleotides) known as short interfering RNAs (siRNAs) by Dicer enzymes (Hannon, 2002). One strand of the siRNA molecule is then bound by an Argonaute protein, the catalytic component of the RNA Induced Silencing Complex (RISC), guiding the protein to mRNA transcripts with matching sequences, which it then cleaves (Hannon, 2002; Ossowski et al., 2008).

Gene silencing by RNAi was first discovered and is best characterised in *Caenorhabditis elegans* (Fire et al., 1998; Hannon, 2002) but can also be applied to plants, most simply by transformation with a construct designed to produce dsRNA (Ossowski et al., 2008). The most efficient method of producing dsRNA from a vector is through the expression of sequences that produce hairpin loops of RNA. The arms of the hairpin are made up of complementary RNA sequences, corresponding to the intended silencing target, that are held together by hydrogen bonding between nucleotides (Waterhouse et al., 1998; Wesley et al., 2001). On the vector, these complementary sequences are separated by a spacer sequence which forms a single stranded loop extending from the annealed duplex, when transcribed. It has been observed that the inclusion of an intron within this loop improves the gene silencing efficiency (Wesley et al., 2001).

The Gateway® compatible, intron-containing, hairpin RNA expression vector pK7GWIWG2(I) (Karimi et al., 2002) was used to reduce the expression of *AtDOA10B* by RNAi. Cloning and agrobacterium-mediated transformation of *Atdoa10a* KO plants with the

RNAi constructs was performed by Dr. Mark Bailey. Four target sequences were selected corresponding to sites along the length of the *AtDOA10B* gene (Figure 27). These sequences were amplified (using primers 202-209) and separately incorporated into pK7GWIWG2(I) by Gateway® cloning via the pENTR™/D-TOPO™ entry vector. T1 selection based on the kanamycin resistance encoded within the inserted T-DNA, was conducted and resistant plants were grown to maturity. T2 seeds were harvested and initially those transformed with RNAi constructs 3 or 4, were screened to identify correct Mendelian segregation on ½ MS plates containing kanamycin. For both RNAi constructs, seeds of five lines exhibiting a survival ratio of approximately 3:1 were then also grown on antibiotic free ½ MS media in order to screen *AtDOA10B* expression levels.

After 14 days of growth on antibiotic free media, RNA was extracted from whole seedlings of correctly segregating lines and was used to analyse gene expression by qRT-PCR. cDNA-specific primers were designed to amplify small regions of DNA close to both the 5' and 3' ends of the *AtDOA10B* gene in order to assess the impact of the silencing on the entire transcript (primers 3+4/1+2). Levels of *AtDOA10B*, determined by qRT-PCR, were normalised to the housekeeping control gene *AtACT7* (primers 78+79). Despite the target sequences being designed to be closer to the 5' end of *AtDOA10B,* in almost all cases the expression level detected at exon 7 was lower than at exon 1, potentially due to proximity to the transcription start site or perhaps simply a result of variations in primer efficiency. Three lines for which both primer sets indicated reduced *AtDOA10B* expression were progressed to the T3 generation: Construct 3 Line 7, Construct 4 line 2 and Construct 4 line 9 (Figure 28).

For each of the selected lines, T2 plants were grown on soil and T3 seeds harvested upon maturity. Seed sets were rescreened on kanamycin ½ MS plates for 100% survival. Homozygous lines were then re-screened by qRT-PCR in order to identify the strongest *AtDOA10B* knockdowns. Due to the presence of heterozygous and WT individuals within T2 populations, knockdown strength was expected to increase in homozygous T3 populations of the same lines. However, despite showing the largest decrease in expression at T2 (≈60%),

T3 plants from six different homozygous Construct 4 Line 9 parents only exhibited a 10-40% reduction in *AtDOA10B* expression (Appendix 6). Stronger knockdowns were observed for Construct 3 Line 7 and Construct 4 line 2 T3 plants. qRT-PCR was performed in three biological replicates on the most promising lines and *AtDOA10B* expression levels were separately normalised to both *AtACT7* and *AtRBCL* housekeeping genes for increased reliability. Overall, reductions in the expression of *AtDOA10B* of 46.0 % for Construct 3 Line 7(1) and 38.5 % for Construct 4 Line 2(2) were observed (based on normalisation to *AtACT7*) (Figure 29). These lines are hereafter referred to as *Atdoa10a/b* (RNAi line 3-7(1)) and *Atdoa10a/b* (RNAi line 4-2(2)).
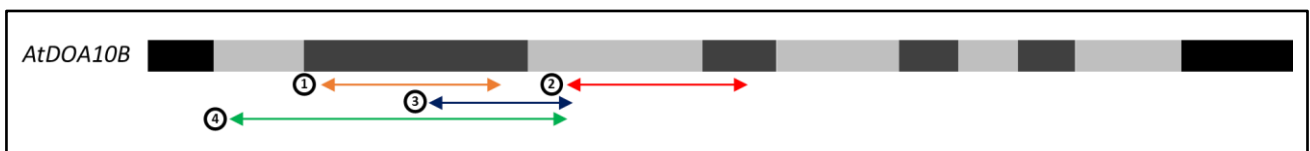
Figure 27 – RNAi target sequences within *AtDOA10B*
Schematic diagram of *AtDOA10B* cDNA structure showing four target sequences covering exons 1 to 4 of *AtDOA10B* that were selected to be incorporated into the RNAi vector pK7GWIWG2(I). Black = UTR, light/dark grey = alternating exons
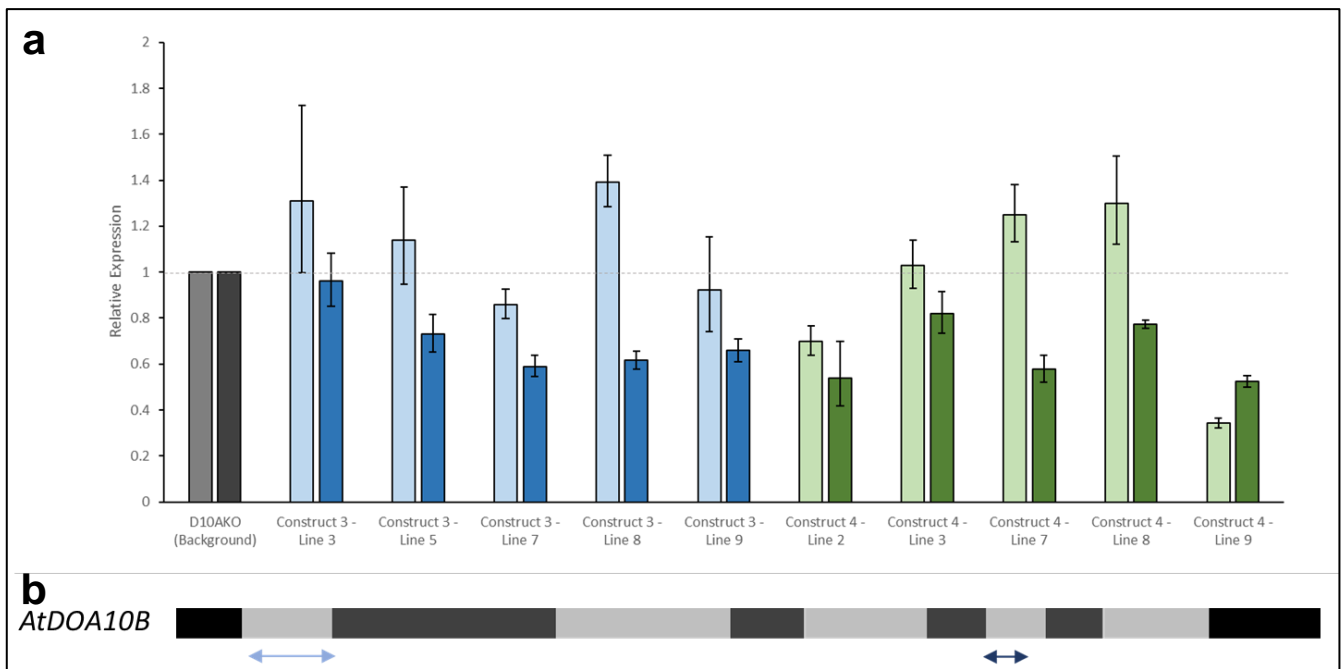
Figure 28 – Screening of T2 *AtDOA10B* RNAi lines
(a) Relative expression of *AtDOA10B* compared to the *Atdoa10a* background, as determined by qRT-PCR (normalised to *AtACT7*). Pale bars correspond to exon 1 amplicons, dark bars correspond to exon 7 amplicons. Values of relative expression are averages of two to three technical replicates ± standard error. (b) Schematic diagram of *AtDOA10B* cDNA structure (black = UTR, light/dark grey = alternating exons) showing locations of primer pairs used for qRT-PCR (exon1 and exon 7).
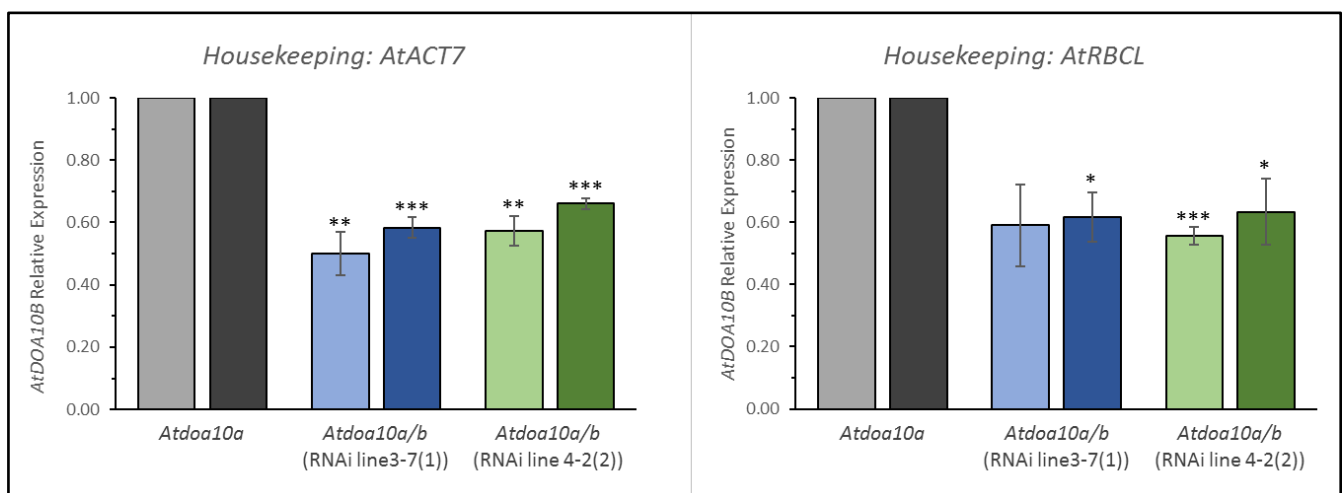
Figure 29 – Screening of T3 *AtDOA10B* RNAi lines
qRT-PCR analysis of *AtDOA10B* expression in RNAi lines relative to *Atdoa10a* background normalised to *AtACT7* and *AtRBCL* housekeeping genes. Data shown are the mean of three biological replicates. Pale bars correspond to exon 1 amplicons, dark bars correspond to exon 7 amplicons. Significant differences were determined by student's t-test.

4.4.3 A CRISPR-Cas9 approach to develop an *Atdoa10ab* double mutant

4.4.3.1 The CRISPR-Cas9 system in *Arabidopsis*

Although significant, the *AtDOA10B* knockdowns observed in the RNAi lines were relatively weak, with neither of the selected lines exhibiting a 50 % reduction in gene expression. Based on this and anticipated difficulties of producing an *Atdoa10ab* double mutant through crossing, an alternative method of double mutant generation via the CRISPR-Cas9 system was pursued. This system aims to utilise the bacterial endonuclease Cas9 to induce double strand breaks (DSBs) in targeted regions of DNA. The Cas9 enzyme, used by bacteria to cleave foreign DNA, can be directed to particular target sequences within the *Arabidopsis* genome using a single guide RNA (sgRNA) in a highly specific manner (Bortesi and Fischer, 2015; Puchta, 2017). A sgRNA contains an RNA sequence specific to the target site (crRNA) as well as a second section of RNA that forms a scaffold to which Cas9 can bind (tracrRNA) (Puchta and Fauser, 2014; Bortesi and Fischer, 2015). Any ≈20 bp sequence within a genome can be selected as a Cas9 cut site using sgRNAs, if they immediately precede a protospacer adjacent motif (PAM) sequence that is specific to the Cas9 of a particular bacterial species (Bortesi and Fischer, 2015). Guided to the target sequence by the bound sgRNA, Cas9 endonucleases can cleave genomic DNA at a position 3 bp upstream of the PAM sequence (Jinek et al., 2012). Target sequences with similarities to multiple sites within a genome can result in the Cas9 being guided to an alternative location, leading to the induction of an off-target DSB and therefore target sequences should be as unique as possible (Voytas, 2013; Bortesi and Fischer, 2015).

DSBs can be repaired by the constitutive cellular processes of either homologous recombination (HR) or, more regularly, non-homologous end joining (NHEJ) (Voytas, 2013; Puchta and Fauser, 2014). These repair mechanisms can be utilised to allow researchers to alter the DNA sequence at a desired location within the *Arabidopsis* genome. At its simplest, the CRISPR-Cas9 system can be used to mutate specific genes of interest due to the error-prone mechanism of NHEJ. NHEJ frequently leads to insertions or deletions (indels) of bases

at the site of the DSB which can lead to a frameshift mutation and hence the production of a premature stop codon and truncated or non-functional protein (Puchta and Fauser, 2014; Bortesi and Fischer, 2015). Indels can vary significantly in length due to the existence of multiple repair subpathways within NHEJ, as well as the random nature of the repairs (Puchta and Fauser, 2014; Osakabe and Osakabe, 2015). As a result of this, targeted genes are commonly mutated in a heterozygous manner, with different indels being generated within the two copies of a gene. Heterozygous mutations can be considered either monoallelic, where one allele contains an indel and the other contains the wild-type sequence, or biallelic, where alleles contain indels of different lengths (Soda et al., 2018; Soyars et al., 2018). Homozygous genes may contain the same indel within both alleles or both alleles may be indel-free. These different combinations of alleles can make genotyping potential mutants more difficult since many screening methods are based on detecting mismatches between the sequences of both alleles. Therefore, they are unable to identify biallelic homozygous mutations or differentiate between monoallelic and biallelic heterozygous mutations (Soda et al., 2018; Soyars et al., 2018). Screening potential mutants is made more difficult still by the potential for the generation of chimeric plants. Driven by a constitutive promoter, Cas9 can generate double strand breaks at any point during the T1 or subsequent generations (Fauser et al., 2014). Mutations introduced to germline cells will be maintained in future generations but somatic mutations can also occur if a recognisable target sequence remains present. This can result in the formation of a chimeric plant potentially containing several mutant alleles, in addition to the WT allele, within different cells and tissues (Fauser et al., 2014; Soyars et al., 2018). Mutations are only considered stable when they are homozygous and once the Cas9 T-DNA has been segregated out by screening and crossing subsequent generations to prevent the production of further DSBs.

4.4.3.2 The *Staphylococcus aureus* CRISPR-Cas9 system

CRISPR mutations are commonly introduced into the *Arabidopsis* genome using the Cas9 enzyme from *Streptococcus pyogenes* (SpCas9) which is able to target DNA sequences

adjacent to the PAM sequence NGG (Mojica et al., 2009). In efforts to allow the simultaneous targeting of multiple sites within the *Arabidopsis* genome with Cas9s designed to perform different catalytic functions, the genome editing potential of Cas9 enzymes from other bacterial species has been explored (Steinert et al., 2015). This work identified that the *Staphylococcus aureus* Cas9 (SaCas9) was able to produce targeted mutations at an efficiency greater than that of SpCas9. SaCas9 has also been shown to maintain this high mutation efficiency in *Nicotiana tabacum* and *Oryza sativa* (Kaya et al., 2016). SaCas9 recognises the PAM sequence NNGRRT; specifically, NNGGGT has been reported to be SaCas9's optimum PAM sequence to maximise mutation efficiency and increase the likelihood of a large deletions (Steinert et al., 2015). The increased specificity of the SaCas9 PAM sequence relative to that of SpCas9 is a drawback of the *Staphylococcus aureus* Cas9 system, since it restricts the number of potential target sites. However, this also results in a reduced probability of off-target mutations (Friedland et al., 2015; Kaya et al., 2016). SaCas9 also has the benefit of being a smaller protein than SpCas9, making transformation and expression easier (Ran et al., 2015; Steinert et al., 2015).

4.4.3.3 Generation and transformation of CRISPR-Cas9 constructs for targeting *AtDOA10B*

Based on the reported efficiency and reduced potential for off-target effects, a *Staphylococcus aureus* CRISPR-Cas9 system was chosen to generate an *Atdoa10ab* double mutant by mutating *AtDOA10B* in the *Atdoa10a* genetic background. The system was developed by the Holger Puchta laboratory, Karlsruhe Institute of Technology, Germany, and was used according to a protocol adapted from the method described for the *Streptococcus pyogenes* system (Schiml et al., 2016). The Gateway® compatible system uses the vectors pEn-Sa-Chimera and pDe-Sa-Cas9 (Steinert et al. 2015) to generate a T-DNA construct encoding both a single guide RNA (sgRNA) driven by the AtU6-26 (RNA polymerase III) promoter and the Cas9 endonuclease driven by the PcUBI4-2 (*Petroselinum crispum* ubiquitin) promoter (Figure 30).

Potential 20bp CRISPR-Cas9 target sites adjacent to the *Staphylococcus aureus* PAM sequence (NNGRRT) were identified in the genomic sequence of *AtDOA10B* using CCTop - CRISPR/Cas9 target online predictor (Stemmer et al., 2015). Four optimum target sites were chosen based on their predicted cutting efficiencies (Labuhn et al., 2018) and locations within exons 1, 2 and 3 of *AtDOA10B* (Figure 31), since the earlier the occurrence of a frameshift within a gene, the more significant the resulting truncation is likely to be. Additional features of target sequences were also considered, including the presence of a leading G that aids transcription driven by the AtU6-26 promoter and the presence of W at position 17 and GG immediately prior to the PAM sequence, both of which have been reported to increase the likelihood of DSBs (Shan et al., 2013; Farboud and Meyer, 2015; Takara Bio, 2018).

Oligonucleotides corresponding to targets 1-4 were designed with 5' overhangs of ATTG and AAAC for the forward and reverse oligonucleotides respectively (primers 47-54). These oligonucleotides were annealed together and cloned into the pEn-Sa-Chimera entry vector using the complementarity of the overhangs produced by digesting the vector with Bbs1-HF (New England Biolabs) and those of the designed oligonucleotides, as described by Schiml et al. (2016). Cloning of the annealed oligonucleotides into pEn-Sa-Chimera produced a complete sgRNA, which included the target sequence, under the control of the AtU6-26 promoter. The promoter and sgRNA for each target sequence were then transferred to the pDe-Sa-Cas9 destination vector using the Gateway® recombination sites.

Agrobacterium tumefaciens was transformed with pDe-Sa-Cas9 vectors containing each of the four target sequences. Following confirmation of successful transformation by colony PCR, *Agrobacterium tumefaciens* cultures were grown and used in agrobacterium-mediated transformation of T0 *Atdoa10a* KO *Arabidopsis thaliana* plants.
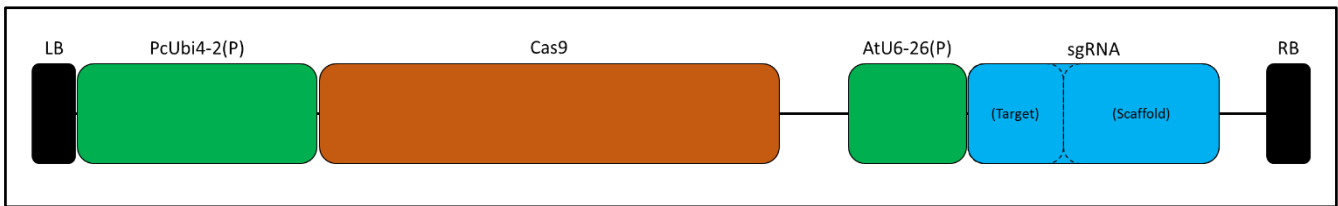
Figure 30 – CRISPR-Cas9 T-DNA construct schematic

The T-DNA insert contained within pDe-Sa-Cas9 encodes both the Cas9 endonuclease and the sgRNA driven by the PcUbi4-2 and AtU6-26 promoters respectively (green). The sgRNA includes both the target sequence and Cas9-binding scaffold sequence. Based on Schiml et al. (2016)
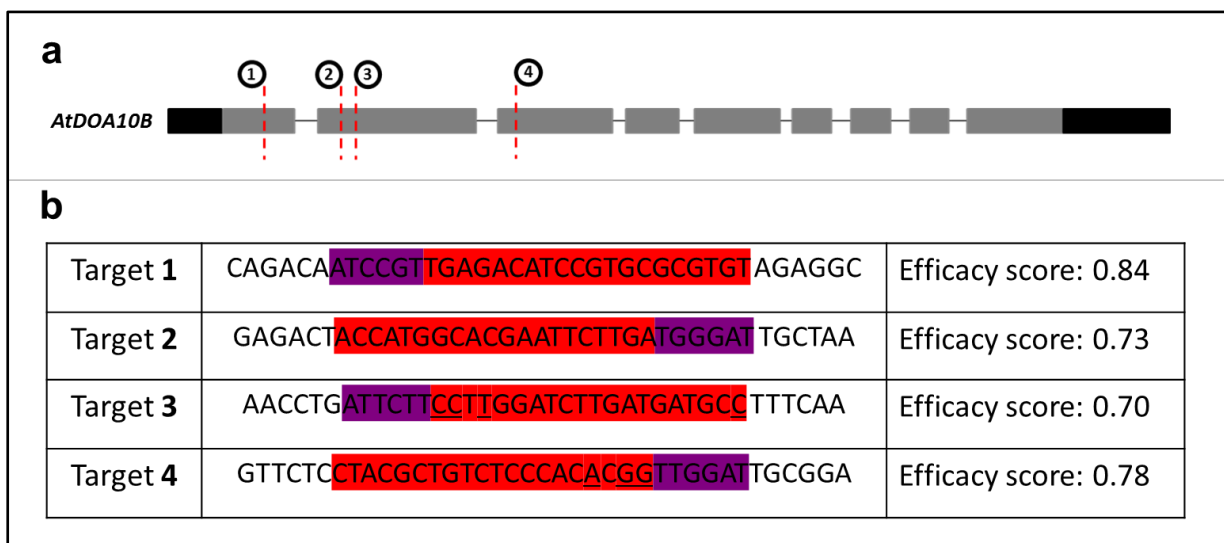
Figure 31 – Selected CRISPR-Cas9 Target sites

(a) Positions of CRISRPR-Cas9 target sites 1-4 in *AtDOA10B*. (b) Target sequences (red) including PAM sequence (purple). DSB efficacy scores listed as predicted by CRISPRater using CCTop - CRISPR/Cas9 target online predictor (Stemmer et al., 2015; Labuhn et al., 2018). Underlined bases correspond to additional desirable features of CRISPR-Cas9 target sites (Shan et al., 2013; Farboud and Meyer, 2015; Takara Bio, 2018).

4.4.3.4 Screening for CRISPR-Cas9 mutations in *AtDOA10B*

Following harvesting, T1 seeds were screened for the kanamycin resistance cassette encoded within the pDe-Sa-Cas9 T-DNA. Initially, only seeds from plants transformed with the two most promising constructs (pDe-Sa-Cas9-Target1 and pDe-Cas9-Target3) were screened. Targets 1 and 3 were selected as most promising based on efficacy score and the identity of nucleotides at significant positions. 40 and 33 resistant Target 1 and Target 3 plants were identified and transplanted onto soil to be grown to the next generation. DNA was extracted from a leaf disc of each kanamycin resistant T1 plant for PCR and high-resolution melt (HRM) analysis.

One way to identify CRISPR-induced indels is to amplify the target gene by PCR and assess changes in the size of the product. As such, PCR and gel electrophoresis was performed using primers that generate ≈500 bp and ≈1000 bp amplicons covering the potential cut sites within *AtDOA10B* (Target 1 Primers 3+51/3+54, Target 3 primers 204+205/3+54). An example gel electrophoresis image is shown overleaf (Figure 32). However, screening by this method is limited by the fact that it is only possible to detect large indels; band shifts caused by small indels, which can also introduce significant frameshifts, are very difficult to detect. Therefore, HRM analysis was employed as a second screening method. HRM analysis was performed by amplifying a ≈100 bp DNA fragment (using primers 105+106/107+108) followed by denaturing, annealing and then incremental heating to identify the temperature at which the two DNA strands separate. Unlike PCR and other screening assays, HRM analysis can identify and distinguish between different mutant alleles (Taylor et al., 2010; Zischewski et al., 2017). Shifts in the peak melting temperature of a PCR product can highlight a change in length of an amplicon as a result of a homozygous mutation. Alternatively, the presence of a secondary peak within the melt curve can indicate the formation of DNA heteroduplexes due to the annealing of amplicons from different alleles (i.e. both mono- and biallelic heterozygous mutations). The melting temperatures of the amplicons including Targets 1 and 3 for control DNA were 81.6 °C and 78.6 °C. DNA samples producing melt peaks shifted from these values,

as well as those producing secondary peaks, were selected for further analysis by Sanger sequencing. Similarly, samples with clear PCR band shifts were also Sanger sequenced.

By using these complementary PCR and HRM approaches over subsequent generations, three Cas9-free, homozygous *Atdoa10ab* double mutants were identified: Target 1 Line 21(11), Target 3 Line 11(2), Target 3 Line 11(11). This demonstrated that *Atdoa10ab* double mutant plants remain viable. Identification and validation of each of these is described below.
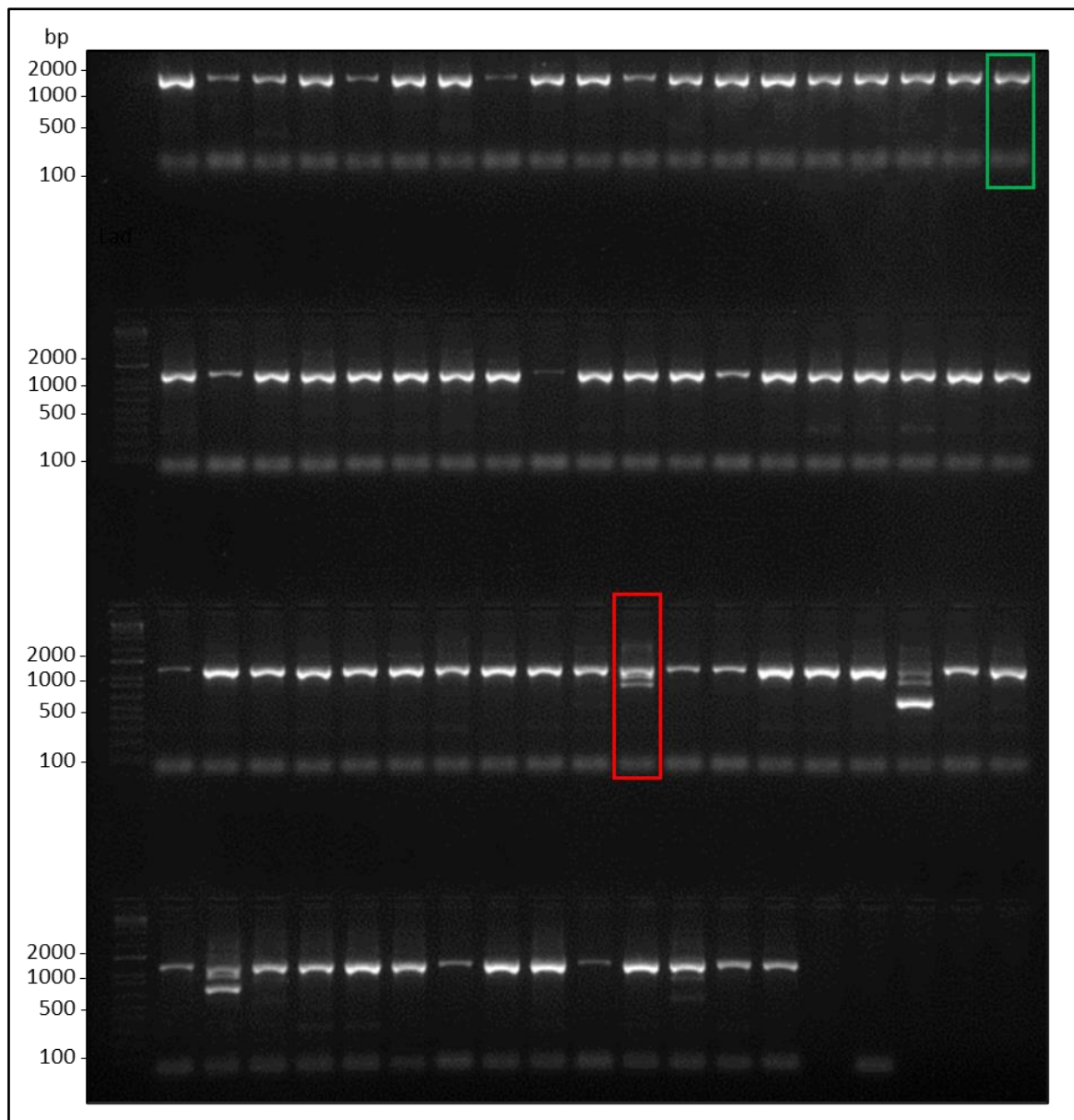
Figure 32 – PCR screening of CRISPR-induced mutations of AtDOA10B in T1 generation
Gel electrophoresis of ≈1 kb amplicons from DNA extracted from T1 CRISPR plants. The amplified region included CRISPR target sites 1 and 3. Multiple bands and band shifts indicated the presence of altered gene alleles arising from indels. Green box = T1 parent of Target 1 Line 21(11). Red box = T1 parent of both Target 3 Line 11(2) and Target 3 Line 11(11).

4.4.3.4.1 Target 1 Line 21(11)

Target 1 Line 21(11) was identified as having a 20 bp deletion within the first exon of *AtDOA10B* (bases 141 to 160). This deletion theoretically results in the production of a heavily truncated protein containing the first 47 amino acids of AtDOA10B, followed by 17 incorrect amino acids before termination due to the formation of a premature stop codon. The complete AtDOA10B protein should contain 860 amino acids.

This line was identified as a double mutant candidate at the T1 generation due to the presence of a secondary peak within the HRM analysis, potentially indicating a heterozygous mutation (Figure 33). The indel was likely to be relatively small due to the absence of a band shift in the gel electrophoresis of the 500 and 1000 bp PCR products covering the target site (Figures 32 and 33).

PCR and HRM analysis were repeated with the T2 generation, where again PCR products from the target site appeared as WT (Figure 33). PCR was also performed to screen for the presence of the Cas9 T-DNA (primers 63+65). 1-21(11) was identified as a line that appeared to be Cas9-free (Figure 33). HRM analysis again highlighted differences between this line and the *Atdoa10a* single mutant. The melt curve of 1-21(11) contained three distinct peaks (Figure 33), again suggesting the presence of a heterozygous mutation. The melting temperature of the primary peak matched that of the *Atdoa10a* control. Secondary and tertiary peaks at lower temperature potentially resulted from the dissociation of mutant DNA strands forming either homoduplexes or heteroduplexes with a WT DNA strand. Sanger sequencing was then performed on the purified ≈500 bp PCR product of 1-21(11) to provide greater clarity on the potential heterozygous mutation. Sequencing in the 3' to 5' direction, the reads generated were consistent with the annotated *AtDOA10B* sequence from base pair 473 to 158. Subsequent reads from base pair 157 towards the start codon were comprised of dual peaks, showing the presence of two alleles (Figure 33). Through analysis of the chromatogram peaks, the two sequences were manually parsed, with one WT and one mutant allele of *AtDOA10B* identified. The isolated mutant sequence appeared to contain a 20 bp deletion.

Analysis was continued with T3 plants, so that an individual carrying a homozygous *AtDOA10B* mutation could be identified. End-point PCR was performed using primers designed to generate a ≈100 bp product (originally intended for HRM analysis), along with high percentage gel electrophoresis to screen for the potential 20 bp deletion. A band shift consistent with such a deletion was observed for some PCR products (Figure 33). Some samples produced double bands; these were assumed to correspond to heterozygous individuals. Those with a shifted single band were assumed to be homozygous for the deletion. Sanger sequencing was then performed on purified ≈500 bp PCR products from an individual ("2B") appearing to be homozygous. The sequencing reads produced consisted of single peaks only, supporting the presence of only one allele, and showed the deletion of base pairs 141 to 160 of *AtDOA10B* (Figure 33). Following this, sequencing was repeated in the same manner but using products purified from the amplification of cDNA. The same 20 bp deletion was detected, confirming that *AtDOA10B* mRNA also carried the mutation.

To fully validate this line as a stable *Atdoa10ab* double mutant T4 seeds were screened for kanamycin resistance. All seedlings were susceptible to kanamycin (Figure 33), proving that the Cas9 T-DNA was no longer present within the genome. PCR was also performed on DNA extracted from eight T4 individuals. A shift in the size of the ≈100 bp product was observed for each of these plants (Figure 33), confirming the homozygosity of the mutation.

To assess any impact of the mutation on the expression of *AtDOA10B*, semi-quantitative RT-PCR was carried out to amplify the ≈1000 bp region using RNA extracted from 1-21(11). A band was detected that matched the size of those generated from Col-0 and *Atdoa10a* RNA but was visibly weaker (Figure 36), suggesting that the mutant transcript may be subject to nonsense mediated decay, a process by which mRNAs containing premature stop codons are degraded (Chang et al., 2007).

Figure 33 – CRISPR-induced *AtDOA10B* mutation screening in line: *Atdoa10ab* 1-11(21)
(T1) ≈1000 bp and ≈500 bp PCR products from CRISPR target site and HRM curves for ≈100 bp PCR products from the candidate plant (blue) and an *Atdoa10a* single mutant control (yellow). (T2) ≈1000 bp and ≈500 bp PCR products, as at T1, and the result of attempted amplification of the Cas9 T-DNA. HRM curves as at T1. Chromatogram shows the sequence of the ≈500 bp amplicon, which contained dual peaks and incorrect base calls (red rather than green letters) from base pair 157 in the 3' to 5' direction. (T3) ≈100 bp PCR product and chromatogram containing only single peaks showing 20 bp deletion from base 141 to 160. (T4) All T4 progenies exhibited a shift in size of the ≈100 bp band confirming homozygosity of the 20 bp deletion. All progenies were also susceptible to 50 ug ml^{-1} kanamycin confirming the absence of the Cas9 T-DNA.

4.4.3.4.2 Target 3 Line 11 (2)

In total, Target 3 Line 11(2) contains 351 deleted base pairs spanning exon 1 and 2. This is made up of 41 missing bases (bases 145 to 185) followed by 5 correct bases and then a 310 bp deletion (bases 191 to 500). The resultant protein sequence is correct for the first 48 amino acids but these are followed by 22 incorrect amino acids prior to a premature stop.

At T1, 3-11 displayed a clear distortion of the ≈1000 bp amplicon, with multiple bands visible, suggesting the presence of several alternative indels (Figures 32 and 34). Unexpectedly the more precise ≈500 bp amplicon produced only a single band matching the *Atdoa10a* control (Figure 34). Additionally, HRM analysis showed a clear decrease in the melt peak of around 1.0 °C, indicative of a homozygous indel (Figure 34). The HRM analysis also showed a very small secondary peak at an even lower melting temperature. This hinted at the presence of some heterozygosity, which implied that this may have been a chimeric individual, which could also contribute to the multiple bands arising from the ≈1000 bp PCR.

Due to the distortions identified, T2 seeds from 3-11 were harvested, grown and the analysis repeated on the next generation. Surprisingly, although changes in relative band intensities were observed, the pattern of PCR product bands seen at T1 persisted into several plants in the T2 generation (Figure 34). This suggested that these bands observed in the T1 parent may not have been due to chimerism. 3-11(2) was one such T2 plant, which also continued to exhibit a WT band for the ≈500 bp amplicon (Figure 34). Surprisingly, this individual plant did not exhibit the same shift in melt temperature as its parent with the melt curve matching that of the *Atdoa10a* control. To clarify the mutation, Sanger sequencing was performed on purified ≈1000 bp PCR products to confirm and characterise any indel present. The resulting chromatogram showed that the PCR product matched the annotated sequence of *AtDOA10B* until base 144, at which point dual peaks were present for each base called, indicating heterozygosity (Figure 34). Using the chromatogram peaks, two sequences were manually parsed which corresponded to a WT and a mutant allele. By ignoring peaks corresponding to WT *AtDOA10B*, the mutant allele appeared to consist of a 351 bp deletion.

PCR was also performed to determine whether the Cas9 T-DNA remained present in this T2 individual. A Cas9 T-DNA band was detected by gel electrophoresis (Figure 34), however the endonuclease T-DNA was confirmed to be heterozygous by observing a 3:1 survival ratio of the T3 progeny on ½ MS + 50 µg ml^{-1} kanamycin plates (not shown).

Due to both the deletion and Cas9 T-DNA appearing to be heterozygous, several T3 progeny from 3-11(2) were grown to maturity and analysed. Gel electrophoresis of the ≈1000 bp PCR amplicons from these individuals yielded three distinct band patterns (Figure 34). Some plants produced a single band corresponding to WT *AtDOA10B*, others re-produced the banding pattern observed at T2 and some showed a single band at ≈700 bp. These bands corresponded to homozygous WT, heterozygous and homozygous *Atdoa10b* mutants respectively, confirming that the T2 parent was indeed heterozygous and showing that the distinctive banding pattern observed was a characteristic of heterozygous mutants. Of the plants recognised as homozygous *Atdoa10b* mutants, three were also identified as Cas9-free by PCR (Figure 34). One of these individuals ("4B") was then sequenced in the same manner as at T2 which generated a chromatogram containing only single peaks, providing further evidence of homozygosity, and showed the presence of a 41 bp and a 310 bp deletion separated by 5 correct nucleotides (145 to 185 and 191 to 500). As with 1-21(11), this deletion was also shown to be present in *AtDOA10B* transcripts in the 3-11(2)"4B" line by sequencing a PCR product derived from cDNA (not shown).

To fully confirm the deletion and absence of Cas9 T-DNA was homozygous, T4 offspring from 3-11(2)"4B" were analysed. All eight seedlings tested, exhibited a clear band shift of the ≈1000 bp PCR product confirming homozygosity (Figure 34). Seedlings were also 100% susceptible to kanamycin, demonstrating that none contained the T-DNA (Figure 34).

Semi-quantitative RT-PCR amplifying the ≈1000 bp amplicon was again performed to assess the levels of the *AtDOA10B* transcript. Compared to Col-0 and *Atdoa10a*, a very weak band was detected (Figure 36). This band was also shifted to a lower molecular weight, consistent with the large deletion identified.
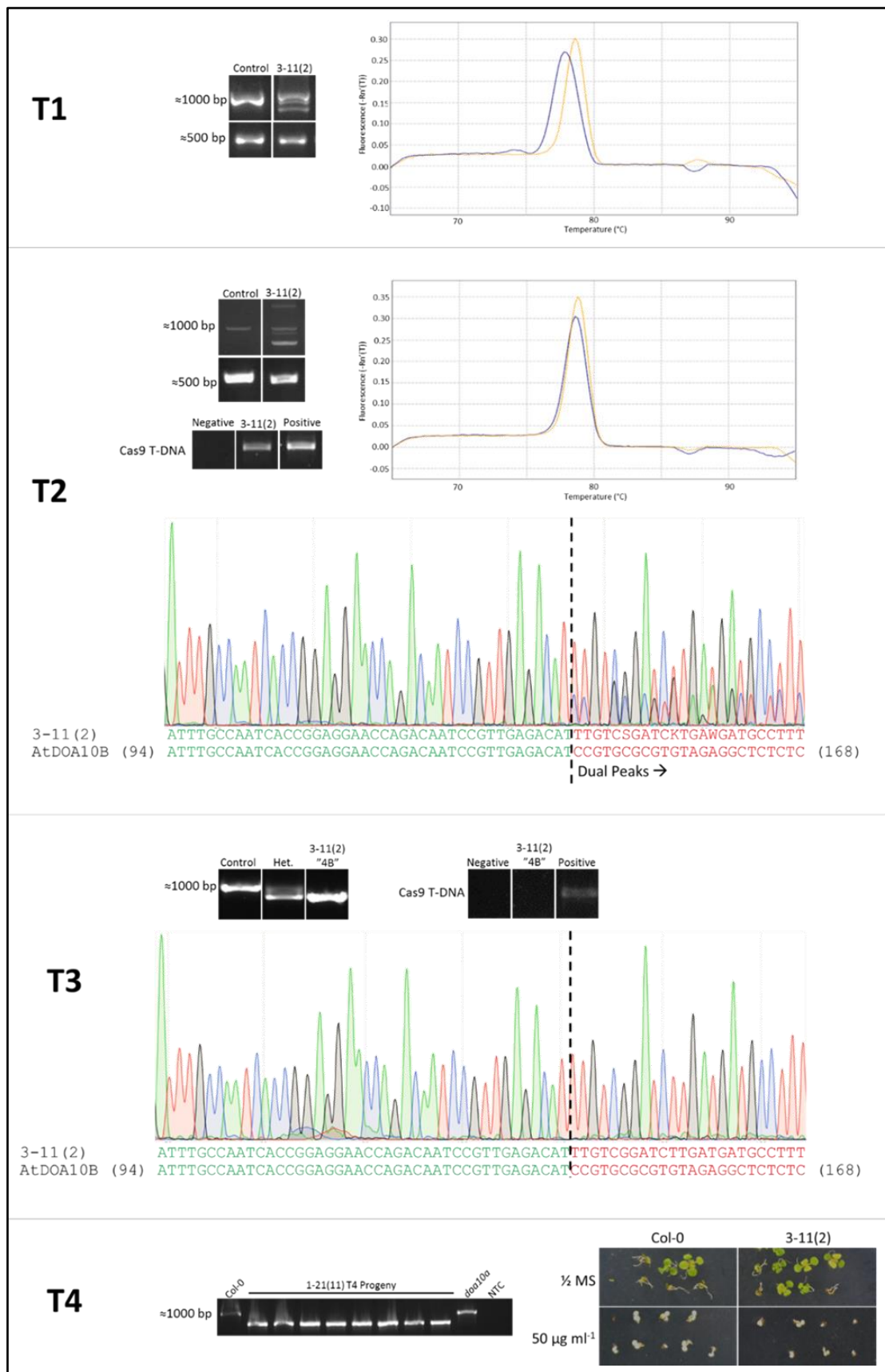
Figure 34 – CRISPR-induced *AtDOA10B* mutation screening in line: *Atdoa10ab* 3-11(2)
(T1) ≈1000 bp and ≈500 bp PCR products from CRISPR target site and HRM curves for ≈100 bp PCR products from the candidate plant (blue) and an *Atdoa10a* single mutant control (yellow). (T2) ≈1000 bp and ≈500 bp PCR products, as at T1, and the result of attempted amplification of the Cas9 T-DNA. HRM curves as at T1. Chromatogram shows sequence of the ≈1000 bp amplicon, which contained dual peaks and incorrect base calls (red rather than green letters) from base pair 145. (T3) ≈1000 bp PCR products from *Atdoa10a* (control), a candidate double mutant (3-11(2)"4B") and a heterozygous individual. 3-11(2)"4B" chromatogram contained only single peaks and showed the deletion of bases 145 to 185 and 191 to 500 (351 bp in total). (T4) All T4 progenies exhibited a shift in size of the ≈1000 bp band confirming homozygosity of the 351 bp deletion. All progenies were also susceptible to 50 ug ml[-1] kanamycin confirming the absence of the Cas9 T-DNA.

4.4.3.4.3 Target 3 Line 11(11)

A 14 bp deletion within exon 2 was identified in Target 3 Line 11(11) (bases 500 to 513). This deletion also translates to a significant protein truncation with 116 correct amino acids preceding 20 incorrect prior to termination of the peptide sequence.

Line 3-11 was identified as a promising candidate in the T1 generation due to abnormalities of its ≈1000 bp PCR product and melt curve (see 4.2.3.2.4.2) (Figure 35). When the same PCRs were performed on the T2 generation, some plants produced the same distorted ≈1000 bp amplicon but others produced a WT band. Plant 3-11(11) showed no abnormalities in either the ≈1000 bp or ≈500 bp PCR product but did appear be free from the Cas9 T-DNA (Figure 35). A sample of purified ≈1000 bp PCR product from 3-11(11) was sequenced alongside several other candidate plants. The sequence generated from 3-11(11) contained a 14 bp deletion, when compared to the annotated *AtDOA10B* sequence (Figure 35). The generated sequence also contained only single peaks, suggesting that the mutation was homozygous. Sequencing was then repeated using a purified PCR product generated from 3-11(11) cDNA, alongside an *Atdoa10a* control. The result showed that the homozygous 14 bp deletion was maintained in *AtDOA10B* mRNA of plant 3-11(11) and was not present in mRNA of the *Atdoa10a* control (not shown).

To confirm that 3-11(11) was a Cas9-free individual bearing a homozygous *Atdoa10b* mutation, eight T3 plants were shown to all be susceptible to kanamycin and to all produce distorted PCR bands (Figure 35). A slight band shift in the ≈500 bp amplicon corresponding to the 14 bp deletion was observed for each of the T3 plants. These plants also failed to generate a ≈100 bp product by PCR due to the deletion removing 11 of the 22 bases complementary to the reverse primer used (primer 108).

RT-PCR was also carried out using RNA extracted from this line. No ≈1000 bp *AtDOA10B* band could be detected for 3-11(11), despite levels of *AtACT7* appearing similar to control samples (Figure 35). This suggested that nonsense mediated decay degrades the majority of mutant *AtDOA10B* transcripts in this line.
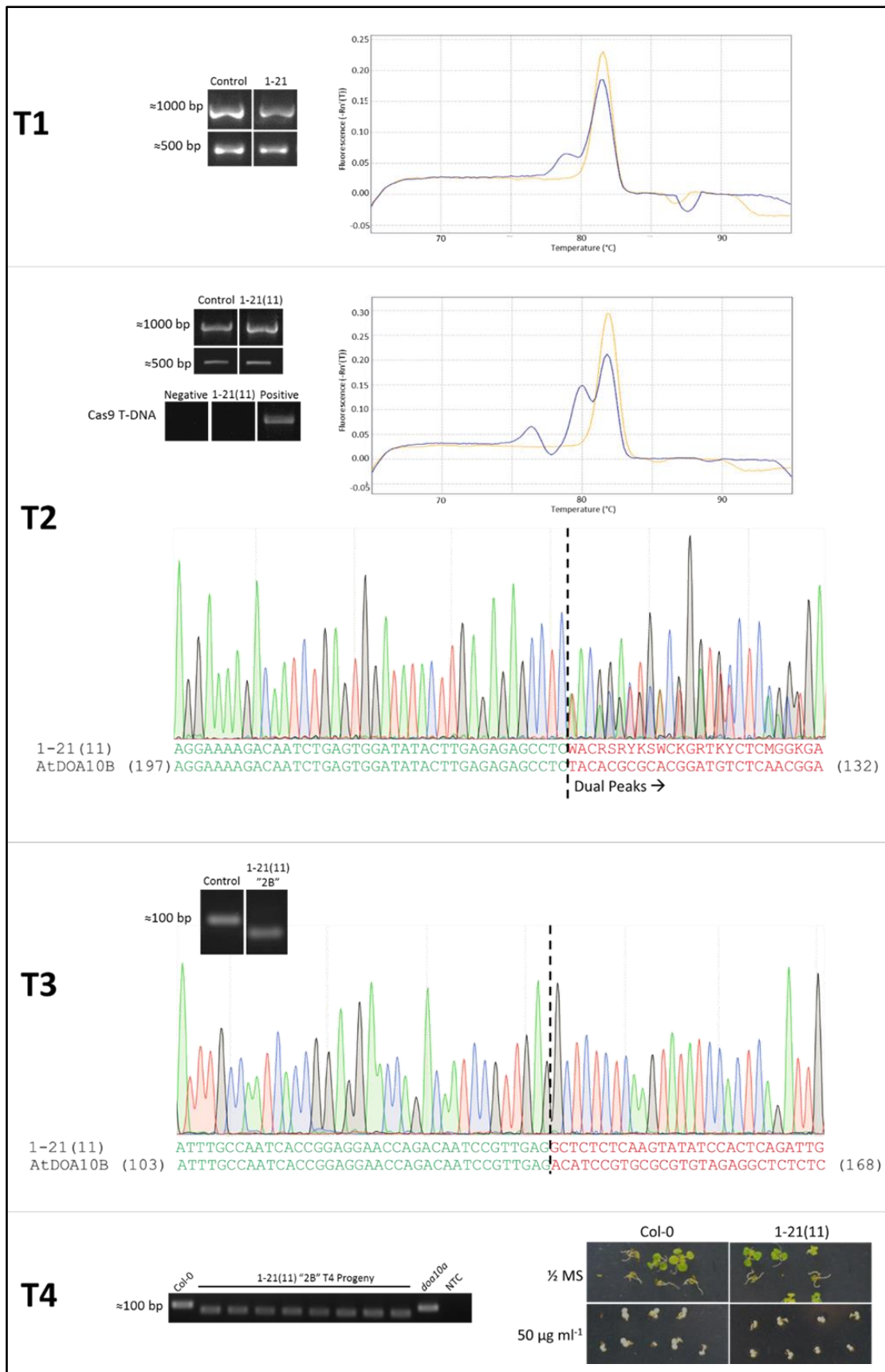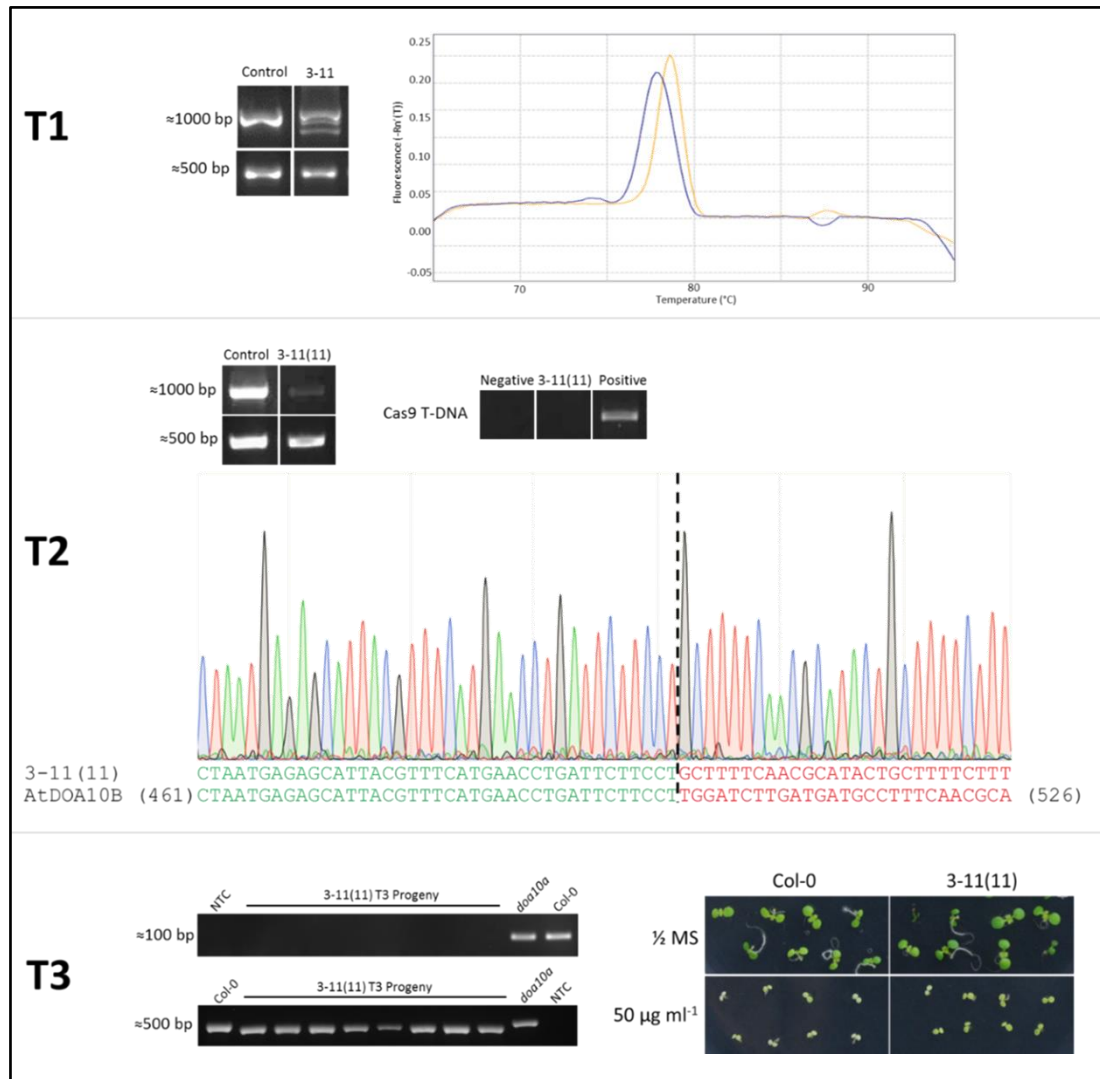
Figure 35 – CRISPR-induced *AtDOA10B* mutation screening in line: *Atdoa10ab* 3-11(11)
(T1) ≈1000 bp and ≈500 bp PCR products from CRISPR target site and high-resolution melt curves for ≈100 bp PCR products from the candidate plant (blue) and an *Atdoa10a* single mutant control (yellow). (T2) ≈1000 bp and ≈500 bp PCR products, as at T1, and the result of attempted amplification of the Cas9 T-DNA. Chromatogram shows the sequence of the ≈1000 bp amplicon which contained only single peaks but incorrect base calls (red rather than green letters) from base pair 500 caused by a 14 bp deletion of bases 500-513. (T3) All T3 progenies failed to produce a ≈100 bp PCR band and exhibited a shift in size of the ≈500 bp band confirming homozygosity of the 14 bp deletion. All progenies were also susceptible to 50 ug ml^{-1} kanamycin confirming the absence of the Cas9 T-DNA.
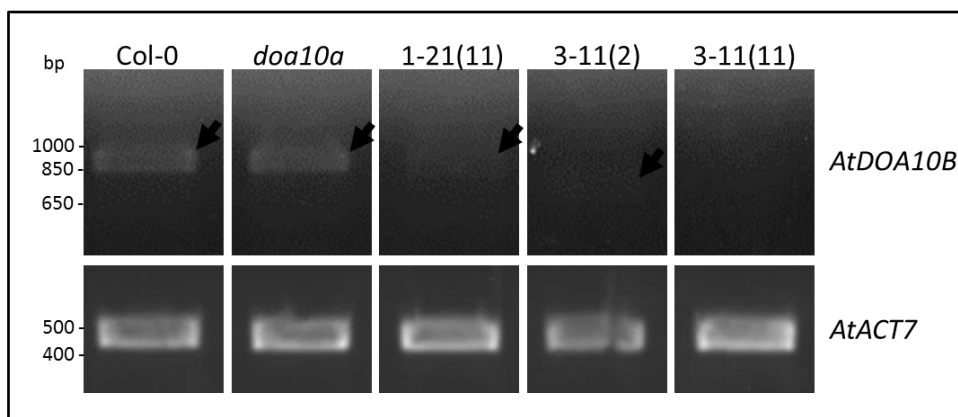
Figure 36 – RT-PCR of CRISPR-targeted *AtDOA10B* in *Atdoa10ab* double mutants
Expression of *AtDOA10B* appeared to be reduced in each of the CRISPR lines. A shift in the size of the RT-PCR product was visible for line 3-11(2). For line 3-11(11) no band could be detected at all.

4.5 Discussion

A range of genetic tools were developed to aid characterisation of AtDOA10A and AtDOA10B in *Arabidopsis*. These included publicly available *Atdoa10a* and *Atdoa10b* T-DNA lines. The *Atdoa10a* mutant line has been studied previously and has a range of reported phenotypes (Jenks et al., 1995; Lu et al., 2012; Doblas et al., 2013; Zhao et al., 2014) (see 1.5.3.2). Further analysis of this line performed here had the potential to give insight into the causes of these phenotypes and also uncover as yet unknown characteristics of this line. No work on *Atdoa10b* has been published. The *Atdoa10b* mutant therefore provided a useful tool for investigating redundancy between the two E3s by assessing any correlations of their phenotypes.

To accompany the T-DNA mutants, GUS- and EYFP-tagged reporter lines were developed. These lines expressed tagged versions of AtDOA10A or AtDOA10B, under the control of their native promoters, in the corresponding T-DNA mutant background, thereby complementing the null mutation. Expression of the recombinant proteins was confirmed by Western blot (Figure 25b). Both AtDOA10A-GUS and AtDOA10B-GUS bands appeared at a higher molecular weight than expected, potentially due to the post translational modifications. Such modifications are likely to have occurred on the GUS tag rather than the E3 since the EYFP tagged proteins appeared at the expected size. The independent generation of complementation lines with two different protein tags ensures results are specific to AtDOA10A or AtDOA10B rather than due to the impact of the tag. Each AtDOA10A reporter line generated, successfully complemented the ABA hypersensitivity phenotype of *Atdoa10a* (Figure 25c), confirming the functionality of the lines and that this phenotype is specifically caused by the absence of AtDOA10A rather than any additional unknown genomic changes. The colorimetric nature of both the GUS and EYFP also allows lines such as these to be used to monitor the localisation of AtDOA10A and AtDOA10B (see 3.5 and 3.6).

The generation of an *Atdoa10ab* double mutant was important due to the possibility of functional redundancy masking the impacts of the individual mutations. A range of methods

were employed to achieve this. Isolating a double mutant from crossing the two single mutants had a low probability of success because of the proximity of the genes on chromosome four. Based on this, and the success of alternative approaches, screening of offspring from an *Atdoa10a-Atdoa10b* cross was not continued beyond the F2 generation. CRISPR-Cas9 and RNAi were used to eliminate and reduce *AtDOA10B* expression respectively. For both techniques, *AtDOA10B* was targeted in the *Atdoa10a* background, since this T-DNA line has been characterised in previous publications (Lu et al., 2012; Zhao et al., 2014) and the insertion lies within an exon (the *Atdoa10b* T-DNA is located in the 5' untranslated region). The RNAi approach yielded independent lines with 46.0 and 38.5 % reductions in *AtDOA10B* expression, but the CRISPR-Cas9 system proved the most effective method of preventing AtDOA10B expression. The use of CRISPR-Cas9 produced an array of AtDOA10B mutations, three of which were fully characterised for use in subsequent experiments. When screening the T1 generation, a larger number of significant PCR band shifts and deviations in melt peak were observed for plants containing Target 3 CRISPR constructs. Target 3 had an inferior DSB efficacy score than Target 1, as predicted by CRISPRater (Labuhn et al., 2018) (Figure 31) but contained several reportedly beneficial sequence features (Shan et al., 2013; Farboud and Meyer, 2015; Takara Bio, 2018). Therefore, in this case, features such as the presence of GG adjacent to the PAM sequence, seemed more important than predicted efficacy score. However, only a selection of individuals were sequenced. It is possible that the Target 1 construct induced more indels, but that these were not detected by PCR or HRM analysis.

Overall, the collection of genetic tools generated provided a key resource for further examination of AtDOA10A and AtDOA10B's *in planta* functions carried out in Chapter V.

Chapter V -

Phenotypic, transcriptomic and proteomic analyses of

AtDOA10 function

5.1 Introduction

Results presented in Chapter III suggested that AtDOA10A and AtDOA10B were involved in ERAD in *Arabidopsis*, but that their overall contribution to this process may be relatively minor. The development of lines described in Chapter IV, provided the tools required to clarify their significance within the ERAD system and on overall plant growth and fitness. Investigating which aspects of development these E3s are involved in will shed light on the roles of ScDoa10 homologues throughout the plant kingdom and has the potential to elucidate the mechanisms underlying the observed drought tolerance and ABA hypersensitivity phenotypes of *Atdoa10a* mutants (see 1.5.3.2). Analysis may also provide explanations for the differences observed between AtDOA10A and AtDOA10B in Chapter III, namely their gene regulation in response to TM (Figure 18) and their ability to functionally complement ScDoa10 (Figure 21). Studies of the developed lines also has the potential to shed light on specific processes regulated by the AtDOA10s and whether they participate in an *Arabidopsis* Ac/N-degron pathway, as has been reported for ScDoa10 in yeast (Hwang et al., 2010). Identifying processes influenced by AtDOA10A and AtDOA10B may additionally aid the identification of ubiquitination targets and therefore potential substrates of the Ac/N-degron pathway.

This chapter describes phenotypic, transcriptomic and proteomic comparisons of the *AtDOA10* genetic tools and WT Col-0 *Arabidopsis* and discusses the significance of differences between them.

5.2 Phenotypic observations of mutant and transgenic plant materials

5.2.1 *Atdoa10* mutants do not display developmental defects in standard conditions

During growth to maturity, *Atdoa10a* did not demonstrate any clear differences in fitness or morphology in comparison to WT Col-0 plants (Figure 37). The GABI_588A06 *Atdoa10a* T-DNA mutant used here, as well as alternative ethyl methanesulfonate mutants, have been reported previously to exhibit a semiglossy inflorescence stem (Koornneef et al., 1989; Lu et al., 2012). This subtle phenotype was not observed, although stem glossiness was not extensively examined. Like *Atdoa10a*, *Atdoa10b* mutants also resembled WT plants throughout development (Figure 37). Unsurprisingly therefore, complementation lines expressing EYFP and GUS-tagged AtDOA10A and AtDOA10B fusion proteins in the corresponding mutant background also appeared to grow as WT (not shown). *Atdoa10ab* double mutants developed using the CRISPR-Cas9 system (see 4.2.3.2) also did not display any clear phenotypes (Figure 37), demonstrating that the lack of phenotypes of either single mutant is not due to functional redundancy between the homologous E3s. Similarly, RNAi lines, in which *AtDOA10B* expression was suppressed in the *Atdoa10a* background did not display any developmental defects when grown under standard conditions (not shown).

Figure 37 – Rosette images of *Arabidopsis Atdoa10* mutants
Representative rosette images showing that both single and double KO mutations of *AtDOA10A* and *AtDOA10B* genes do not have a significant effect on development throughout the *Arabidopsis* lifecycle. Photographs taken 25 days after sowing. *Atdoa10ab* is CRISPR line 3-11(2).

5.2.2 Protein misfolding stress-associated phenotypes of *Atdoa10* mutants

Due to the role ERAD E3 ligases play in the targeted degradation of misfolded proteins, it is perhaps not surprising that no obvious phenotypes were observed for *Atdoa10* mutants under non-stressed conditions. To induce protein misfolding, and therefore increase demand on the ERAD system, Col-0 and *Atdoa10* mutant seedlings were exposed to a range of chemical stresses and their responses were assessed. The first of these was azetidine-2-carboxylic acid (AZC), a proline analogue that generates misfolded proteins in *Arabidopsis* through its incorporation within polypeptides, resulting in a change in the relative angle of the peptide bonds formed (Fowden and Richmond, 1963; Trotter et al., 2002; Sugio et al., 2009). Secondly, the reducing agent dithiothreitol (DTT), which denatures proteins though the reduction of disulfide bonds between cysteine residues that are required to maintain protein tertiary structure, was also used (Jasma et al., 1994; Kamauchi et al., 2005). Both stresses were applied through their direct addition to ½ MS growing media and plant growth was monitored for 3 weeks.

Plants grown on AZC and on DTT exhibited delayed development and reduced growth in comparison to those grown on ½ MS media control plates (Figure 38). Relative to Col-0 plants, neither *Atdoa10a* or *Atdoa10b* T-DNA insertion lines demonstrated any increased tolerance or resistance to the ERAD stresses. The growth of each line was similarly inhibited by the treatments (Figure 38).

Next we examined responses to the alternative protein misfolding agent geldanamycin. This antibiotic inhibits the function of Heat Shock Protein 90 (HSP90), a protein chaperone that prevents misfolding, particularly in response to heat stress (Roe et al., 1999; Queitsch et al., 2002). Geldanamycin treatments were performed transiently, due to the high concentration required and the instability of the compound. *Arabidopsis* seedlings were grown on vertical ½ MS plates for six days, at which point they were transferred to water supplemented with 150 µM geldanamycin or an equivalent volume of DMSO (control) and agitated for 5 hours. Individual seedlings were then removed from the treatment and mounted onto fresh vertical ½

MS plates and left to grow for a further eight days. In general, the treatment appeared to cause the death of most Col-0 seedlings. *Atdoa10a* seedlings showed a higher rate of survival, although this was variable (Figure 39). *Atdoa10b* seedlings responded differently in multiple biological replicates of the treatment, ranging from Col-0-like susceptibility to almost complete survival (Figure 39). *Atdoa10a/b* (RNAi line 4-2(2)) was also tested, with a slight resistance observed, similar to *Atdoa10a* (Figure 39).

Despite preliminary evidence of an altered geldanamycin response in the *Atdoa10* mutants, the inconsistency of the observed phenotypes prompted the use of tunicamycin (TM) as a further alternative ERAD stressor. A transient TM treatment was previously observed to significantly upregulate *AtDOA10B* but downregulate *AtDOA10A* in Col-0 seedlings (Figure 18). TM causes ER-specific protein misfolding by inhibiting N-linked glycosylation (Koizumi et al., 1999; Chen and Brandizzi, 2013). Like AZC and DTT, TM was sufficiently stable to affect seedling development when added directly into ½ MS growth media. Here, as with the geldanamycin treatment, a subtle TM-resistance phenotype was observed in the *Atdoa10a* mutant. This resistance was not observed in *Atdoa10b* (Figure 40). An *Atdoa10ab* double mutant (CRISPR line 3(11)-2), was also assessed for TM-associated phenotypes. Like *Atdoa10a*, *Atdoa10ab* also demonstrated a resistance to the ERAD stressor, particularly at a concentration of 75 ng ml^{-1} (Figure 40).

The subtleness of the resistance phenotypes observed supports the idea that AtDOA10A and AtDOA10B only make a minor contribution to the overall ERAD system (see Chapter 3). However, the mild TM and geldanamycin insensitivity of *Atdoa10a* seedlings was perhaps surprising given that the loss of this E3 would be expected to disrupt normal ERAD function and hence potentially cause increased susceptibility. This partial resistance to ERAD-elicitors may arise from the upregulation of alternative ERAD pathways due to the *Atdoa10a* mutation. Alternatively, the absence of AtDOA10A may permit misfolded but still functional proteins to escape degradation, potentially allowing them to fulfil their role and contribute to plant fitness.

Enzymes of the sterol synthesis pathway, such as HMGR and SQE are known to be degraded by the ERAD system (see 1.2.1 and 1.5.3). A range of compounds are generated during sterol biosynthesis in plants which regulate signalling, defence, membrane fluidity and more (Hartmann, 1998; Rogowska and Szakiel, 2020; De Vriese et al., 2021). The sterol synthesis pathway also produces precursors for many phytohormones, particularly brassinosteroids (Fujioka and Yokota, 2003). Brassinosteroids play a key role in the salt (NaCl) stress response (Krishna, 2003; Kagale et al., 2007) and removal of AtUBC32, a ScUbc6 homologue known to interact with AtDOA10B, has been shown to increase brassinosteroid responses and enhance salinity tolerance (Cui et al., 2012a). As high salinity also acts as an ERAD stress (Liu et al., 2011), the responses of the *Atdoa10* mutants to increasing concentrations of salt were tested. Here, increasing concentrations of salt inhibited the growth of seedlings but none of the *Atdoa10* mutant lines exhibited any increase in resistance or sensitivity to the treatment relative to Col-0 (Figure 41). This adds further support to the conclusion that AtDOA10A and AtDOA10B are of relatively low significance within the ERAD system.
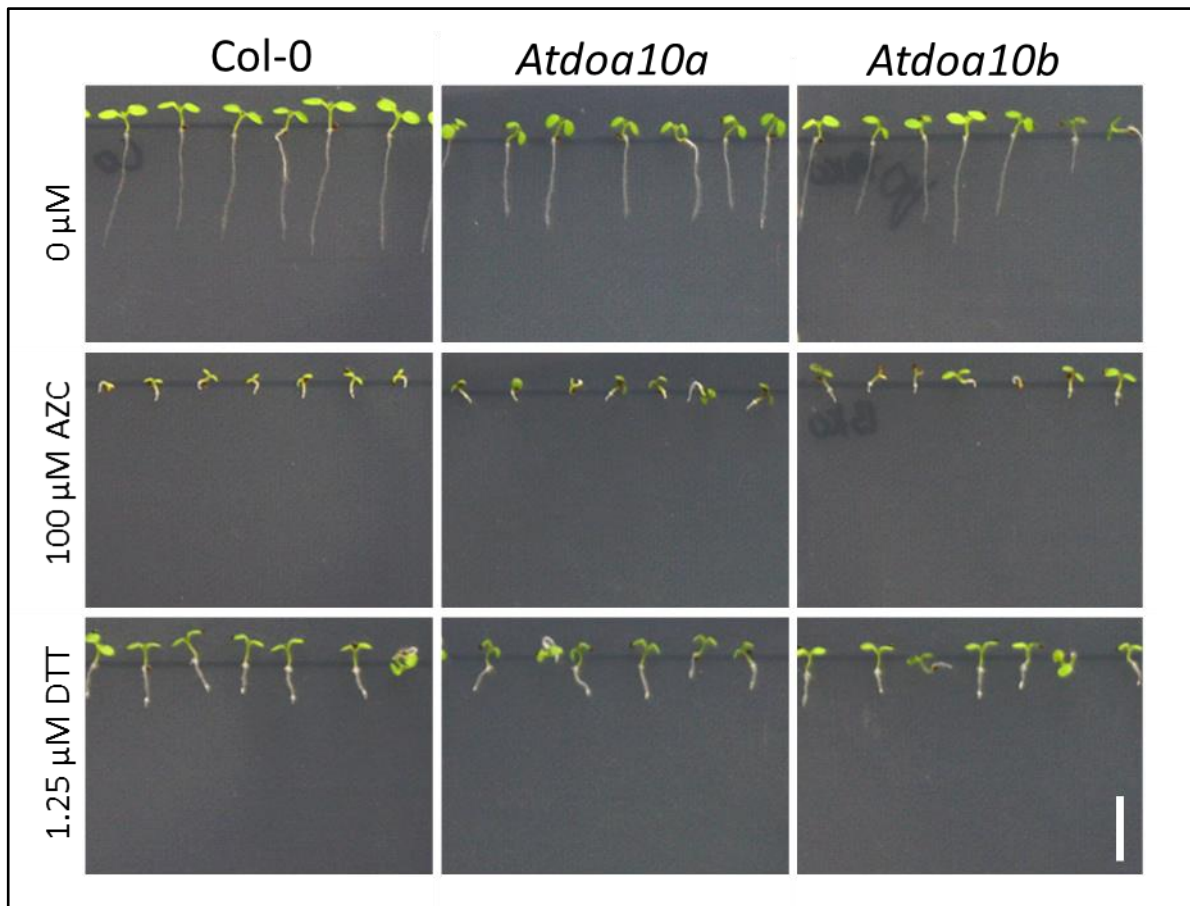
**Figure 38 – Azetidine-2-carboxylic acid (AZC) and dithiothreitol (DTT) treatments of
Arabidopsis doa10 mutants**

Atdoa10a and *Atdoa10b* responded as WT (Col-0) when grown on ½ MS media containing 100 μM
AZC and 1.25 μM DTT. Representative images of 6-day old seedlings. Scale bar = 5mm.
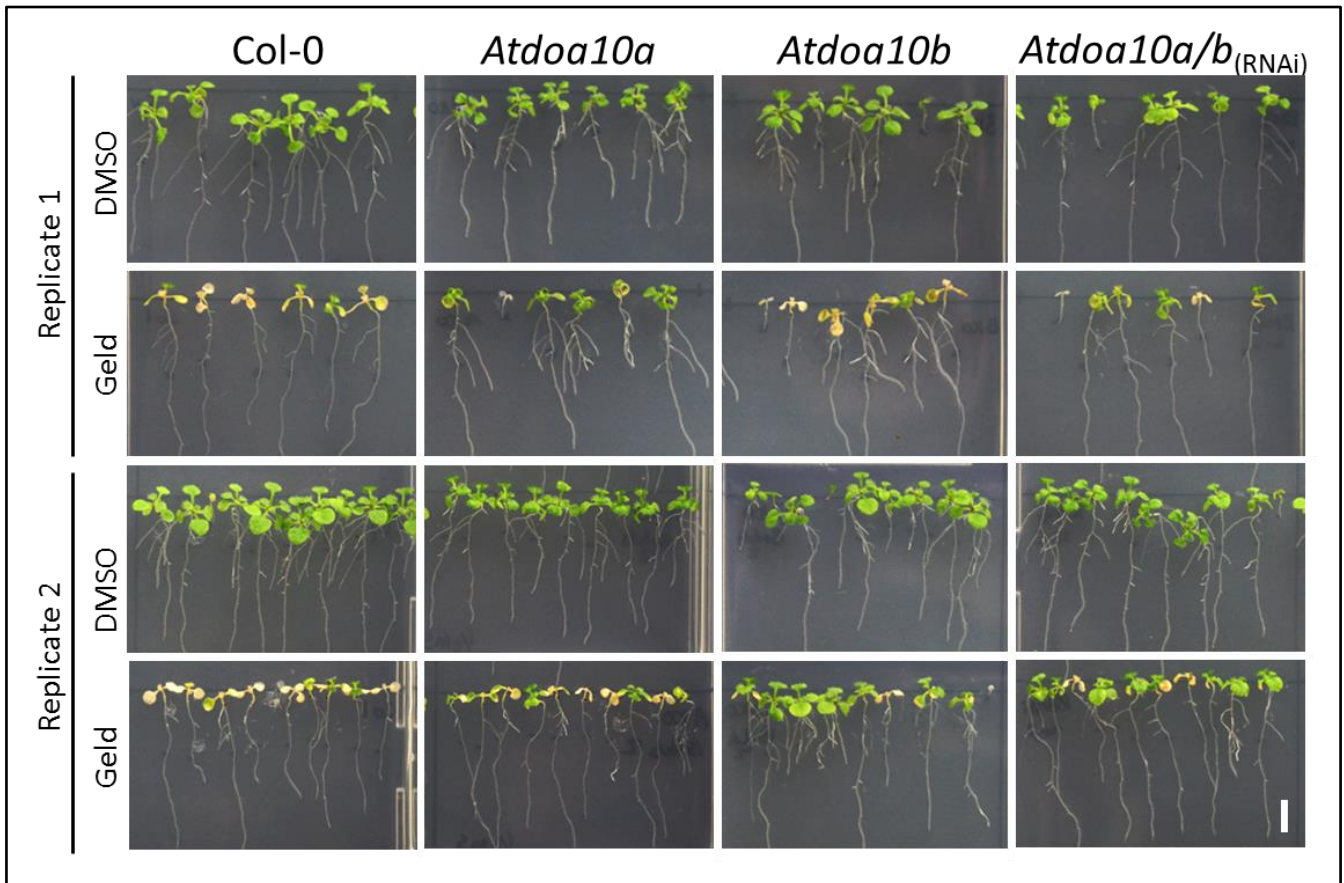
Figure 39 – Geldanamycin treatment of *Arabidopsis doa10* mutants

Atdoa10a and *Atdoa10a/b* (RNAi) showed slight resistance to geldanamycin (Geld) treatment. The responses of *Atdoa10b* to geldanamycin were highly variable between replicates. Images show 14-day old seedlings from two biological replicates that were treated with 150 µM geldanamycin or an equivalent volume of DMSO (control) for 5 hours on day 6. *Atdoa10a/b* is RNAi line 4-2(2). Scale bar = 5mm.
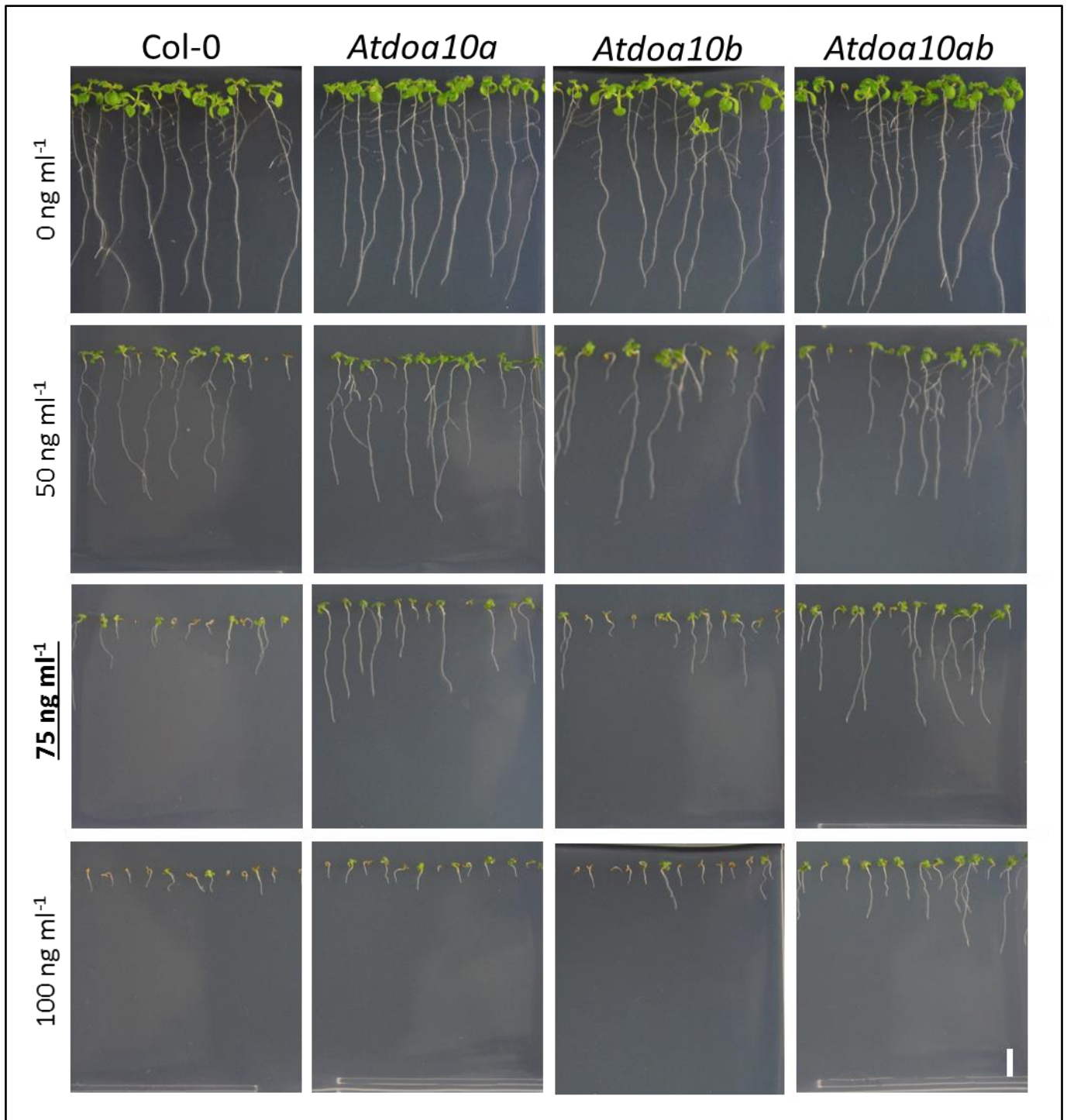
Figure 40 – Tunicamycin (TM) treatment of *Arabidopsis doa10* mutants

TM inhibited the growth of *Atdoa10a* and *Atdoa10ab* mutants less than WT (Col-0) and *Atdoa10b,* especially at a concentration of 75 ng ml^{-1}. Representative images of 10-day old seedlings grown on ½ MS media containing 50-100 ng ml^{-1} of TM. *Atdoa10ab* is CRISPR line 3(11)-2. The resistance phenotype was also observed with other CRISPR *Atdoa10ab* double mutants. Scale bar = 5mm.
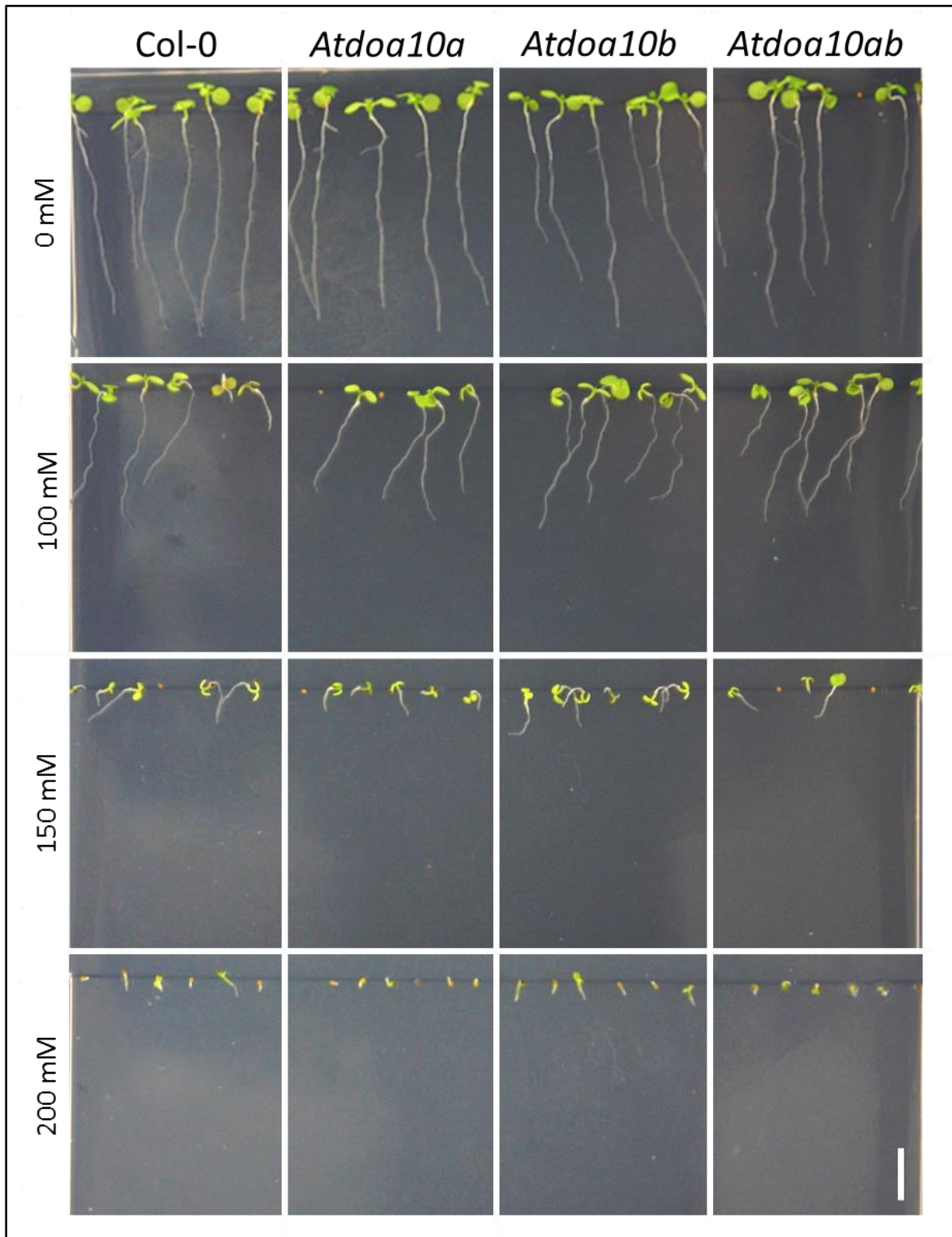
Figure 41 – Salt (NaCl) stress treatment of *Arabidopsis doa10* mutants
Increasing concentrations of NaCl inhibited seedling growth similarly in all lines, independent of *Atdoa10a* and *Atdoa10b* mutations. Representative images of 7-day old seedlings grown on ½ MS media containing 100-200 mM NaCl. *Atdoa10ab* is CRISPR line 3(11)-2. Scale bar = 5mm.

5.2.3 Hormone-associated phenotypes of *Atdoa10* mutants

Given their established links to the sterol synthesis pathway and the known ABA hypersensitivity of *Atdoa10a*, next hormone-associated phenotypes of the *Atdoa10* mutants were examined.

Initially, the responses of the mutants to ABA were investigated. *Atdoa10a* was already known to exhibit ABA hypersensitivity, most likely by acting as a negative regulator of ABA biosynthesis (see 1.5.3.2) (Zhao et al., 2014). The effect of ABA on *Atdoa10b* remained unknown. As expected *Atdoa10a* and *Atdoa10a/b* (RNAi line 4-2(2)) were hypersensitive to ABA (Figure 42), as were *Atdoa10ab* CRISPR double mutants (not shown). Complementation of *Atdoa10a* with AtDOA10A-EYFP restored normal ABA sensitivity (Figure 42), confirming that the loss of AtDOA10A is responsible for the ABA phenotype. Unlike *Atdoa10a*, *Atdoa10b* single mutants did not show an altered ABA response (Figure 42). The contrasting phenotypes of *Atdoa10a* and *Atdoa10b* showed that any redundancy in the functions of the two E3s does not extend to their involvement in ABA biosynthesis or signalling.

Next, the response of the mutant lines to gibberellins (GAs), was assessed. Like ABA, GAs are synthesised from precursors generated by the plastid specific methylerythritol phosphate (MEP) pathway, an equivalent pathway to the cytosolic mevalonate (MVA) pathway, which results in the formation of the same precursor compounds (De Vriese et al., 2021). The MVA pathway is potentially influenced by the AtDOA10s in a dual manner, by affecting the expression of AtHMGR, the rate limiting enzyme of the MVA pathway (Brown and Goldstein, 1980), and by regulating the sterol biosynthesis pathway, a consumer of the end products of the MVA pathway. Crosstalk between the analogous MEP and MVA pathways has been reported (Laule et al., 2003; Mendoza-Poudereux et al., 2015; Movahedi et al., 2021), providing a theoretical link between AtDOA10s and ABA and GA biosynthesis.

Seedlings were grown on ½ MS media with 5 µM gibberellic acid (GA_3) added. GA treatment caused an observable increase in aerial tissue in Col-0. This was also observed for *Atdoa10b* but was less obvious for *Atdoa10a* and *Atdoa10ab* (Figure 43). The reduced

response to GA was consistent with the reported increase in ABA levels in *Atdoa10a* (Zhao et al., 2014). Therefore, the observed differences in GA response may be the result of excess ABA, rather than an indicator of AtDOA10A being involved in GA biosynthesis or signalling.

To further dissect a role of AtDOA10A in GA biosynthesis, separate from its impact on ABA levels, the responses of the mutant lines to the GA biosynthesis inhibitor paclobutrazol (PAC) (Rademacher, 2000) were examined. The addition of 1 µM PAC to the ½ MS growth medium significantly reduced shoot and root growth of all the lines tested. Root lengths were shortest for *Atdoa10a* and *Atdoa10ab*, although overall health appeared similar (Figure 43). Since differences between the lines with and without a functional AtDOA10A were relatively minor, they could, again, have been the result of variation in ABA level rather than GA-specific responses. As such, the results were inconclusive with regards to a role of AtDOA10A in GA biosynthesis, however, the relatively understated phenotypes observed would suggest that any influence on this process is modest.

The responses of the *Atdoa10* mutants to 24-epibrassinolide (EBR) were also investigated because brassinosteroids are the most direct hormone product of the sterol synthesis pathway. EBR, a stereoisomer of the first characterised brassinosteroid (Grove et al., 1979), has wide ranging impacts including promoting cell elongation and division, controlling vascular differentiation and alleviating abiotic stresses, particularly salt stress (Clouse and Sasse, 1998; Tanveer et al., 2018). When grown on ½ MS containing 100 mM EBR, seedlings displayed a characteristic EBR-induced wavy-root phenotype. This phenotype is thought to be the result of alterations of the actin cytoskeleton and distribution of AtPIN2 auxin transport proteins in response to the EBR (Lanza et al., 2012). The wavy-root phenotype, along with an overall reduction in primary root length, was observed to a similar extent in all lines (Figure 44), suggesting that AtDOA10A and AtDOA10B activity is not linked to brassinosteroid biosynthesis or response. This was consistent with the lack of a salt stress phenotype of the *Atdoa10* mutants (Figure 41), since brassinosteroids enhance salinity tolerance (see 4.3.3). A subtle difference in the regularity of root waves could be observed

between lines carrying mutations of *Atdoa10a* and those with WT *AtDOA10A* alleles, potentially indicating a peripheral role of this E3 in the establishment of this phenotype. *Atdoa10a* and *Atdoa10ab* (CRISPR line 3-11(11)) appeared to exhibit a stronger gravitropic response with more regular waves, whereas Col-0 and *Atdoa10b* seemed hyper-curled and to have lost almost all directionality. Whether a more regular waved root was indicative of a stronger or weaker EBR response would require further investigation.

As the most significant of the plant hormones in regulating growth and development, auxin was also employed to screen for any altered responses in the *Atdoa10* mutant lines. Auxins are a class of plant hormones that, amongst many other things, control shoot architecture, root branching, tropisms and flowering (Woodward and Bartel, 2005). Application of the synthetic auxin 1-Naphthaleneacetic acid (NAA) to the growth media at a range of concentrations decreased primary root length and caused a large increase in the number of lateral roots. Each *Atdoa10* mutant tested responded like WT at all the NAA concentrations (Figure 45), showing that AtDOA10A and AtDOA10B do not affect auxin responses.
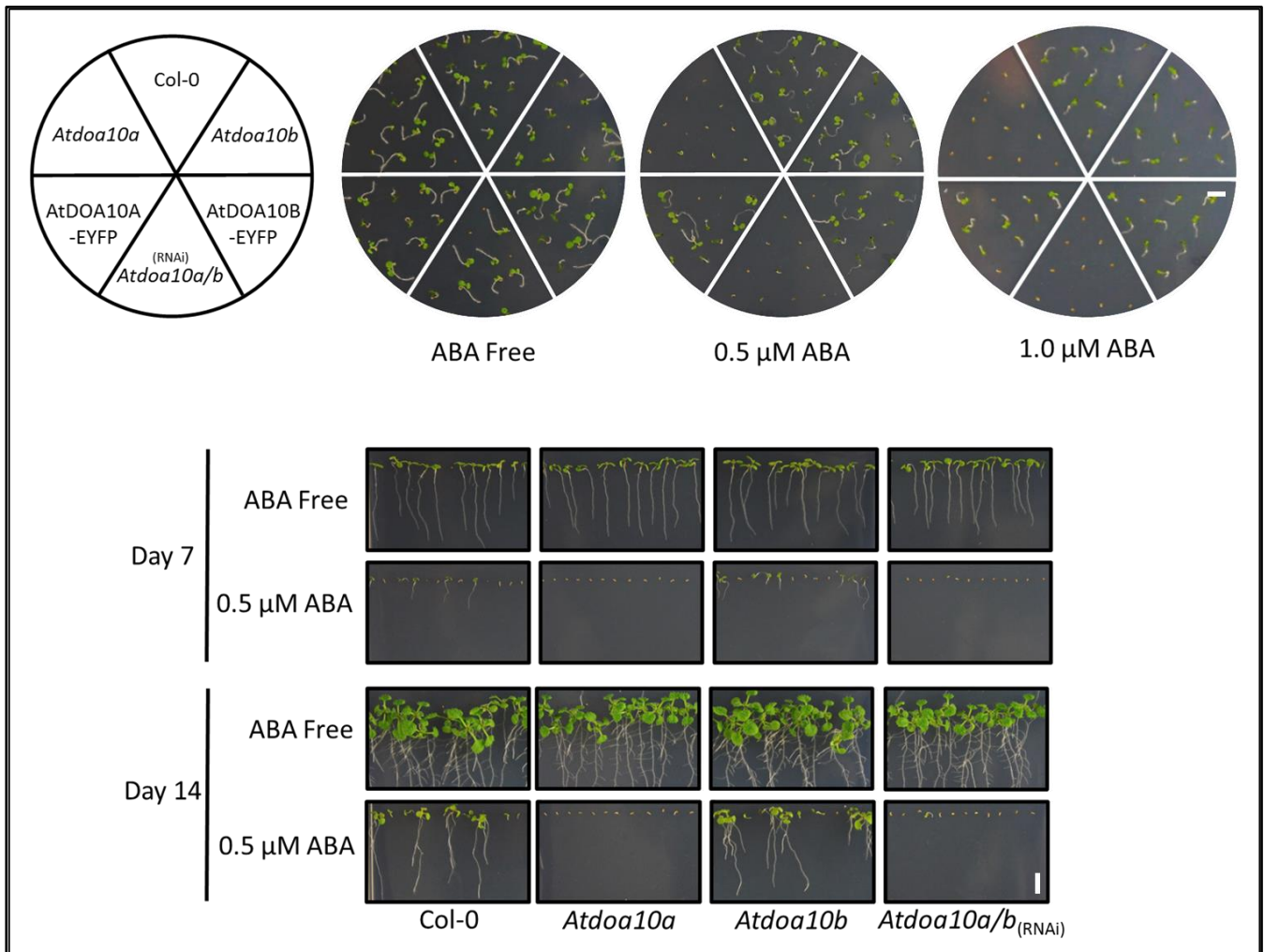
Figure 42 – Abscisic acid (ABA) treatment of *Arabidopsis doa10* mutants

ABA treatment inhibited post germination growth of *Atdoa10a* mutants when grown on horizontal (upper) and vertical (lower) ½ MS plates. *Atdoa10b* mutants exhibited a WT response. Complementation of *Atdoa10a* with AtDOA10A-EYFP was able to reduce ABA sensitivity to a WT level. Representative images of 7-day old seedlings grown on horizontal ½ MS plates with 0, 0.5 and 1 µM ABA and 7- and 14-day old seedlings grown on vertical 1/2MS plates ± 0.5 µM ABA. AtDOA10A-EYFP is line 2-2, AtDOA10B-EYFP is line 4-2, Atd*oa10a/b* is RNAi line 4-2(2). Scale bars = 5mm.
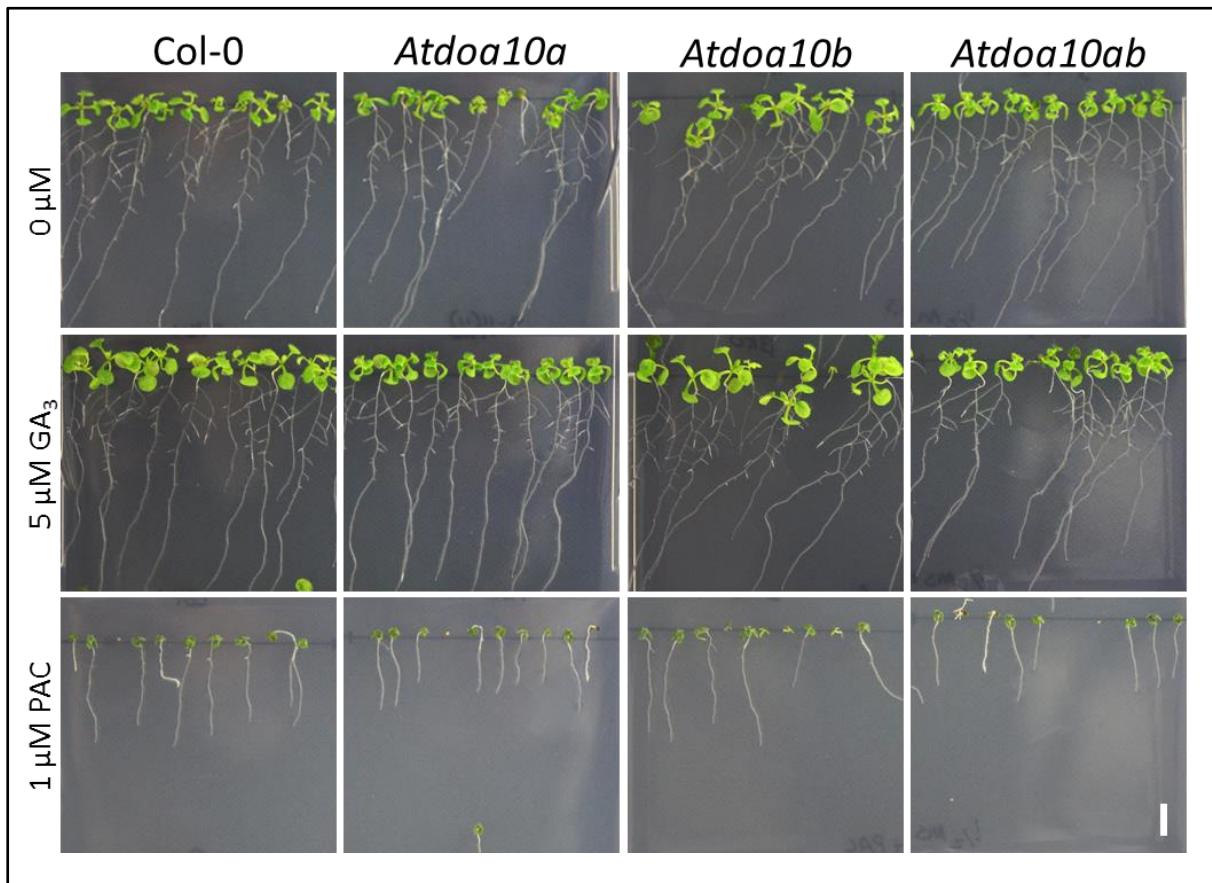
Figure 43 – Gibberellic acid (GA₃) and paclobutrazol (PAC) treatments of *Arabidopsis doa10* mutants

No strong variations between lines were observed following GA₃ and PAC treatments. Col-0 and *Atdoa10b* seedlings exhibited a slight increase in growth of aerial tissues in response to GA₃ which *Atdoa10a* and *Atdoa10ab* mutants did not. Root growth of *Atdoa10a* and *Atdoa10ab* was also reduced slightly more strongly in response to the PAC treatment compared to Col-0 and *Atdoa10b*. Representative images of 10-day old seedlings grown on ½ MS plates containing 5 µM GA₃ and 1 µM PAC. *Atdoa10ab* is CRISPR line 3(11)-11. Scale bar = 5mm.
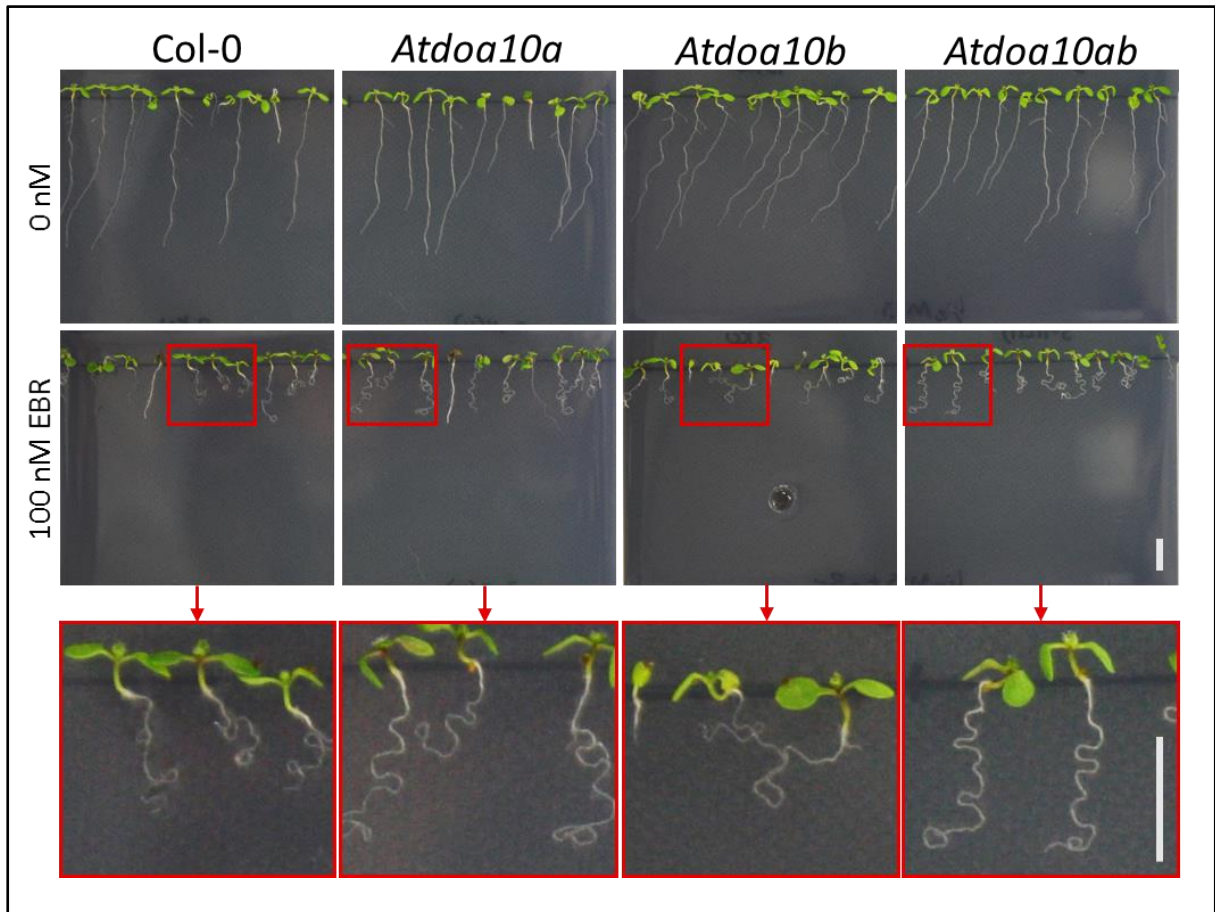
Figure 44 – 24-epibrassinolide (EBR) treatment of *Arabidopsis doa10* mutants
EBR affected the growth of all lines similarly. EBR-induced root waves were marginally more regular in *Atdoa10a* and *Atdoa10ab*. Red boxes show magnified areas (bottom row). Representative images of 7-day old seedlings grown on ½ MS plates containing 100 nM EBR. *doa10ab* is CRISPR line 3(11)-11. Scale bars = 5mm.
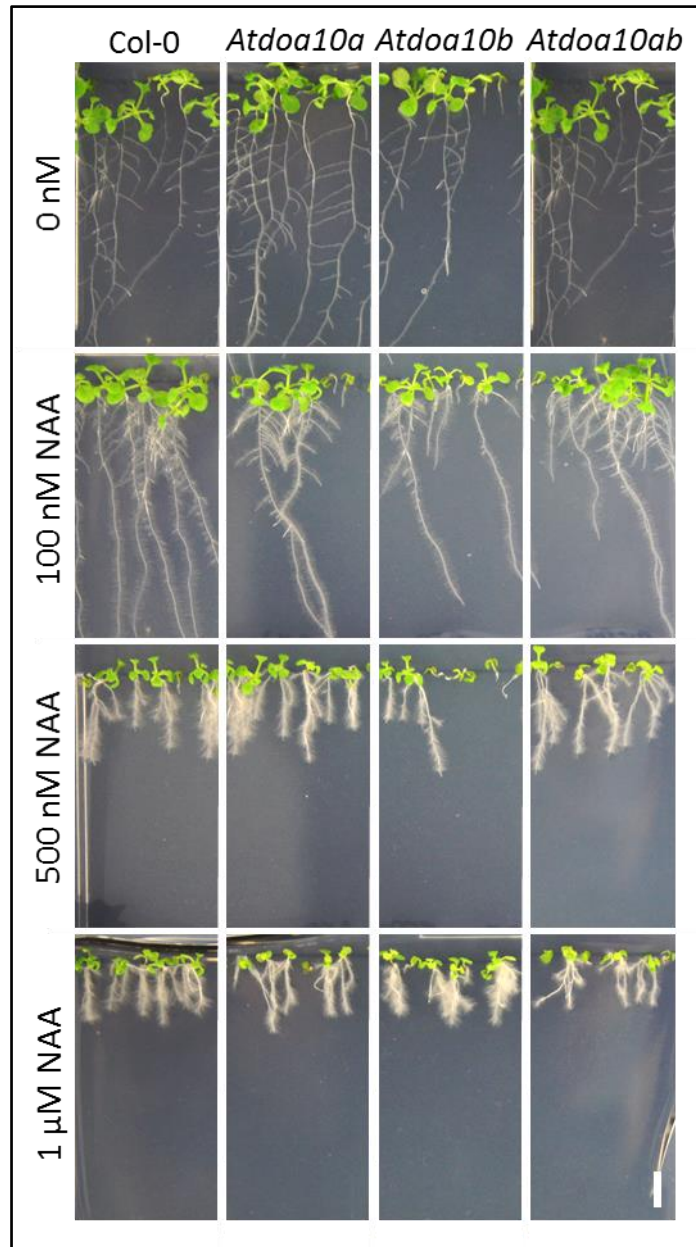
Figure 45 - 1-Naphthaleneacetic acid (NAA) treatment of *Arabidopsis doa10* mutants
No differences in the NAA response of the lines tested were observed. Representative images of 14-
day old seedlings grown on ½ MS plates containing 0 nM, 100 nM, 500 nM and 1 µM NAA.
Atdoa10ab is CRISPR line 3(11)-11. Scale bar = 5mm.

5.2.4 Sterol synthesis-associated phenotypes of *Atdoa10* mutants

The responses of *Atdoa10a, Atdoa10b* and *Atdoa10ab* mutants to brassinosteroids and the other phytohormones tested demonstrated that the role of AtDOA10s in regulating sterol biosynthesis does not significantly impact phytohormone signalling, even for sterol-derived phytohormones. To investigate this further, the mutants were screened for abnormal responses to the sterol biosynthesis inhibitors: terbinafine and lauryldimethylamine oxide (LDAO). Terbinafine is an inhibitor of squalene epoxidase (SQE) enzymes, which catalyse the oxidation of squalene to 2,3-oxidosqualene at the start of the biosynthesis pathway (De Vriese et al., 2021). Terbinafine molecules can bind to SQE enzymes and induce a conformational change, preventing the binding of squalene to the active site in a non-competitive manner (Ryder, 1992; Nowosieski et al., 2011). SQE is known to be ubiquitinated by DOA10 in yeast and humans and has been linked to AtDOA10A in plants (Doblas et al., 2013; Foresti et al., 2013; Zelcer et al., 2014). LDAO acts as a dual inhibitor of both the cycloartenol synthase (CAS) enzymes, which cyclise 2,3-oxidosqualene to produce cycloartenol, and cycloeucalenol cycloisomerase (CPI) enzymes, which convert cycloeucalenol into obtusifoliol (Darnet et al., 2020). The structure of LDAO is thought to mimic those of intermediates within these reactions, leading to competitive inhibition.

Seedlings were grown on ½ MS media with added terbinafine and LDAO. On vertical plates, the terbinafine treatment stunted the growth of *Arabidopsis* seedlings and had a strong impact on root morphology, although shoots appeared relatively unaffected (Figure 46). On horizontal plates, root growth was again affected, and seedlings were slower to produce cotyledons, but no differences in sensitivity were observed between Col-0 and any of the *Atdoa10* mutants (Figure 46). The observed responses were more uniform than previously reported (Pose et al., 2009; Laranjeira et al., 2015), with all seedlings appearing to respond similarly, rather than a proportion of seedlings being killed by the treatment.

The LDAO treatment induced a more substantial effect. At 10 µg ml^{-1} seedlings were very slow to develop and when roots were eventually formed, they were short and highly

branched; similar to the roots of terbinafine-treated seedlings. Seedlings did not survive a 20 µg ml^{-1} treatment (Figure 47). As with the terbinafine treatment, all lines appeared to respond similarly to the LDAO, suggesting that AtDOA10A and AtDOA10B may not exert a significant influence on the sterol biosynthesis pathway.
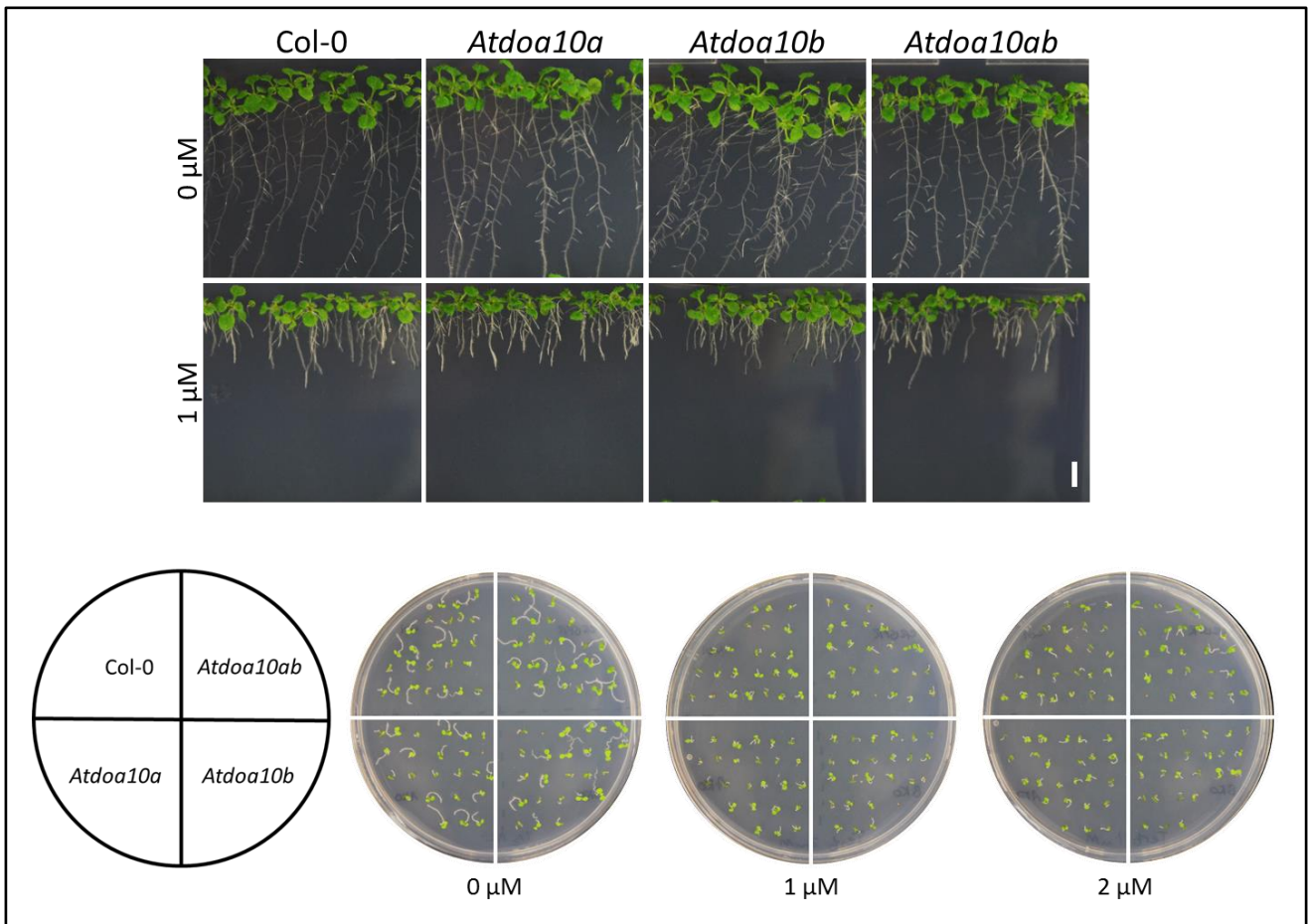
Figure 46 – Terbinafine treatment of *Arabidopsis doa10* mutants
No differences in the susceptibility of the lines tested were observed in response to terbinafine treatment. Representative images of 14-day old seedlings grown on vertical plates (upper) and 7-day old seedlings grown on horizontal plates (lower). *Atdoa10ab* is CRISPR line 3(11)-2. Scale bar = 5mm.
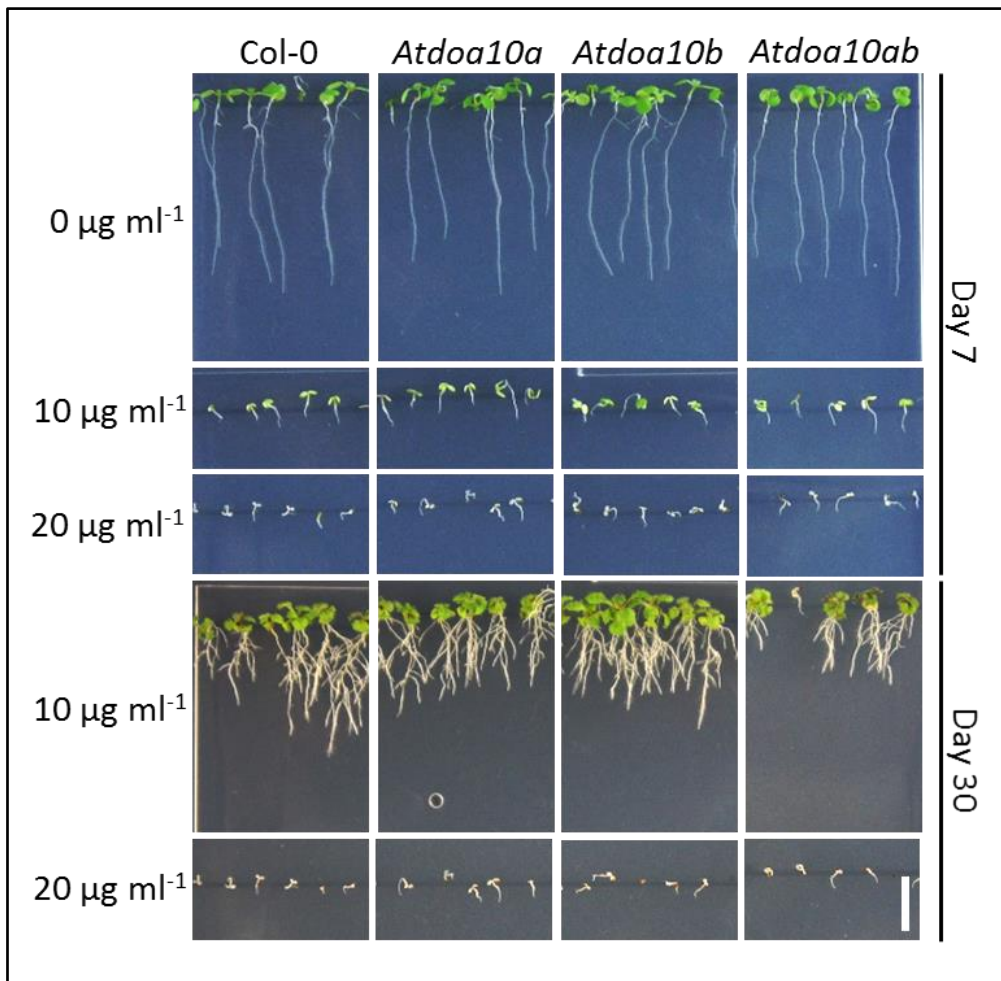
Figure 47 – LDAO treatment of *Arabidopsis doa10* mutants

No differences in the susceptibility of the lines tested were observed in response to LDAO treatment. Representative images of 7- and 30-day old seedlings grown on vertical ½ MS plates containing 10-20 μg ml^{-1}. *Atdoa10ab* is CRISPR line 3(11)-2. Scale bar = 5mm.

5.3 Multiomic analyses of an *Arabidopsis Atdoa10a/b* RNAi line

The phenotypic analyses outlined above provided little insight into the roles of the *Arabidopsis* DOA10s beyond potentially minor functions in ERAD and some hormone responses. Therefore, several omic analyses were conducted on *Atdoa10a/b* (RNAi line 4-2(2)) in order to identify additional potential functions. This included transcriptome analysis (RNA-seq), a global proteome analysis, and Nt-acetylation profiling.

5.3.1 Transcriptome analysis of *Atdoa10a/b* (RNAi line 4-2(2))

The transcriptome of *Atdoa10a/b* (RNAi line 4-2(2)) was characterised by RNA-sequencing (RNA-seq). RNA was extracted from Col-0 and *Atdoa10a/b* (RNAi line 4-2(2)) seedlings grown vertically on ½ MS media for 10 days. The sequencing was performed by Novogene (HK) Company Limited (Hong Kong) (see 2.7.7).

The transcript analysis identified 446 differentially expressed genes (DEGs) (216 up, 230 down) between the Col-0 and *Atdoa10a/b* (RNAi line 4-2(2)) samples (FDR adjusted p-value (q)<0.05) (Figure 48a). This included *AtDOA10A*, the transcript of which was shown to be 63% lower in *Atdoa10a/b* (RNAi line 4-2(2)) 4-2(2) (q=7.85×10^{-51}). This suggested that *AtDOA10A* expression was not completely eliminated in the line, but the presence of the T-DNA insertion still prevented the production of a fully functioning AtDOA10A protein, as evidenced by the observed ABA hypersensitivity of this line (Figure 42). *AtDOA10B* was also identified as a DEG. Surprisingly, *AtDOA10B* transcripts were reported to be more abundant in *Atdoa10a/b* (RNAi line 4-2(2)) (FC=1.81, q=1.14×10^{-9}). However, this was likely to be a result of expression of the RNAi construct containing a portion of the *AtDOA10B* sequence, since qRT-PCR analysis had confirmed a significant reduction in *AtDOA10B* mRNA (see 4.2.3.3). The most significant change in expression found was for gene AT3G18535 which encodes a tubulin-tyrosine ligase that was not expressed at all in Col-0.

Gene ontology (GO) term analysis was performed on the complete list of 446 DEGs using the AgriGO analysis tool (Tian et al., 2017). Although little correlation was observed between the DEGs and those identified by previous microarray analysis of *Atdoa10a* performed by Lu et al. (2012), the GO terms associated with both sets of DEGs were more similar. Several of the enriched GO terms were consistent with AtDOA10A and AtDOA10B being involved in ERAD, such as "Response to Stress", "Membrane" and "Protein Folding" (Figure 48b-e). Interestingly each DEG annotated with the "Protein Folding" GO term was upregulated, albeit to a relatively small extent (Figure 48e), potentially indicating an increase in the number of misfolded proteins in *Atdoa10a/b* (RNAi line 4-2(2)). This also correlated with the hypothesised roles of AtDOA10A and AtDOA10B as Ac/N-recognins that target misfolded proteins with exposed acetylated N-termini for degradation. Another strongly enriched GO term was "Monooxygenase". Six of the seven DEGs associated with the this GO term were downregulated in the *Atdoa10a/b* (RNAi line 4-2(2)) (Figure 48f). This included *AtSQE5* (FC=0.64, q=1.74×10^{-10}), an *Arabidopsis* homologue of the human and yeast DOA10 substrate SQE (Foresti et al., 2013; Zelcer et al., 2014). The broad downregulation of monooxygenase genes could indicate an accumulation of the corresponding proteins, potentially highlighting a class of proteins that are targeted by the AtDOA10s (see 6.1).

RNA-seq analysis was also conducted on the *Atnaa20* mutant lacking the catalytic subunit of NatB and on an AtNAA10 depleted line (named ami*AtNAA10*) developed using RNAi by the Markus Wirtz Laboratory, Universität Heidelberg, Germany (Linster et al., 2015). Two of the 446 DEGs identified in *Atdoa10a/b* (RNAi line 4-2(2)) lacked ATG numbers. Of the remaining 444 DEGs, 306 were also identified as differentially expressed in at least one of the NAT-depleted lines (Figure 49a). Such a significant overlap of DEGs would be expected if AtDOA10A and AtDOA10B were to act as Ac/N-recognins, since the loss of any enzyme within the Ac/N-degron pathway would result in stabilisation of the same set of substrates. GO term analysis was then performed on the lists of DEGs shared between *Atdoa10a/b* (RNAi line 4-2(2)) and each NAT mutant individually, in order to provide insight into potential classes of

substrate. The analysis yielded several enriched GO terms, and highlighted some differences between the lists of shared DEGs (Figure 49b). For example, the "Cell Wall" GO term was much more strongly enriched within the shared DEGs of *Atdoa10a/b* (RNAi line 4-2(2)) and *Atnaa20* than it was for the DEGs shared between *Atdoa10a/b* (RNAi line 4-2(2)) and *Atnaa10.* This provided a preliminary indication that many cell wall-associated proteins could be substrates of a NatB-mediated Ac/N-degron pathway. Similarly, proteins involved in the "Response to Water Deprivation" may be useful to investigate as potential substrates of a NatA-mediated Ac/N-degron pathway. The significant enrichment of drought response DEGs (Figure 49) was consistent with the observed ABA hypersensitivity and drought tolerance of *Atdoa10a* mutants (Lu et al., 2012; Zhao et al., 2014) and the ABA-mediated drought tolerance observed for ami*AtNAA10* (Linster 2015).
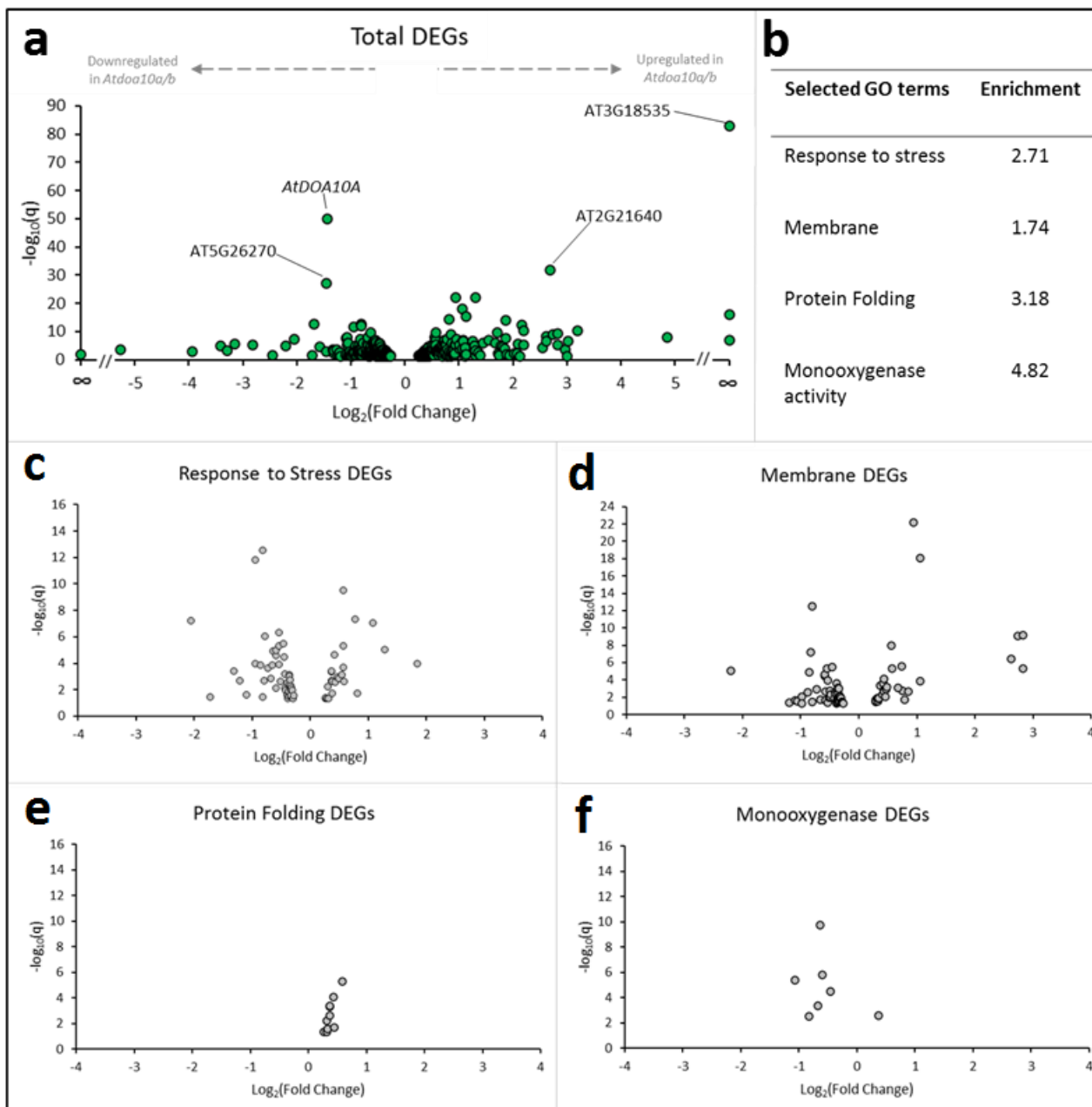
Figure 48 – Transcriptome analysis of *Atdoa10a/b* (RNAi line 4-2(2))
(a) Volcano plot of the 446 genes that were identified as differentially expressed in *Atdoa10a/b* (RNAi line 4-2(2)) compared to Col-0 (q<0.05) and (b) the level of enrichment of selected GO terms within the list of DEGs. Volcano plots of DEGs within the enriched GO terms (c) "Response to Stress", (d) "Membrane", (e) "Protein Folding" and (f) "Monooxygenase Activity" showing gene distribution. All "Protein Folding" genes were upregulated, whereas all but one "Monooxygenase" gene were downregulated.
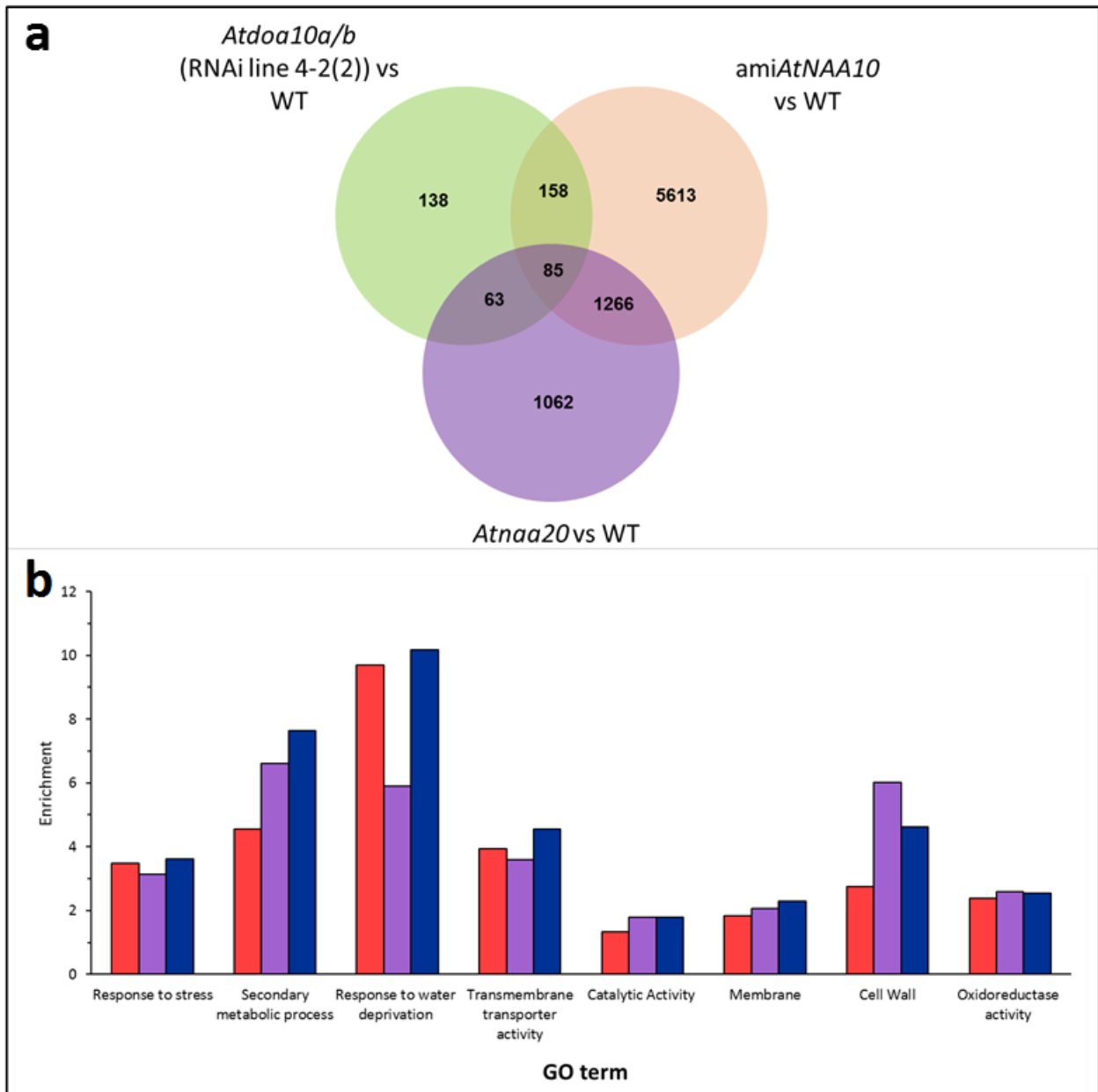
Figure 49 – Comparison of *Atdoa10a/b* (RNAi line 4-2(2)), *Atnaa10* and *Atnaa20* mutant transcriptomes
(a) Venn diagram showing overlap of DEGs relative to WT identified within the transcriptomes of *Atdoa10a/b* (RNAi line 4-2(2)), ami*AtNAA10* and *Atnaa20*. Over half of the genes differentially expressed in *Atdoa10a/b* (RNAi line 4-2(2)) overlap with DEGs in the NAT mutants. 243 differentially expressed genes are shared between *Atdoa10a/b* (RNAi line 4-2(2)) and ami*AtNAA10*, 148 differentially expressed genes are shared between *Atdoa10a/b* (RNAi line 4-2(2)) and *Atnaa20*, 85 genes are differentially expressed in all three lines. (b) Selected GO term enrichment levels of overlapping DEGs. Red = *Atdoa10a/b* (RNAi line 4-2(2)) and ami*AtNAA10*, Purple = *Atdoa10a/b* (RNAi line 4-2(2)) and *Atnaa20*, Blue = all three lines.

5.3.2 Proteome analysis of *Atdoa10a/b* (RNAi line 4-2(2))

To accompany the RNA-seq analysis, the proteome of *Atdoa10a/b* (RNAi line 4-2(2)) was also examined to look at differentially accumulated proteins. Samples were the same as those used for RNA-seq analysis (10-day old seedlings grown vertically on ½ MS), so that direct comparisons could be made between the datasets. Protein levels were quantified using mass spectrometry, carried out by Carmela Giglione and the protein maturation, cell fate and therapeutics lab at the Institute for Integrative Biology of the Cell (I2BC) (Gif-sur-Yvette cedex, France) (see 2.9.9).

From the proteome analysis 1594 proteins were detected in at least one replicate for both Col-0 and *Atdoa10a/b* (RNAi line 4-2(2)) and 1400 were present in multiple replicates of each genotype. Of these, 101 proteins with significant differences (q<0.05) in abundance between Col-0 and *Atdoa10a/b* (RNAi line 4-2(2)) plants were identified (46 up, 55 down) (Figure 50a). Within this set of differentially expressed proteins (DEPs), the enzyme LIPOXYGENASE2 (AtLOX2, AT3G45140) accumulated to higher levels than any other protein (FC=4.27, q=0.0028), and therefore was the most promising candidate substrate of the AtDOA10s. AtLOX2 is one of six *Arabidopsis* lipoxygenases that catalyse the dioxygenation of polyunsaturated fatty acids in processes such as the biosynthesis of jasmonic acid and green leaf volatiles (Bell et al., 1995; Mochizuki et al., 2016). The strong stabilisation of a dioxygenase was interesting due to the downregulation of several monooxygenase transcripts observed in the RNA-seq analysis (Figure 48f) and the enrichment of the "Oxidoreductase Activity" GO term within the DEGs shared between *Atdoa10a/b* (RNAi line 4-2(2)) and NAT mutants (Figure 49b). The "Oxidoreductase Activity" GO term was also enriched within the list of *Atdoa10a/b* (RNAi line 4-2(2)) DEPs, although no overall trend of increased or decreased abundance was observed (Figure 50f). As with the transcriptome, the "Response to Stress" and "Membrane" GO terms were enriched within the DEPs, again correlating with roles in ERAD (Figure 50b-d). The same GO terms were also enriched in proteomic analyses carried out by Zhao et al. (2014), although changes in the levels of specific proteins were variable.

Substrates of AtDOA10A and AtDOA10B would be expected to be more abundant in *Atdoa10a/b* (RNAi line 4-2(2)), potentially without an associated increase in gene expression. By comparing the 444 annotated DEGs with the full list of 1594 protein detected by mass spectrometry, 48 genes/proteins were found to be present in both datasets, allowing changes in protein and gene expression to be compared. Based on this, 23 proteins were classed as stabilised in *Atdoa10a/b* (RNAi line 4-2(2)) (increase in protein > increase in transcript), whilst 25 were classed as destabilised (increase in protein < increase in transcript) (Figure 51a). Seven proteins were identified as strongly stabilised, most significantly the catalase AtCAT3 (Figure 51b). The N-terminus of all seven of these proteins consisted of amino acid residues that permit Nt-acetylation (Figure 51c) (see 1.3, Aksnes et al., 2016); a result potentially compatible with AtDOA10A and AtDOA10B functioning as Ac/N-recognins. GO analysis of the 23 stabilised proteins was consistent with previous results, with "Response to Stress", "Response to Water Deprivation" and "Oxidoreductase Activity" again being enriched (Figure 51d).
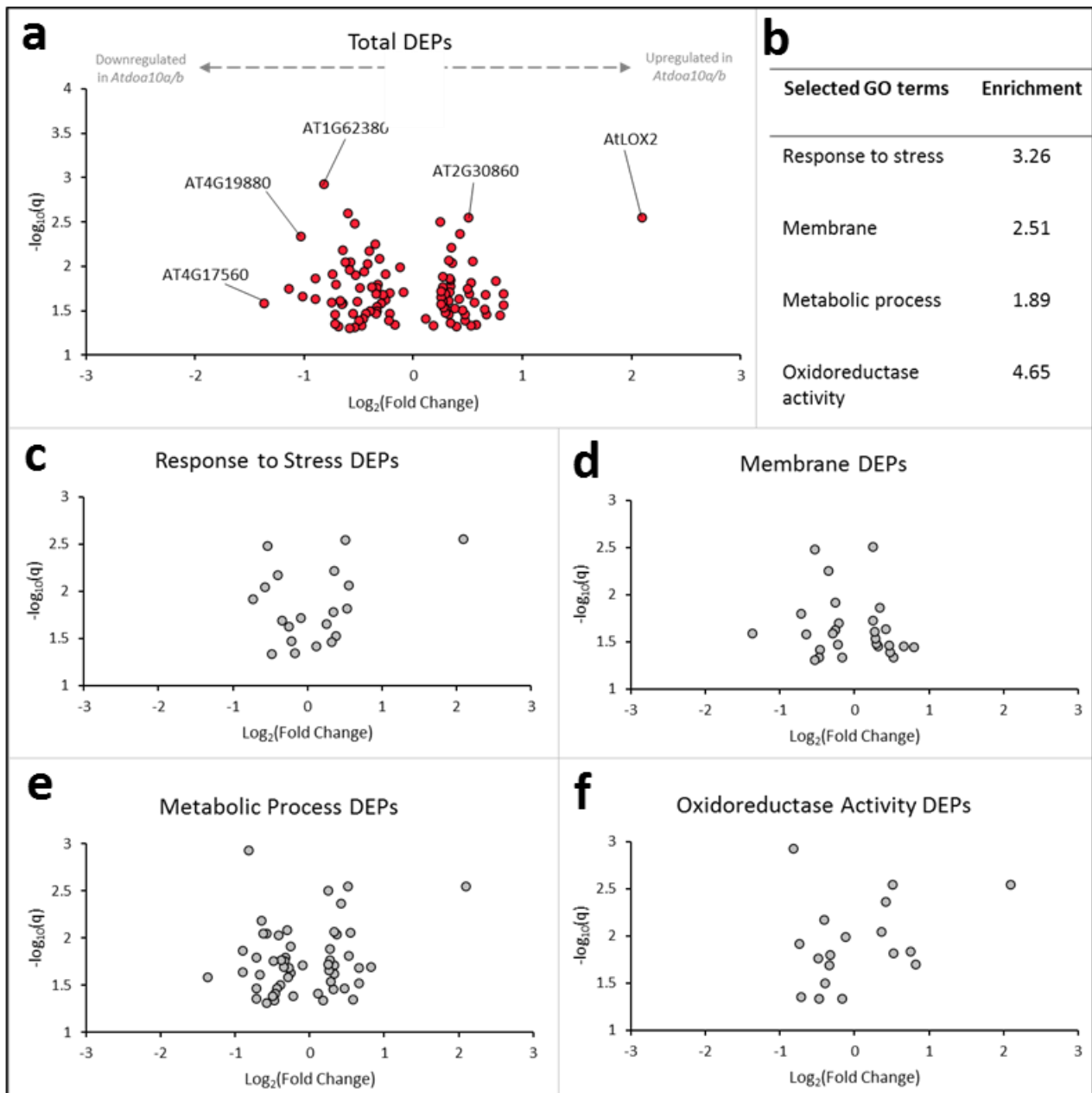
Figure 50 – Proteome analysis of *Atdoa10a/b* (RNAi line 4-2(2))
(a) Volcano plot of the 101 proteins with significantly different abundances in *Atdoa10a/b* (RNAi line 4-2(2)) compared to Col-0 (q<0.05) and (b) the level of enrichment of selected GO terms within the list of DEPs. Volcano plots of DEGs within the enriched GO terms (c) "Response to Stress", (d) "Membrane", (e) "Metabolic Process" and (f) "Oxidoreductase Activity" showing protein distribution.
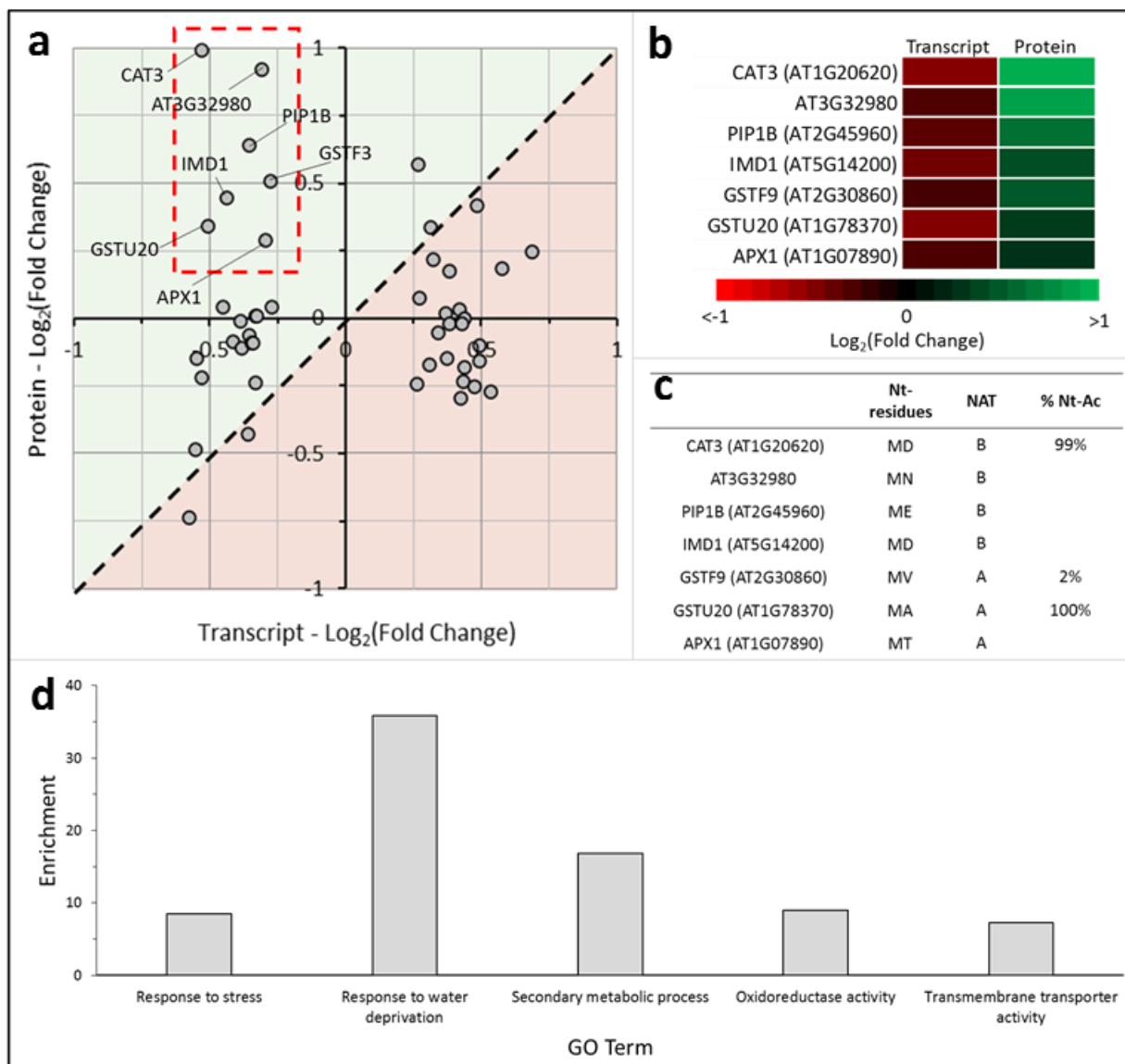
Figure 51 – Stability analysis of differentially expressed proteins in _Atdoa10a/b_ (RNAi line 4-2(2)) (a) Scatter plot of changes in transcription (x-axis) and protein abundance (y-axis) in _Atdoa10a/b_ (RNAi line 4-2(2)) vs Col-0 showing the 48 of genes/proteins that were both present in the proteome dataset and differentially expressed in the RNA-seq. (b) Heat map of seven strongly stabilised proteins (increased protein, decreased transcript) from red box in (a). (c) Details of the amino acids present at the N-terminus of the seven stabilised proteins and the corresponding NAT complex they are theoretically acetylated by. The percentage Nt-acetylation of the protein is also given, if known, based on data from Linster et al. (2015). (d) Selected enriched GO terms within the 23 stabilised proteins (increase in protein > increase in transcript; green shading in (a)).

5.3.3 Nt-acetylome analysis of *Atdoa10a/b* (RNAi line 4-2(2))

In addition to global analysis of the proteome, an Nt-acetylome was also complied for 10-day old Col-0 and *Atdoa10a/b* (RNAi line 4-2(2)) by mass spectrometry at the protein maturation, cell fate and therapeutics lab (I2BC) (see 2.9.9). The percentage acetylation of the N-termini of *Arabidopsis* proteins in Col-0 and in *Atdoa10a/b* (RNAi line 4-2(2)) was determined to specifically investigate potential roles for AtDOA10s as Ac/N-recognins. It was hypothesised that if AtDOA10s function as Ac/N-recognins, then their depletion in *Atdoa10a/b* (RNAi line 4-2(2)) may lead to an overaccumulation of the Nt-acetylated form of a protein, relative to Col-0 (due to a lack of subsequent ubiquitination and degradation).

To allow comparison the same N-termini needed to be detected in at least one of three replicates for both Col-0 and *Atdoa10a/b* (RNAi line 4-2(2)) samples. This was the case for 191 different N-termini, but a significant difference ($p < 0.05$) in the level of Nt-acetylation between the two genotypes was only detected for two N-termini (Table 18). These N-termini were 99.93 and 99.75% Nt-acetylated in Col-0 and, despite the statistical significance, this only dropped to 99.81 and 98.88% respectively in *Atdoa10a/b* (RNAi line 4-2(2)). Therefore these proteins are unlikely to represent potential Ac/N-degron pathway substrates. A further six N-termini exhibited a change in Nt-acetylation at the $p < 0.1$ significance level but differences in Nt-acetylation percentage were again relatively small (Table 18). Overall, when considering the average percentage Nt-acetylation of all 191 N-termini, the level of Nt-acetylation actually decreased in *Atdoa10a/b* (RNAi line 4-2(2)) in comparison to Col-0 (Table 19). The lack of an accumulation of Nt-acetylated substrates suggested that either the Ac/N-degron pathway does not have a major role in global protein turnover or that the AtDOA10s are not components of the pathway. The lack of impact of the AtDOA10s depletion on Nt-acetylation levels was also confirmed for specific groups of N-termini bearing the same Nt-residues. No increase in percentage Nt-acetylation greater than 0.42% was observed for any particular Nt-residue (Table 19).

Further evidence of the AtDOA10-Ac/N-degron pathway's insignificance in global protein degradation was provided by the absence of a correlation between changes in protein abundance and changes in Nt-acetylation in *Atdoa10a/b* (RNAi line 4-2(2)). Such data was available for 131 proteins that were present within the proteome and Nt-acetylome datasets but only 26 showed an increase in both protein abundance and Nt-acetylation in *Atdoa10a/b* (RNAi line 4-2(2)) relative to Col-0 (Figure 52a). The strongest correlation was observed for the proteins Serine Hydroxymethyltransferase 1 (AtSHM1), Hypersensitive To High Light 1 (AtHHL1) and Photosystem II light harvesting complex protein B1B2 (AtLHB1B2) (Figure 52), but in each case the variation between *Atdoa10a/b* (RNAi line 4-2(2)) and Col-0 was insignificant in all respects (p/q<0.05).

When considering only the 101 proteins for which a significant difference in quantity was observed (see 5.3.2), information about the Nt-acetylation status of peptides corresponding to eight of these proteins was obtained. Changes in the Nt-acetylation and protein abundance of each protein were assessed to check for the influence of the AtDOA10-Ac/N-degron pathway at an individual protein level (Figure 52b). Only one of the eight proteins, PHOTOSYSTEM II SUBUNIT Q-1 (AtPSBQ-1, AT4G21280), showed a slight increase in percentage Nt-acetylation and protein abundance, that would be expected of Ac/N-degron pathway substrates.

Table 18 – Status of the eight proteins with significant changes (p<0.1) to their Nt-acetylation level detected within Nt-acetylome

ATG Number	ID (if known)	Col-0			Atdoa10a/b (RNAi line 4-2(2))			FC (d10/Col)	P-value
		Reps	% NTA	SD	Reps	% NTA	SD		
AT1G76180	ERD14	3	99.93	0.03	3	99.81	0.05	0.999	0.0217
AT4615630	CASPL1E1	2	99.75	0.05	2	98.88	0.16	0.991	0.0482
AT2G37660		3	4.11	1.07	3	2.28	0.52	0.555	0.0610
AT4G37930	SHM1	3	0.21	0.07	2	0.66	0.12	3.176	0.0611
AT4G01310	PRPL5	3	1.65	0.52	3	3.08	0.90	1.869	0.0703
ATCG00560	PSBL	2	26.46	0.38	2	22.18	1.18	0.838	0.0723
AT1G54780	TLP18.3	2	1.25	0.22	2	0.51	0.05	0.408	0.0841
AT1G35720	ANNAT1	3	99.85	0.07	3	99.95	0.02	1.001	0.0916
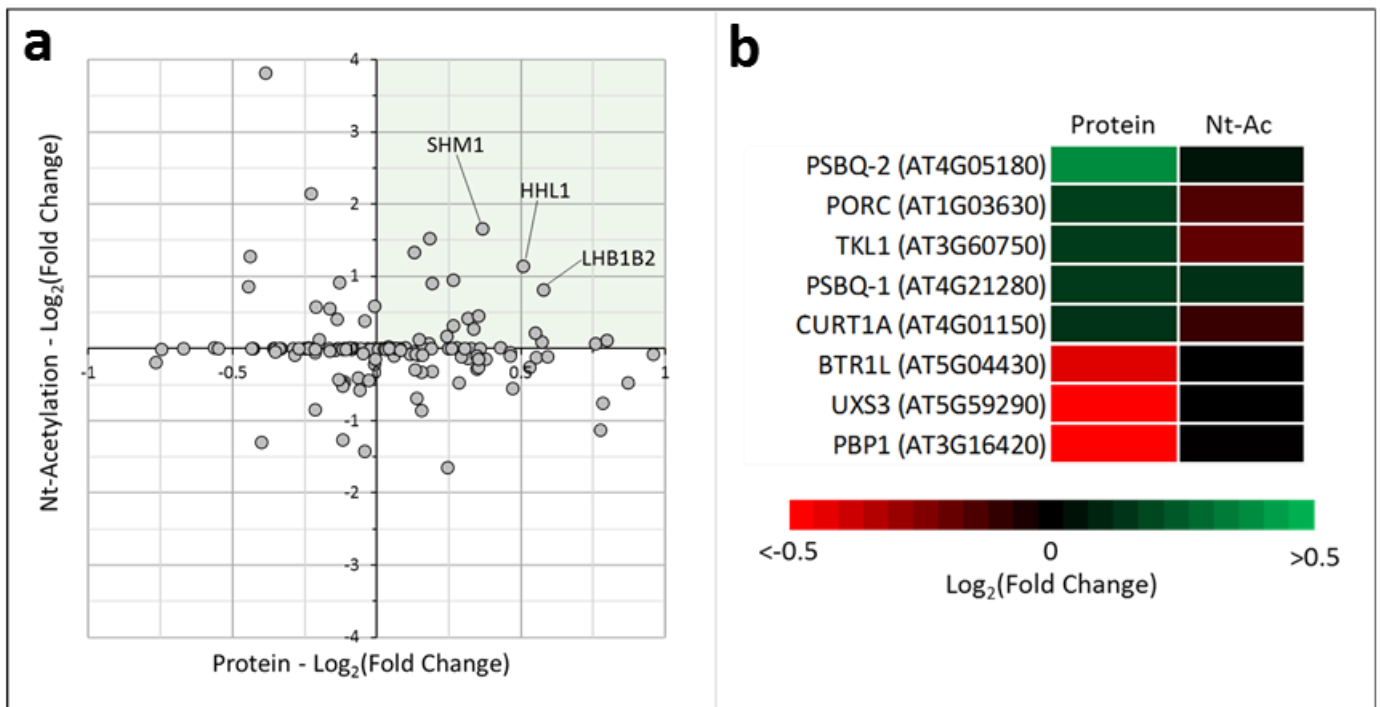
Figure 52 – Comparison of Nt-acetylation changes and protein accumulation in *Atdoa10a*/b (RNAi line 4-2(2))

(a) No correlation was detected between the changes in Nt-acetylation level (y-axis) and changes in abundance of individual proteins (x-axis) in *Atdoa10a/b* (RNAi line 4-2(2)) relative to Col-0. Only 26 of the 131 proteins for which data was available, exhibited an increase in both quantity and Nt-acetylation level (green shading) in *Atdoa10a/b* (RNAi line 4-2(2)) but none of these proteins were both significantly differentially acetylated and accumulated. (b) Heat maps of the eight significant DEPs for which Nt-acetylation data was obtained. Only AtPSBQ-1 shows a slight increase in both protein abundance and Nt-acetylation consistent with what would be expected for a substrate of the AtDOA10-Ac/N-degron pathway.

Table 19 – Average Nt-acetylation percentages in Col-0 and in *Atdoa10a/b* (RNAi line 4-2(2)) for all 191 peptides detected and groups of peptides with the same Nt-amino acid

	Average % Nt-acetylation	
	Col	*Atdoa10a/b*
ALL (191 peptides)	46.28	45.87
Nt-Ala (94 peptides)	54.52	54.21
Nt-Asp (2 peptides)	1.29	1.71
Nt-Glu (6 peptides)	0.99	1.07
Nt-Gly (7 peptides)	34.27	33.66
Nt-Lys (6 peptides)	4.95	3.85
Nt-Leu (1-peptide)	0.53	0.39
Nt-Met (20 peptides)	85.26	85.30
Nt-Pro (5 peptides)	2.36	1.69
Nt-Ser (28 peptides)	47.38	46.58
Nt-Thr (4 peptides)	39.13	35.42
Nt-Val (18 peptides)	13.09	13.13

5.4 Discussion

Analyses described in Chapter III confirmed AtDOA10A and AtDOA10B as homologues of ScDoa10 due to their similarities in amino acid sequence, transmembrane structure and subcellular localisation. Their ubiquitous gene expression patterns, responsiveness to TM and AtDOA10A's ability to functionally complement ScDoa10 also suggested that their functions within *Arabidopsis* were likely to be similar to their yeast homologue, namely as components of the ERAD system and potentially as Ac/N-recognins. However, the lack of impact on the magnitude of the TM-induced UPR suggested that AtDOA10A and AtDOA10B did not play a central role in the degradation of misfolded proteins at the ER.

Phenotypic analysis of various *Arabidopsis doa10* mutants was consistent with this conclusion, as *Atdoa10* mutants grew and behaved similarly to WT under standard conditions and when subjected to most ERAD stresses (Figures 37-41). The inclusion of *Atdoa10ab* CRISPR lines and *Atdoa10a*/b RNAi lines in these experiments confirmed that the lack of impact of the single mutations was not the result of functional redundancy between the E3s. A small increase in the resistance to TM of lines carrying the *Atdoa10a* mutation was observed (Figure 40), indicating some influence on ERAD. However, the subtlety of this phenotype provided further evidence that the *Arabidopsis* DOA10s appear to have only a peripheral role within the ERAD system. The loss of ERAD components might be expected to cause increased susceptibility to protein misfolding stress rather than the resistance observed here. This paradoxical phenotype may be the result of constitutive upregulation of alternative ERAD components, for example, 10 genes annotated with the "Protein Folding" GO term were identified as significantly upregulated within the transcriptome of *Atdoa10a/b* (RNAi line 4-2(2)) (Figure 48). Alternatively, the resistance may be the result of misfolded, but still functional, substrates of AtDOA10A escaping degradation. Such a phenomenon has been observed for the brassinosteroid receptor AtBRI1, whereby the loss of the E2 enzyme AtUBC32 allows functional AtBRI1 proteins bearing a point mutation to escape degradation

and contribute to salt tolerance (Cui et al., 2012a). The absence of even subtle AtDOA10B-associated phenotypes aligns with observations that AtDOA10A is much more strongly expressed at both the gene and protein level (see 3.5 and 3.6) and is therefore likely to have a larger influence on cellular processes. However, the WT-like response of *Atdoa10b* to TM was particularly surprising due to the observed 5-fold upregulation of *AtDOA10B* in response to the same chemical (Figure 18). The significance of this upregulation therefore remains unknown.

In addition to a role in ERAD, AtDOA10A is also thought to act as a negative regulator of ABA biosynthesis, with *Atdoa10a* mutants displaying hypersensitivity to the phytohormone (Zhao et al., 2014). Given this, the responses of the *Atdoa10* mutants to ABA and a range of other hormone-related chemicals were investigated. As expected, *Atdoa10a* mutants failed to develop beyond radicle emergence when grown on 0.5 µM ABA (Figure 42). Conversely, the ABA response of *Atdoa10b* mutants was unaffected, demonstrating that the regulation of ABA biosynthesis is a role specific to AtDOA10A. This provided further evidence of the diversification of function between AtDOA10A and AtDOA10B, in addition to differences in TM response (Figures 18 and 40) and their ability to complement ScDOA10 (Figure 21).

Components of the sterol synthesis pathway are known to be ubiquitinated by HsMARCH6 in humans and ScDOA10 in yeast (Foresti et al., 2013; Zelcer et al., 2014; Scott et al., 2020) and, as such, AtDOA10A and AtDOA10B are anticipated to also regulate the *Arabidopsis* sterol synthesis pathway. AtDOA10A has been identified as a positive regulator of AtHMGR in *Arabidopsis*, with mutations of AtDOA10A able to rescue the drought hypersensitivity and stunted root defect of *Atsqe1* sterol synthesis mutants by reducing *AtHMGR* expression (Doblas et al., 2013). AtHMGR is the rate-limiting enzyme of the MVA pathway that precedes sterol biosynthesis, and the stability of its human homologue is also known to be influenced by HsMARCH6 (Zelcer et al., 2014). Due to the existence of crosstalk between the MVA pathway and the MEP pathway, from which ABA is synthesised, it was plausible that regulation of AtHMGR and AtSQE1 by AtDOA10A may have explained the ABA

hypersensitivity phenotype of *Atdoa10a* mutants. However, the lack of any strong phenotypes in response to the other hormones tested, including brassinosteroids (EBR) which are produced directly from phytosterols (Vriet et al., 2013), suggests that this is not the case (Figures 42-45). This raises the possibility that AtDOA10A may target an as yet undiscovered protein that is specific to ABA biosynthesis. Despite their strong links to the MVA and sterol biosynthesis pathways, the unaltered responses of the AtDOA10 mutants to most hormones implied that they were not of critical importance to the functioning of these pathways. This was reinforced by the WT-like responses of *Atdoa10a*, *Atdoa10b* and *Atdoa10ab* to the sterol synthesis inhibitors terbinafine and LDAO (Figures 46 and 47).

In addition to phenotypic analysis, the impacts of *Atdoa10* mutations on the *Arabidopsis* transcriptome, proteome and Nt-acetylome were investigated by studying *Atdoa10a/b* (RNAi line 4-2(2)). Despite the absence of any strong protein misfolding stress phenotypes, GO analysis of DEGs and DEPs did highlight enrichment of ERAD-associated GO terms supporting the suggestion that the AtDOA10s have a peripheral involvement in ERAD. Adding further support to this, the enriched GO terms identified were also similar to those enriched in previous analyses of the transcriptome and proteome of the *Atdoa10a* single mutant (Lu et al., 2012; Zhao et al., 2014). However, there was not always a strong correlation between genes and proteins identified in the datasets generated here, and those previously published. For example, here the transcript of *AtPBS3* was downregulated in *Atdoa10a/b* (RNAi line 4-2(2)) (FC=0.243), whereas Lu et al. (2012) reported an upregulation in *Atdoa10a* compared to WT (FC=5.347). Similarly, only eight DEPs matched those identified from proteomics carried out by Zhao et al. (2014) and five of these showed opposing abundance changes. Such inconsistencies may be the result of the additional knockdown of *AtDOA10B* in the *Atdoa10a* background, or, more likely are caused by variation in the growth stage of the plant material analysed. Lu et al. (2012) studied 5-week-old plants and Zhao et al. (2014) carried out their analysis on seeds.

As well as ERAD, enrichment of GO terms like "Protein Folding" was also consistent with potential roles as Ac/N-recognins that target misfolded proteins due to the exposure of their acetylated N-terminus. The observation that the seven most strongly stabilised proteins possess Nt-amino acids that permit acetylation also supports this (Figure 51). However, analysis of global levels of Nt-acetylated proteins provided no evidence of this, with the average % Nt-acetylation unchanged between Col-0 and *Atdoa10a/b* (RNAi line 4-2(2)) (Table 19). This strongly implies that Nt-acetylation does not act as a universal degradation signal for targeting by AtDOA10s, although only a small number (191) of N-termini were detected, and a distinct set of Nt-acetylated substrates remains possible. Additionally, the fact that 68.9% of DEGs identified in *Atdoa10a/b* (RNAi line 4-2(2)) overlapped with those detected in either ami*AtNAA10* or *Atnaa20* (Figure 46) suggests that the AtDOA10s are, in some way, linked to Nt-acetylation.

The combined omics datasets did not rule out that AtDOA10A and AtDOA10B may target proteins with acetylated N-termini as part of the ERAD system but if this is the case, then the number of these proteins is likely to be relatively few. The datasets also provided useful information for the identification of these potential substrates. Such candidates included the AtSQE family of enzymes, since they are homologues of yeast and human DOA10 substrates and *AtSQE5* was identified as one of six significantly downregulated monooxygenases (See Chapter 6); AtLOX2, which showed the largest significant increase in protein level in *Atdoa10a/b* (RNAi line 4-2(2)); and AtCAT3, which was identified as having a strong increase in protein abundance despite a decrease in expression in the mutant line. Interestingly, all these candidates are enzymes annotated with the "Oxidoreductase Activity" GO term, potentially highlighting a specific class of proteins that are AtDOA10 substrates. The identification of an *Arabidopsis* DOA10 substrate and establishment of its mechanism of recognition would provide the strongest evidence for or against the existence of an *Arabidopsis* Ac/N-degron pathway.

Chapter VI -

AtSQE1 is a likely substrate of AtDOA10A

6.1 Introduction

Since the identification of protein substrates was required to confirm whether AtDOA10A and AtDOA10B could target proteins based on their Nt-acetylation status, efforts were made to identify a ubiquitination target. Transcriptome, proteome and Nt-acetylome analyses in Chapter V highlighted several candidate substrates worthy of investigation. Of particular interest were the *Arabidopsis* SQE family of enzymes due to their human and yeast homologues being known targets of native DOA10s (Foresti et al., 2013; Zelcer et al., 2014). Significantly, the degradation of SQE in humans has been shown to be dependent on the presence of the N-terminal 100 amino acids (Gill et al., 2011), potentially implicating the N-degron pathways. The transcript of *AtSQE5* was identified as significantly downregulated in *Atdoa10a/b* (RNAi line 4-2(2)) during RNA-seq analysis (see 5.4.1), perhaps suggesting stabilisation at the protein level. AtSQE enzymes are also all annotated with the "Monooxygenase" and "Oxidoreductase" GO terms, which were consistently enriched in the transcriptome and proteome datasets (see 5.4).

SQE enzymes catalyse the rate-limiting first oxidation step of the sterol synthesis pathway, in eukaryotes, whereby squalene is monooxygenated to produce 2,3-oxidosqualene (Phillips et al., 2006) (Figure 53). In plants, products of this pathway include the phytosterols β-sitosterol, stigmasterol and campesterol which contribute to the structure and fluidity of cell membranes (De Vriese et al., 2021). Campesterol is also a precursor of brassinolide (see 5.3.4, Figure 53) (Fujioka and Yokota, 2003; Clouse, 2011). The synthesis of triterpenes, from which the defence compounds saponins are derived, is also dependent on the monooxygenation of squalene by AtSQEs (Moses et al., 2014; Thimmappa et al., 2014). In addition to the products, intermediates produced during sterol biosynthesis can act as key signalling molecules (Vriet et al., 2013) demonstrating the importance of the sterol biosynthesis pathway in numerous aspects of plant development. The rate of flux through the pathway is also regulated by a second conserved enzyme: HMGR. HMGR catalyses the reduction of 3-hydroxy-3-methylglutaryl-CoA to produce mevalonate, which occurs upstream

of SQE, within the MVA pathway (De Vriese et al., 2021) (Figure 53). Like SQE, the stability and activity of HMGR has been shown to be influenced by DOA10 (Doblas et al., 2013; Zelcer et al., 2014). In humans, HsMARCH6 has also been shown to target the downstream sterol synthesis enzyme lanosterol 14α-demethylase, making HsMARCH6 the first the E3 ligase to be known to target multiple enzymes within a single biosynthetic pathway (Scott et al., 2020)

The genomes of dicot plants, like *Arabidopsis*, often contain many *SQE* genes. This is unlike the genomes of primitive land plants such as *Physcomitrella patens* which encode only one *SQE* gene, as is the case in yeast and humans. Monocots possess one or two *SQE* genes (Laranjeira et al., 2015). The *Arabidopsis* genome encodes six *SQE*s (*AtSQE1-6*). Three of these six *SQE* genes (*AtSQE1-3*) have been shown to be able to functionally complement the sole yeast SQE enzyme ScERG1, to permit growth in the absence of exogenous sterols (Rasbery et al., 2007). *AtSQE4-6* failed to complement the yeast homologue and, as such, *AtSQE4-6* are sometimes referred to as *SQE-like* genes. These *SQE-like* genes are specific to the *Brassicaceae* family and are found in tandem on chromosome 5, suggesting that they result from gene duplication events (Laranjeira et al., 2015). The expression patterns of *AtSQE1-6* show some transcription of each gene in the majority of tissues but significant variation in the level of expression between genes and tissues (Rasbery et al., 2007; Pose et al., 2009; Laranjeira et al., 2015). For example, *AtSQE5* is much more strongly expressed than *AtSQE4* in most tissues, but *AtSQE4* expression exceeds that of *AtSQE5* in roots (Rasbery et al., 2007).

Although RNA-seq analysis identified significant downregulation of *AtSQE5* in *Atdoa10a/b* (RNAi line 4-2(2)) (see 5.4.1), the degradation of AtSQE1 was chosen for further investigation because mutations of *Atdoa10a* are known to supress the drought sensitive phenotype and severe root defects caused by a point mutation in *Atsqe1* (Doblas et al., 2013). The extreme phenotypes caused by an *AtSQE1* point mutation also indicate a major non-redundant function of AtSQE1; such clear phenotypes have not been reported for mutations of other AtSQEs (Pose et al., 2009; Laranjeira et al., 2015). Suppression of these phenotypes

through *Atdoa10a* mutation theoretically correlated with AtDOA10A targeting AtSQE1 for degradation, by allowing an accumulation of AtSQE1 to compensate for a potential reduction in activity caused by the point mutation. However, suppression of the phenotype was instead assigned to a decrease in AtHMGR activity in the *Atdoa10a* mutant. This reduction in AtHMGR activity was presumed to reduce the accumulation of a toxic mobile intermediate that amassed in *Atsqe1* mutants, since *Atdoa10a* mutations appeared to have no effect on the overall sterol composition of the mutant (Doblas et al., 2013).

Given that yeast and mammalian SQEs are proteolytic targets of DOA10, and the report that *Atdoa10a* mutations suppress *Atsqe1* mutant phenotypes, work was undertaken to determine whether AtSQE1 was targeted for ubiquitination by AtDOA10A and AtDOA10B. Confirmation of AtSQE1 as an AtDOA10 substrate would shed light on the Ac/N-degron pathway, with AtSQE1 having the potential to be the first substrate of the pathway to be identified in plants.

Figure 53 – The sterol synthesis pathway in *Arabidopsis*
Schematic diagram of simplified *Arabidopsis* sterol biosynthetic pathway showing the roles of the rate limiting HMGR and SQE enzymes. HMGR is the rate limiting step of the cytosolic mevalonate (MVA) pathway which produces IPP and DMAPP. These products can also be generated via the chloroplastic methylerythritol phosphate (MEP) pathway. SQE enzymes catalyse the oxidation of squalene to 2,3-oxidosqualene from which all plant sterols are derived. Campesterol is a brassinolide precursor. "P" within molecule structures represents a phosphate group. Double arrows represent multiple enzymatic steps. Abbreviations: CoA = Coenzyme A; HMG = 3-Hydroxy-3-methylglutaryl; HMGR = HMG-CoA reductase; IPP = isopentenyl pyrophosphate; DMAPP = dimethylallyl pyrophosphate; FPP = farnesyl pyrophosphate; SQE = squalene epoxidase. Based on De Vriese et al (2021).
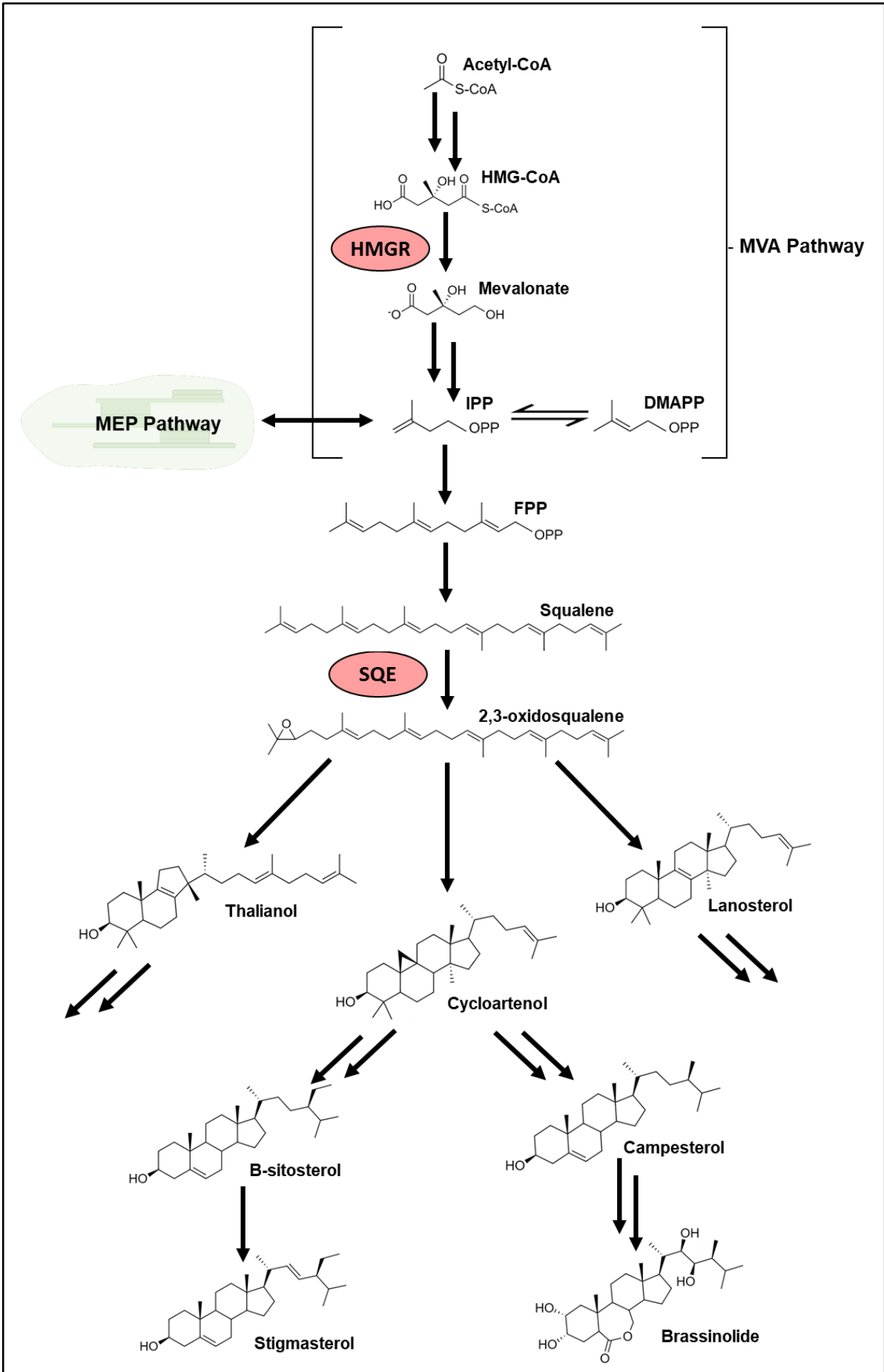
6.2 AtSQE1 stability analysis in yeast

First, analysis of AtSQE1's stability was carried out through expression of the *Arabidopsis* protein in *S. cerevisiae*; whilst this is a heterologous system, it allowed the rapid assessment of protein degradation and avoided the difficulties associated with the cell-free expression of complex membrane proteins.

6.2.1 Genotyping of yeast mutants

As well as WT yeast, AtSQE1 was also transformed into *ScDoa10Δ, ScNaa10Δ* and *ScNaa20Δ* yeast mutants, so that the role of the yeast Ac/N-degron pathway in AtSQE1 degradation could be studied. These mutant strains were obtained from the Dharmacon yeast KO collection (Horizon Discovery) (see 2.2.1.1). The *ScDoa10Δ* mutant was previously genotyped by PCR during analysis of its hygromycin resistance phenotype (see 3.9). PCR was repeated here, to genotype the *ScNaa10Δ* and *ScNaa20Δ* mutants. Primers 211-218 were used to show the absence of the genes within genomic DNA extracted from the mutants and primers 45+46 used to confirm the presence of the gene cassette, including *KanMX4*, in their place. PCR fully confirmed the genotype of *ScNaa20Δ* (Figure 54), however a binary PCR product including the 5' end of the *KanMX4* gene could not be amplified from *ScNaa10Δ* (Figure 55). Whilst this may indicate that the complete gene cassette was not present at the *ScNaa10* locus, the absence of gene-specific products and successful amplification from the 3' end of the *KanMX4* gene confirmed that *ScNaa10* gene was disrupted.
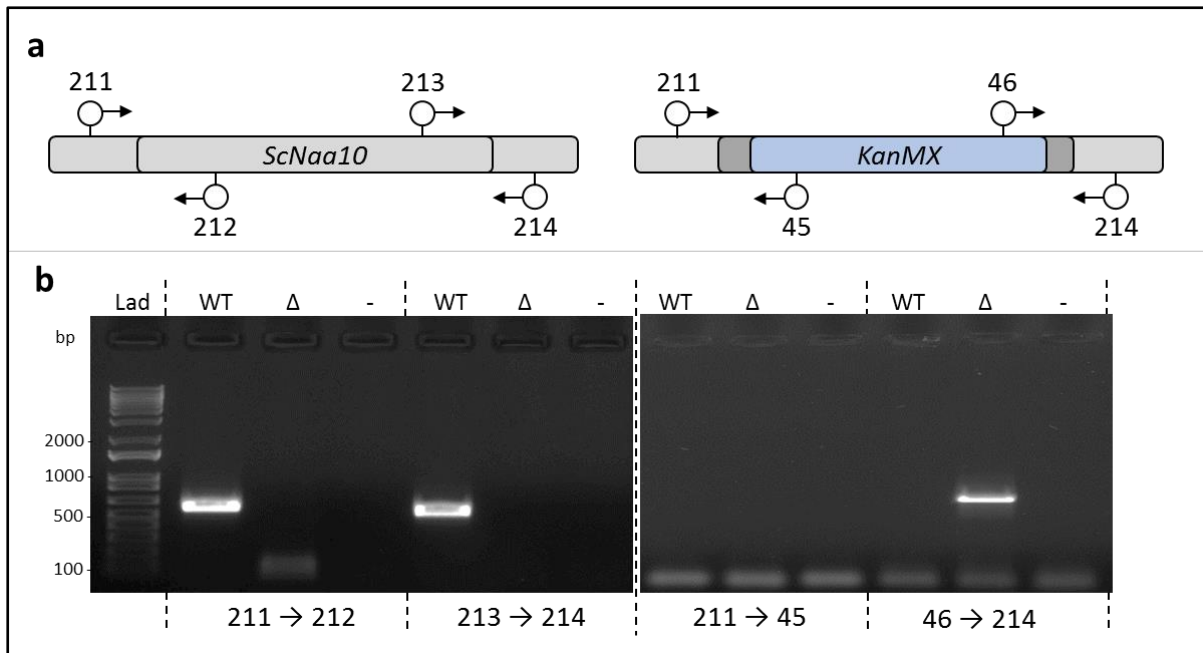
Figure 54 – PCR genotyping of *ScNaa10Δ*
(a) Schematic diagram of *ScNaa10* locus in WT (left) and *ScNaa10Δ* (right) yeast including genotyping primer positions. *ScNaa10* is replaced by a gene cassette including the *KanMX4* gene in *ScNaa10Δ.* (b) Gel electrophoresis of PCR products resulting from amplification of DNA from WT and *ScNaa10Δ* (Δ) with specified primers. No template control PCR products also included (-).
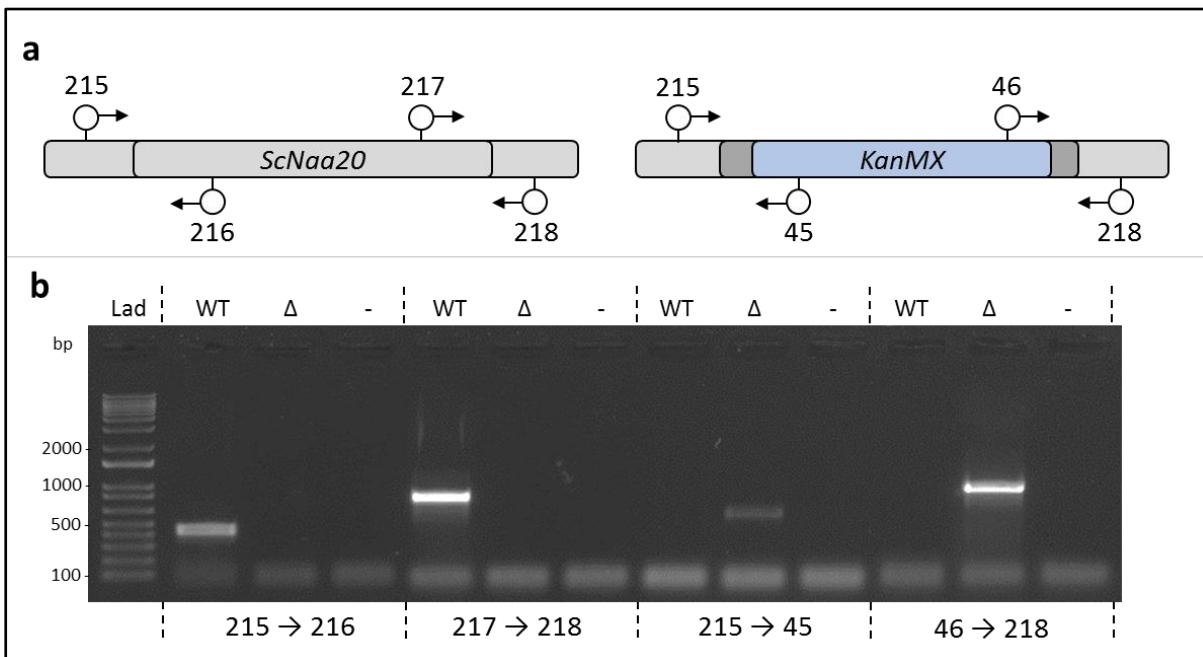
Figure 55 – PCR genotyping of *ScNaa20Δ*
(a) Schematic diagram of *ScNaa20* locus in WT (left) and *ScNaa20Δ* (right) yeast including genotyping primer positions. *ScNaa20* is replaced by a gene cassette including the *KanMX4* gene in *ScNaa20Δ.* (b) Gel electrophoresis of PCR products resulting from amplification of DNA from WT and *ScNaa20Δ* (Δ) with specified primers. No template control PCR products also included (-).

6.2.2 AtSQE1-HA is stabilised in *ScDoa10Δ* and *ScNaa20Δ*

Yeast were transformed with the *AtSQE1* CDS (according to 2.11.3), so that degradation of the protein could be investigated. The sequence was amplified from *Arabidopsis* cDNA and cloned into the pENTR™/D-TOPO™ entry vector (Invitrogen) followed by the pAG413GPD-ccdB-HA destination vector for transformation into yeast.

Expression of *AtSQE1* was driven by GAP promoter within the pAG413GPD-ccdB-HA vector, which also encodes a C-terminal HA-tag (x3). This tag allowed the levels of the protein in yeast to be assessed via Western blot. When expressed in *ScDoa10Δ* yeast, AtSQE1-HA accumulated to higher steady state levels than in WT yeast (Figure 56). Whilst consistent with AtSQE1-HA being more stable in *ScDoa10Δ*, this did not take differential levels of expression into account. Therefore, to assess degradation independently of transcription and translation, cycloheximide (CHX) chases were used to track the level of AtSQE1-HA over time.

CHX inhibits translation in eukaryotes by interacting with the E-site of the 60S ribosomal subunit (Pestova and Hellen, 2003; Schneider-Poetsch et al., 2010) and hence halts protein synthesis. The depletion of a protein can then be tracked, as it is turned over without being replenished. CHX was added to cultures of transformed yeast up to a concentration of 250 µg ml^{-1} with samples of the culture removed at 0-, 30- and 60-minute time points (see 2.9.7). Following the addition of CHX, the abundance of AtSQE1-HA was observed to decrease strongly in WT yeast with just 27% remaining after 30 minutes (Figure 57). In *ScDoa10Δ* yeast, however, the rate of depletion was much slower and 79 % of AtSQE1-HA remained after 30 minutes (Figure 57). This result confirmed that differences in stability did contribute to the increased steady state abundance of AtSQE1-HA in *ScDoa10Δ* relative to WT, potentially due to ubiquitination of AtSQE1-HA by ScDoa10.

Due to its Met-Glu- (ME-) N-terminus, AtSQE1-HA is theoretically subject to Nt-acetylation by NatB and therefore degradation via the yeast Ac/N-degron pathway would be dependent on both ScNaa20 and ScDoa10. The stability of AtSQE1-HA in *ScNaa20Δ* yeast

was therefore also investigated by CHX chase. As with *ScDoa10Δ*, when expressed in *ScNaa20Δ*, AtSQE1-HA was stabilised relative to WT. After 60 minutes, the levels of AtSQE1-HA had only decreased to 58 % of their starting amount (Figure 57). The observed stabilisation of AtSQE1-HA in yeast lacking either ScDoa10 or ScNaa20, was consistent with the monooxygenase being degraded by the Ac/N-degron pathway.
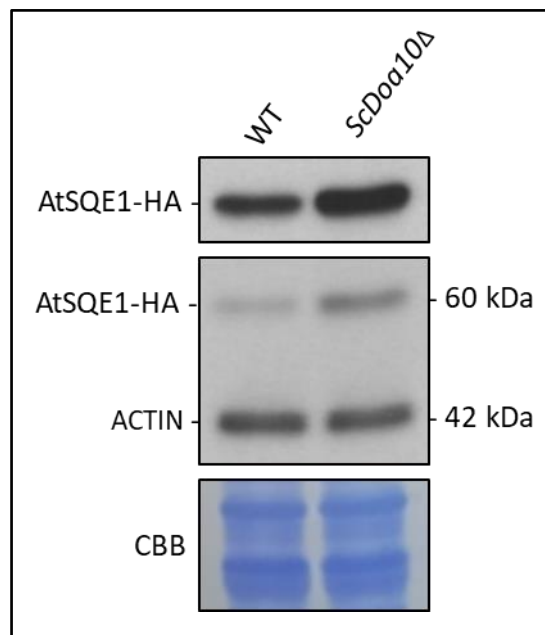
Figure 56 – Steady state levels of AtSQE1-HA in WT and *ScDoa10Δ* yeast
AtSQE1-HA accumulated to a higher level in *ScDoa10Δ* relative to WT. Upper panel shows bands detected by anti-HA antibody (see 2.4.1). AtSQE1-HA bands remain visible when re-probed with anti-β-Actin to show equal loading (middle panel). Lower panel shows staining of membrane with Coomassie Brilliant Blue (CBB). Representative image from three biological replicates.
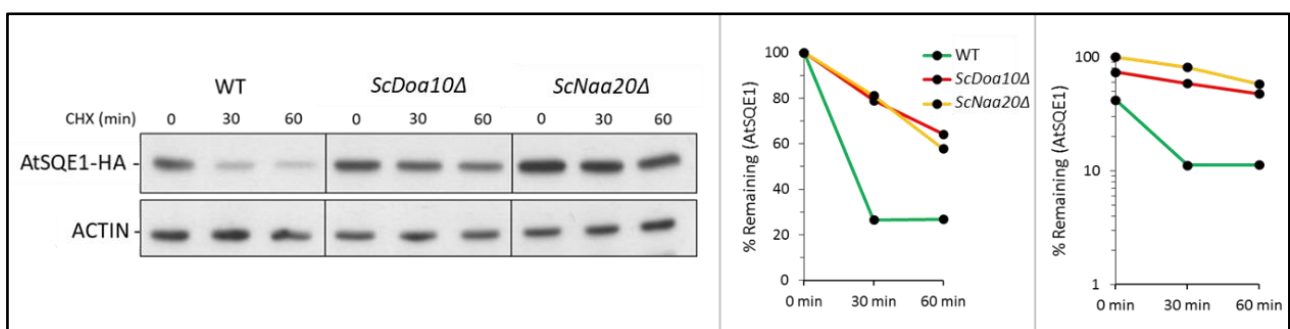
Figure 57 – CHX chase stability analysis of AtSQE1 in *ScDoa10Δ* and *ScNaa20Δ*
Degradation of AtSQE1-HA was slower in *ScDoa10Δ* and *ScNaa20Δ* yeast in comparison to WT. AtSQE1-HA and ACTIN bands were detected at 60 and 42 kDa respectively, using anti-HA and anti-β-Actin antibodies (see 2.4.1). Graphs show quantification of bands detected by Western blot, presented on both a linear and $\log_{10}$ axes. Band intensities were measured using ImageJ, normalised to the corresponding ACTIN band strength and quantified relative to the 0-minute time point (for the linear scale) or the most intense starting band (for the $\log_{10}$ scale).

6.2.3 AtDOA10A and AtNAA20 accelerate the rate of AtSQE1 degradation

The results above suggested that AtSQE1 may be degraded by the Ac/N-degron pathway in yeast, requiring Nt-acetylation by ScNaa20 followed by ubiquitination by ScDoa10. To determine whether the *Arabidopsis* homologues of these enzymes could also facilitate AtSQE1-HA degradation, *ScNaa20Δ* and *ScDoa10Δ* mutants were complemented with the equivalent *Arabidopsis* genes.

The CDS of *AtDOA10A*, *AtDOA10B* and *AtNAA20* were cloned into pDONR™221 (*AtDOA10A*) or pENTR™/D-TOPO™ (*AtDOA10B* and *AtNAA20*) entry vectors (Invitrogen) and then transformed into yeast using the pAG416GPD-ccdB-EGFP destination vector which also uses the GAP promoter to drive expression. Expression of the constructs, including the C-terminal EGFP-tag, was confirmed by RT-PCR (Figure 58).

First, analysis of steady state levels of AtSQE1-HA by Western blot was repeated. As previously observed, AtSQE1-HA accumulated in *ScDoa10Δ* in comparison to WT (Figure 59). Concurrent expression of AtDOA10B-EGFP had no effect on the accumulation of AtSQE1-HA in *ScDoa10Δ*. However, co-expression of AtDOA10A-EGFP caused a clear reduction in the level of AtSQE1-HA producing a protein band of similar intensity to WT (Figure 59). This suggested that both ScDOA10 and AtDOA10A-EGFP may be capable of ubiquitinating AtSQE1-HA for degradation, providing evidence of conservation of function, similar to the complementation of the yeast hygromycin phenotype (see 3.9). AtDOA10B-EGFP appeared unlikely to target AtSQE1-HA suggesting functional divergence, again matching results observed during the yeast hygromycin response. RT-PCR was also performed on the yeast cultures to determine the influence of transcriptional differences. *AtSQE1-HA* expression levels were relatively similar (Figure 59), indicating that the observed differences in protein level were a consequence of post translational control.

To further confirm that differences in stability were responsible for the variation observed, CHX chases were performed and were consistent with previous replicates in that

AtSQE1-HA was degraded more quickly in WT than in *ScDoa10Δ* and *ScNaa20Δ*. (Figure 60 and 61). Complementation of the appropriate mutant with AtDOA10A-EGFP or AtNAA20-EGFP each produced an intermediate degradation rate, demonstrating a partial rescue. After 60 minutes AtSQE1-HA abundance was reduced to 17 % in *ScDoa10Δ+AtDOA10A-EGFP,* similar to the 12 % observed in WT. As seen previously, AtSQE1-HA abundance decreased more slowly in *ScDoa10Δ*, with 35 % remaining, indicating increased stability (Figure 60). Complementation of *ScNaa20Δ* (*ScNaa20Δ+AtNAA20-EGFP*) also destabilised AtSQE1-HA. In this replicate, AtSQE1-HA levels fell to 28 % over the course of 60 minutes in WT and to 69 % in *ScNaa20Δ*. In *ScNaa20Δ+AtNAA20-EGFP,* AtSQE1-HA abundance decreased to 38 %, a clear destabilisation relative to the uncomplemented mutant (Figure 61).
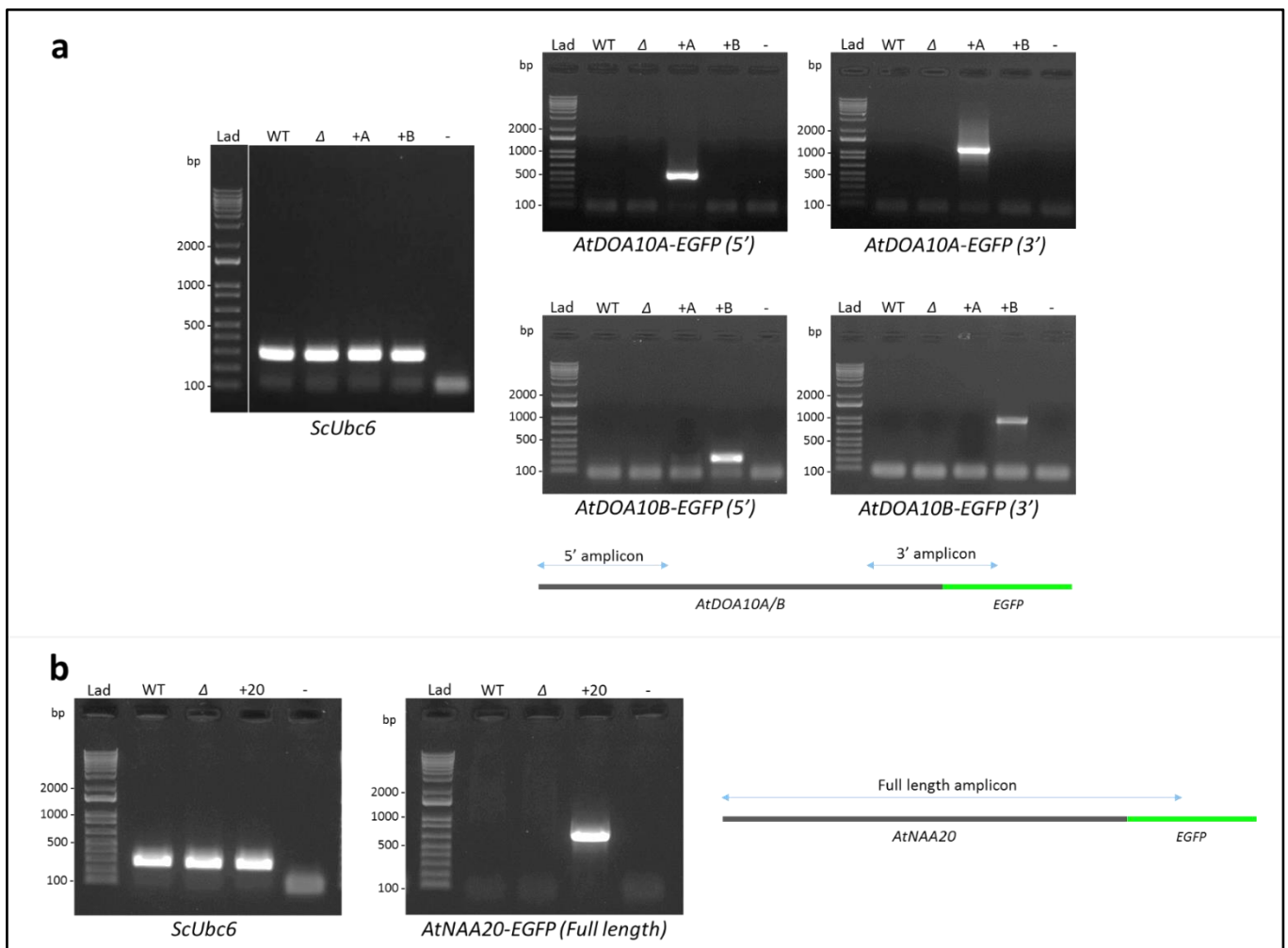
Figure 58 – Confirmation of heterologous expression of *AtDOA10A*, *AtDOA10B* and *AtNAA20*(*-EGFP*) in yeast
(a) RT-PCR of amplicons from 5' and 3' ends confirmed that both *AtDOA10A* and *AtDOA10B* were being successfully transcribed (primers 4, 9, 10, 12, 18, 19, 100). 3' amplicons were generated from primers specific to both the gene and C-terminal EGFP tag. No template control PCR products also included (-). Expression of the constitutive *ScUBC6* was consistent (primers 110+111). (b) RT-PCR of full length *AtNAA20*, including a portion of C-terminal EGFP tag (primers 12+92), and *ScUBC6* control. No template control PCR products also included (-).
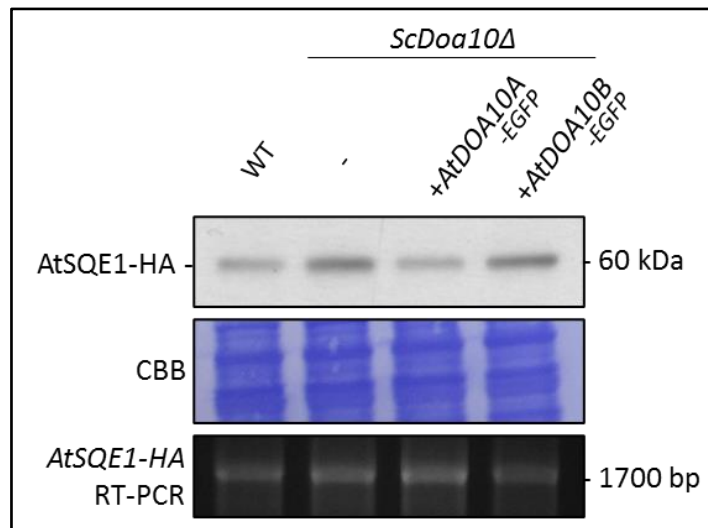
Figure 59 – Steady state levels of AtSQE1-HA following complementation of *ScDoa10Δ* with AtDOA10A/B-EGFP

AtSQE1-HA accumulates to a higher level in *ScDoa10Δ* relative to WT but complementation with AtDOA10A-EGFP reduces AtSQE1-HA abundance to a WT level. Complementation with AtDOA10B-EGFP does not change AtSQE1-HA abundance. RT-PCR (primers 112+116) shows that differences are not due to variation in expression level. Upper panel shows bands detected by anti-HA antibody (see 2.4.1). Middle panel shows equal loading through staining of membrane with Coomassie Brilliant Blue (CBB). Representative image from three biological replicates.
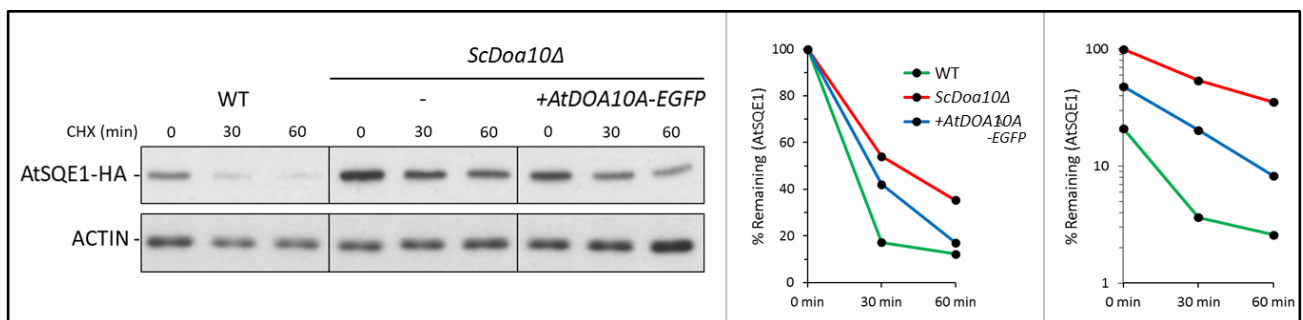
Figure 60 – CHX chase stability analysis of AtSQE1-HA in *ScDoa10Δ* and *ScDoa10Δ+AtDOA10A-EGFP*

AtSQE1-HA is degraded more slowly in *ScDoa10Δ* than in WT but complementation of *ScDoa10Δ* with AtDOA10A-EGFP increases this rate of degradation. AtSQE1-HA and ACTIN bands were detected at 60 and 42 kDa respectively, using anti-HA and anti-β-Actin antibodies (see 2.4.1). Graphs show quantification of bands detected by Western blot, presented on both a linear and $\log_{10}$ axes. Band intensities were measured using ImageJ, normalised to the corresponding ACTIN band strength and quantified relative to the 0-minute time point (for the linear scale) or the most intense starting band (for the $\log_{10}$ scale).
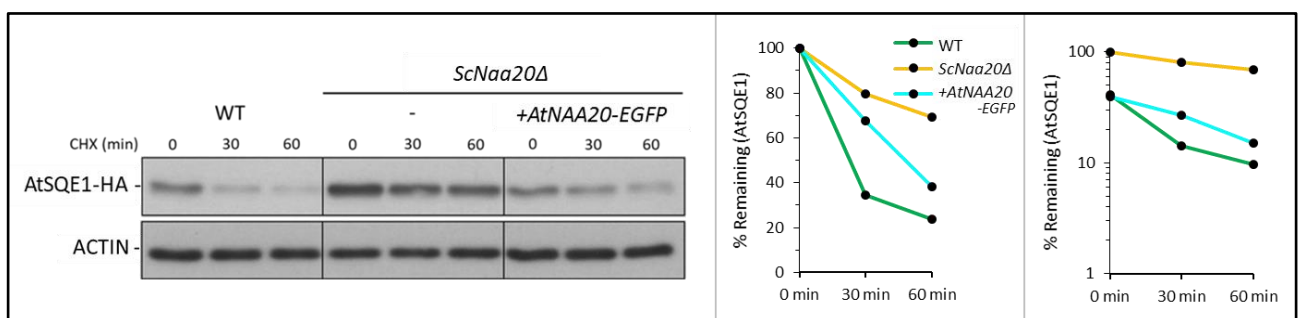
Figure 61 – CHX chase stability analysis of AtSQE1-HA in *ScNaa20Δ* and *ScNaa20Δ+AtNAA20-EGFP*

AtSQE1-HA is degraded more slowly in *ScNaa20Δ* than in WT but complementation of *ScNaa20Δ* with AtNAA20-EGFP increases this rate of degradation. AtSQE1-HA and ACTIN bands were detected at 60 and 42 kDa respectively, using anti-HA and anti-β-Actin antibodies (see 2.4.1). Graphs show quantification of bands detected by Western blot, presented on both a linear and $\log_{10}$ axes. Band intensities were measured using ImageJ, normalised to the corresponding ACTIN band strength and quantified relative to the 0-minute time point (for the linear scale) or the most intense starting band (for the $\log_{10}$ scale).

6.2.4 AtSQE1-HA is stabilised in *ScNaa10Δ*

Evidence thus far supported the potential degradation of AtSQE1-HA via the Ac/N-degron pathway in both yeast and *Arabidopsis* due to the stabilisation of the protein in the absence of DOA10 or NAA20 homologues. To confirm that degradation of AtSQE1-HA was dependent specifically on NAA20, which, as the catalytic subunit of NatB, acetylates ME-N-termini, the stability of AtSQE1-HA in an alternative yeast NAT mutant was tested. ScNaa10 is the catalytic subunit of the NatA complex and is responsible for the acetylation of Ala-, Ser-, Thr-, Cys-, Gly- and Val- N-terminal amino acids following Met cleavage (Linster and Wirtz, 2018; Aksnes et al., 2019). *ScNaa10* mutations, therefore, should not prevent degradation of proteins with ME- N-termini, such as AtSQE1-HA, via the Ac/N-degron pathway.

Unexpectedly, as in *ScDoa10Δ* and *ScNaa20Δ,* AtSQE1-HA was found to be stabilised in *ScNaa10Δ* yeast (Figure 62). As previously observed, when expressed in WT yeast, AtSQE1-HA was strongly depleted after 60 minutes (36 % remaining). In *ScNaa10Δ,* AtSQE1-HA levels only fell to 76 % after 60 minutes (Figure 62). Stabilisation in *ScNaa10Δ* demonstrated that AtSQE1-HA's degradation was not specifically dependent on NatB activity, suggesting that Nt-acetylation of AtSQE1-HA may not be necessary for its degradation and that a universal reduction in NAT activity may actually be responsible. However, it remained possible that AtSQE1-HA could be degraded by the Ac/N-degron pathway mediated by NatB and DOA10 and that the absence of ScNaa10 caused stabilisation through an alternative indirect mechanism.
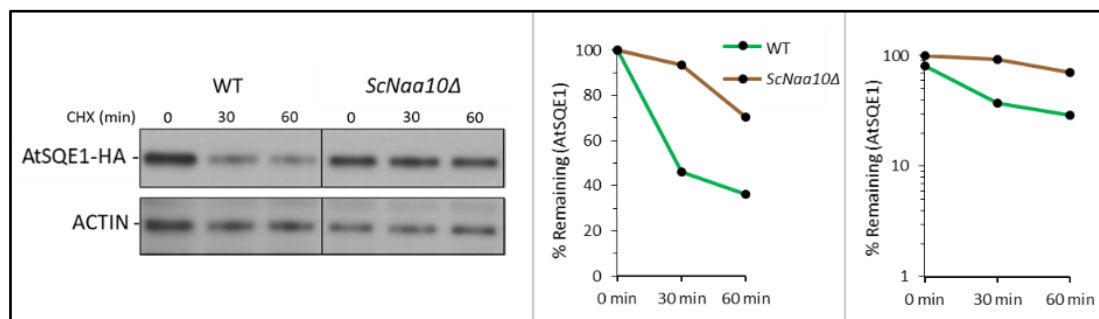
Figure 62 – CHX chase stability analysis of AtSQE1-HA in *ScNaa10Δ*
AtSQE1-HA is stabilised in *ScNaa10Δ* yeast relative to WT. AtSQE1-HA and ACTIN bands were detected at 60 and 42 kDa respectively, using anti-HA and anti-β-Actin antibodies (see 2.4.1). Graphs show quantification of bands detected by Western blot, presented on both a linear and $\log_{10}$ axes. Band intensities were measured using ImageJ, normalised to the corresponding ACTIN band strength and quantified relative to the 0-minute time point (for the linear scale) or the most intense starting band (for the $\log_{10}$ scale).

6.2.5 Prevention of Nt-acetylation does not stabilise AtSQE1-HA in yeast

To further clarify whether Nt-acetylation of AtSQE1-HA was required for its degradation, next, variants of AtSQE1-HA were generated bearing altered Nt-residues. Site directed mutagenesis (see 2.7.11) was used to change the DNA sequence of *AtSQE1* within the pENTR™/D-TOPO™ entry vector in order to produce AtSQE1-HA proteins beginning with Met-Lys- (MK-) and Met-Pro- (MP-) amino acids, when expressed in yeast. MP-N-termini are not acetylated in any species and acetylation of MK-N-termini is known to be very rare in yeast (Polevoda and Sherman, 2003; Arnesen et al., 2009; Goetze et al., 2009), thereby theoretically blocking Nt-acetylation of AtSQE1-HA (Table 20). As a potential substrate of the Ac/N-degron pathway, preclusion of Nt-acetylation would be expected to stabilise AtSQE1-HA by preventing recognition by its N-recognin.

The ME- (native), MK- and MP-AtSQE1-HA proteins were expressed in WT and *ScDoa10Δ* yeast and their steady state levels assessed by Western blot. An increased abundance of each variant was observed in *ScDoa10Δ* (Figure 63a), potentially indicating that all three remain substrates of ScDoa10. In both yeast genotypes however, mutations of AtSQE1-HA's N-terminus did not cause accumulation of the protein relative to WT ME-AtSQE1-HA, as would be expected if an Ac/N-degron target. The same pattern of band strength was observed for both WT and *ScDoa10Δ*, with ME-AtSQE1-HA clearly more abundant than both MK-AtSQE1-HA and MP-AtSQE1-HA. This result suggested that Nt-acetylation was not required for AtSQE1-HA degradation, and that blocking Nt-acetylation led to a reduced accumulation of steady state protein levels.

To analyse this further, without the influence of transcription and translation, CHX chases were again performed, now to compare the depletion of ME-AtSQE1-HA and MP-AtSQE1-HA. Over 60 minutes the degradation rates of the two proteins were relatively similar. In fact, the abundance of MP-AtSQE1-HA decreased slightly faster than ME-AtSQE1-HA, with only 17 % remaining after 60 minutes in comparison to 26 % (Figure 63b). The lack of stabilisation of MP-AtSQE1-HA confirmed that the protein did not require Nt-acetylation in

order to be degraded and was therefore unlikely to be an Ac/N-degron pathway substrate. Moreover, the enhanced abundance of all Nt-mutated variants in *ScDoa10Δ* relative to WT indicated that ScDoa10 targets AtSQE1 for degradation independent of the nature of its N-terminus.

Since preventing Nt-acetylation did not preclude AtSQE1-HA degradation, the stabilisation of AtSQE1-HA in *ScNaa20Δ* seemed likely to be an indirect effect of reduced global levels of Nt-acetylation, especially since *ScNaa10Δ* mutations had the same impact. To confirm this, an additional N-terminal variant of AtSQE1-HA was produced by site directed mutagenesis with a Met-Ala- (MA-) N-terminus. Peptide sequences beginning with the residues MA- are targeted by either NatA or NatE, depending on whether the initiator methionine is cleaved (Linster and Wirtz, 2018; Aksnes et al., 2019) (Table 20). MA-AtSQE1-HA was expressed in WT and *ScNaa20Δ* yeast to determine whether the altered N-terminus affected the stabilisation previously observed in *ScNaa20Δ*. As peptide sequences beginning with the residues MA- are targeted by either NatA or NatE, the loss of ScNaa20 should not affect the degradation of MA-AtSQE1-HA via the Ac/N-degron pathway. However, CHX chase analysis revealed that MA-AtSQE1-HA levels dropped to 16 % in WT after 60 minutes but only to 49 % in *ScNaa20Δ* (Figure 64), showing that MA-AtSQE1-HA was stabilised in *ScNaa20Δ*. This result provided further evidence that ScNaa20 facilitates AtSQE1-HA degradation in an indirect manner rather than via the Ac/N-degron pathway because AtSQE1-HA is degraded more quickly in the presence of ScNaa20, even when NatB is not required for Nt-acetylation of the protein.

Table 20 – NAT specificity of AtSQE1 N-terminal variants

AtSQE1 N-terminus	NAT specificity
ME- (Native)	NatB (NAA20)
MA-	NatA (NAA10) / NatE (NAA50)
MK-	Rarely acetylated
MP-	Not acetylated
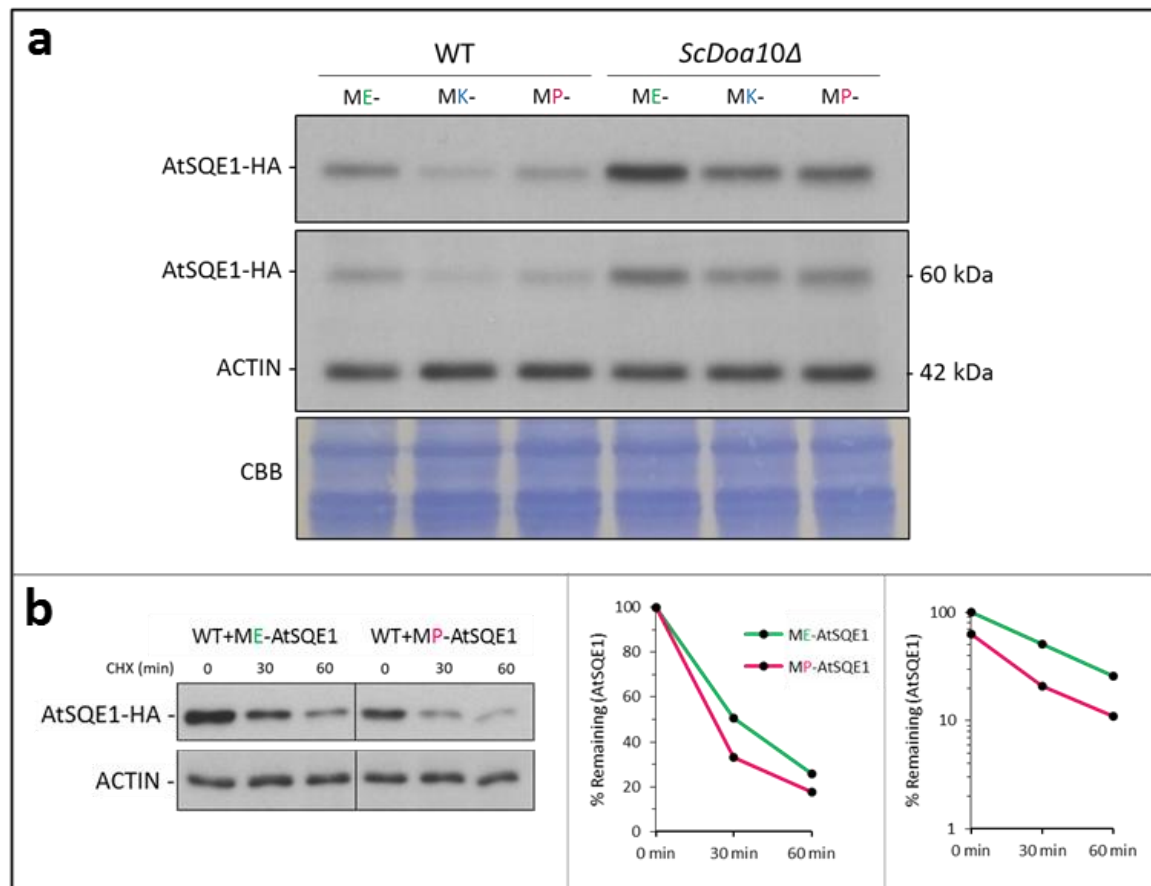
Figure 63 – Steady state and CHX chase stability analysis of unacetylatable AtSQE1-HA Nt-mutants
(a) Steady state levels of unacetylatable MK- and MP-AtSQE1-HA are lower than ME-AtSQE1 in both WT
ScDoa10Δ yeast. Each variant accumulates to a higher level in *ScDoa10Δ* relative to WT. (b) CHX chase
analysis shows that MP-AtSQE1-HA is degraded slightly faster than ME-AtSQE1-HA. AtSQE1-HA and
ACTIN bands were detected at 60 and 42 kDa respectively, using anti-HA and anti-β-Actin antibodies (see
2.4.1). Graphs show quantification of bands detected by Western blot, presented on both a linear and
log10 axes. Band intensities were measured using ImageJ, normalised to the corresponding ACTIN band
strength and quantified relative to the 0-minute time point (for the linear scale) or the most intense starting
band (for the log10 scale).
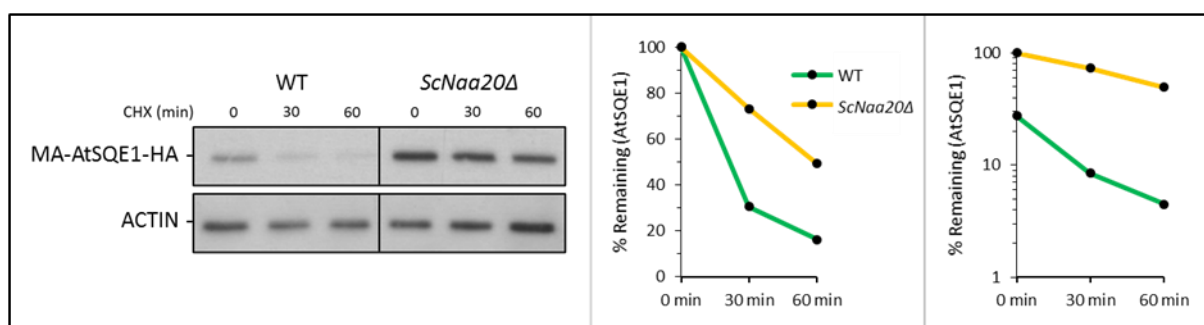
Figure 64 – CHX chase stability analysis of MA-AtSQE1-HA in *ScNaa20Δ*
Targeting AtSQE1-HA to NatA/E through mutation of its N-terminus does not eliminate stability in
ScNaa20Δ. AtSQE1-HA and ACTIN bands were detected at 60 and 42 kDa respectively, using anti-HA
and anti-β-Actin antibodies (see 2.4.1). Graphs show quantification of bands detected by Western blot,
presented on both a linear and $\log_{10}$ axes. Band intensities were measured using ImageJ, normalised
to the corresponding ACTIN band strength and quantified relative to the 0-minute time point (for the
linear scale) or the most intense starting band (for the $\log_{10}$ scale).

6.3 AtSQE1 stability analysis in *Arabidopsis*

Analysis of AtSQE1-HA's stability in yeast suggested that it was a substrate of both *ScDoa10* and *AtDOA10A* but not of the Ac/N-degron pathway, although its degradation was indirectly dependent on NAT activity. Although heterologous expression in yeast was a useful system for monitoring these dynamics, it was important to study AtSQE1's *in planta* degradation since yeast and plant Nt-acetylation systems are not identical (Linster and Wirtz, 2018) and results are more relevant to the real world. To establish whether AtSQE1 was degraded by the same mechanism in plants ME-AtSQE1 was transformed into a variety of *Arabidopsis* mutant lines (characterised in Chapter IV), so that degradation could be tracked *in planta*. MP-AtSQE1 was also transformed into Col-0 *Arabidopsis*, to determine the significance of Nt-acetylation.

6.3.1 Transformation of *Arabidopsis* with ME- and MP-AtSQE1-Myc

The CDSs of ME- and MP-AtSQE1 in pENTR™/D-TOPO™ (Invitrogen) (see 6.2.2 and 6.2.5) were transferred to the binary pGWB17 destination vector (Nakagawa et al., 2007), which drives gene expression using the cauliflower mosaic virus 35S promoter and contains four C-terminal Myc tags. This vector was then used for *Agrobacterium*-mediated transformation of *Arabidopsis* (see 2.11.6) to produce transgenic plants expressing 35S::ME-/MP-AtSQE1-Myc. Successfully transformed *Arabidopsis* plants were selected based on their resistance to hygromycin, encoded within the T-DNA insert of pGWB17. By the T3 generation, lines homozygous for the transgenes were identified by their 100% hygromycin resistance.

6.3.2 Confirmation of AtSQE1-Myc production and analysis of expression level

To confirm that the homozygous lines selected successfully expressed *AtSQE1-Myc* and to determine the strength of this expression, qRT-PCR was performed on RNA extracted from 7-day old T3 seedlings. The amplification of native *AtSQE1* was avoided through the amplification of a small DNA fragment (141 bp) including portions of *AtSQE1*'s C-terminus and the *Myc* tag (primers 159+160). Expression of the transgene was detected in each of the lines

but a significant level of variation was observed (Figure 65). For Col-0, both a strongly and a weakly expressing line was identified.

The expression of AtSQE1-Myc was also assessed at the protein level by Western blot. A band corresponding to the recombinant protein was present for each line (Figure 65). In general, band strength strongly correlated with the gene expression level determined by qRT-PCR. For example, the strongest expressors of *AtSQE1-Myc*, Col-0 (8-4) and *Atdoa10a/b* (RNAi line 4-2(2)) (8-4), also produced the most intense protein bands. Similarly, Col-0 (6-1), the weakest *AtSQE1-Myc* expressor, produced the least intense protein band. This observed correlation highlighted that differences in gene expression had a substantial influence on AtSQE1-Myc abundance and that CHX chases would again be required in order to study protein stability independently.
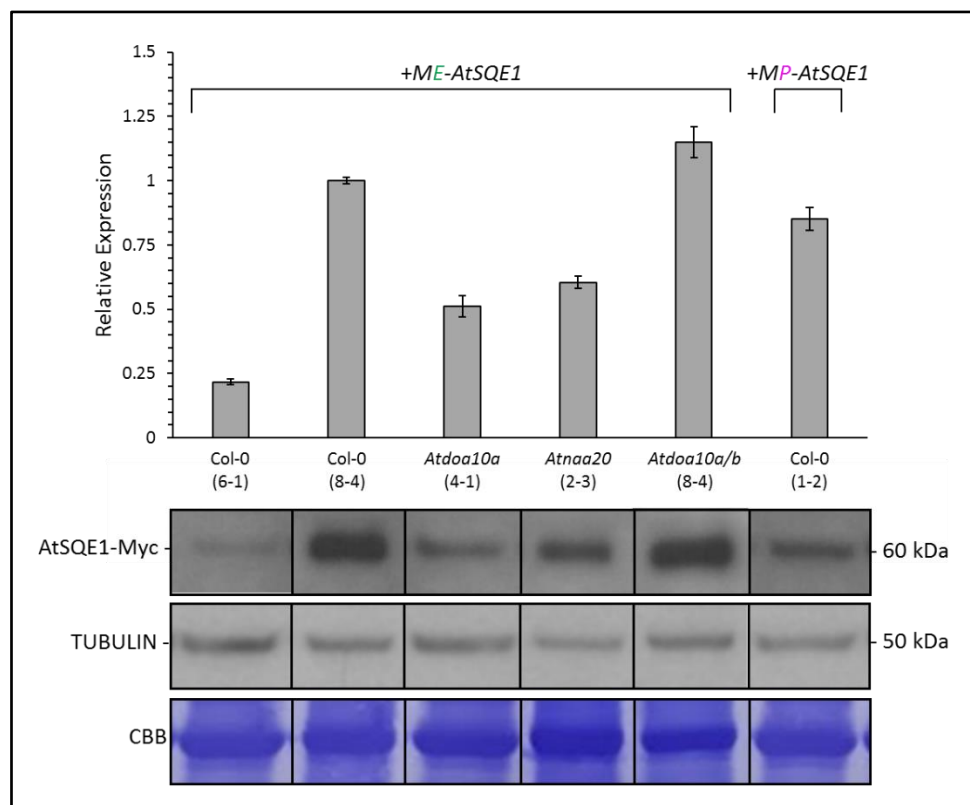
Figure 65 – Gene expression and protein steady-state analysis of *AtSQE1-Myc* transgenic lines
AtSQE1-Myc was expressed to varying degrees across the generated lines. Differences in steady-state protein abundance detected by Western blot correlate strongly with the levels of gene expression determined by qRT-PCR (primers 159+160). AtSQE1-Myc and TUBULIN bands were detected, using anti-Myc and anti-β-Tubulin antibodies (see 2.4.1). Expression of *AtSQE1-Myc* was calculated according to the delta delta C_t method, relative to Col-0 (8-4). Values of relative expression are the average of two technical replicates (±SE). Both RNA and protein samples were extracted from 7-day old *Arabidopsis* seedlings.

6.3.3 Mutation of the AtSQE1-Myc's N-terminus does not alter its cellular localisation

The observation that MP-AtSQE1-HA was degraded at a similar rate to ME-AtSQE1-HA when expressed in yeast was considered to provide evidence that Nt-acetylation was not required for AtSQE1-Myc degradation (see 6.2.5). However, blocking Nt-acetylation could also result in a change in protein localisation, since acetylation affects a protein's affinity for plasma membranes and their translocation to the ER (see 1.3.2.2). Additionally, sequence changes close to the N-terminus may have a direct impact on signal peptides, although AtSQE1 is not predicted to contain such a sequence according to SignalP-5.0 (http://www.cbs.dtu.dk/services/SignalP/). With any potential change in localisation likely to have a significant impact on potential degradation pathways and rates, it was important to determine whether ME- and MP-AtSQE1-Myc are found within the same cellular compartments.

Cell fractionation was used to separate organelles by according to size (see 2.9.8), as was performed for AtDOA10A- and AtDOA10B-GUS fusion proteins (see 3.6) (Abas and Luschnig, 2010). Following preclearance to remove cell debris and nuclei, the microsomal fraction was isolated from the soluble fraction by centrifugation at 21000 x g for 2 hours to produce a pellet. GFP-tagged AtSQE1 has been previously shown to co-localise with an ER marker (Laranjeira et al., 2015). Consistent with this observation, ME-AtSQE1-Myc was detected in the microsomal fraction, as were the ER marker proteins AtCNX1/2 (Figure 66). MP-AtSQE1-Myc was also detected in the microsomal fraction, demonstrating that the N-terminal mutation did not affect localisation. The cytosolic marker AtUGPase was present in the soluble fraction (Figure 66). The ER localisation of both ME- and MP-SQE1-Myc matches that of AtDOA10A (see 3.6), confirming the feasibility of AtSQE1 ubiquitination by AtDOA10A.
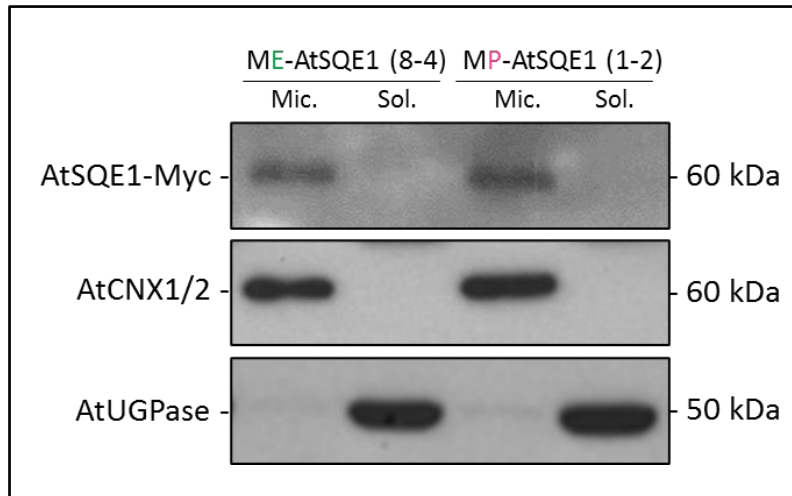
Figure 66 – Assessment of ME- and MP-AtSQE1-Myc localisation by cell fractionation
ME- and MP-AtSQE1-Myc were both detected within the microsomal fraction of cell lysates following cell fractionation, as were the ER markers AtCNX1/2. AtUGPase remained in the soluble fraction.

6.3.4 AtSQE1-Myc is stabilised in the absence of AtDOA10A

As was performed for analysis in yeast, and because of the significant differences in the expression level of *AtSQE1-Myc* observed (see 6.3.2), CHX chases were again performed to assess the degradation rate of the AtSQE1-Myc in the various backgrounds. Chases were carried out by adding seven-day old seedlings to 300 µg ml^{-1} CHX. Seedlings were then removed from the CHX treatment at 0-, 3- and 6-hour time points (see 2.9.7.2).

First the stability of AtSQE1-Myc was compared in Col-0 and *Atdoa10a*. Consistent with observations in yeast, AtSQE1-Myc was stabilised in the absence of AtDOA10A. Levels of AtSQE1-Myc fell to 30 % in Col-0 after 3 hours, whereas in *Atdoa10a*, 81% remained at the same time point (Figure 67). Unexpectedly the AtSQE1-Myc level increased over the next 3 hours in Col-0, but this was most likely due to biological variation and technical errors; relative abundance still remained below the level of AtSQE1-Myc in *Atdoa10a* and such a recovery was not observed for Col-0 seedlings in any other comparison.

To further confirm that AtSQE1-Myc degradation was dependent of AtDOA10A, CHX chases were performed with the two strongest expressing lines: a second (independent) Col-

0 line and *Atdoa10a/b* (RNAi line 4-2(2)). Again, AtSQE1-Myc was stabilised in a line lacking

AtDOA10A, as, after 6 hours, AtSQE1-Myc abundance remained at 95 % of the starting

amount in *Atdoa10a/b* (RNAi line 4-2(2)) (Figure 68). In Col-0 only 17 % of AtSQE1-Myc

remained after 6 hours showing that AtSQE1-Myc was strongly stabilised in *Atdoa10a/b* (RNAi

line 4-2(2)), as it was in *Atdoa10a*.
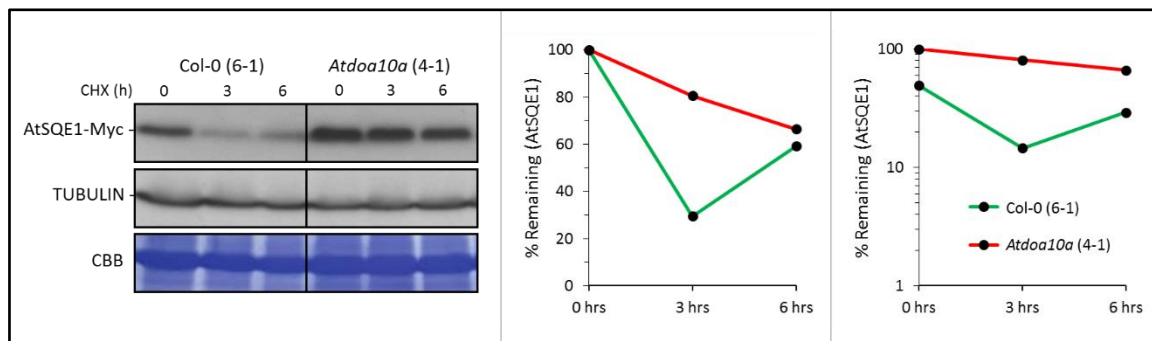
Figure 67 – CHX chase stability analysis of AtSQE1-Myc in *Atdoa10a*
AtSQE1-Myc is stabilised in *Atdoa10a* relative to Col-0. AtSQE1-Myc and TUBULIN bands were
detected at 60 and 50 kDa respectively, using anti-HA and anti-β-Tubulin antibodies (see 2.4.1).
Graphs show quantification of bands detected by Western blot, presented on both a linear and $\log_{10}$
axes. Band intensities were measured using ImageJ, normalised to the corresponding TUBULIN band
strength and quantified relative to the 0-minute time point (for the linear scale) or the most intense
starting band (for the $\log_{10}$ scale).
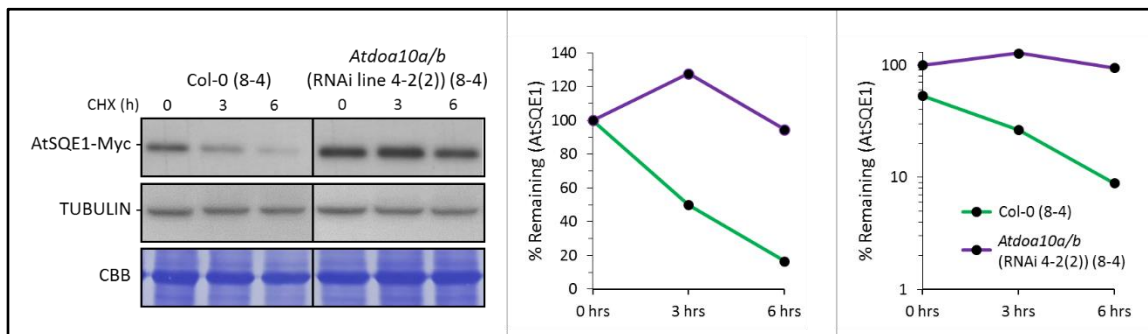
Figure 68 – CHX chase stability analysis of AtSQE1-Myc in *Atdoa10a/b* (RNAi line 4-2(2))
AtSQE1-Myc is stabilised in *Atdoa10a/b* (RNAi line 4-2(2) relative to Col-0. AtSQE1-Myc and TUBULIN
bands were detected at 60 and 50 kDa respectively, using anti-HA and anti-β-Tubulin antibodies (see
2.4.1). Graphs show quantification of bands detected by Western blot, presented on both a linear and
$\log_{10}$ axes. Band intensities were measured using ImageJ, normalised to the corresponding TUBULIN
band strength and quantified relative to the 0-minute time point (for the linear scale) or the most
intense starting band (for the $\log_{10}$ scale).

6.3.5 AtSQE1-Myc is stabilised in the absence of AtNAA20

Degradation of AtSQE1-HA expressed in yeast appeared to be independent of its Nt-acetylation (see 6.2.5). Despite this, removal of ScNaa20, the catalytic subunit of its cognate NAT (NatB), was shown to stabilise AtSQE1-HA (see 6.2.2). To investigate the relationship between AtSQE1 and NatB in *Arabidopsis*, AtSQE1-Myc degradation was studied in the *Atnaa20* background.

Due to technical difficulties, AtSQE1-Myc bands could only be very weakly detected in this comparison, but quantification of bands was still able to be performed following contrast enhancement. The initial abundance of AtSQE1-Myc was lower in *Atnaa20* compared to Col-0, but levels of the protein only dropped to 84 % over the course of 6 hours. Alternatively, in Col-0, AtSQE1-Myc band strength decreased to 23 %, showing that AtNAA20 accelerates AtSQE1-Myc degradation in *Arabidopsis,* as well as in yeast (Figure 69).

Given these differences in degradation rate, it was surprising to see that the steady state level (0-hour time point) of AtSQE1-Myc was markedly higher in Col-0 than in *Atnaa20*. Whilst expression of *AtSQE1-Myc* was observed to be lower in the *Atnaa20* line used (*Atnaa20*+ME-AtSQE1-Myc (2-3) vs Col-0+ME-AtSQE1-Myc (8-4)) (see 6.3.2), the clear reduction in degradation rate might be expected to significantly increase the levels of AtSQE1-Myc at equilibrium. The relatively low steady state level of AtSQE1-Myc in *Atnaa20* may therefore indicate further complexities in the link between the two enzymes, potentially consistent with the indirect relationship observed in the yeast system. Further replicates are required to explore this further and to fully confirm the effect of the *Atnaa20* mutation on AtSQE1-Myc stability, due to the problems detecting protein bands.
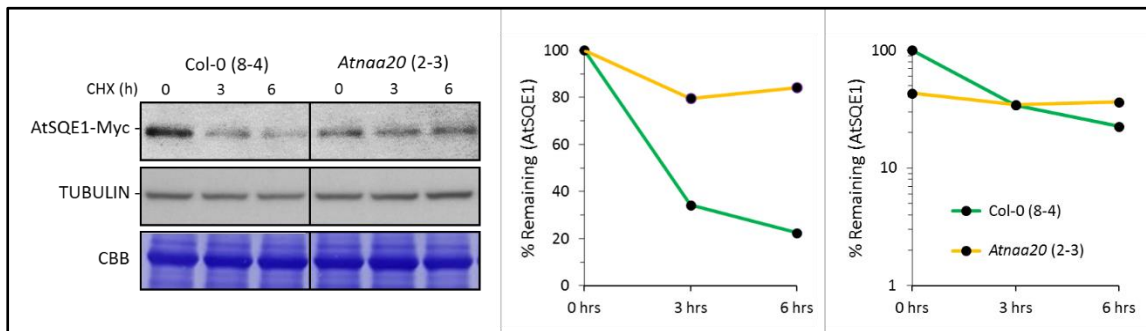
Figure 69 – CHX chase stability analysis of AtSQE1-Myc in *AtNaa20*
AtSQE1-Myc is stabilised in *Atnaa20* relative to Col-0 despite a reduced steady state level (0-hour time point). AtSQE1-Myc and TUBULIN bands were detected at 60 and 50 kDa respectively, using anti-HA and anti-β-Tubulin antibodies (see 2.4.1). Graphs show quantification of bands detected by Western blot, presented on both a linear and log_{10} axes. Band intensities were measured using ImageJ, normalised to the corresponding TUBULIN band strength and quantified relative to the 0-minute time point (for the linear scale) or the most intense starting band (for the log_{10} scale).

6.3.6 Prevention of Nt-acetylation does not stabilise AtSQE1-Myc in *Arabidopsis*

To further examine this potential indirect relationship, the *in planta* stability of the unacetylatable MP-AtSQE1-Myc variant was studied, as had been performed in yeast (see 6.2.5). Similar to observations in yeast, little difference in the stability of the native ME- and mutant MP-AtSQE1-Myc variants was observed in Col-0 backgrounds, although technical problems again prevented the detection of strong bands (Figure 70). After 6 hours the amount of ME-AtSQE1-Myc decreased to 48 % of its starting quantity and MP-AtSQE1-Myc levels fell to 52 %. Both ME- and MP-AtSQE1-Myc were shown to be ER localised (see 6.3.3) and therefore degradation rates were not influenced by potential Nt-acetylation dependent changes in localisation. In yeast MP-AtSQE1-HA was depleted marginally faster than ME-AtSQE1-HA (Figure 63b) but here MP-AtSQE1-Myc appeared slightly more stable. Notwithstanding this minor difference, the depletion of MP-AtSQE1-Myc indicated that Nt-acetylation of AtSQE1-Myc is not required for its degradation. This result, combined with the increased stability of AtSQE1-Myc observed in *Atdoa10a*, *Atdoa10a/b* (RNAi line 4-2(2)) and *Atnaa20* showed that the degradation dynamics in *Arabidopsis* were generally similar to the heterologous yeast system.
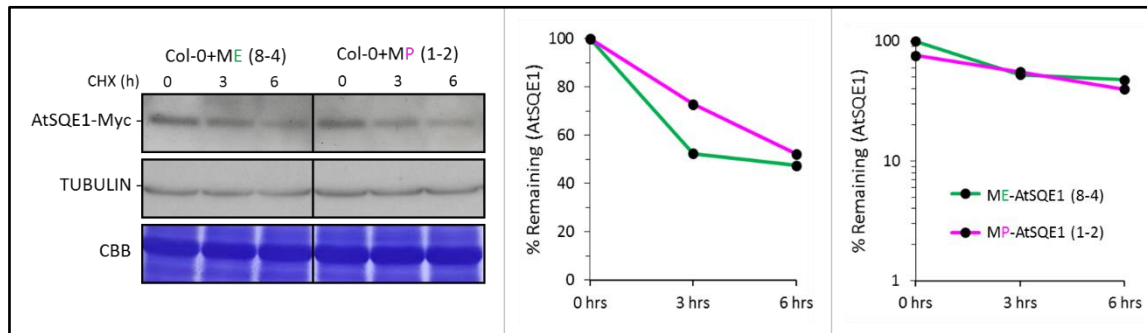
Figure 70 – CHX chase stability analysis of ME-AtSQE1-Myc and the unacetylatable MP-AtSQE1-Myc mutant

Native ME-AtSQE1-Myc is degraded at a similar rate to mutant MP-AtSQE1-Myc despite the prevention of Nt-acetylation. AtSQE1-Myc and TUBULIN bands were detected at 60 and 50 kDa respectively, using anti-HA and anti-β-Tubulin antibodies (see 2.4.1). Graphs show quantification of bands detected by Western blot, presented on both a linear and log_{10} axes. Band intensities were measured using ImageJ, normalised to the corresponding TUBULIN band strength and quantified relative to the 0-minute time point (for the linear scale) or the most intense starting band (for the log_{10} scale).

6.3.7 AtSQE1-Myc is degraded by the proteasome

As an E3 ligase, AtDOA10A tags substrates for degradation through the addition of ubiquitin (see 1.1). Polyubiquitnated proteins are recognised and degraded by the 26S proteasome, and therefore if AtSQE1 is a true substrate of AtDOA10A, its degradation would also be dependent on the functioning of the proteasome. To confirm that AtSQE1's degradation did require the 26S proteasome, its rate of degradation was studied in the presence of the chemical bortezomib. Bortezomib is a proteasome inhibitor which works by binding to active sites within the 20S core of the 26S proteasome to prevent chymotrypsin-like and caspase-like activity, thereby halting proteolysis (Groll et al., 2006; Kisselev et al., 2012).

Treatment with bortezomib prevented the CHX-induced degradation of AtSQE1-Myc; no decrease in band strength was observed after 6 hours (103 % remaining) (Figure 71). Seedlings treated with CHX and DMSO (control) exhibited AtSQE1-Myc depletion over 6 hours, with only 35 % remaining, matching previous chases. Inhibition of AtSQE1-Myc depletion by bortezomib confirmed that the degradation occurred via the UPS, consistent with targeting by AtDOA10A.
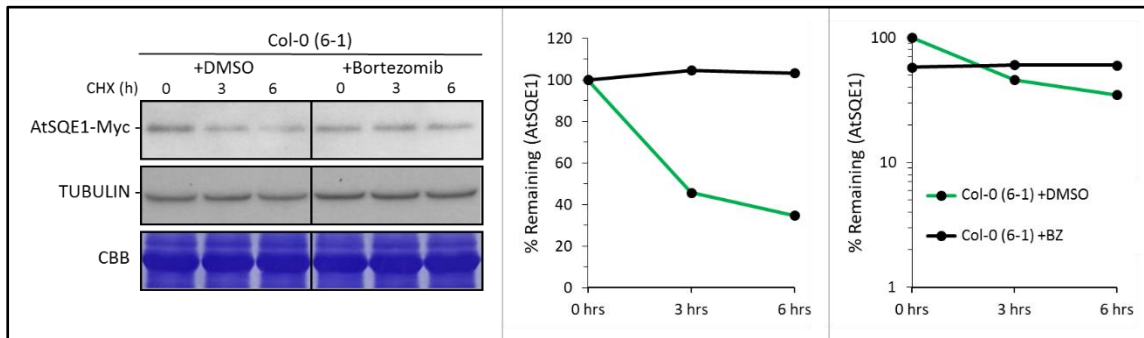
Figure 71 – The effect of bortezomib on CHX-induced degradation of AtSQE1-Myc in Col-0
CHX-induced degradation of AtSQE1-Myc was inhibited by the addition of 50 µM bortezomib (BZ), showing that it occurs via the UPS. AtSQE1-Myc and TUBULIN bands were detected at 60 and 50 kDa respectively, using anti-HA and anti-β-Tubulin antibodies (see 2.4.1). Graphs show quantification of bands detected by Western blot, presented on both a linear and $\log_{10}$ axes. Band intensities were measured using ImageJ, normalised to the corresponding TUBULIN band strength and quantified relative to the 0-minute time point (for the linear scale) or the most intense starting band (for the $\log_{10}$ scale).

6.3.8 LDAO treatment induces AtSQE1-Myc degradation independently of AtDOA10A and AtNAA20

Precise control of the enzymes of the sterol synthesis pathway is required to regulate and maintain appropriate levels of phytosterols and their intermediates (De Vriese et al., 2021). In human and yeast systems, SQE expression and stability have been shown to be influenced by the accumulation or consumption of different sterols and intermediates of the pathway (Leber et al., 2001; Gill et al., 2011; Foresti et al., 2013; Howe et al., 2017; Scott et al., 2020). For human cells, applications of cholesterol, the major end product of sterol biosynthesis have been shown to cause proteolysis of the human SQE (known as squalene monooxygenase, HsSM) (Gill et al., 2011). By contrast, ScErg1 degradation can be induced through the build-up of the alternative sterol lanosterol, an ergosterol precursor, following inhibition of the downstream demethylase ScErg11 either through treatment with the chemical fluconazole or by mutation (Foresti et al., 2013).

The key sterol(s) or intermediate(s) that feedback into AtSQE1 regulation in *Arabidopsis* remain unknown. To try and identify these compounds, and to determine whether AtDOA10A may be involved in sterol-induced degradation of AtSQE1, we examined the

impact of LDAO on AtSQE1-Myc stability. As an inhibitor of both CAS and CPI enzymes of the sterol synthesis pathway (see 5.3.5) (Darnet et al., 2020), LDAO could potentially cause the accumulation of multiple products downstream of AtSQE1, increasing the likelihood of the treatment having an impact. Treatments were performed by adding seven-day old seedlings to 100 µg ml^{-1} LDAO.

When expressed in Col-0, AtSQE1-Myc was degraded rapidly in response to the LDAO treatment, with band strength falling to 14 % over 60 minutes (Figure 72). A similar rapid depletion was observed for *Atdoa10a*/b (RNAi line 4-2(2)), where only 11% of AtSQE1-Myc proteins remained after 60 minutes (Figure 72), demonstrating that AtDOA10A is not responsible for this response. LDAO is a surfactant and therefore the strong reduction in AtSQE1-Myc during the treatment could have been caused by disruption of plasma membranes rather than the accumulation of a downstream sterol. To test this, the effect of the treatment on the ER-markers AtCNX1/2 was assessed. AtCNX1/2 levels remained relatively constant, confirming that the ER integrity was maintained (Figure 72). Like AtDOA10A, AtNAA20 was also not required for LDAO-induced degradation of AtSQE1-Myc. LDAO caused AtSQE1-Myc levels to decrease to 19 % in *Atnaa20* after 60 minutes during a separate treatment (Figure 72).

To confirm that AtSQE1-Myc depletion was LDAO specific, rather than a submergence response, a water-only control treatment was also performed. No decrease in AtSQE1-Myc was observed in this treatment (Figure 73). Therefore, application of LDAO was sufficient to specifically trigger AtSQE1-Myc degradation, indicating that, like yeast and human sterol synthesis pathways, accumulation of downstream intermediates feeds back on upstream enzymes to reduce flux through the pathway.
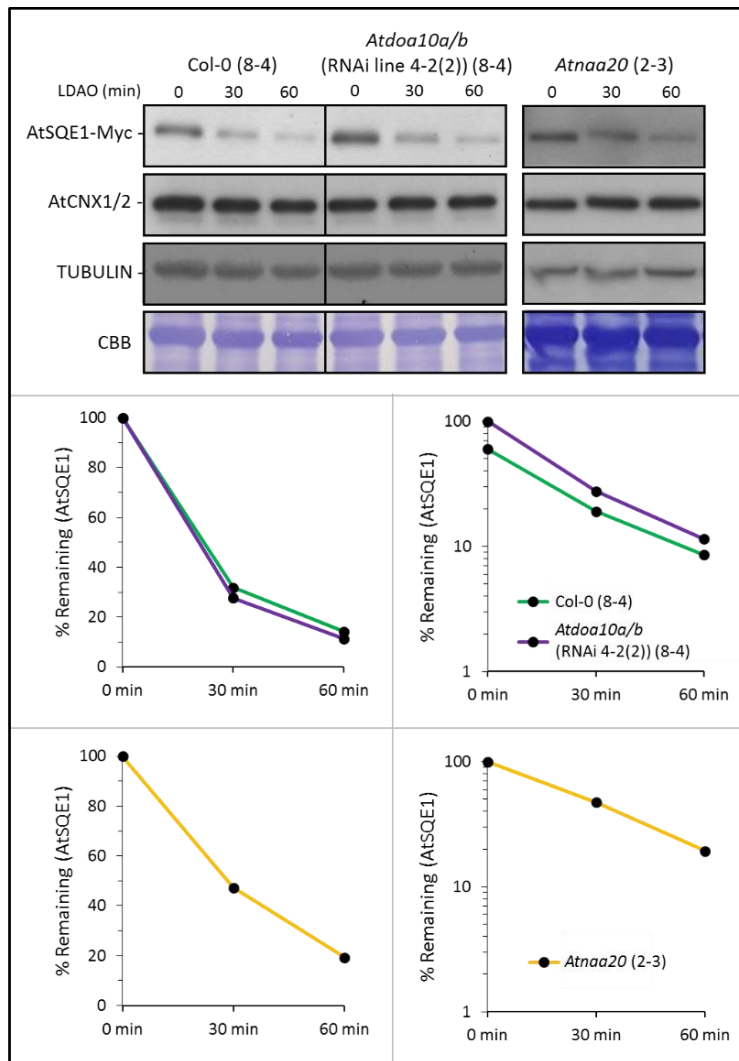
Figure 72 – LDAO treatment stability analysis of AtSQE1-Myc in Col-0, *Atdoa10a/b* (RNAi line 4-2(2)) and *AtNaa20*

AtSQE1-Myc was rapidly degraded in response to LDAO in all genetic backgrounds. The stability of AtCNX1/2 during the treatment confirmed that plasma membranes were not significantly disrupted. AtSQE1-Myc and AtCNX1/2 bands were detected at 60 kDa and TUBULIN bands were detected at 50 kDa, using anti-HA, anti-AtCNX1/2 and anti-β-Tubulin antibodies (see 2.4.1). Graphs show quantification of bands detected by Western blot, presented on both a linear and $\log_{10}$ axes. Band intensities were measured using ImageJ, normalised to the corresponding TUBULIN band strength and quantified relative to the 0-minute time point (for the linear scale) or the most intense starting band (for the $\log_{10}$ scale).
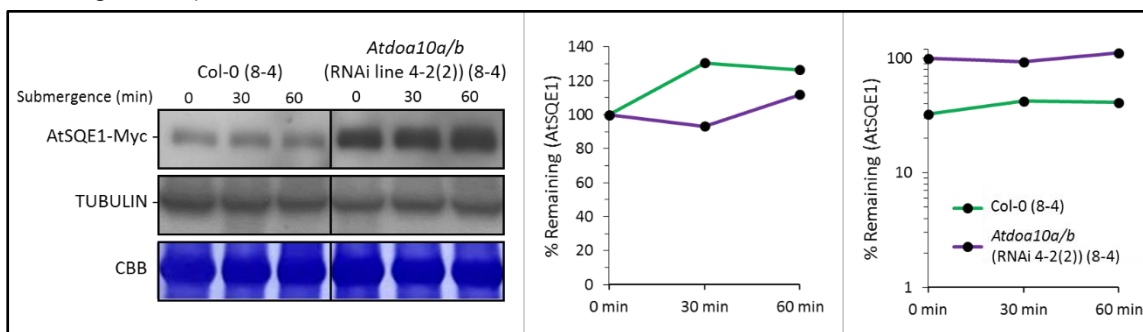
Figure 73 – Submergence treatment stability analysis of AtSQE1-Myc

AtSQE1-Myc was not degraded when seedlings were placed in additive-free water. AtSQE1-Myc and TUBULIN bands were detected at 60 and 50 kDa respectively, using anti-HA and anti-β-Tubulin antibodies (see 2.4.1). Graphs show quantification of bands detected by Western blot, presented on both a linear and $\log_{10}$ axes. Band intensities were measured using ImageJ, normalised to the corresponding TUBULIN band strength and quantified relative to the 0-minute time point (for the linear scale) or the most intense starting band (for the $\log_{10}$ scale).

6.3.9 LDAO induces non-proteasomal degradation of AtSQE1-Myc

As AtDOA10A appeared not to be required for LDAO-induced degradation of AtSQE1-Myc, bortezomib was used (as in 6.3.7) to establish whether its degradation was proteasomal and therefore potentially caused by the action of an alternative E3 ligase. However, inhibition of the proteasome with bortezomib failed to prevent the strong depletion of AtSQE1-Myc caused by the LDAO treatment (Figure 74). During this treatment AtSQE1-Myc levels fell to 39 % in 60 minutes in the presence of LDAO and DMSO (control). The addition of 50 μM bortezomib did not stabilise AtSQE1-Myc, with only 15 % remaining after 60 minutes of the LDAO treatment.

Applications of LDAO therefore appear to stimulate AtSQE1-Myc via an alternative mechanism to the constitutive UPS-mediated degradation observed during the CHX treatments. Further work is required to elucidate this mechanism but the stability of AtCNX1/2 proteins, which are also ER-membrane bound would suggest that this non-proteasomal degradation of AtSQE1-Myc remains selective.
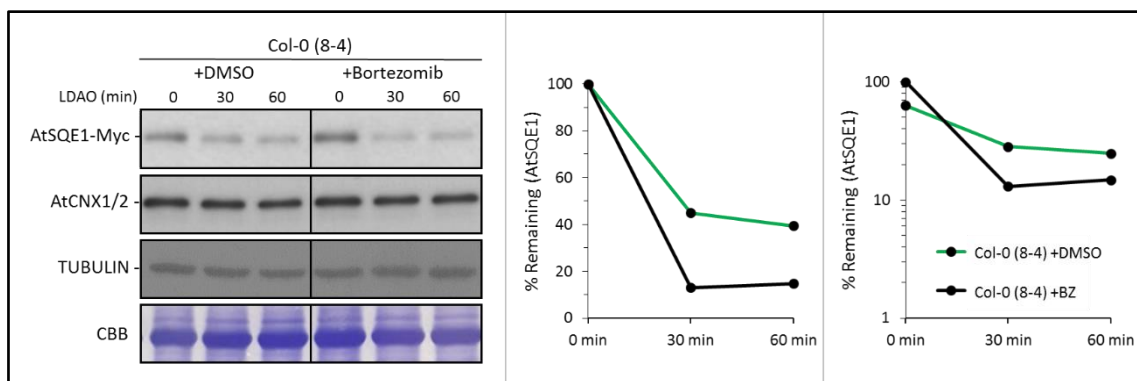
Figure 74 – The effect of bortezomib on LDAO-induced degradation of AtSQE1-Myc in Col-0
LDAO-induced degradation of AtSQE1-Myc was not prevented by the addition of 50 μM bortezomib (BZ), showing that it occurs via a non-proteasomal mechanism. The stability of AtCNX1/2 during the treatment confirmed that plasma membranes were not significantly disrupted. AtSQE1-Myc and AtCNX1/2 bands were detected at 60 kDa and TUBULIN bands were detected at 50 kDa, using anti-HA, anti-AtCNX1/2 and anti-β-Tubulin antibodies (see 2.4.1). Graphs show quantification of bands detected by Western blot, presented on both a linear and $\log_{10}$ axes. Band intensities were measured using ImageJ, normalised to the corresponding TUBULIN band strength and quantified relative to the 0-minute time point (for the linear scale) or the most intense starting band (for the $\log_{10}$ scale).

6.3.10 Squalene epoxidase genes are downregulated in *Atdoa10a/b* (RNAi line 4-2(2))

The stability of AtSQE1-Myc observed in both *Atdoa10a* and *Atdoa10a*/b (RNAi line 4-2(2)) following CHX-treatment demonstrated that AtDOA10A contributes to the basal turnover of AtSQE1. The increased stability of the protein might be expected to reduce the level of *AtSQE1* gene expression by negative feedback, to prevent its overaccumulation. However, AtSQE1 was not identified as a DEG during RNA-seq analysis of *Atdoa10a*/b (RNAi line 4-2(2)) (see 4.4.1). On the other hand, differential expression of the alternative *Arabidopsis* squalene epoxidase gene *AtSQE5* was detected (FC=0.64, q=1.74×10^{-10}). To further analyse the effect of AtSQE1 stabilisation on the family of squalene epoxidase enzymes in *Arabidopsis* qRT-PCR was performed to assess the level of expression of each of the six genes.

qRT-PCR was performed on the same RNA samples that were analysed by RNA-seq (see 4.4.1). As previously reported (Rasbery et al., 2007), a large amount of variation in the level of expression was observed between the *AtSQE* genes, with *AtSQE1*, *AtSQE3* and *AtSQE5* appearing to be responsible for more than 95 % of the *AtSQE* transcripts detected in Col-0 (Figure 75). This comparison of different genes within the same background relies on the assumption that the amplification efficiency of each primer set used is the same. This was not tested but the expression variation observed closely matched the eFP Browser 2.0 database (http://bar.utoronto.ca/efp2/Arabidopsis/Arabidopsis_eFPBrowser2.html) (Winter et al., 2007), supporting the reliability of the result.

The expression of each *AtSQE* gene was then measured in *Atdoa10a*/b (RNAi line 4-2(2)), in comparison to Col-0. Despite the observed stabilisation of AtSQE1-Myc at the protein level in this line (see 6.3.4), the level of *AtSQE1* gene expression in *Atdoa10a*/b (RNAi line 4-2(2)) was almost identical to Col-0 (Figure 76). However, a stronger reduction in expression of each of the other *AtSQE*s was identified in *Atdoa10a*/b (RNAi line 4-2(2)), although this was only statistically significant for *AtSQE3* (p=0.037). A similar observation was made by Pose et al. (2009), whereby a hypomorphic mutation of AtSQE1 led to a slight upregulation of *AtSQE3*

but not *AtSQE1*. This may indicate that *AtSQE1* is less responsive to overall cellular squalene epoxidase levels than other *AtSQE* genes, although it should be noted that the abundance of AtSQE2-6 proteins in *Atdoa10a*/b (RNAi line 4-2(2)) has not been studied and may also vary significantly, potentially influencing *AtSQE* gene expression. Unlike in the RNA-seq dataset, the downregulation of *AtSQE5* was not statistically significant (p=0.127).
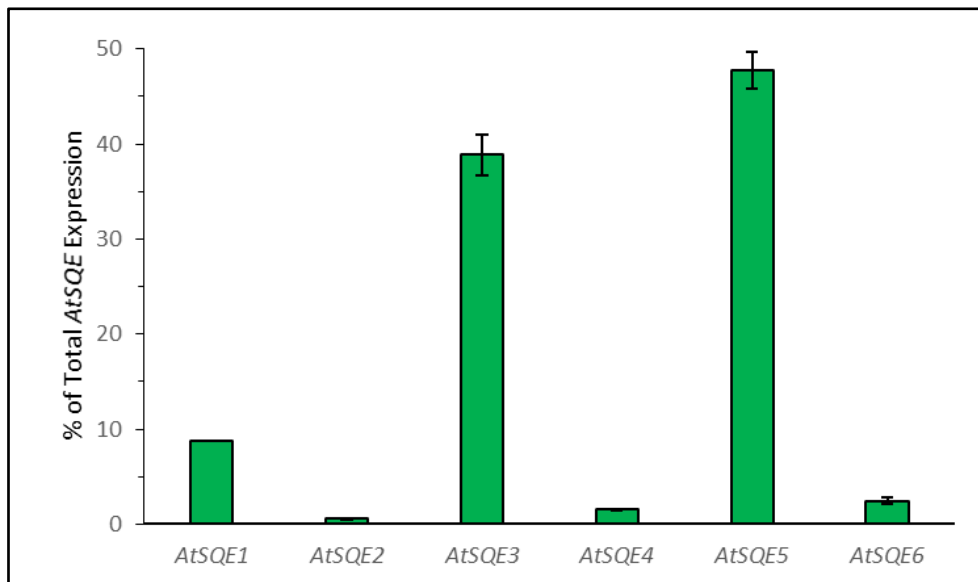
Figure 75 – Gene expression level of *AtSQE1-6* in WT *Arabidopsis* seedlings
Abundance of *AtSQE1-6* transcripts measured by qRT-PCR as a percentage of total *AtSQE* expression (primers 121+122, 131+132, 135+136, 139+140, 143+144, 147+148, 78+79). Relative expression in comparison to *AtSQE1* was calculated according to the delta delta C_t method and results are the average of three biological replicates (±SE) from RNA extracted from 10-day old Col-0 *Arabidopsis* seedlings.
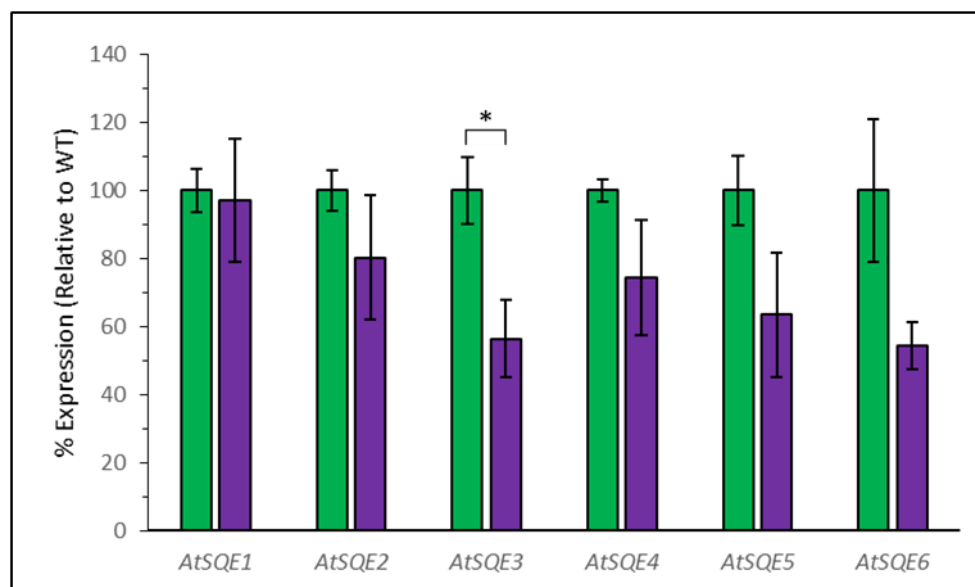
Figure 76 – *AtSQE1-6* gene expression changes in *Atdoa10a/b* (RNAi line 4-2(2))
Comparison of *AtSQE1-6* expression in Col-0 (green) and *Atdoa10a/b* (RNAi line 4-2(2)) (purple) by qRT-PCR (primers as Figure 75). Expression of *AtSQE1-6* in *Atdoa10a/b* (RNAi line 4-2(2)) relative to Col-0 calculated according to the delta delta C_t method. Results are the average of three biological replicates (±SE) from RNA extracted from 10-day old *Arabidopsis* seedlings (* = p<0.05, student's t-test).

6.4 Discussion

Squalene epoxidase is a conserved substrate of both ScDoa10 in yeast and HsMARCH6 in humans (Foresti et al., 2013; Zelcer et al., 2014). Combined with the observation that the loss of AtDOA10A rescues the pleiotropic mutant phenotypes of *Atsqe1* mutants (Doblas et al., 2013), it was hypothesised that AtSQE1 might be a ubiquitination target of AtDOA10A. However, the expansion of the squalene epoxidase family of enzymes in *Arabidopsis* (Rasbery et al., 2007) represented a significant divergence that may have extended to degradation mechanisms. Here, through analysis of AtSQE1 degradation when ectopically expressed in both yeast and *Arabidopsis*, the strongest evidence to date of AtDOA10A targeting AtSQE1 for degradation is reported. In yeast AtSQE1-HA was observed to accumulate in *ScDoa10Δ* cells, an effect that was eliminated through complementation with AtDOA10A (Figure 56 and 59). CHX chases confirmed that ScDoa10 or AtDOA10A is required for AtSQE1-HA degradation (Figure 57 and 60). CHX chases also revealed a slower rate of AtSQE1-Myc degradation in *Arabidopsis* lines lacking AtDOA10A (Figures 67 and 68), showing that the E3 contributes to AtSQE1 degradation *in planta*.

As a component of the ERAD system (see 1.2), AtDOA10A is thought to target misfolded proteins. In the present analysis, AtSQE1 was degraded by AtDOA10A despite presumably being in its native conformation. Distortion of AtSQE1's protein structure may produce even greater differences in stability between Col-0 and *Atdoa10a*. Indeed, the *Atsqe1* mutant phenotypes observed by Doblas et al. (2013) and Pose et al. (2009) may have been caused by rapid degradation of the misfolded protein, since the mutant studied contained only a single amino acid substitution in a region away from the enzyme's active site (G183R). This could also explain the suppression of these phenotypes through mutation of AtDOA10A (Doblas et al., 2013).

To definitively prove that AtSQE1 is a substrate of AtDOA10A, evidence of an interaction between the two proteins is required. Efforts were made to test protein-protein interactions via the mating-based Split-Ubiquitin System (Grefen et al., 2009) but no

interaction could be detected due to technical limitations of this assay (Appendix 7). Only a small amount of growth was observed for the positive control reaction suggesting that the absence of yeast growth during the interaction screen was more likely to be due to technical difficulties rather than the lack of an interaction. ScErg1 is generally accepted as a substrate of ScDoa10, but a physical interaction between these two proteins has also not been confirmed. An interaction has been shown between the HsSM and HsMARCH6 in human cells via a co-IP (Zelcer et al., 2014).

In addition to roles in ERAD, ScDoa10 and HsMARCH6 are reported to be E3 ligases of the Ac/N-degron pathway. Whilst appearing to be a substrate of the potential N-recognin AtDOA10A, AtSQE1 does not seem to undergo degradation via the Ac/N-degron pathway. Although AtSQE1 was stabilised in *Atnaa20* (Figure 67) and in *ScNaa20Δ* yeast (Figures 59 and 61), an organism with a characterised Ac/N-degron pathway (Hwang et al., 2010; Shemorry et al., 2013; Kim et al., 2014), this was independent of its Nt-acetylation. Mutation of AtSQE1's second amino acid to an unacetylatable proline residue did not yield increased stability in either *Arabidopsis* or yeast (Figures 63 and 70), showing that its degradation did not require acetylation of the N-terminus. It is possible that the N-terminal mutation may have opened up MP-AtSQE1-HA/-Myc to other degradative pathways (e.g. the Pro/N-degron pathway), with the alternative mode of degradation potentially masking the prevention of proteolysis via the Ac/N-degron pathway. However, the confirmation that the mutation has no effect on AtSQE1-Myc's localisation (Figure 66) and that the similar MK- mutant variant is also unstable in yeast (Figure 63), suggests that this is unlikely. This is further supported by the observation that MA-AtSQE1-HA was still stabilised in *ScNaa20Δ* despite being a substrate of NatA rather than NatB (Figure 64), and that the loss of ScNaa10 also stabilised AtSQE1-HA suggesting that its degradation may be influenced by global levels of Nt-acetylation rather than the Ac/N-degron pathway. The extent of Nt-acetylation of AtSQE1 is also an important consideration when contemplating its influence on AtSQE1 degradation. AtSQE1 is assumed to be broadly Nt-acetylated, because only 1% of ME- N-termini are not acetylated in humans

(Aksnes et al., 2016) and Nt-acetylation prevalence and machinery are fairly well conserved in humans and plants (Linster and Wirtz, 2018; Aksnes et al., 2019) but this is yet to be confirmed. The acetylation status of AtSQE1-Myc's N-terminus could be determined through isolation of the protein followed by MS, however, this may prove difficult due to the hydrophobic nature of AtSQE1's N-terminus (Giglione and Meinnel, 2020, personal communication).

AtSQE1 is one of six squalene epoxidase enzymes in *Arabidopsis* (see 6.1). It would be interesting to investigate whether AtDOA10A is able to target each member of this family, or perhaps only those which are functional homologues of ScErg1 (AtSQE1-3), thereby making it a key regulator or sterol production during all aspects of plant development. Natural variations in the N-termini of the AtSQEs, and therefore their levels of Nt-acetylation, could also provide insight into the role of the N-terminus in their degradation. Gene expression analysis of *AtSQE1-6* suggested that AtDOA10A may indeed reduce the stability of AtSQE2-6 proteins, in addition to AtSQE1, since each of these genes was downregulated in *Atdoa10a/b* (RNAi line 4-2(2)) (Figure 76). Alternatively, this may simply show that *AtSQE2-6* expression is more responsive to perturbations of the sterol synthesis pathway caused by stabilisation of AtSQE1 than the *AtSQE1* gene itself, perhaps due to expression of *AtSQE1* in a tissue of particular importance. AtSQE1 is reported to be the primary AtSQE responsible for squalene oxidation in the roots (Rasbery et al., 2007; Pose et al., 2009; Doblas et al., 2013). It is tempting to speculate that AtSQE1-3 may be targeted by AtDOA10A whilst the *Brassicaceae*-specific SQE-like proteins (AtSQE4-6) may be substrates of AtDOA10B, as it also appears to be specific to *Brassicaceae* (see 3.4). Having said that, AtSQE1's degradation rate was slower in *Atdoa10a*/b (RNAi line 4-2(2)) than in *Atdoa10a* (Figure 67 and 68), so an influence of AtDOA10B on the degradation of AtSQE1 cannot be ruled out, although the inability of AtDOA10B to prevent AtSQE1 accumulation in yeast suggests this is unlikely (see 6.2.3). Biological and technical variation between experiments is more plausible.

Degradation of SQEs is induced by the accumulation of downstream sterol products and intermediates in yeast and humans (Gill et al., 2011; Foresti et al., 2013). Here, treatment of *Arabidopsis* seedlings with the chemical LDAO was observed to induce rapid degradation of AtSQE1 (Figure 72). LDAO is an inhibitor of CAS and CPI enzymes of the sterol synthesis pathway (Darnet et al., 2020). Inhibition of these enzymes leads to a build-up of several sterol intermediates downstream of AtSQEs including their substrates 2,3-oxidosqualene and cycloeucalenol respectively (Darnet et al., 2020), suggesting that accumulation of one of these compounds may trigger AtSQE1 degradation. This is the first evidence of squalene epoxidase degradation in plants being regulated by negative feedback from a downstream product, as it is in other organisms. Degradation of squalene epoxidase in yeast and humans is thought to be primarily influenced by one particular downstream product (lanosterol in yeast and cholesterol in humans), as applications of several other sterols and intermediates have not induced degradation (Gill et al., 2011; Foresti et al., 2013). This may indicate that either 2,3-oxidosqualene or cycloeucalenol is the key compound for this negative feedback in plants, although, it remains possible that squalene epoxidase regulation in plants may have evolved to be responsive more than one downstream sterol.

Whilst both appear to be regulated by negative feedback, a difference between the sterol-induced degradation of squalene epoxidases in *Arabidopsis* and yeast is that AtSQE1 is degraded independently of the UPS, whereas ScErg1 degradation has been shown to require ScDOA10 (Foresti et al., 2013). In humans the depletion of HsSM caused by cholesterol treatment has been shown to be accompanied by an increase in HsMARCH6 stability, most likely through reduced autoubiquitination, but has not yet been proven to specifically require HsMARCH6 (Sharpe et al., 2019). Although the LDAO treatment induced non-proteasomal destruction of AtSQE1, the stability of the alternative ER-resident proteins AtCNX1/2 observed (Figures 72 and 74) confirmed that the degradation still occurs by a selective mechanism. In humans, cholesterol induces a conformational change in HsSM that is not induced by other sterols. This change in conformation occurs at the N-terminus of HsSM

and involves the ejection of a membrane-embedded amphipathic helix during high cholesterol conditions, that is thought to lead to its degradation (Howe et al., 2015; Chua et al., 2017). This structural change is also likely to be the reason why HsSM degradation is specifically induced by cholesterol, as whilst other sterols can cause HsMARCH6 accumulation, they do not affect the HsSM's structure in the same manner (Howe et al., 2015; Sharpe et al., 2019). Although this conformational change in HsSM occurs in an Nt-extension, that is not conserved in AtSQE1, a sterol-induced conformational change resulting in ejection from the ER membrane and cytosolic autophagy is a possible explanation for the non-proteasomal degradation observed. Such a mechanism of degradation would also be consistent with LDAO's surfactant properties causing a minor disruption to the ER-membrane; sufficient to permit the release of AtSQE1 into the cytosol but not strong enough to induce AtCNX1/2 depletion. This reasoning could also be applied to the apparent influence of global levels of Nt-acetylation on AtSQE1 stability. Since Nt-acetylation can increase the membrane-affinity of proteins (see 1.3.2.2), NAT mutations may lead to altered membrane properties, thereby potentially affecting AtSQE1's structure and stability.

In conclusion, AtDOA10A is highly likely to target AtSQE1 for degradation, based on stability analysis presented here in *Arabidopsis* as well as in other species (Foresti et al., 2013; Zelcer et al., 2014) and their reported genetic interactions (Doblas et al., 2013). AtDOA10B appeared unlikely to ubiquitinate AtSQE1, a further example of divergence of function between the two E3s, in addition to their ability to complement yeast hygromycin susceptibility, their differential gene regulation in response to ER-stress and their variations in mutant phenotype (see Chapters 3 and 5). It would be interesting to discover whether either AtDOA10A or AtDOA10B can ubiquitinate any other members of the squalene epoxidase family in *Arabidopsis* and also to investigate their relationship with other sterol synthesis enzymes to fully understand their regulation of the pathway. HsMARCH6 also ubiquitinates lanosterol 14α-demethylase and affects HsHMGR stability to control human sterol biosynthesis at multiple stages (Zelcer et al., 2014; Scott et al., 2020). Mutation of AtDOA10A has already

been shown to lead to a reduction in AtHMGR activity (Doblas et al., 2013) in addition to targeting AtSQE1 suggesting that this could also be the case in plants and revealing that sterol biosynthesis is complex, with feedback interventions at various stages.

Chapter VII –

General Discussion

7.1 Introduction

The degradation of misfolded proteins is an important cellular process, required to prevent the accumulation of harmful aggregates which, in humans, can lead to many diseases, particularly neurodegenerative disorders such as Parkinson's and Alzheimer's diseases (Ogen-Shtern et al., 2016). In plants, protein quality control systems are crucial for normal development and mitigating the effects of abiotic and biotic stresses (Howell, 2013; Chen et al., 2020). Understanding these systems therefore potentially enables the improvement of crop stress tolerance, hence minimising yield losses associated with environmental stresses.

The E3 ligase DOA10 is implicated in protein quality control on two fronts: as part of the ERAD system and as a potential N-recognin of the Ac/N-degron pathway (Hirsch et al., 2009; Hwang et al., 2010). The ERAD system is responsible for the extraction and degradation of misfolded ER-associated proteins via the UPS (Strasser, 2018). ScDoa10 is one of two ERAD E3s in yeast (the other being ScHrd1) and has been shown to be responsible for the ubiquitination of proteins containing cytosolic degrons (ERAD-C) as well as some with degrons within the lipid bilayer (ERAD-M) (Hirsch et al., 2009; Habeck et al., 2015; Wu and Rapoport, 2018). ScDoa10 was also identified as an E3 ligase capable of specifically recognising acetylated N-termini, as part of the Ac/N-degron pathway (Hwang et al., 2010). Like ERAD, the Ac/N-degron pathway is hypothesised to be involved in the degradation of misfolded proteins, with acetylated Nt-termini only being accessible when proteins are folded incorrectly or are not in complex with binding partners (Shemorry et al., 2013). HsMARCH6 has also been shown to be able to target proteins bearing acetylated N-termini in humans (Park et al., 2015).

This work aimed to validate the putative *Arabidopsis* DOA10s (previously identified by Liu et al. (2011)) as homologues of ScDoa10 and to determine whether their functions within the ERAD system and as Ac/N-recognins were conserved in plants. The study also focussed on establishing the significance of their contribution to ERAD and a potential plant Ac/N-degron pathway on overall plant health and development. Results were presented in Chapters 3 to 6 and are discussed below.

7.2 AtDOA10A and AtDOA10B as part of the *Arabidopsis* ERAD system

AtDOA10A and AtDOA10B were identified as *Arabidopsis* homologues of ScDoa10 through BLAST searches performed by Liu et al. (2011). Subsequent analysis highlighted a high degree of similarity between AtDOA10A and both ScDoa10 and HsMARCH6 in the RING-CH and TD domains as well as in the protein's overall transmembrane structure (Lu et al., 2012; Doblas et al., 2013). Here, analysis was expanded to include AtDOA10B and to assess the conservation of a C-terminal motif thought to influence substrate recognition (Zattas et al., 2016) as well as the full-length protein sequences. In agreement with previous studies, AtDOA10A was found to be well conserved with ScDoa10 both in terms of sequence and structure (see 3.2 – 3.3). Likewise, AtDOA10B exhibited a high degree of similarity to ScDoa10, however this conservation was weaker than for AtDOA10A in almost all cases.

AtDOA10A was shown to be ER-localised (see 3.6) and universally expressed (see 3.5) fitting the profile of an ERAD component. *Atdoa10a* mutants also exhibited enrichment of "protein folding" genes, that were all upregulated (see 5.3.1), and a differential response to TM, a protein misfolding agent (see 5.2.2). AtDOA10A was also shown to be able to complement ScDoa10 to facilitate the degradation of an ER-resident protein (AtSQE1-HA) (see 6.2). Combined, these results show that ERAD function is conserved in AtDOA10A. Similarly, complementation of *ScDoa10Δ* with AtDOA10A restored WT sensitivity to hygromycin; although whether this complementation depends on AtDOA10A specifically performing ERAD is unknown. Whilst AtDOA10B was also ER-localised and expressed in a universal manner, suggesting that it too functions within the ERAD system, it failed to complement ScDoa10 in the degradation of AtSQE1-HA or the hygromycin phenotype. *Atdoa10b* mutants also exhibited WT responses to all applied ERAD stresses, demonstrating a level of diversification from the functions of ScDoa10 and AtDOA10A.

A further example of diversification between the AtDOA10s was the differing gene expression changes in response to TM. Unexpectedly, *AtDOA10A* expression was observed

to decrease in response to TM treatment. Alternatively, AtDOA10B expression increased 5-fold (see 3.7), indicating that it likely remains part of the ERAD system despite its inability to complement ScDoa10. The unanticipated drop in *AtDOA10A* transcripts following TM treatment may show that its role in *Arabidopsis* is only in the basal turnover of ER proteins. *ScDoa10* is also not strongly induced by TM (Travers et al., 2000; Leber et al., 2004); the significant upregulation of *AtDOA10B* may therefore indicate diversification to become a protein specifically expressed to perform ERAD when misfolded proteins are accumulating

All analyses therefore supported the classification of AtDOA10A and AtDOA10B as homologues of the ERAD E3 ligase ScDoa10, albeit with some diversification of AtDOA10B. To assess their importance in *Arabidopsis* ERAD, the effect of a range of *Atdoa10* mutations on the magnitude of the UPR was assessed (see 3.8). Unexpectedly, the mutations had no significant impact on UPR gene expression, even following treatment with TM. This was surprising since ScDoa10 is one of only two ERAD E3 ligases known in yeast (Hirsch et al., 2009) and because *AtDOA10B* expression is strongly induced by TM. Similar experiments performed with plants lacking both *Arabidopsis* homologues of the other yeast ERAD E3 (ScHrd1) also revealed no significant differences from WT in the induction of UPR genes following a heat stress (Li et al., 2017). However, Li et al. (2017) did observe a significant UPR induction in *Atdoa10a Athrd1a Athrd1b* triple mutants. These triple mutants also displayed an enhanced thermotolerance suggesting that a degree of redundancy may exist between the ERAD E3s.

Functional redundancy between these E3s has also been observed in yeast, where a *ScDoa10Δ ScHrd1Δ* double mutant exhibited a significant increase in susceptibility to cadmium and a UPR induction, in comparison to the single mutants (Swanson et al., 2001). The existence of functional redundancy opposes the reported distinct roles of DOA10 (primarily in ERAD-C) and HRD1 (in ERAD-L and ERAD-M) and many substrates have been shown to be targeted specifically by one of these E3s and not the other (Vembar and Brodsky, 2008). However, a mutated form of the yeast membrane transporter ScSte6, which is stable

in *ScDoa10Δ* but not in *ScHrd1Δ* can be further stabilised in *ScDoa10Δ ScHrd1Δ* double mutants, potentially indicating some form of facultative redundancy (Huyer et al., 2004). The cytosolic Arg/N-recognin ScUbr1 has also been shown to contribute to the ERAD of mutated ScSte6 (Stolz et al., 2013), demonstrating that enzymes that are not ERAD-specific are able to contribute to the degradation of proteins at the ER. Further evidence of this is the dual targeting of the lipid droplet protein ScPgc1 by ScDoa10 and ScUbr1 (Ruggiano et al., 2016), the overlap of functions between ScHrd1 and the metalloprotease ScSte24 (Runnebohm et al., 2020), and that many ERAD substrates are not fully stabilised in *ScDoa10Δ ScHrd1Δ* double mutants. Dual targeting of HsMARCH6 substrates has also been observed in combination with HsTRC8, an E3 ligase without a yeast equivalent (Stefanovic-Barrett et al., 2018), demonstrating expansion of the ERAD system in higher eukaryotes.

The increased complexity of the ERAD system in plants compared to yeast expands the possibility of functional redundancy accounting for the lack of impact of *Atdoa10* mutations on the UPR. AtEBS7, AtPAWH1 and AtPAWH2 are plant-specific components of the HRD1 complex and the discovery of many more plant-specific ERAD components is expected (Liu et al., 2015; Lin et al., 2019). Paradoxically, mutation of these components, and those of other HRD1 subunits (Liu et al., 2011; Su et al., 2011; Hüttner et al., 2012) often appear to cause a more significant impact on stress tolerance and the UPR than mutations of AtHRD1A and AtHRD1B E3 ligases themselves (Li et al., 2017).

7.3 Diversification of AtDOA10B

The examples of functional redundancy between unrelated enzymes outlined above, made the apparent variation in function of AtDOA10A and AtDOA10B particularly surprising. In *Arabidopsis, Atdoa10a* mutants are drought tolerant (Lu et al., 2012), ABA hypersensitive (Zhao et al., 2014) and TM resistant (this study, see 5.2.2), but none of these phenotypes are shared by *Atdoa10b* mutants. Whilst this could be explained by the much lower expression level of *AtDOA10B* in *Arabidopsis* compared to *AtDOA10A* (see 3.5), this cannot account for AtDOA10B's failure to complement ScDoa10 or differences in the TM responses of the genes.

AtDOA10B appeared to be specific to *Brassicaceae* (see 3.4). Whilst the genomes of many plant species encode multiple homologues of ScDoa10, in all of the non-*Brassicaceae* species analysed, each homologue resembled AtDOA10A rather than AtDOA10B. Plants within the *Brassicaceae* family, however, contained proteins with sequences that matched AtDOA10B more closely than AtDOA10A, showing that the evolution of AtDOA10B likely occurred within this family. Having said that, even within *Brassicaceae*, AtDOA10B-like proteins are much less well conserved than those that are AtDOA10A-like (see 3.4).

With no phenotypes associated with *Atdoa10b* mutations, the role of this E3 in plant development remains elusive. Additionally, *Atdoa10ab* double mutants and RNAi lines were indistinguishable from *Atdoa10a* plants (see 5.2), indicating that redundant targeting of substrates is unlikely to be the reason for the lack of phenotypes of *Atdoa10b*. The amino acid sequence of AtDOA10B showed a higher degree of divergence to ScDoa10, than AtDOA10A did to the yeast E3. This was particularly true for the TD-domain and C-terminal motif, both of which are thought to be important for substrate recognition (Kreft and Hochstrasser, 2011; Zattas et al., 2016). Conversely, the RING-CH domains of AtDOA10A and AtDOA10B were equally similar to ScDoa10 (see 3.2). This conservation of the RING-CH domain suggests that the E3 ligase function of AtDOA10B is maintained, but variations in the TD and C-terminal domains may indicate a change in the set of substrates it targets. It could therefore be inferred that the *Atdoa10a* specific phenotypes and *ScDoa10Δ*'s hygromycin response may be the result of stabilisation of particular substrates that AtDOA10B is unable to target, as appears to be the case for AtSQE1 (see 6.2). It would be interesting to determine whether the SQE-like proteins (AtSQE4-6), that are also specific to *Brassicaceae*, are ERAD substrates of AtDOA10B and to analyse whether the evolution of the E3 ligase and these potential substrates coincided. These enzymes differ in structure from AtSQE1-3, including in their number of transmembrane domains (Laranjeira et al., 2015), potentially preventing recognition by AtDOA10A.

7.4 The role of AtDOA10A in plant development

Analysis suggested that AtDOA10 homologues were part of the *Arabidopsis* ERAD system, but that their overall contribution to this process was either small or functionally redundant with other enzymes. Unlike *Atdoa10b*, however, *Atdoa10a* exhibited some mutant phenotypes, such as its drought and ABA responses (Lu et al., 2012; Zhao et al., 2014), suggesting that AtDOA10A has a significant role in these processes. Through the use of complementation lines, which restored the WT phenotype, the ABA hypersensitivity of *Atdoa10a* was confirmed to be specifically caused by the absence of AtDOA10A (see 4.3 and 5.2.3). These phenotypes are likely due to the targeting of substrate(s) that are specific to drought and ABA responses, since plants grow as WT in standard conditions. AtSQE1 could be such a substrate (see 6.2 and 6.3) as it is has been previously linked to both drought and ABA. *Atsqe1* mutants are drought hypersensitive and insensitive to ABA (Pose et al., 2009) and, significantly, both of these phenotypes are suppressed by mutation of AtDOA10A (Doblas et al., 2013). However, the ABA hypersensitivity of *Atdoa10a* is thought to be caused by over-production of the hormone (Zhao et al., 2014). AtSQE1's involvement in ABA biosynthesis is indirect (via crosstalk between the MVA and MEP pathways) and therefore it is unlikely that its stabilisation in *Atdoa10a* would have such a strong impact on ABA sensitivity, especially since responses to brassinosteroids (EBR), hormones produced directly from sterols, were not significantly altered in *Atdoa10a* (see 5.2.3).

In addition to these phenotypes, given its perceived role in ERAD, the response of the *Atdoa10a* mutant to protein misfolding stresses was analysed. In general, *Atdoa10a* did not differ from WT in its response to these stresses, however, a subtle resistance to TM was observed (see 5.2.2), suggesting a minor role in stress responses. Mutation of plant ERAD components has been previously reported to lead to stress tolerance, in some cases, rather than susceptibility, as might be expected. *Atdoa10a Athrd1a Athrd1b* triple mutants were tolerant to a 45 °C heat shock (Li et al., 2017) and *Atubc32* mutants were resistant to TM and salt stress, whereas *AtUBC32* overexpressors displayed increased sensitivity to these

stresses (Cui et al., 2012). The salt stress tolerance of *Atubc32* mutants was shown to be consistent with misfolded AtBRI1 escaping ERAD and functioning as a brassinosteroid receptor to convey tolerance to salt stress (Cui et al., 2012). These reports and observations in this study perhaps indicate that the protein quality control system is naturally over-stringent.

Given AtDOA10A's influence on AtSQE1 and AtHMGR and the ABA-hypersensitivity of plants lacking the E3, the responses of *Atdoa10a* to sterol synthesis inhibitors and other hormones were also investigated but no differential effects relative to WT were observed (see 5.2.3 and 5.2.4).

7.5 Targeting of AtSQE1 by AtDOA10A

Despite not observing differences in sensitivity to sterol synthesis inhibitors, the enrichment of monooxygenases in the RNA-seq analysis of *Atdoa10a*/b (RNAi line 4-2(2)) (including the differential expression of AtSQE5, see 5.3.1) in addition to the known targeting of SQEs by DOA10s in yeast and humans (Foresti et al., 2013; Zelcer et al., 2014) prompted investigation of AtSQE1 as a candidate substrate of AtDOA10A. AtSQE1 was selected for analysis, ahead of the other AtSQE homologues because of the published genetic interaction between AtDOA10A and AtSQE1 (Doblas et al., 2013). When heterologously expressed in a yeast system, AtSQE1-HA was stabilised in *ScDoa10Δ* (see 6.2). This stabilisation was reduced when AtDOA10A-EGFP was co-expressed, implying that both ScDoa10 and AtDOA10A were able to ubiquitinate AtSQE1-HA in the yeast system, leading to its degradation. Similarly, in *Arabidopsis,* AtSQE1's degradation rate was reduced in lines lacking AtDOA10A (see 6.3), suggesting that AtDOA10A also targets AtSQE1 in *Arabidopsis*.

The degradation of AtSQE1 was also affected by mutations of the NAT machinery. Due to its ME- N-terminus AtSQE1 was assumed to be subject to Nt-acetylation because 99 % of proteins with ME- N-termini are acetylated in humans (Aksnes et al., 2016). In the yeast system, AtSQE1-HA was stabilised by the deletion of ScNaa20 or ScNaa10 (see 6.2). This was also the case in *Arabidopsis*, where AtSQE1-Myc seemed to be stabilised in *Atnaa20*

(see 6.3). In spite of this, AtSQE1's degradation was independent of its own Nt-acetylation, as proven by the consistent degradation of unacetylatable Nt-mutants of AtSQE1 in both yeast and *Arabidopsis*. This confirmed that AtSQE1 degradation did not occur via the Ac/N-degron pathway and that AtDOA10A was not functioning as an N-recognin in this case.

NAT mutations therefore affected AtSQE1's degradation in an indirect manner. Nt-acetylation is such a prevalent co- or post-translational modification impacting a broad range of processes that there are numerous potential explanations for this indirect influence. One such explanation is that changes in Nt-acetylation rates may affect the targeting of proteins to plasma membranes and hence alter membrane properties. HsSM (the human SQE homologue) is known to be ejected from the ER membrane and degraded in response to applications of membrane-stiffening cholesterol, which induce a conformational change in the monooxygenase (Gill et al., 2011; Howe et al., 2015; Chua et al., 2017). Altered membrane properties in NAT mutants may therefore also affect AtSQE1's membrane affinity and impact its degradation rate. Another plausible hypothesis is that components of the ERAD machinery may require Nt-acetylation in order to function. ScDer1, a component of the HRD1 complex in yeast, has been shown to be highly unstable when not Nt-acetylated by NatB (Zattas et al., 2013). This led to the stabilisation of a known ScHrd1 substrate in *ScNaa20Δ* and *ScNaa25Δ* KOs. Disruption of ERAD-machinery in *ScNaa10Δ, ScNaa20Δ*, and *Atnaa20* may also explain the stabilisation of AtSQE1 observed here.

It would be interesting to decipher whether *ScNaa10Δ* and *ScNaa20Δ* mutations stabilise AtSQE1-HA via the same mechanism, for example by independently destabilising ERAD components. Recent reports have outlined distinct physiological roles for NatA and NatB, thereby suggesting this may not be the case. In yeast, transcriptome and proteome analyses have revealed a low level of overlap in the changes induced by NatA and NatB mutations (Friedrich et al., 2021). The loss of NatA affected genome maintenance and levels of gene expression and silencing, whereas, alternatively, NatB mutations impacted the folding and aggregation of proteins. Distinct roles for NatA and NatB have also been observed in

plants. Nt-acetylation by NatA is involved in drought stress responses (Linster et al., 2015), whereas NatB is not involved in this process and *Atnaa20/25* mutations specifically impact salt, osmotic and reductive stresses (Huber et al., 2020; Huber et al., 2022). RNA-seq analysis of *amiAtNAA10* and *Atnaa20* mutants also showed that the majority of DEGs were specific to one of these genotypes rather than shared (see 5.3.1).

However AtNAA20 and AtDOA10A facilitate degradation of AtSQE1, this mechanism is not required for LDAO-induced degradation of the monooxygenase (see 6.3.8). Application of this downstream inhibitor caused rapid depletion of AtSQE1-Myc even in the absence of AtDOA10A, AtNAA20 or a functioning proteasome (bortezomib inhibition), indicating that alternative non-proteasomal mechanisms can be utilised to quickly change the abundance of AtSQE1. This may suggest that AtDOA10A and AtNAA20 are only involved in the basal turnover of AtSQE1 rather than active proteolysis in response to a stimulus.

7.6 DOA10s as Ac/N-recognins

Results here showed that both ScDoa10 and AtDOA10A were able to target AtSQE1 for degradation independently of the substrate's Nt-acetylation. This therefore demonstrated that these DOA10s were able to recognise AtSQE1, without the generation of an Ac/N-degron and hence they did not function as N-recognins in this case. During the discovery of the Ac/N-degron pathway ScDoa10 was reported to act as an N-recognin in the degradation of several Nt-acetylated substrates (Hwang et al., 2010). Both artificial substrates and overexpressed yeast proteins could be stabilised by deletion of ScDoa10, the appropriate NAT or mutation of the substrate's N-terminus to prevent Nt-acetylation (see 1.4.3). These contradictory results confirm that whilst DOA10s may function as Ac/N-recognins in the degradation of certain substrates, they can also recognise alternative degrons in other circumstances (see 1.5.2).

Other recent investigations have shed light on the extent to which ScDoa10 functions as an Ac/N-recognin. High throughput screening of the stabilities of artificial N-termini expressed in yeast performed by Kats et al. (2018), revealed that mutation of *ScDoa10* was

sufficient to stabilise 41 % of reporters that were unstable in WT yeast. This suggests that a large number of yeast proteins are potentially subject to ubiquitination by ScDoa10, since the reporters were based on 15 aa sequences corresponding to the N-termini of all known yeast proteins. Alternatively, the degradation of only 9 % of these unstable reporters appeared to be consistent with the Ac/N-degron pathway (e.g. stabilised in NAT mutants). It was also observed that the stability of individual reporters did not correlate well with known stabilities of the corresponding full-length yeast protein. This highlights the significance of internal degrons in determining stability and suggests that N-degrons may be regularly inaccessible. Overall, this further confirms that ScDoa10 can target substrates independently of the Ac/N-degron pathway and shows that functioning as an N-recognin may only be a small part of its overall role. This correlates with the degradation of AtSQE1 investigated here and also with the ubiquitination of the ERAD-M substrate ScSbh2, which is targeted by ScDoa10 based on a degron within an internal transmembrane domain, rather than at the N-terminus (Habeck et al., 2015). It is also in agreement with Nt-acetylomics performed here (see 5.3.3), where no accumulation of Nt-acetylated proteins was observed in *Atdoa10a*/b (RNAi line 4-2(2)).

The potential ScDoa10 substrates identified by Kats et al. (2018) (unstable reporters that were stabilised in *ScDoa10Δ*), appeared to be defined by their high hydrophobicity. The majority of these potential substrates were not stabilised in *ScNaa10Δ*, *ScNaa20Δ* or *ScNaa30Δ* mutants and therefore it appeared to be hydrophobicity that was the main determinant of their destruction via ScDoa10, rather than Nt-acetylation. The degradation rate of an artificial HsMARCH6 substrate has also been shown to increase with hydrophobicity (Stefanovic-Barrett et al., 2018). This hypothesis does not exclude the possibility of DOA10s acting as Ac/N-recognins in certain situations; acetylation can increase the hydrophobicity of N-termini and therefore may indeed cause the targeting of some substrates by DOA10s.

These findings led Kats et al. (2018) to propose that "Scdoa10 could be involved in clearing mislocalised secretory proteins that fail to insert into the ER". Such a hypothesis fits nicely with the degradation of AtSQE1 analysed here, with the hydrophobic nature of AtSQE1

facilitating its targeting by ScDoa10 or AtDOA10A, potentially following ejection from the ER due to changes in membrane properties.

7.7 The impact of Nt-acetylation on proteins

Nt-acetylation is reported to affect numerous features of cellular proteins including their localisation, interactions and folding (see 1.3.2). The effect of Nt-acetylation on stability is subject to conflicting reports, with the discovery of the Ac/N-degron pathway challenging the existing view that Nt-acetylation increased protein stability (Hershko et al., 1984; Hwang et al., 2010; Aksnes et al., 2019). Evidence here, in the form of the Nt-acetylome of AtDOA10 depleted plants, suggested that the Ac/N-degron pathway did not contribute significantly to global protein degradation in plants, or at least that AtDOA10A/B-mediated degradation of Nt-acetylated proteins was insignificant. This was backed up by the Nt-acetylation independent degradation of AtSQE1 via AtDOA10A.

Further large-scale analyses of yeast proteins has shown that the loss of NatA or NatB does not cause an increase in the abundance of their potential substrates (based on reported specificities) or global protein levels, although protein aggregation did increase, particularly in yeast lacking NatB (Friedrich et al., 2021). Similarly, further studies of artificial reporters in yeast, this time only varying in their two Nt-residues, revealed that potential NatA and NatB substrates were generally stable in WT yeast (Kats et al., 2018). This was in contrast to reporters bearing N-termini compatible with the Arg/N-degron pathway which were largely unstable. Mutations of NatA and NatB also failed to stabilise the potential Ac/N-degron pathway substrates any further. The unstructured nature of these reporters ruled out the possibility that shielding of Ac/N-degrons provided this stability. This observation correlated with the historical view of Nt-acetylation increasing protein half-life. In fact, Kats et al. (2018) identified two contexts in which Nt-acetylation can protect substrates from degradation. Acetylation of Nt-Glu was observed to block arginylation and acetylation of Nt-Met-Asn prevented Met cleavage, in both cases precluding degradation via the Arg/N-degron pathway. Similar protection from the Arg/N-degron pathway has been reported for several endogenous

Met-Φ-proteins (where Φ is a hydrophobic residue, see 1.4.4). Conversely, potential substrates of NatC were mostly unstable in WT yeast and stabilised by the *ScNaa30Δ* mutation, consistent with the Ac/N-degron pathway (Kats et al., 2018). However, when studying the 15 residue reporters, mutations of NatA, NatB and NatC only caused the stabilisation of a small proportion of their potential substrates that were unstable in WT. When considering all reporters (not just those that were unstable in WT), mutations of NatA and NatC did lead to enrichment of their potential substrates, suggesting that the degradation of certain proteins is dependent on Nt-acetylation. For NatB mutants though, around 75% of reporters that increased in abundance were not potential targets of NatB. This provides evidence that NAT mutations can cause indirect stabilisation of proteins and this seems particularly true for mutations of NatB, correlating with AtSQE1's stabilisation in *ScNaa20Δ* and *Atnaa20*.

Taken together, these findings suggest that Nt-acetylation can influence protein stability but that in the majority of cases, alternative degradation signals are more significant. It seems likely that the effect of Nt-acetylation on protein stability is highly dependent on context. Whilst in some cases, Nt-acetylation may increase hydrophobicity and drive targeting by E3 ligases like DOA10, in others acetylation may protect proteins from other degradative pathways or be required for correct localisation of a protein, with mislocalisation potentially leading to rapid degradation. Even in the case of reporters bearing the same degron, Deg1 (the N-terminal 67 amino acids of the canonical ScDoa10 substrate ScMATα2), different effects of Nt-acetylation on stability have been reported. The degradation of Deg1 reporters has been observed to be both dependent and independent of Nt-acetylation by ScNaa20 (Hwang et al., 2010; Kim et al., 2013; Zattas et al., 2013)*.* This discrepancy is presumably context dependent and caused by differences in the protein fusions and tags used in the expression of the peptide.

Overall if Nt-acetylation is the primary determinant of stability for some proteins, it appears that these proteins are very few and not representative of the proteome as a whole. The number of proteins for which this is the case remains to be determined. Proteins

previously identified as Ac/N-degron pathway substrates perhaps represent some members of this small subset (Hwang et al., 2010; Shemorry et al., 2013; Park et al., 2015; Nguyen et al., 2019).

7.8 Future perspectives

This doctoral work provides evidence that the DOA10 homologues AtDOA10A and AtDOA10B remain part of the ERAD system in *Arabidopsis* and that whilst AtDOA10B has partially diversified, probably to facilitate the recognition of plant-specific substrates, AtDOA10A retains the functions of ScDoa10. This was observed through the complementation of *ScDoa10Δ* with AtDOA10A to restore WT yeast responses to hygromycin and to accelerate the degradation of a substrate (AtSQE1). Despite this conservation of function, no evidence of AtDOA10A functioning as an Ac/N-recognin was observed. This was the case both at a global level, where no increase in Nt-acetylated proteins was observed in the absence of AtDOA10A, and for analysis of an individual AtDOA10A substrate, where preventing the Nt-acetylation of AtSQE1 did not halt its degradation.

The study raises several questions that could be the subject of future work and could further clarify the roles of plant DOA10s. Primarily, further replicates of AtSQE1-Myc CHX-chases in *Arabidopsis* are required to fully confirm the reported differences in stability. This is particularly true for comparisons of AtSQE1-Myc's stability in WT vs *Atnaa20* and for assessment of the degradation of the ME- and MP- N-terminal variants, for which technical issues prevented the detection of strong bands during Western blot analysis. Additionally, to fully confirm the targeting of AtSQE1 by AtDOA10A, detection of an interaction between the two proteins is required; due to difficulties with the yeast mating-based Split-Ubiquitin System, this could not be shown here (Appendix 7).

It would also be interesting to further investigate the roles of AtDOA10s in regulating sterol biosynthesis, potentially providing a physiological relevance to the targeting of AtSQE1 by AtDOA10A. AtSQE1 was shown to be rapidly depleted in response to downstream

inhibition of the pathway by LDAO, but this was independent of AtDOA10A. The impact of other factors such as hormones, environmental stresses or alternative inhibitors on AtSQE1's stability and thereby the sterol synthesis pathway could easily be investigated using the tools generated in this study to determine whether AtDOA10A controls AtSQE1 abundance in response to these alternative stimuli. The *Arabidopsis* AtSQE1-Myc expressing lines also provide a potentially useful way of investigating whether AtDOA10A's targeting of AtSQE1 is responsible for the drought resistance and ABA hypersensitivity of *Atdoa10a* mutants. Future work could also consider whether other members of the AtSQE family of proteins are also substrates of AtDOA10s. Particulaly, establishing whether AtSQE4-6 are ubiquitinated by AtDOA10B is of interest due to the *Brassicaceae*-specific nature of these enzymes.

In addition to AtSQE1, the data within the transcriptome, proteome and Nt-acetylome provide a useful resource for the identification of further candidate substrates of AtDOA10s. For example, this multiomic approach, identified AtCAT3 as a candidate that was both stabilised at the protein level and downregulated at the transcript level in *Atdoa10a*/b (RNAi line 4-2(2)). Another potential candidate, AtLOX2, was the protein most strongly stabilised in *Atdoa10a*/b (RNAi line 4-2(2)). Whilst no change in gene expression was observed and the protein was not detected during Nt-acetylomics, *AtLOX2* has been reported to be significantly downregulated in ami*AtNAA10* (Linster et al., 2015), as were the alternative lipoxygenases *AtLOX3* and *AtLOX4*. *AtLOX3* and *AtLOX4* transcripts were also significantly reduced in *Atnaa20* (Huber et al., 2020), indicating that, like AtSQE1, the expression and stability of some AtLOX proteins is influenced by *Arabidopsis* DOA10s and NATs, although this does not always conform with NAT specificities. Interestingly AtLOXs share some similarity with AtSQEs in that they exist as a six-member family and they catalyse the oxygenation of long unsaturated hydrocarbons. Analysis of AtLOX's could provide a convenient method of determining whether AtDOA10s potentially target multiple enzymes with oxygenation activity and whether Nt-acetylation affects the degradation of other proteins in the same way as AtSQE1. Publicly available plant lipoxygenase antibodies also provide a useful tool to investigate this.

The identification of AtSQE1 as the first substrate of AtDOA10A in plants provides insight into the role of Nt-acetylation in protein degradation at a substrate-specific level and shows that AtDOA10A does not solely recognise acetylated N-termini. On a broader scale, the lack of impact of AtDOA10 depletion on global proteostasis suggests that degradation via DOA10-mediated Ac/N-degron pathway may only occur for a small subset of proteins in plants. The identification of further AtDOA10 substrates will provide even greater clarity into the role of AtDOA10s in ERAD and help to determine whether their targeting requires Nt-acetylation under any circumstances.

Chapter VIII -

References

8.1 List of references

Abas, L. and C. Luschnig (2010). Maximum yields of microsomal-type membranes from small amounts of plant material without requiring ultracentrifugation. **Analytical Biochemistry** 401(2): 217-227.

Ahn, M. Y., T. R. Oh, D. H. Seo, J. H. Kim, N. H. Cho and W. T. Kim (2018). *Arabidopsis* group XIV ubiquitin-conjugating enzymes AtUBC32, AtUBC33, and AtUBC34 play negative roles in drought stress response. **Journal of Plant Physiology** 230: 73-79.

Aksnes, H., A. Drazic, M. Marie and T. Arnesen (2016). First Things First: Vital Protein Marks by N-Terminal Acetyltransferases. **Trends in biochemical sciences** 41(9): 746-760.

Aksnes, H., M. Goris, Ø. Strømland, A. Drazic, Q. Waheed, N. Reuter and T. Arnesen (2017). Molecular determinants of the N-terminal acetyltransferase Naa60 anchoring to the Golgi membrane. **Journal of Biological Chemistry** 292(16): 6821-6837.

Aksnes, H., K. Hole and T. Arnesen (2015). Molecular, Cellular, and Physiological Significance of N-Terminal Acetylation. **International Review of Cell and Molecular Biology** 316: 267-305.

Aksnes, H., R. Ree and T. Arnesen (2019). Co-translational, Post-translational, and Non-catalytic Roles of N-Terminal Acetyltransferases. **Molecular cell** 73(6): 1097-1114.

Aksnes, H., P. Van Damme, M. Goris, K. K. Starheim, M. Marie, S. I. Støve, C. Hoel, T. V. Kalvik, K. Hole, N. Glomnes, C. Furnes, S. Ljostveit, M. Ziegler, M. Niere, K. Gevaert and T. Arnesen (2015). An organellar nα-acetyltransferase, naa60, acetylates cytosolic N termini of transmembrane proteins and maintains Golgi integrity. **Cell Rep** 10(8): 1362-1374.

Akutsu, M., I. Dikic and A. Bremm (2016). Ubiquitin chain diversity at a glance. **Journal of Cell Science** 129(5): 875-880.

Aranda, P. S., D. M. LaJoie and C. L. Jorcyk (2012). Bleach gel: A simple agarose gel for analyzing RNA quality. **Electrophoresis** 33(2): 366-369.

Ardley, H. C. and P. A. Robinson (2005). E3 ubiquitin ligases. **Essays in Biochemistry** 41: 15-30.

Armbruster, L., E. Linster, J.-B. Boyer, A. Brünje, J. Eirich, I. Stephan, W. V. Bienvenut, J. Weidenhausen, T. Meinnel, R. Hell, I. Sinning, I. Finkemeier, C. Giglione and M. Wirtz (2020). NAA50 Is an Enzymatically Active N α-Acetyltransferase That Is Crucial for Development and Regulation of Stress Responses **Plant Physiology** 183(4): 1502-1516.

Arnesen, T., P. Van Damme, B. Polevoda, K. Helsens, R. Evjenth, N. Colaert, J. E. Varhaug, J. Vandekerckhove, J. R. Lillehaug, F. Sherman and K. Gevaert (2009). Proteomics analyses reveal the evolutionary conservation and divergence of N-terminal acetyltransferases from yeast and humans. **Proceedings of the National Academy of Sciences of the United States of America** 106(20): 8157-8162.

Bachmair, A., D. Finley and A. Varshavsky (1986). In Vivo Half-Life of a Protein is a Function of its Amino-Terminal Residue. **Science** 234(4773): 179-186.

Bartels, T., N. C. Kim, E. S. Luth and D. J. Selkoe (2014). N-Alpha-Acetylation of alpha-Synuclein Increases Its Helical Folding Propensity, GM1 Binding Specificity and Resistance to Aggregation. **Plos One** 9(7): e103727.

Bell, E., R. A. Creelman and J. E. Mullet (1995). A Chloroplast Lipoxygenase is Required for Wound-Induced Jasmonic Acid Accumulation in *Arabidopsis*. **Proceedings of the National Academy of Sciences of the United States of America** 92(19): 8675-8679.

Benjamini, Y. and Y. Hochberg (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. **Journal of the Royal Statistical Society: Series B (Methodological)** 57(1): 289-300.

Berndsen, C. E. and C. Wolberger (2014). New insights into ubiquitin E3 ligase mechanism. **Nature Structural & Molecular Biology** 21(4): 301-307.

Bersuker, K. and J. A. Olzmann (2017). Establishing the lipid droplet proteome: Mechanisms of lipid droplet protein targeting and degradation. **Biochimica et Biophysica Acta-Molecular and Cell Biology of Lipids** 1862(10): 1166-1177.

Bienvenut, W. V., A. Brünje, J.-B. Boyer, J. S. Mühlenbeck, G. Bernal, I. Lassowskat, C. Dian, E. Linster, T. V. Dinh, M. M. Koskela, V. Jung, J. Seidel, L. K. Schyrba, A. Ivanauskaite, J. Eirich, R. Hell, D. Schwarzer, P. Mulo, M. Wirtz, T. Meinnel, C. Giglione and I. Finkemeier (2020). Dual lysine and N-terminal acetyltransferases reveal the complexity underpinning protein acetylation. **Molecular Systems Biology** 16(7): e9464.

Bienvenut, W. V., J. P. Scarpelli, J. Dumestier, T. Meinnel and C. Giglione (2017). EnCOUNTer: a parsing tool to uncover the mature N-terminus of organelle-targeted proteins in complex samples. **BMC bioinformatics** 18(1): 182

Bienvenut, W. V., D. Sumpton, A. Martinez, S. Lilla, C. Espagne, T. Meinnel and C. Giglione (2012). Comparative Large Scale Characterization of Plant versus Mammal Proteins Reveals Similar and Idiosyncratic N-alpha-Acetylation Features. **Molecular & Cellular Proteomics** 11(6): M111.015131

Bockhorn, J., B. Balar, D. He, E. Seitomer, P. R. Copeland and T. G. Kinzy (2008). Genome-wide screen of Saccharomyces cerevisiae null allele strains identifies genes involved in selenomethionine resistance. **Proceedings of the National Academy of Sciences of the United States of America** 105(46): 17682-17687.

Bortesi, L. and R. Fischer (2015). The CRISPR/Cas9 system for plant genome editing and beyond. **Biotechnology Advances** 33(1): 41-52.

Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. **Analytical Biochemistry** 72(1): 248-254.

Brown, M. S. and J. L. Goldstein (1980). Multivalent Feedback-Regulation of Hmg Coa Reductase, a Control Mechanism Coordinating Isoprenoid Synthesis and Cell-Growth. **Journal of lipid research** 21(5): 505-517.

Buchanan, B. W., M. E. Lloyd, S. M. Engle and E. M. Rubenstein (2016). Cycloheximide Chase Analysis of Protein Degradation in Saccharomyces cerevisiae. **Journal of visualized experiments : JoVE**(110): 53975.

Callis, J. (2014). The ubiquitination machinery of the ubiquitin system. **The *Arabidopsis* Book** 12: e0174.

Chang, Y.-F., J. S. Imam and M. E. Wilkinson (2007). The nonsense-mediated decay RNA surveillance pathway. **Annual Review of Biochemistry** 76: 51-74.

Chatr-Aryamontri, A., A. van der Sloot and M. Tyers (2018). At Long Last, a C-Terminal Bookend for the Ubiquitin Code. **Molecular cell** 70(4): 568-571.

Chen, L. and H. Hellmann (2013). Plant E3 Ligases: Flexible Enzymes in a Sessile World. **Molecular Plant** 6(5): 1388-1404.

Chen, Q., R. Liu, Q. Wang and Q. Xie (2017). ERAD Tuning of the HRD1 Complex Component AtOS9 Is Modulated by an ER-Bound E2, UBC32. **Molecular Plant** 10(6): 891-894.

Chen, Q., F. Yu and Q. Xie (2020). Insights into endoplasmic reticulum-associated degradation in plants. **New Phytologist** 226(2): 345-350.

Chen, Q., Y. Zhong, Y. Wu, L. Liu, P. Wang, R. Liu, F. Cui, Q. Li, X. Yang, S. Fang and Q. Xie (2016). HRD1-mediated ERAD tuning of ER-bound E2 is conserved between plants and mammals. **Nature Plants** 2(7): 16094-16094.

Chen, S.-J., X. Wu, B. Wadas, J.-H. Oh and A. Varshavsky (2017). An N-end rule pathway that recognizes proline and destroys gluconeogenic enzymes. **Science** 355(6323): eaal3655.

Chen, Y. and F. Brandizzi (2013). Analysis of Unfolded Protein Response in *Arabidopsis*. **Methods in Molecular Biology** 1043: 73-80.

Cheng, M.-C., E.-J. Hsieh, J.-H. Chen, H.-Y. Chen and T.-P. Lin (2012). *Arabidopsis* RGLG2, Functioning as a RING E3 Ligase, Interacts with AtERF53 and Negatively Regulates the Plant Drought Stress Response. **Plant Physiology** 158(1): 363-375.

Choi, C. M., W. M. Gray, S. Mooney and H. Hellmann (2014). Composition, roles, and regulation of cullin-based ubiquitin e3 ligases. **The *Arabidopsis* Book** 12: e0175.

Chua, N. K., V. Howe, N. Jatana, L. Thukral and A. J. Brown (2017). A conserved degron containing an amphipathic helix regulates the cholesterol-mediated turnover of human squalene monooxygenase, a rate-limiting enzyme in cholesterol synthesis. **Journal of Biological Chemistry** 292(49): 19959-19973.

Ciechanover, A. (1994). The ubiquitin-proteasome proteolytic pathway. **Cell** 79(1): 13-21.

Clague, M. J. and S. Urbe (2010). Ubiquitin: Same Molecule, Different Degradation Pathways. **Cell** 143(5): 682-685.

Clough, S. J. and A. F. Bent (1998). Floral dip: a simplified method for Agrobacterium-mediated transformation of *Arabidopsis thaliana*. **Plant Journal** 16(6): 735-743.

Clouse, S. D. (2011). Brassinosteroids. **The *Arabidopsis* Book** 9: e0151.

Clouse, S. D. and J. M. Sasse (1998). Brassinosteroids: Essential regulators of plant growth and development. **Annual Review of Plant Physiology and Plant Molecular Biology** 49: 427-451.

Coleman, D., A. Kawamura, M. Ikeuchi, D. S. Favero, A. Lambolez, B. Rymen, A. Iwase, T. Suzuki and K. Sugimoto (2020). The SUMO E3 Ligase SIZ1 Negatively Regulates Shoot Regeneration. **Plant Physiology** 184(1): 330-344.

Conte, M., C. Franceschi, M. Sandri and S. Salvioli (2016). Perilipin 2 and Age-Related Metabolic Diseases: A New Perspective. **Trends in Endocrinology and Metabolism** 27(12): 893-903.

Copeland, C. (2020). Starting Off Right: N-Terminal Acetylation Stabilizes an Immune-Activating Protein. **Plant Physiology** 183(1): 35-36.

Crowder, J. J., M. Geigges, R. T. Gibson, E. S. Fults, B. W. Buchanan, N. Sachs, A. Schink, S. G. Kreft and E. M. Rubenstein (2015). Rkr1/Ltn1 Ubiquitin Ligase-mediated Degradation of Translationally Stalled Endoplasmic Reticulum Proteins. **Journal of Biological Chemistry** 290(30): 18454-18466.

Cuesta-Marban, A., J. Botet, O. Czyz, L. M. Cacharro, C. Gajate, V. Hornillos, J. Delgado, H. Zhang, F. Amat-Guerri, A. Ulises Acuna, C. R. McMaster, J. Luis Revuelta, V. Zaremberg and F. Mollinedo (2013). Drug Uptake, Lipid Rafts, and Vesicle Trafficking Modulate Resistance to an Anticancer Lysophosphatidylcholine Analogue in Yeast. **Journal of Biological Chemistry** 288(12): 8405-8418.

Cui, F., L. Liu, Q. Li, C. Yang and Q. Xie (2012b). UBC32 Mediated Oxidative Tolerance in *Arabidopsis*. **Journal of Genetics and Genomics** 39(8): 415-417.

Cui, F., L. Liu, Q. Zhao, Z. Zhang, Q. Li, B. Lin, Y. Wu, S. Tang and Q. Xie (2012a). *Arabidopsis* Ubiquitin Conjugase UBC32 Is an ERAD Component That Functions in Brassinosteroid-Mediated Salt Stress Tolerance. **Plant Cell** 24(1): 233-244.

Darnet, S., L. B. B. Martin, P. Mercier, F. Bracher, P. Geoffroy and H. Schaller (2020). Inhibition of Phytosterol Biosynthesis by Azasterols. **Molecules** 25(5): 1111-1111.

de Bie, P. and A. Ciechanover (2011). Ubiquitination of E3 ligases: self-regulation of the ubiquitin system via proteolytic and non-proteolytic mechanisms. **Cell death and differentiation** 18(9): 1393-1402.

De Vriese, K., J. Pollier, A. Goossens, T. Beeckman and S. Vanneste (2021). Dissecting cholesterol and phytosterol biosynthesis via mutants and inhibitors. **Journal of experimental botany** 72(2): 241-253.

Deng, F., T. Guo, M. Lefebvre, S. Scaglione, C. J. Antico, T. Jing, X. Yang, W. Shan and K. M. Ramonell (2017). Expression and regulation of ATL9, an E3 ubiquitin ligase involved in plant defense. **Plos One** 12(11): e0188458.

Deng, M. and M. Hochstrasser (2006). Spatially regulated ubiquitin ligation by an ER/nuclear membrane ligase. **Nature** 443(7113): 827-831.

Deng, S., R. S. Magin, X. Wei, B. Pan, E. J. Petersson and R. Marmorstein (2019). Structure and Mechanism of Acetylation by the N-Terminal Dual Enzyme NatA/Naa50 Complex. **Structure** 27(7): 1057-1070.e1054.

Deshaies, R. J. and C. A. P. Joazeiro (2009). RING Domain E3 Ubiquitin Ligases. **Annual Review of Biochemistry** 78: 399-434.

Dharmasiri, N., S. Dharmasiri and M. Estelle (2005). The F-box protein TIR1 is an auxin receptor. **Nature** 435(7041): 441-445.

Dikiy, I. and D. Eliezer (2014). N-terminal Acetylation Stabilizes N-terminal Helicity in Lipid- and Micelle-bound alpha-Synuclein and Increases Its Affinity for Physiological Membranes. **Journal of Biological Chemistry** 289(6): 3652-3665.

Dinh, T. V., W. V. Bienvenut, E. Linster, A. Feldman-Salit, V. A. Jung, T. Meinnel, R. Hell, C. Giglione and M. Wirtz (2015). Molecular identification and functional characterization of the first N-acetyltransferase in plastids by global acetylome profiling. **Proteomics** 15(14): 2426-2435.

Doblas, V. G., V. Amorim-Silva, D. Pose, A. Rosado, A. Esteban, M. Arro, H. Azevedo, A. Bombarely, O. Borsani, V. Valpuesta, A. Ferrer, R. M. Tavares and M. A. Botella (2013). The SUD1 Gene Encodes a Putative E3 Ubiquitin Ligase and Is a Positive Regulator of 3-Hydroxy-3-Methylglutaryl Coenzyme A Reductase Activity in *Arabidopsis*. **Plant Cell** 25(2): 728-743.

Dougan, D. A., D. Micevski and K. N. Truscott (2012). The N-end rule pathway: From recognition by N-recognins, to destruction by AAA+ proteases. **Biochimica et Biophysica Acta-Molecular Cell Research** 1823(1): 83-91.

Dougan, D. A. and A. Varshavsky (2018). Understanding the Pro/N-end rule pathway. **Nature Chemical Biology** 14(5): 415-416.

Drazic, A., H. Aksnes, M. Marie, M. Boczkowska, S. Varland, E. Timmerman, H. Foyn, N. Glomnes, G. Rebowski, F. Impens, K. Gevaert, R. Dominguez and T. Arnesen (2018). NAA80 is actin's N-terminal acetyltransferase and regulates cytoskeleton assembly and cell motility. **Proceedings of the National Academy of Sciences of the United States of America** 115(17): 4399-4404.

Earley, K. W., J. R. Haag, O. Pontes, K. Opper, T. Juehne, K. M. Song and C. S. Pikaard (2006). Gateway-compatible vectors for plant functional genomics and proteomics. **Plant Journal** 45(4): 616-629.

Eldeeb, M. A., R. P. Fahlman, M. A. Ragheb and M. Esmaili (2019). Does N-Terminal Protein Acetylation Lead to Protein Degradation? **Bioessays** 41(11): 1800167-1800167.

Farboud, B. and B. J. Meyer (2015). Dramatic Enhancement of Genome Editing by CRISPR/Cas9 Through Improved Guide RNA Design. **Genetics** 199(4): 959-971.

Fauser, F., S. Schiml and H. Puchta (2014). Both CRISPR/Cas-based nucleases and nickases can be used efficiently for genome engineering in *Arabidopsis thaliana*. **Plant Journal** 79(2): 348-359.

Feng, H., S. Wang, D. Dong, R. Zhou and H. Wang (2020). *Arabidopsis* Ubiquitin-Conjugating Enzymes UBC7, UBC13, and UBC14 Are Required in Plant Responses to Multiple Stress Conditions. **Plants** 9(6): e723.

Feng, J., J. Hu, Y. Li, R. Li, H. Yu and L. Ma (2020). The N-Terminal Acetyltransferase Naa50 Regulates *Arabidopsis* Growth and Osmotic Stress Response. **Plant and Cell Physiology** 61(9): 1565-1575.

Feng, J., R. Li, J. Yu, S. Ma, C. Wu, Y. Li, Y. Cao and L. Ma (2016). Protein N-terminal acetylation is required for embryogenesis in *Arabidopsis*. **Journal of Experimental Botany** 67(15): 4779-4789.

Ferrandez-Ayela, A., R. Micol-Ponce, A. Belen Sanchez-Garcia, M. Magdalena Alonso-Peral, J. L. Micol and M. R. Ponce (2013). Mutation of an *Arabidopsis* NatB N-Alpha-Terminal Acetylation Complex Component Causes Pleiotropic Developmental Defects. **Plos One** 8(11): e80697.

Finley, D., H. D. Ulrich, T. Sommer and P. Kaiser (2012). The Ubiquitin-Proteasome System of Saccharomyces cerevisiae. **Genetics** 192(2): 319-360.

Fire, A., S. Q. Xu, M. K. Montgomery, S. A. Kostas, S. E. Driver and C. C. Mello (1998). Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. **Nature** 391(6669): 806-811.

Foresti, O., A. Ruggiano, H. K. Hannibal-Bach, C. S. Ejsing and P. Carvalho (2013). Sterol homeostasis requires regulated degradation of squalene monooxygenase by the ubiquitin ligase Doa10/Teb4. **eLife** 2: e00953.

Forte, G. M. A., M. R. Pool and C. J. Stirling (2011). N-Terminal Acetylation Inhibits Protein Targeting to the Endoplasmic Reticulum. **Plos Biology** 9(5): e1001073.

Fowden, L. and M. H. Richmond (1963). Replacement of Proline by Azetidine-2-Carbolic Acid during Biosynthesis of Protein. **Biochimica et Biophysica Acta** 71(2): 459-461.

Freemont, P. S., I. M. Hanson and J. Trowsdale (1991). A Novel Cysteine-Rich Sequence Motif. **Cell** 64(3): 483-484.

Friedland, A. E., R. Baral, P. Singhal, K. Loveluck, S. Shen, M. Sanchez, E. Marco, G. M. Gotta, M. L. Maeder, E. M. Kennedy, A. V. R. Kornepati, A. Sousa, M. A. Collins, H. Jayaram, B. R. Cullen and D. Bumcrot (2015). Characterization of Staphylococcus aureus Cas9: a smaller Cas9 for all-in-one adeno-associated virus delivery and paired nickase applications. **Genome Biology** 16: 257.

Friedrich, U. A., M. Zedan, B. Hessling, K. Fenzl, L. Gillet, J. Barry, M. Knop, G. Kramer and B. Bukau (2021). N(α)-terminal acetylation of proteins by NatA and NatB serves distinct physiological roles in Saccharomyces cerevisiae. **Cell reports** 34(5): 108711.

Fujioka, S. and T. Yokota (2003). Biosynthesis and metabolism of brassinosteroids. **Annual Review of Plant Biology** 54: 137-164.

Garzon, M., K. Eifler, A. Faust, H. Scheel, K. Hofmann, C. Koncz, A. Yephremov and A. Bachmair (2007). PRT6/At5g02310 encodes an *Arabidopsis* ubiquitin ligase of the N-end rule pathway with arginine specificity and is not the CER3 locus. **FEBS letters** 581(17): 3189-3196.

Gasch, A. P., P. T. Spellman, C. M. Kao, O. Carmel-Harel, M. B. Eisen, G. Storz, D. Botstein and P. O. Brown (2000). Genomic expression programs in the response of yeast cells to environmental changes. **Molecular Biology of the Cell** 11(12): 4241-4257.

Gibbs, D. J. (2015). Emerging Functions for N-Terminal Protein Acetylation in Plants. **Trends in Plant Science** 20(10): 599-601.

Gibbs, D. J., J. Bacardit, A. Bachmair and M. J. Holdsworth (2014a). The eukaryotic N-end rule pathway: conserved mechanisms and diverse functions. **Trends in Cell Biology** 24(10): 603-611.

Gibbs, D. J., M. Bailey, H. M. Tedds and M. J. Holdsworth (2016). From start to finish: amino-terminal protein modifications as degradation signals in plants. **New Phytologist** 211(4): 1188-1194.

Gibbs, D. J., N. M. Isa, M. Movahedi, J. Lozano-Juste, G. M. Mendiondo, S. Berckhan, N. Marin-de la Rosa, J. V. Conde, C. S. Correia, S. P. Pearce, G. W. Bassel, B. Hamali, P. Talloji, D. F. A. Tome, A. Coego, J. Beynon, D. Alabadi, A. Bachmair, J. Leon, J. E. Gray, F. L. Theodoulou and M. J. Holdsworth (2014b). Nitric Oxide Sensing in Plants Is Mediated by Proteolytic Control of Group VII ERF Transcription Factors. **Molecular cell** 53(3): 369-379.

Gibbs, D. J., S. C. Lee, N. M. Isa, S. Gramuglia, T. Fukao, G. W. Bassel, C. S. Correia, F. Corbineau, F. L. Theodoulou, J. Bailey-Serres and M. J. Holdsworth (2011). Homeostatic response to hypoxia is regulated by the N-end rule pathway in plants. **Nature** 479(7373): 415-418.

Gibbs, D. J., H. M. Tedds, A.-M. Labandera, M. Bailey, M. D. White, S. Hartman, C. Sprigg, S. L. Mogg, R. Osborne, C. Dambire, T. Boeckx, Z. Paling, L. A. C. J. Voesenek, E. Flashman and M. J. Holdsworth (2018). Oxygen-dependent proteolysis regulates the stability of angiosperm polycomb repressive complex 2 subunit VERNALIZATION 2. **Nature Communications** 9: 5438-5438.

Giglione, C., S. Fieulaine and T. Meinnel (2015). N-terminal protein modifications: Bringing back into play the ribosome. **Biochimie** 114: 134-146.

Giglione, C. and T. Meinnel (2021). Evolution-Driven Versatility of N Terminal Acetylation in Photoautotrophs. **Trends in Plant Science** 26(4): 375-391.

Gill, S., J. Stevenson, I. Kristiana and A. J. Brown (2011). Cholesterol-Dependent Degradation of Squalene Monooxygenase, a Control Point in Cholesterol Synthesis beyond HMG-CoA Reductase. **Cell Metabolism** 13(3): 260-273.

Glickman, M. H. and A. Ciechanover (2002). The ubiquitin-proteasome proteolytic pathway: Destruction for the sake of construction. **Physiological Reviews** 82(2): 373-428.

Goetze, S., E. Qeli, C. Mosimann, A. Staes, B. Gerrits, B. Roschitzki, S. Mohanty, E. M. Niederer, E. Laczko, E. Timmerman, V. Lange, E. Hafen, R. Aebersold, J. Vandekerckhove, K. Basler, C. H. Ahrens, K. Gevaert and E. Brunner (2009). Identification and Functional Characterization of N-Terminally Acetylated Proteins in Drosophila melanogaster. **Plos Biology** 7(11): e1000236.

Goris, M., R. S. Magin, H. Foyn, L. M. Myklebust, S. Varland, R. Ree, A. Drazic, P. Bhambra, S. I. Stove, M. Baumann, B. E. Haug, R. Marmorstein and T. Arnesen (2018). Structural determinants and cellular environment define processed actin as the sole substrate of the N-terminal acetyltransferase NAA80. **Proceedings of the National Academy of Sciences of the United States of America** 115(17): 4405-4410.

Grefen, C. and M. R. Blatt (2012). Do Calcineurin B-Like Proteins Interact Independently of the Serine Threonine Kinase CIPK23 with the K+ Channel AKT1? Lessons Learned from a Menage a Trois. **Plant Physiology** 159(3): 915-919.

Grefen, C., S. Lalonde and P. Obrdlik (2007). Split-ubiquitin system for identifying protein-protein interactions in membrane and full-length proteins. **Current Protocols in Neuroscience** Chapter 5: Unit 5.27.

Grefen, C., P. Obrdlik and K. Harter (2009). The determination of protein-protein interactions by the mating-based split-ubiquitin system (mbSUS). **Methods in Molecular Biology** 479: 217-233.

Groll, M., C. R. Berkers, H. L. Ploegh and H. Ovaa (2006). Crystal structure of the boronic acid-based proteasome inhibitor bortezomib in complex with the yeast 20S proteasome. **Structure** 14(3): 451-456.

Grove, M. D., G. F. Spencer, W. K. Rohwedder, N. Mandava, J. F. Worley, J. D. Warthen, G. L. Steffens, J. L. Flippen-Anderson and J. C. Cook (1979). Brassinolide, a Plant Growth-Promoting Steroid Isolated from Brassica-Napus Pollen. **Nature** 281(5728): 216-217.

Habeck, G., F. A. Ebner, H. Shimada-Kreft and S. G. Kreft (2015). The yeast ERAD-C ubiquitin ligase Doa10 recognizes an intramembrane degron. **Journal of Cell Biology** 209(2): 261-273.

Han, L. Q., M. Mason, E. P. Risseeuw, W. L. Crosby and D. E. Somers (2004). Formation of an SCFZTL complex is required for proper regulation of circadian timing. **Plant Journal** 40(2): 291-301.

Hannon, G. J. (2002). RNA interference. **Nature** 418(6894): 244-251.

Hartmann, M. A. (1998). Plant sterols and the membrane environment. **Trends in Plant Science** 3(5): 170-175.

Hassink, G., M. Kikkert, S. van Voorden, S. J. Lee, R. Spaapen, T. van Laar, C. S. Coleman, E. Bartee, K. Fruh, V. Chau and E. Wiertz (2005). TEB4 is a C4HC3 RING finger-containing ubiquitin ligase of the endoplasmic reticulum. **Biochemical Journal** 388: 647-655.

Helsens, K., P. Van Damme, S. Degroeve, L. Martens, T. Arnesen, J. Vandekerckhove and K. Gevaert (2011). Bioinformatics Analysis of a Saccharomyces cerevisiae N-Terminal Proteome Provides Evidence of Alternative Translation Initiation and Post-Translational N-Terminal Acetylation. **Journal of Proteome Research** 10(8): 3578-3589.

Hershko, A. and A. Ciechanover (1998). The ubiquitin system. **Annual Review of Biochemistry** 67: 425-479.

Hershko, A., H. Heller, E. Eytan, G. Kaklij and I. A. Rose (1984). Role of the Alpha-Amino Group of Protein in Ubiquitin-Mediated Protein Breakdown. **Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences** 81(22): 7021-7025.

Hinkson, I. V. and J. E. Elias (2011). The dynamic state of protein turnover: It's about time. **Trends in Cell Biology** 21(5): 293-303.

Hirsch, C., R. Gauss, S. C. Horn, O. Neuber and T. Sommer (2009). The ubiquitylation machinery of the endoplasmic reticulum. **Nature** 458(7237): 453-460.

Hofmann, K. and W. Stoffel (1993). TMbase - A database of membrane spanning proteins segments. **Biological Chemistry Hoppe-Seyler** 374(166).

Holdsworth, M. J., J. Vicente, G. Sharma, M. Abbas and A. Zubrycka (2019). The plant N-degron pathways of ubiquitin-mediated proteolysis. **Journal of Integrative Plant Biology** 62(1): 70-89.

Holman, T. J., P. D. Jones, L. Russell, A. Medhurst, S. U. Tomas, P. Talloji, J. Marquez, H. Schmuths, S.-A. Tung, I. Taylor, S. Footitt, A. Bachmair, F. L. Theodoulou and M. J. Holdsworth (2009). The N-end rule pathway promotes seed germination and establishment through removal of ABA sensitivity in *Arabidopsis*. **Proceedings of the National Academy of Sciences of the United States of America** 106(11): 4549-4554.

Holmes, W. M., B. K. Mannakee, R. N. Gutenkunst and T. R. Serio (2014). Loss of amino-terminal acetylation suppresses a prion phenotype by modulating global protein folding. **Nature Communications** 5: 4383-4383.

Howe, V., N. K. Chua, J. Stevenson and A. J. Brown (2015). The Regulatory Domain of Squalene Monooxygenase Contains a Re-entrant Loop and Senses Cholesterol via a Conformational Change*. **Journal of Biological Chemistry** 290(46): 27533-27544.

Howe, V., L. J. Sharpe, A. V. Prabhu and A. J. Brown (2017). New insights into cellular cholesterol acquisition: promoter analysis of human HMGCR and SQLE, two key control enzymes in cholesterol synthesis. **Biochimica et Biophysica Acta - Molecular and Cell Biology of Lipids** 1862(7): 647-657.

Howell, S. H. (2013). Endoplasmic Reticulum Stress Responses in Plants. **Annual Review of Plant Biology** 64: 477-499.

Hu, R. G., J. Sheng, X. Qi, Z. M. Xu, T. T. Takahashi and A. Varshavsky (2005). The N-end rule pathway as a nitric oxide sensor controlling the levels of multiple regulators. **Nature** 437(7061): 981-986.

Hua, Z. and R. D. Vierstra (2011). The Cullin-RING Ubiquitin-Protein Ligases. **Annual Review of Plant Biology** 62: 299-334.

Huber, M., L. Armbruster, R. D. Etherington, C. De La Torre, M. J. Hawkesford, C. Sticht, D. J. Gibbs, R. Hell and M. Wirtz (2022). Disruption of the Nα-acetyltransferase NatB causes sensitivity to reductive stress in *Arabidopsis thaliana*. **Frontiers in Plant Science** 12

Huber, M., W. V. Bienvenut, E. Linster, I. Stephan, L. Armbruster, C. Sticht, D. Layer, K. Lapouge, T. Meinnel, I. Sinning, C. Giglione, R. Hell and M. Wirtz (2020). NatB-Mediated N-Terminal Acetylation Affects Growth and Biotic Stress Responses. **Plant Physiology** 182(2): 792-806.

Hüttner, S., C. Veit, J. Schoberer, J. Grass and R. Strasser (2012). Unraveling the function of *Arabidopsis thaliana* OS9 in the endoplasmic reticulum-associated degradation of glycoproteins. **Plant Molecular Biology** 79(1-2): 21-33.

Huyer, G., W. F. Piluek, Z. Fansler, S. G. Kreft, M. Hochstrasser, J. L. Brodsky and S. Michaelis (2004). Distinct machinery is required in Saccharomyces cerevisiae for the endoplasmic reticulum-associated degradation of a multispanning membrane protein and a soluble luminal protein. **Journal of Biological Chemistry** 279(37): 38369-38378.

Hwang, C.-S., A. Shemorry and A. Varshavsky (2010). N-Terminal Acetylation of Cellular Proteins Creates Specific Degradation Signals. **Science** 327(5968): 973-977.

Hwang, C.-S., M. Sukalo, O. Batygin, M.-C. Addor, H. Brunner, A. Perez Aytes, J. Mayerle, H. K. Song, A. Varshavsky and M. Zenker (2011). Ubiquitin Ligases of the N-End Rule Pathway: Assessment of Mutations in UBR1 That Cause the Johanson-Blizzard Syndrome. **Plos One** 6(9): e24925.

Jasma, E., M. Simonen and M. Makarow (1994). Selective Retention of Secretory Proteins in the Yeast Endoplasmic-Reticulum by Treatment of Cells with a Reducing Agent. **Yeast** 10(3): 355-370.

Jenks, M. A., H. A. Tuttle, S. D. Eigenbrode and K. A. Feldmann (1995). Leaf Epicuticular Waxes of the Eceriferum Mutants in *Arabidopsis*. **Plant Physiology** 108(1): 369-377.

Jinek, M., K. Chylinski, I. Fonfara, M. Hauer, J. A. Doudna and E. Charpentier (2012). A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. **Science** 337(6096): 816-821.

Jo, Y. and R. A. DeBose-Boyd (2010). Control of cholesterol synthesis through regulated ER-associated degradation of HMG CoA reductase. **Critical Reviews in Biochemistry and Molecular Biology** 45(3): 185-198.

Johnston, A. J., P. Meier, J. Gheyselinck, S. E. J. Wuest, M. Federer, E. Schlagenhauf, J. D. Becker and U. Grossniklaus (2007). Genetic subtraction profiling identifies genes essential for *Arabidopsis* reproduction and reveals interaction between the female gametophyte and the maternal sporophyte. **Genome Biology** 8(10): R204.

Kagale, S., U. K. Divi, J. E. Krochko, W. A. Keller and P. Krishna (2007). Brassinosteroid confers tolerance in *Arabidopsis thaliana* and Brassica napus to a range of abiotic stresses. **Planta** 225(2): 353-364.

Kall, L., A. Krogh and E. L. L. Sonnhammer (2004). A combined transmembrane topology and signal peptide prediction method. **Journal of Molecular Biology** 338(5): 1027-1036.

Kalvik, T. V. and T. Arnesen (2013). Protein N-terminal acetyltransferases in cancer. **Oncogene** 32(3): 269-276.

Kamauchi, S., H. Nakatani, C. Nakano and R. Urade (2005). Gene expression in response to endoplasmic reticulum stress in *Arabidopsis thaliana*. **FEBS Journal** 272(13): 3461-3476.

Karimi, M., D. Inze and A. Depicker (2002). GATEWAY((TM)) vectors for Agrobacterium-mediated plant transformation. **Trends in Plant Science** 7(5): 193-195.

Kats, I., A. Khmelinskii, M. Kschonsak, F. Huber, R. A. Kniess, A. Bartosik and M. Knop (2018). Mapping Degradation Signals and Pathways in a Eukaryotic N-terminome. **Molecular cell** 70(3): 488-501.

Kaya, H., M. Mikami, A. Endo, M. Endo and S. Toki (2016). Highly specific targeted mutagenesis in plants using Staphylococcus aureus Cas9. **Scientific Reports** 6: 26871-26871.

Kepinski, S. and O. Leyser (2005). The *Arabidopsis* F-box protein TIR1 is an auxin receptor. **Nature** 435(7041): 446-451.

Kim, H.-K., R.-R. Kim, J.-H. Oh, H. Cho, A. Varshavsky and C.-S. Hwang (2014). The N-Terminal Methionine of Cellular Proteins as a Degradation Signal. **Cell** 156(1-2): 158-169.

Kim, I., C. R. Miller, D. L. Young and S. Fields (2013). High-throughput analysis of in vivo protein stability. **Molecular & Cellular Proteomics : MCP** 12(11): 3370-3378.

Kim, J. H., S. K. Cho, T. R. Oh, M. Y. Ryu, S. W. Yang and W. T. Kim (2017). MPSR1 is a cytoplasmic PQC E3 ligase for eliminating emergent misfolded proteins in *Arabidopsis thaliana*. **Proceedings of the National Academy of Sciences of the United States of America** 114(46): E10009-E10017.

Kim, J.-M. (2019). N-terminal formylmethionine as a novel initiator and N-degron of eukaryotic proteins. **BMB Reports** 52(3): 163-164.

Kim, J.-M. and C.-S. Hwang (2014). Crosstalk between the Arg/N-end and Ac/N-end rule. **Cell Cycle** 13(9): 1366-1367.

Kim, J.-M., O.-H. Seok, S. Ju, J.-E. Heo, J. Yeom, D.-S. Kim, J.-Y. Yoo, A. Varshavsky, C. Lee and C.-S. Hwang (2018). Formyl-methionine as an N-degron of a eukaryotic N-end rule pathway. **Science** 362(6418): eaat0174.

Kisselev, A. F., W. A. van der Linden and H. S. Overkleeft (2012). Proteasome inhibitors: an expanding army attacking a unique target. **Chemistry & biology** 19(1): 99-115.

Klepikova, A. V., A. S. Kasianov, E. S. Gerasimov, M. D. Logacheva and A. A. Penin (2016). A high resolution map of the *Arabidopsis thaliana* developmental transcriptome based on RNA-seq profiling. **Plant Journal** 88(6): 1058-1070.

Koizumi, N., T. Ujino, H. Sano and M. J. Chrispeels (1999). Overexpression of a gene that encodes the first enzyme in the biosynthesis of asparagine-linked glycans makes plants resistant to tunicamycin and obviates the tunicamycin-induced unfolded protein response. **Plant Physiology** 121(2): 353-361.

Komander, D. and M. Rape (2012). The Ubiquitin Code. **Annual Review of Biochemistry** 81: 203-229.

Koornneef, M., C. J. Hanhart and F. Thiel (1989). A Genetic and Phenotypic Description of Eceriferum (Cer) Mutants in *Arabidopsis-Thaliana*. **Journal of Heredity** 80(2): 118-122.

Koren, I., R. T. Timms, T. Kula, Q. Xu, M. Z. Li and S. J. Elledge (2018). The Eukaryotic Proteome Is Shaped by E3 Ubiquitin Ligases Targeting C-Terminal Degrons. **Cell** 173(7): 1622-1635.

Koskela, M. M., A. Brünje, A. Ivanauskaite, M. Grabsztunowicz, I. Lassowskat, U. Neumann, T. V. Dinh, J. Sindlinger, D. Schwarzer, M. Wirtz, E. Tyystjärvi, I. Finkemeier and P. Mulo (2018). Chloroplast Acetyltransferase NSI Is Required for State Transitions in *Arabidopsis thaliana*. **The Plant Cell** 30(8): 1695-1709.

Kostova, Z., Y. C. Tsai and A. M. Weissman (2007). Ubiquitin ligases, critical mediators of endoplasmic reticulum-associated degradation. **Seminars in Cell & Developmental Biology** 18(6): 770-779.

Kreft, S. G. and M. Hochstrasser (2011). An Unusual Transmembrane Helix in the Endoplasmic Reticulum Ubiquitin Ligase Doa10 Modulates Degradation of Its Cognate E2 Enzyme. **Journal of Biological Chemistry** 286(23): 20163-20174.

Kreft, S. G., L. Wang and M. Hochstrasser (2006). Membrane topology of the yeast endoplasmic reticulum-localized ubiquitin ligase Doa10 and comparison with its human ortholog TEB4 (MARCH-VI). **Journal of Biological Chemistry** 281(8): 4646-4653.

Krishna, P. (2003). Brassinosteroid-mediated stress responses. **Journal of Plant Growth Regulation** 22(4): 289-297.

Krogh, A., B. Larsson, G. von Heijne and E. L. L. Sonnhammer (2001). Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. **Journal of Molecular Biology** 305(3): 567-580.

Kunst, L. and L. Samuels (2009). Plant cuticles shine: advances in wax biosynthesis and export. **Current Opinion in Plant Biology** 12(6): 721-727.

Kwon, Y. T., Z. Xia, I. V. Davydov, S. H. Lecker and A. Varshavsky (2001). Construction and Analysis of Mouse Strains Lacking the Ubiquitin Ligase UBR1 (E3α) of the N-End Rule Pathway. **Molecular and Cellular Biology** 21(23): 8007-8021.

Labuhn, M., F. F. Adams, M. Ng, S. Knoess, A. Schambach, E. M. Charpentier, A. Schwarzer, J. L. Mateo, J. H. Klusmann and D. Heckl (2018). Refined sgRNA efficacy prediction improves large- and small-scale CRISPR-Cas9 applications. **Nucleic Acids Research** 46(3): 1375-1385.

Lanza, M., B. Garcia-Ponce, G. Castrillo, P. Catarecha, M. Sauer, M. Rodriguez-Serrano, A. Paez-Garcia, E. Sanchez-Bermejo, T. C. Mohan, Y. Leo del Puerto, L. Maria Sandalio, J. Paz-Ares and A. Leyva (2012). Role of Actin Cytoskeleton in Brassinosteroid Signaling and in Its Integration with the Auxin Response in Plants. **Developmental Cell** 22(6): 1275-1285.

Laranjeira, S., V. Amorim-Silva, A. Esteban, M. Arro, A. Ferrer, R. Manuel Tavares, M. Angel Botella, A. Rosado and H. Azevedo (2015). *Arabidopsis* Squalene Epoxidase 3 (SQE3) Complements SQE1 and Is Important for Embryo Development and Bulk Squalene Epoxidase Activity. **Molecular Plant** 8(7): 1090-1102.

Laule, O., A. Furholz, H. S. Chang, T. Zhu, X. Wang, P. B. Heifetz, W. Gruissem and B. M. Lange (2003). Crosstalk between cytosolic and plastidial pathways of isoprenoid biosynthesis in *Arabidopsis thaliana*. **Proceedings of the National Academy of Sciences of the United States of America** 100(11): 6866-6871.

Leber, J. H., S. Bernales and P. Walter (2004). IRE1-independent gain control of the unfolded protein response. **Plos Biology** 2(8): 1197-1207.

Leber, R., R. Zenz, K. Schröttner, S. Fuchsbichler, B. Pühringer and F. Turnowsky (2001). A novel sequence element is involved in the transcriptional regulation of expression of the ERG1 (squalene epoxidase) gene in Saccharomyces cerevisiae. **European Journal of Biochemistry** 268(4): 914-924.

Lee, K.-E., J.-E. Heo, J.-M. Kim and C.-S. Hwang (2016). N-Terminal Acetylation-Targeted N-End Rule Proteolytic System: The Ac/N-End Rule Pathway. **Molecules and cells** 39(3): 169-178.

Letunic, I. and P. Bork (2018). 20 years of the SMART protein domain annotation resource. **Nucleic Acids Research** 46(D1): D493-D496.

Li, L.-M., S.-Y. Lu and R.-J. Li (2017). The *Arabidopsis* endoplasmic reticulum associated degradation pathways are involved in the regulation of heat stress response. **Biochemical and Biophysical Research Communications** 487(2): 362-367.

Li, Z., V. Dogra, K. P. Lee, R. Li, M. Li, M. Li and C. Kim (2020). N-Terminal Acetylation Stabilizes SIGMA FACTOR BINDING PROTEIN1 Involved in Salicylic Acid-Primed Cell Death. **Plant Physiology** 183(1): 358-370.

Licausi, F., M. Kosmacz, D. A. Weits, B. Giuntoli, F. M. Giorgi, L. A. C. J. Voesenek, P. Perata and J. T. van Dongen (2011). Oxygen sensing in plants is mediated by an N-end rule pathway for protein destabilization. **Nature** 479(7373): 419-422.

Lin, H.-C., C.-W. Yeh, Y.-F. Chen, T.-T. Lee, P.-Y. Hsieh, D. V. Rusnac, S.-Y. Lin, S. J. Elledge, N. Zheng and H.-C. S. Yen (2018). C-Terminal End-Directed Protein Elimination by CRL2 Ubiquitin Ligases. **Molecular Cell** 70(4): 602-613.

Lin, L., C. Zhang, Y. Chen, Y. Wang, D. Wang, X. Liu, M. Wang, J. Mao, J. Zhang, W. Xing, L. Liu and J. Li (2019). PAWH1 and PAWH2 are plant-specific components of an *Arabidopsis* endoplasmic reticulum-associated degradation complex. **Nature Communications** 10: 3492-3492.

Linster, E., D. Layer, W. V. Bienvenut, T. V. Dinh, F. A. Weyer, W. Leemhuis, A. Brünje, M. Hoffrichter, P. Miklankova, J. Kopp, K. Lapouge, J. Sindlinger, D. Schwarzer, T. Meinnel, I. Finkemeier, C. Giglione, R. Hell, I. Sinning and M. Wirtz (2020). The *Arabidopsis* Nα-acetyltransferase NAA60 locates to the plasma membrane and is vital for the high salt stress response. **New Phytologist** 228(2): 554-569.

Linster, E., I. Stephan, W. V. Bienvenut, J. Maple-Grodem, L. M. Myklebust, M. Huber, M. Reichelt, C. Sticht, S. G. Moller, T. Meinnel, T. Arnesen, C. Giglione, R. Hell and M. Wirtz (2015). Downregulation of N-terminal acetylation triggers ABA-mediated drought responses in *Arabidopsis*. **Nature Communications** 6: 7640-7640.

Linster, E. and M. Wirtz (2018). N-terminal acetylation: an essential protein modification emerges as an important regulator of stress responses. **Journal of Experimental Botany** 69(19): 4555-4568.

Liszczak, G., J. M. Goldberg, H. Foyn, E. J. Petersson, T. Arnesen and R. Marmorstein (2013). Molecular basis for N-terminal acetylation by the heterodimeric NatA complex. **Nature Structural & Molecular Biology** 20(9): 1098-1105.

Liu, H.-Q., Y. Zou, X.-F. Li, L. Wu and G.-Q. Guo (2021). Stablization of ACOs by NatB mediated N-terminal acetylation is required for ethylene homeostasis. **BMC Plant Biology** 21: 320

Liu, J.-X. and S. H. Howell (2016). Managing the protein folding demands in the endoplasmic reticulum of plants. **New Phytologist** 211(2): 418-428.

Liu, L., F. Cui, Q. Li, B. Yin, H. Zhang, B. Lin, Y. Wu, R. Xia, S. Tang and Q. Xie (2011). The endoplasmic reticulum-associated degradation is necessary for plant salt tolerance. **Cell Research** 21(6): 957-969.

Liu, Y. and J. Li (2014). Endoplasmic reticulum-mediated protein quality control in *Arabidopsis*. **Frontiers in Plant Science** 5: 162.

Liu, Y., C. Zhang, D. Wang, W. Su, L. Liu, M. Wang and J. Li (2015). EBS7 is a plant-specific component of a highly conserved endoplasmic reticulum-associated degradation system in *Arabidopsis*. **Proceedings of the National Academy of Sciences of the United States of America** 112(39): 12205-12210.

Livak, K. J. and T. D. Schmittgen (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. **Methods** 25(4): 402-408.

Loertscher, J., L. L. Larson, C. K. Matson, M. L. Parrish, A. Felthauser, A. Sturm, C. Tachibana, M. Bard and R. Wright (2006). Endoplasmic reticulum-associated degradation is required for cold adaptation and regulation of sterol biosynthesis in the yeast Saccharomyces cerevisiae. **Eukaryotic Cell** 5(4): 712-722.

Lu, S., H. Zhao, D. L. Des Marais, E. P. Parsons, X. Wen, X. Xu, D. K. Bangarusamy, G. Wang, O. Rowland, T. Juenger, R. A. Bressan and M. A. Jenks (2012). *Arabidopsis* ECERIFERUM9 Involvement in Cuticle Formation and Maintenance of Plant Water Status. **Plant Physiology** 159(3): 930-944.

Mao, Y., Z. Zhang, Z. Feng, P. Wei, H. Zhang, J. R. Botella and J.-K. Zhu (2016). Development of germ-line-specific CRISPR-Cas9 systems to improve the production of heritable gene modifications in *Arabidopsis*. **Plant Biotechnology Journal** 14(2): 519-532.

Masson, N., T. P. Keeley, B. Giuntoli, M. D. White, M. L. Puerta, P. Perata, R. J. Hopkinson, E. Flashman, F. Licausi and P. J. Ratcliffe (2019). Conserved N-terminal cysteine dioxygenases transduce responses to hypoxia in animals and plants. **Science** 365(6448): 65-69.

Matlack, K. E. S., W. Mothes and T. A. Rapoport (1998). Protein translocation: Tunnel vision. **Cell** 92(3): 381-390.

McDowell, J. M., Y. Q. An, S. R. Huang, E. C. McKinney and R. B. Meagher (1996). The *Arabidopsis* ACT7 actin gene is expressed in rapidly developing tissues and responds to several external stimuli. **Plant Physiology** 111(3): 699-711.

Mendoza-Poudereux, I., E. Kutzner, C. Huber, J. Segura, W. Eisenreich and I. Arrillaga (2015). Metabolic cross-talk between pathways of terpenoid backbone biosynthesis in spike lavender. **Plant Physiology and Biochemistry** 95: 113-120.

Metzger, M. B., M. J. Maurer, B. M. Dancy and S. Michaelis (2008). Degradation of a Cytosolic Protein Requires Endoplasmic Reticulum-associated Degradation Machinery. **Journal of Biological Chemistry** 283(47): 32302-32316.

Metzger, M. B., J. N. Pruneda, R. E. Klevit and A. M. Weissman (2014). RING-type E3 ligases: Master manipulators of E2 ubiquitin-conjugating enzymes and ubiquitination. **Biochimica et Biophysica Acta-Molecular Cell Research** 1843(1): 47-60.

Meusser, B., C. Hirsch, E. Jarosch and T. Sommer (2005). ERAD: the long road to destruction. **Nature Cell Biology** 7(8): 766-772.

Millar, A. H., J. L. Heazlewood, C. Giglione, M. J. Holdsworth, A. Bachmair and W. X. Schulze (2019). The Scope, Functions, and Dynamics of Posttranslational Protein Modifications. **Annual Review of Plant Biology** 70: 119-151.

Mochizuki, S., K. Sugimoto, T. Koeduka and K. Matsui (2016). *Arabidopsis* lipoxygenase 2 is essential for formation of green leaf volatiles and five-carbon volatiles. **FEBS letters** 590(7): 1017-1027.

Mojica, F. J. M., C. Diez-Villasenor, J. Garcia-Martinez and C. Almendros (2009). Short motif sequences determine the targets of the prokaryotic CRISPR defence system. **Microbiology-SGM** 155: 733-740.

Monda, J. K., D. C. Scott, D. J. Miller, J. Lydeard, D. King, J. W. Harper, E. J. Bennett and B. A. Schulman (2013). Structural Conservation of Distinctive N-terminal Acetylation-Dependent Interactions across a Family of Mammalian NEDD8 Ligation Enzymes. **Structure** 21(1): 42-53.

Morreale, F. E. and H. Walden (2016). SnapShot: Types of Ubiquitin Ligases. **Cell** 165(1): 248.

Moses, T., K. K. Papadopoulou and A. Osbourn (2014). Metabolic and functional diversity of saponins, biosynthetic intermediates and semi-synthetic derivatives. **Critical Reviews in Biochemistry and Molecular Biology** 49(6): 439-462.

Movahedi, A., H. Wei, B. Pucker, W. Sun, D. Li, L. Yang and Q. Zhuge (2021). Regulation of poplar isoprenoid biosynthesis by methylerythritol phosphate and mevalonic acid pathways interactions. **bioRxiv**: 2020.2007.2022.216804.

Myklebust, L. M., P. Van Damme, S. I. Stove, M. J. Doerfel, A. Abboud, T. V. Kalvik, C. Grauffel, V. Jonckheere, Y. Wu, J. Swensen, H. Kaasa, G. Liszczak, R. Marmorstein, N. Reuter, G. J. Lyon, K. Gevaert and T. Arnesen (2015). Biochemical and cellular analysis of Ogden syndrome reveals downstream Nt-acetylation defects. **Human Molecular Genetics** 24(7): 1956-1976.

Nakagawa, T., T. Suzuki, S. Murata, S. Nakamura, T. Hino, K. Maeo, R. Tabata, T. Kawai, K. Tanaka, Y. Niwa, Y. Watanabe, K. Nakamura, T. Kimura and S. Ishiguro (2007). Improved gateway binary vectors: High-performance vectors for creation of fusion constructs in Transgenic analysis of plants. **Bioscience Biotechnology and Biochemistry** 71(8): 2095-2100.

Neubauer, M. and R. W. Innes (2020). Loss of the Acetyltransferase NAA50 Induces Endoplasmic Reticulum Stress and Immune Responses and Suppresses Growth. Plant **Physiology** 183(4): 1838-1854.

Nguyen K. T., C.-S. Lee, S.-H. Mun, T. Nhung Thimy, S. K. Park and C.-S. Hwang (2019). N-terminal acetylation and the N-end rule pathway control degradation of the lipid droplet protein PLIN2. **Journal of Biological Chemistry** 294(1): 379-388.

Nguyen K. T., S.-H. Mun, C.-S. Lee and C.-S. Hwang (2018). Control of protein degradation by N-terminal acetylation and the N-end rule pathway. **Experimental and Molecular Medicine** 50: 91.

Nowosieski, M., M. Hoffmann, L. S. Wyrwicz, P. Stepniak, D. M. Plewczynski, M. Lazniewski, K. Ginalski and L. Rychlewski (2011). Detailed Mechanism of Squalene Epoxidase Inhibition by Terbinafine. **Journal of Chemical Information and Modeling** 51(2): 455-462.

Obrdlik, P., M. El-Bakkoury, T. Hamacher, C. Cappellaro, C. Vilarino, C. Fleischer, H. Ellerbrok, R. Kamuzinzi, V. Ledent, D. Blaudez, D. Sanders, J. L. Revuelta, E. Boles, B. Andre and W. B. Frommer (2004). K+ channel interactions detected by a genetic system optimized for systematic studies of membrane protein interactions. **Proceedings of the National Academy of Sciences of the United States of America** 101(33): 12242-12247.

Ogen-Shtern, N., T. Ben David and G. Z. Lederkremer (2016). Protein aggregation and ER stress. **Brain Research** 1648: 658-666.

Ohi, M. D., C. W. Vander Kooi, J. A. Rosenberg, W. J. Chazin and K. L. Gould (2003). Structural insights into the U-box, a domain associated with multi-ubiquitination. **Nature Structural Biology** 10(4): 250-255.

Osakabe, Y. and K. Osakabe (2015). Genome Editing with Engineered Nucleases in Plants. **Plant and Cell Physiology** 56(3): 389-400.

Ossowski, S., R. Schwab and D. Weigel (2008). Gene silencing in plants using artificial microRNAs and other small RNAs. **Plant Journal** 53(4): 674-690.

Park, S.-E., J.-M. Kim, O.-H. Seok, H. Cho, B. Wadas, S.-Y. Kim, A. Varshavsky and C.-S. Hwang (2015). Control of mammalian G protein signaling by N-terminal acetylation and the N-end rule pathway. **Science** 347(6227): 1249-1252.

Parra-Rojas, J., A. A. Moreno, I. Mitina and A. Orellana (2015). The Dynamic of the Splicing of bZIP60 and the Proteins Encoded by the Spliced and Unspliced mRNAs Reveals Some Unique Features during the Activation of UPR in *Arabidopsis thaliana*. **Plos One** 10(4): e0122936.

Pegg, A. E. (2006). Regulation of ornithine decarboxylase. **Journal of Biological Chemistry** 281(21): 14529-14532.

Pesaresi, P., N. A. Gardner, S. Masiero, A. Dietzmann, L. Eichacker, R. Wickner, F. Salamini and D. Leister (2003). Cytoplasmic N-terminal protein acetylation is required for efficient photosynthesis in *Arabidopsis*. **Plant Cell** 15(8): 1817-1832.

Pestova, T. V. and C. U. Hellen (2003). Translation elongation after assembly of ribosomes on the Cricket paralysis virus internal ribosomal entry site without initiation factors or initiator tRNA. **Genes & Development** 17(2): 181-186.

Petroski, M. D. and R. J. Deshaies (2005). Function and regulation of Cullin-RING ubiquitin ligases. **Nature Reviews Molecular Cell Biology** 6(1): 9-20.

Phillips, D. R., J. M. Rasbery, B. Bartel and S. P. Matsuda (2006). Biosynthetic diversity in plant triterpene cyclization. **Current Opinion in Plant Biology** 9(3): 305-314.

Piatkov, K. I., T. T. M. Vu, C.-S. Hwang and A. Varshavsky (2015). Formyl-methionine as a degradation signal at the N-termini of bacterial proteins. **Microbial Cell** 2(10): 376-393.

Pickart, C. M. and M. J. Eddins (2004). Ubiquitin: structures, functions, mechanisms. **Biochimica et Biophysica Acta-Molecular Cell Research** 1695(1-3): 55-72.

Polevoda, B., S. Brown, T. S. Cardillo, S. Rigby and F. Sherman (2008). Yeast N-alpha-terminal acetyltransferases are associated with ribosomes. **Journal of Cellular Biochemistry** 103(2): 492-508.

Polevoda, B. and F. Sherman (2003). N-terminal acetyltransferases and sequence requirements for N-terminal acetylation of eukaryotic proteins. **Journal of Molecular Biology** 325(4): 595-622.

Pose, D., I. Castanedo, O. Borsani, B. Nieto, A. Rosado, L. Taconnat, A. Ferrer, L. Dolan, V. Valpuesta and M. A. Botella (2009). Identification of the *Arabidopsis* dry2/sqe1-5 mutant reveals a central role for sterols in drought tolerance and regulation of reactive oxygen species. **Plant Journal** 59(1): 63-76.

Potuschak, T., S. Stary, P. Schlogelhofer, F. Becker, V. Nejinskaia and A. Bachmair (1998). PRTI of *Arabidopsis thaliana* encodes a component of the plant N-end rule pathway. **Proceedings of the National Academy of Sciences of the United States of America** 95(14): 7904-7908.

Puchta, H. (2017). Applying CRISPR/Cas for genome engineering in plants: the best is yet to come. **Current Opinion in Plant Biology** 36: 1-8.

Puchta, H. and F. Fauser (2014). Synthetic nucleases for genome engineering in plants: prospects for a bright future. **Plant Journal** 78(5): 727-741.

Queitsch, C., T. A. Sangster and S. Lindquist (2002). Hsp90 as a capacitor of phenotypic variation. **Nature** 417(6889): 618-624.

Rademacher, W. (2000). Growth retardants: Effects on gibberellin biosynthesis and other metabolic pathways. **Annual Review of Plant Physiology and Plant Molecular Biology** 51: 501-531.

Ran, F. A., L. Cong, W. X. Yan, D. A. Scott, J. S. Gootenberg, A. J. Kriz, B. Zetsche, O. Shalem, X. Wu, K. S. Makarova, E. V. Koonin, P. A. Sharp and F. Zhang (2015). In vivo genome editing using Staphylococcus aureus Cas9. **Nature** 520(7546): 186-191.

Rasbery, J. M., H. Shan, R. J. LeClair, M. Norman, S. P. T. Matsuda and B. Bartel (2007). *Arabidopsis thaliana* squalene epoxidase 1 is essential for root and seed development. **Journal of Biological Chemistry** 282(23): 17002-17013.

Ravid, T. and M. Hochstrasser (2008). Diversity of degradation signals in the ubiquitin-proteasome system. **Nature Reviews Molecular Cell Biology** 9(9): 679-689.

Ravid, T., S. G. Kreft and M. Hochstrasser (2006). Membrane and soluble substrates of the Doa10 ubiquitin ligase are degraded by distinct pathways. **Embo Journal** 25(3): 533-543.

Reisberg, E. E., U. Hildebrandt, M. Riederer and U. Hentschel (2013). Distinct Phyllosphere Bacterial Communities on *Arabidopsis* Wax Mutant Leaves. **Plos One** 8(11): e78613.

Roe, S. M., C. Prodromou, R. O'Brien, J. E. Ladbury, P. W. Piper and L. H. Pearl (1999). Structural basis for inhibition of the Hsp90 molecular chaperone by the antitumor antibiotics radicicol and geldanamycin. **Journal of Medicinal Chemistry** 42(2): 260-266.

Roemisch, K. (2017). A Case for Sec61 Channel Involvement in ERAD. **Trends In Biochemical Sciences** 42(3): 171-179.

Rogowska, A. and A. Szakiel (2020). The role of sterols in plant response to abiotic stress. **Phytochemistry Reviews** 19(6): 1525-1538.

Rope, A. F., K. Wang, R. Evjenth, J. Xing, J. J. Johnston, J. J. Swensen, W. E. Johnson, B. Moore, C. D. Huff, L. M. Bird, J. C. Carey, J. M. Opitz, C. A. Stevens, T. Jiang, C. Schank, H. D. Fain, R. Robison, B. Dalley, S. Chin, S. T. South, T. J. Pysher, L. B. Jorde, H. Hakonarson, J. R. Lillehaug, L. G. Biesecker, M. Yandell, T. Arnesen and G. J. Lyon (2011). Using VAAST to Identify an X-Linked Disorder Resulting in Lethality in Male Infants Due to N-Terminal Acetyltransferase Deficiency. **American Journal of Human Genetics** 89(1): 28-43.

Rowe, S. M., S. Miller and E. J. Sorscher (2005). Mechanisms of disease: Cystic fibrosis. **New England Journal of Medicine** 352(19): 1992-2001.

Ruan, Y.-L., J. W. Patrick, M. Bouzayen, S. Osorio and A. R. Fernie (2012). Molecular regulation of seed and fruit set. **Trends in Plant Science** 17(11): 656-665.

Ruggiano, A., G. Mora, L. Buxó and P. Carvalho (2016). Spatial control of lipid droplet proteins by the ERAD ubiquitin ligase Doa10. **The EMBO Journal** 35(15): 1644-1655.

Runnebohm, A. M., K. A. Richards, C. B. Irelan, S. M. Turk, H. E. Vitali, C. J. Indovina and E. M. Rubenstein (2020). Overlapping function of Hrd1 and Ste24 in translocon quality control provides robust channel surveillance. **Journal of Biological Chemistry** 295(47): 16113-16120.

Ryder, N. S. (1992). Terbinafine - Mode of Action and Properties of the Squalene Epoxidase Inhibition. **British Journal of Dermatology** 126: 2-7.

Schiml, S., F. Fauser and H. Puchta (2016). CRISPR/Cas-Mediated Site-Specific Mutagenesis in *Arabidopsis thaliana* Using Cas9 Nucleases and Paired Nickases. **Methods in Molecular Biology** 1469: 111-122.

Schindelin, J., I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J.-Y. Tinevez, D. J. White, V. Hartenstein, K. Eliceiri, P. Tomancak and A. Cardona (2012). Fiji: an open-source platform for biological-image analysis. **Nature Methods** 9(7): 676-682.

Schmidt, C. C., V. Vasic and A. Stein (2020). Doa10 is a membrane protein retrotranslocase in ER-associated protein degradation. **eLife** 9: e56945.

Schneider-Poetsch, T., J. Ju, D. E. Eyler, Y. Dang, S. Bhat, W. C. Merrick, R. Green, B. Shen and J. O. Liu (2010). Inhibition of eukaryotic translation elongation by cycloheximide and lactimidomycin. **Nature Chemical Biology** 6(3): 209-217.

Scott, D. C., J. K. Monda, E. J. Bennett, J. W. Harper and B. A. Schulman (2011). N-Terminal Acetylation Acts as an Avidity Enhancer Within an Interconnected Multiprotein Complex. **Science** 334(6056): 674-678.

Scott, N. A., L. J. Sharpe, I. M. Capell-Hattam, S. J. Gullo, W. Luu and A. J. Brown (2020). The cholesterol synthesis enzyme lanosterol 14 alpha-demethylase is post-translationally regulated by the E3 ubiquitin ligase MARCH6. **Biochemical Journal** 477: 541-555.

Sena, F., M. Sotelo-Silveira, S. Astrada, M. A. Botella, L. Malacrida and O. Borsani (2017). Spectral phasor analysis reveals altered membrane order and function of root hair cells in *Arabidopsis* dry2/sqe1-5 drought hypersensitive mutant. **Plant Physiology and Biochemistry** 119: 224-231.

Shan, Q., Y. Wang, J. Li, Y. Zhang, K. Chen, Z. Liang, K. Zhang, J. Liu, J. J. Xi, J.-L. Qiu and C. Gao (2013). Targeted genome modification of crop plants using a CRISPR-Cas system. **Nature Biotechnology** 31(8): 686-688.

Sharpe, L. J., V. Howe, N. A. Scott, W. Luu, L. Phan, J. M. Berk, M. Hochstrasser and A. J. Brown (2019). Cholesterol increases protein levels of the E3 ligase MARCH6 and thereby stimulates protein degradation. **The Journal of Biological Chemistry** 294(7): 2436-2448.

Shemorry, A., C.-S. Hwang and A. Varshavsky (2013). Control of Protein Quality and Stoichiometries by N-Terminal Acetylation and the N-End Rule Pathway. **Molecular Cell** 50(4): 540-551.

Schulman, B. A. and J. W. Harper (2009). Ubiquitin-like protein activation by E1 enzymes: the apex for downstream signalling pathways. Nature reviews. **Molecular Cell Biology** 10(5): 319-331.

Sievers, F., A. Wilm, D. Dineen, T. J. Gibson, K. Karplus, W. Li, R. Lopez, H. McWilliam, M. Remmert, J. Soeding, J. D. Thompson and D. G. Higgins (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. **Molecular Systems Biology** 7: 539-539.

Silva, P. A., J. C. F. Silva, H. D. N. Caetano, J. P. B. Machado, G. C. Mendes, P. A. B. Reis, O. J. B. Brustolini, M. Dal-Bianco and E. P. B. Fontes (2015). Comprehensive analysis of the endoplasmic reticulum stress response in the soybean genome: conserved and plant-specific features. **BMC Genomics** 16: 783-783.

Smyth, D. R., J. L. Bowman and E. M. Meyerowitz (1990). Early Flower Development in *Arabidopsis*. **Plant Cell** 2(8): 755-767.

Soda, N., L. Verma and J. Giri (2018). CRISPR-Cas9 based plant genome editing: Significance, opportunities and recent advances. **Plant Physiology and Biochemistry** 131: 2-11.

Soyars, C. L., B. A. Peterson, C. A. Burr and Z. L. Nimchuk (2018). Cutting Edge Genetics: CRISPR/Cas9 Editing of Plant Genomes. **Plant and Cell Physiology** 59(8): 1608-1620.

Spratt, D. E., H. Walden and G. S. Shaw (2014). RBR E3 ubiquitin ligases: new structures, new insights, new questions. **Biochemical Journal** 458: 421-437.

Sriram, S. M., B. Y. Kim and Y. T. Kwon (2011). The N-end rule pathway: emerging functions and molecular principles of substrate recognition. **Nature Reviews Molecular Cell Biology** 12(11): 735-747.

Starheim, K. K., K. Gevaert and T. Arnesen (2012). Protein N-terminal acetyltransferases: when the start matters. **Trends in Biochemical Sciences** 37(4): 152-161.

Stary, S., X. J. Yin, T. Potuschak, P. Schlogelhofer, V. Nizhynska and A. Bachmair (2003). PRT1 of *Arabidopsis* is a ubiquitin protein ligase of the plant N-end rule pathway with specificity for aromatic amino-terminal residues. **Plant Physiology** 133(3): 1360-1366.

Stefanovic-Barrett, S., A. S. Dickson, S. P. Burr, J. C. Williamson, I. T. Lobb, D. J. van den Boomen, P. J. Lehner and J. A. Nathan (2018). MARCH6 and TRC8 facilitate the quality control of cytosolic and tail-anchored proteins. **EMBO Reports** 19(5): e45603.

Steinert, J., S. Schiml, F. Fauser and H. Puchta (2015). Highly efficient heritable plant genome engineering using Cas9 orthologues from Streptococcus thermophilus and Staphylococcus aureus. **Plant Journal** 84(6): 1295-1305.

Stemmer, M., T. Thumberger, M. d. S. Keyer, J. Wittbrodt and J. L. Mateo (2015). CCTop: An Intuitive, Flexible and Reliable CRISPR/Cas9 Target Prediction Tool. **Plos One** 10(4): e0124633.

Stolz, A., S. Besser, H. Hottmann and D. H. Wolf (2013). Previously unknown role for the ubiquitin ligase Ubr1 in endoplasmic reticulum-associated protein degradation. **Proceedings of the National Academy of Sciences of the United States of America** 110(38): 15271-15276.

Strasser, R. (2018). Protein Quality Control in the Endoplasmic Reticulum of Plants. **Annual Review of Plant Biology** 69: 147-172.

Su, W., Y. Liu, Y. Xia, Z. Hong and J. Li (2011). Conserved endoplasmic reticulum-associated degradation system to eliminate mutated receptor-like kinases in *Arabidopsis*. **Proceedings of the National Academy of Sciences of the United States of America** 108(2): 870-875.

Sugio, A., R. Dreos, F. Aparicio and A. J. Maule (2009). The Cytosolic Protein Response as a Subcomponent of the Wider Heat Shock Response in *Arabidopsis*. **Plant Cell** 21(2): 642-654.

Sukalo, M., A. Fiedler, C. Guzman, S. Spranger, M.-C. Addor, J. N. McHeik, M. O. Benavent, J. M. Cobben, L. A. Gillis, A. G. Shealy, C. Deshpande, B. Bozorgmehr, D. B. Everman, E.-L. Stattin, J. Liebelt, K.-M. Keller, D. R. Bertola, C. D. M. van Karnebeek, C. Bergmann, Z. Liu, G. Dueker, N. Rezaei, F. S. Alkuraya, G. Ogur, A. Alrajoudi, C. A. Venegas-Vega, N. E. Verbeek, E. J. Richmond, O. Kirbiyik, P. Ranganath, A. Singh, K. Godbole, F. A. M. Ali, C. Alves, J. Mayerle, M. M. Lerch, H. Witt and M. Zenker (2014). Mutations in the Human UBR1 Gene and the Associated Phenotypic Spectrum. **Human Mutation** 35(5): 521-531.

Swanson, R., M. Locher and M. Hochstrasser (2001). A conserved ubiquitin ligase of the nuclear envelope/endoplasmic reticulum that functions in both ER-associated and Mat alpha 2 repressor degradation. **Genes & Development** 15(20): 2660-2674.

Takahashi, Y., A. Shinoda, H. Kamada, M. Shimizu, J. Inoue and R. Sato (2016). Perilipin2 plays a positive role in adipocytes during lipolysis by escaping proteasomal degradation. **Scientific Reports** 6: 20975.

Takara Bio Inc. (No Date). How to design sgRNA sequences. Available at: https://www.takarabio.com/learning-centers/gene-function/gene-editing/gene-editing-tools-and-information/how-to-design-sgrna-sequences.

Takase, T., Y. Nishiyama, H. Tanihigashi, Y. Ogura, Y. Miyazaki, Y. Yamada and T. Kiyosue (2011). LOV KELCH PROTEIN2 and ZEITLUPE repress *Arabidopsis* photoperiodic flowering under non-inductive conditions, dependent on FLAVIN-BINDING KELCH REPEAT F-BOX1. **Plant Journal** 67(4): 608-621.

Tanveer, M., B. Shahzad, A. Sharma, S. Biju and R. Bhardwaj (2018). 24-Epibrassinolide; an active brassinolide and its role in salt stress tolerance in plants: A review. **Plant Physiology and Biochemistry** 130: 69-79.

Tasaki, T. and Y. T. Kwon (2007). The mammalian N-end rule pathway: new insights into its components and physiological roles. **Trends in Biochemical Sciences** 32(11): 520-528.

Tasaki, T., S. M. Sriram, K. S. Park and Y. T. Kwon (2012). The N-End Rule Pathway. **Annual Review of Biochemistry** 81: 261-289.

Taylor, S., R. Scott, R. Kurtz, C. Fisher, V. Patel and F. Bizouarn (2010). A Practical Guide to High Resolution Melt Analysis Genotyping. **Bio-Rad Laboratories, Inc.** Tech Note 6004

Theodoraki, M. A., N. B. Nillegoda, J. Saini and A. J. Caplan (2012). A Network of Ubiquitin Ligases Is Important for the Dynamics of Misfolded Protein Aggregates in Yeast. **Journal of Biological Chemistry** 287(28): 23911-23922.

Thimmappa, R., K. Geisler, T. Louveau, P. O'Maille and A. Osbourn (2014). Triterpene biosynthesis in plants. **Annual Review of Plant Biology** 65: 225-257.

Tian, T., Y. Liu, H. Yan, Q. You, X. Yi, Z. Du, W. Xu and Z. Su (2017). agriGO v2.0: a GO analysis toolkit for the agricultural community, 2017 update. **Nucleic Acids Research** 45(W1): W122-W129.

Timms, R. T. and I. Koren (2020). Tying up loose ends: the N-degron and C-degron pathways of protein degradation. **Biochemical Society Transactions** 48(4): 1557-1567.

Timms, R. T., Z. Zhang, D. Y. Rhee, J. W. Harper, I. Koren and S. J. Elledge (2019). A glycine-specific N-degron pathway mediates the quality control of protein N-myristoylation. **Science** 365(6448): eaaw4912.

Toyama, B. H. and M. W. Hetzer (2013). Protein homeostasis: live long, won't prosper. **Nature Reviews.Molecular Cell Biology** 14(1): 55-61.

Travers, K. J., C. K. Patil, L. Wodicka, D. J. Lockhart, J. S. Weissman and P. Walter (2000). Functional and genomic analyses reveal an essential coordination between the unfolded protein response and ER-associated degradation. **Cell** 101(3): 249-258.

Trotter, E. W., C. M. F. Kao, L. Berenfeld, D. Botstein, G. A. Petsko and J. V. Gray (2002). Misfolded proteins are competent to mediate a subset of the responses to heat shock in Saccharomyces cerevisiae. **Journal of Biological Chemistry** 277(47): 44817-44825.

Van Damme, P., R. Evjenth, H. Foyn, K. Demeyer, P.-J. De Bock, J. R. Lillehaug, J. Vandekerckhove, T. Arnesen and K. Gevaert (2011). Proteome-derived Peptide Libraries Allow Detailed Analysis of the Substrate Specificities of N-alpha-acetyltransferases and Point to hNaa10p as the Post-translational Actin N-alpha-acetyltransferase. **Molecular & Cellular Proteomics** 10(5): M110.004580

Van Damme, P., T. V. Kalvik, K. K. Starheim, V. Jonckheere, L. M. Myklebust, G. Menschaert, J. E. Varhaug, K. Gevaert and T. Arnesen (2016). A Role for Human N-alpha Acetyltransferase 30 (Naa30) in Maintaining Mitochondrial Integrity. **Molecular & Cellular Proteomics** 15(11): 3361-3372.

Varland, S., C. Osberg and T. Arnesen (2015). N-terminal modifications of cellular proteins: The enzymes involved, their substrate specificities and biological effects. **Proteomics** 15(14): 2385-2401.

Varshavsky, A. (2011). The N-end rule pathway and regulation by proteolysis. **Protein Science** 20(8): 1298-1345.

Varshavsky, A. (2019). N-degron and C-degron pathways of protein degradation. **Proceedings of the National Academy of Sciences of the United States of America** 116(2): 358-366.

Vembar, S. S. and J. L. Brodsky (2008). One step at a time: endoplasmic reticulum-associated degradation. **Nature Reviews Molecular Cell Biology** 9(12): 944-957.

Vierstra, R. D. (2009). The ubiquitin-26S proteasome system at the nexus of plant biology. **Nature Reviews Molecular Cell Biology** 10(6): 385-397.

Vijay-Kumar, S., C. E. Bugg, K. D. Wilkinson, R. D. Vierstra, P. M. Hatfield and W. J. Cook (1987). Comparison of the three-dimensional structures of human, yeast, and oat ubiquitin. **The Journal of Biological Chemistry** 262(13): 6396-6399.

Voytas, D. F. (2013). Plant Genome Engineering with Sequence-Specific Nucleases. **Annual Review of Plant Biology** 64: 327-350.

Vriet, C., E. Russinova and C. Reuzeau (2013). From Squalene to Brassinolide: The Steroid Metabolic and Signaling Pathways across the Plant Kingdom. **Molecular Plant** 6(6): 1738-1757.

Wadas, B., J. Borjigin, Z. Huang, J.-H. Oh, C.-S. Hwang and A. Varshavsky (2016). Degradation of Serotonin N-Acetyltransferase, a Circadian Regulator, by the N-end Rule Pathway. **Journal of Biological Chemistry** 291(33): 17178-17196.

Waterhouse, P. M., H. W. Graham and M. B. Wang (1998). Virus resistance and gene silencing in plants can be induced by simultaneous expression of sense and antisense RNA. **Proceedings of the National Academy of Sciences of the United States of America** 95(23): 13959-13964.

Weber, H., L. Borisjuk and U. Wobus (2005). Molecular physiology of legume seed development. **Annual Review of Plant Biology** 56: 253-279.

Weber, H. and H. Hellmann (2009). *Arabidopsis thaliana* BTB/POZ-MATH proteins interact with members of the ERF/AP2 transcription factor family. **FEBS Journal** 276(22): 6624-6635.

Weigel, D. and J. Glazebrook (2006). Transformation of agrobacterium using the freeze-thaw method. **CSH protocols** 2006(7): 1031-1036

Weits, D. A., B. Giuntoli, M. Kosmacz, S. Parlanti, H.-M. Hubberten, H. Riegler, R. Hoefgen, P. Perata, J. T. van Dongen and F. Licausi (2014). Plant cysteine oxidases control the oxygen-dependent branch of the N-end-rule pathway. **Nature Communications** 5: 3425-3425.

Weits, D. A., A. B. Kunkowska, N. C. W. Kamps, K. M. S. Portz, N. K. Packbier, Z. N. Venza, C. Gaillochet, J. U. Lohmann, O. Pedersen, J. T. van Dongen and F. Licausi (2019). An apical hypoxic niche sets the pace of shoot meristem activity. **Nature** 569(7758): 714-717.

Wesley, S. V., C. A. Helliwell, N. A. Smith, M. B. Wang, D. T. Rouse, Q. Liu, P. S. Gooding, S. P. Singh, D. Abbott, P. A. Stoutjesdijk, S. P. Robinson, A. P. Gleave, A. G. Green and P. M. Waterhouse (2001). Construct design for efficient, effective and high-throughput gene silencing in plants. **Plant Journal** 27(6): 581-590.

White, M. D., M. Klecker, R. J. Hopkinson, D. A. Weits, C. Mueller, C. Naumann, R. O'Neill, J. Wickens, J. Yang, J. C. Brooks-Bartlett, E. F. Garman, T. N. Grossmann, N. Dissmeyer and E. Flashman (2017). Plant cysteine oxidases are dioxygenases that directly enable arginyl transferase-catalysed arginylation of N-end rule targets. **Nature Communications** 8: 14690.

Wiame, E., G. Tahay, D. Tyteca, D. Vertommen, V. Stroobant, G. T. Bommer and E. Van Schaftingen (2018). NAT6 acetylates the N-terminus of different forms of actin. **FEBS Journal** 285(17): 3299-3316.

Winter, D., B. Vinegar, H. Nahal, R. Ammar, G. V. Wilson and N. J. Provart (2007). An "Electronic Fluorescent Pictograph" Browser for Exploring and Analyzing Large-Scale Biological Data Sets. **Plos One** 2(8): e718.

Woodward, A. W. and B. Bartel (2005). Auxin: Regulation, action, and interaction. **Annals of Botany** 95(5): 707-735.

Wu, X. and T. A. Rapoport (2018). Mechanistic insights into ER-associated protein degradation. **Current Opinion in Cell Biology** 53: 22-28.

Wu, X., M. Siggel, S. Ovchinnikov, W. Mi, V. Svetlov, E. Nudler, M. Liao, G. Hummer and T. A. Rapoport (2020). Structural basis of ER-associated protein degradation mediated by the Hrd1 ubiquitin ligase complex. **Science** 368(6489): eaaz2449.

Xia, Z., A. Webster, F. Du, K. Piatkov, M. Ghislain and A. Varshavsky (2008). Substrate-binding sites of UBR1, the ubiquitin ligase of the N-end rule pathway. **Journal of Biological Chemistry** 283(35): 24011-24028.

Xu, F., Y. Huang, L. Li, P. Gannon, E. Linster, M. Huber, P. Kapos, W. Bienvenut, B. Polevoda, T. Meinnel, R. Hell, C. Giglione, Y. Zhang, M. Wirtz, S. Chen and X. Li (2015). Two N-Terminal Acetyltransferases Antagonistically Regulate the Stability of a Nod-Like Receptor in *Arabidopsis*. **Plant Cell** 27(5): 1547-1562.

Xu, K., X. Xu, T. Fukao, P. Canlas, R. Maghirang-Rodriguez, S. Heuer, A. M. Ismail, J. Bailey-Serres, P. C. Ronald and D. J. Mackill (2006). Sub1A is an ethylene-response-factor-like gene that confers submergence tolerance to rice. **Nature** 442(7103): 705-708.

Xu, P., D. M. Duong, N. T. Seyfried, D. Cheng, Y. Xie, J. Robert, J. Rush, M. Hochstrasser, D. Finley and J. Peng (2009). Quantitative Proteomics Reveals the Function of Unconventional Ubiquitin Chains in Proteasomal Degradation. **Cell** 137(1): 133-145.

Xu, Z., R. Payoe and R. P. Fahlman (2012). The C-terminal Proteolytic Fragment of the Breast Cancer Susceptibility Type 1 Protein (BRCA1) Is Degraded by the N-end Rule Pathway. **Journal of Biological Chemistry** 287(10): 7495-7502.

Yamano, K. and R. J. Youle (2013). PINK1 is degraded through the N-end rule pathway. **Autophagy** 9(11): 1758-1769.

Yang, J., K. Kamide, Y. Kokubo, S. Takiuchi, C. Tanaka, M. Banno, Y. Miwa, M. Yoshii, T. Horio, A. Okayama, H. Tomoike, Y. Kawano and T. Miyata (2005). Genetic variations of regulator of G-protein signaling 2 in hypertensive patients and in the general population. **Journal of Hypertension** 23(8): 1497-1505.

Zattas, D., D. J. Adle, E. M. Rubenstein and M. Hochstrasser (2013). N-terminal acetylation of the yeast Derlin Der1 is essential for Hrd1 ubiquitin-ligase activity toward luminal ER substrates. **Molecular Biology of the Cell** 24(7): 890-900.

Zattas, D., J. M. Berk, S. G. Kreft and M. Hochstrasser (2016). A Conserved C-terminal Element in the Yeast Doa10 and Human MARCH6 Ubiquitin Ligases Required for Selective Substrate Degradation. **Journal of Biological Chemistry** 291(23): 12105-12118.

Zattas, D. and M. Hochstrasser (2015). Ubiquitin-dependent protein degradation at the yeast endoplasmic reticulum and nuclear envelope. **Critical Reviews in Biochemistry and Molecular Biology** 50(1): 1-17.

Zelcer, N., L. J. Sharpe, A. Loregger, I. Kristiana, E. C. L. Cook, L. Phan, J. Stevenson and A. J. Brown (2014). The E3 Ubiquitin Ligase MARCH6 Degrades Squalene Monooxygenase and Affects 3-Hydroxy-3-Methyl-Glutaryl Coenzyme A Reductase and the Cholesterol Synthesis Pathway. **Molecular and Cellular Biology** 34(7): 1262-1270.

Zenker, M., J. Mayerle, M. M. Lerch, A. Tagariello, K. Zerres, P. R. Durie, M. Beier, G. Hulskamp, C. Guzman, H. Rehder, F. A. Beemer, B. Hamel, P. Vanlieferinghen, R. Gershoni-Baruch, M. W. Vieira, M. Dumic, R. Auslender, V. L. Gil-da-Silva-Lopes, S. Steinlicht, M. Rauh, S. A. Shalev, C. Thiel, A. Winterpacht, Y. T. Kwon, A. Varshavsky and A. Reis (2005). Deficiency of UBR1, a ubiquitin ligase of the N-end rule pathway, causes pancreatic dysfunction, malformations and mental retardation (Johanson-Blizzard syndrome). **Nature Genetics** 37(12): 1345-1350.

Zhang, J., B. Wei, R. Yuan, J. Wang, M. Ding, Z. Chen, H. Yu and G. Qin (2017). The *Arabidopsis* RING-Type E3 Ligase TEAR1 Controls Leaf Development by Targeting the TIE1 Transcriptional Repressor for Degradation. **Plant Cell** 29(2): 243-259.

Zhang, X., R. Henriques, S.-S. Lin, Q.-W. Niu and N.-H. Chua (2006). Agrobacterium-mediated transformation of *Arabidopsis thaliana* using the floral dip method. **Nature Protocols** 1(2): 641-646.

Zhang, Z., K. Kulkarni, S. J. Hanrahan, A. J. Thompson and D. Barford (2010). The APC/C subunit Cdc16/Cut9 is a contiguous tetratricopeptide repeat superhelix with a homo-dimer interface similar to Cdc27. **Embo Journal** 29(21): 3733-3744.

Zhao, H., H. Zhang, P. Cui, F. Ding, G. Wang, R. Li, M. A. Jenks, S. Lu and L. Xiong (2014). The Putative E3 Ubiquitin Ligase ECERIFERUM9 Regulates Abscisic Acid Biosynthesis and Response during Seed Germination and Postgermination Growth in *Arabidopsis*. **Plant Physiology** 165(3): 1255-1268.

Zischewski, J., R. Fischer and L. Bortesi (2017). Detection of on-target and off-target mutations generated by CRISPR/Cas9 and other sequence-specific nucleases. **Biotechnology Advances** 35(1): 95-104.

<u>Appendix</u>

Appendix 1 – List of primers used

Primer Number	Primer Name	Primer Sequence
1	DOA10B Exon 7F	GGACTCAAACACGACGATCT
2	DOA10B Exon 7R	AGCAAATCGGTTCTTCCCGTT
3	DOA10B Exon 1F	GGCGGAAGACAAACTCGTTG
4	DOA10B Exon 1R	TGCAAATCTCGCAATGGTTG
9	DOA10A Cloning F	CACCATGGAGATTTCCCCGGC
10	DOA10B Cloning F	CACCATGGAGATTTCTCCGGCG
11	DOA10B Cloning R (Open)	CTCGAGATCTTCAGTGAAATCG
12	eG/YFP R	GCTGAACTTGTGGCCGTTTA
13	GWB433/533 (GUS) R	TGCCCAACCTTTCGGTATAA
14	Act7 exonJ F	ATGGCCGATGGTGAGGATAT
15	Act7 exonJ R	GAGCACAATACCGGTTGTACG
16	gDOA10A Cloning F	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCGCATAATATACAAAAGGTCGTCC
18	DOA10A internal 4 F	TCCGTGCTGGGAATGTCA
19	DOA10B internal 3 F	TATGTTTCTTGTTAGTACAACCTTCATG
21	DOA10A flanking primer F	AATTTCTCCTCTGGCAAGCTC
22	DOA10A flanking primer R	AAGAGTCACCCATGCAACAAG
23	GABI-KAT T-DNA border primer	GGGCTACACTGAATTGGTAGCTC
24	DOA10B flanking primer F	GGGGTGTCTCCTTAAAAGCAC
25	DOA10B flanking primer R	TTTCCTTTTCTTTTGCGTGAG
26	SAIL T-DNA border primer	TGGTTCACGTAGTGGGCCATCG
27	Naa20 flanking primer F	TTTGTGTGGAAGAAAACGAAAC
28	Naa20 flanking primer R	TCTTGAGAATCCAACGTAAGTGAC
29	SALK T-DNA border primer	ATTTTGCCGATTTCGGAAC
37	ScDOA10 5' F	AGTAGTTTTTCCAACGAAATAGCCT
38	ScDOA10 5' R	GTTTTGAATTGAATGGGATAGTGAC
39	ScDOA10 3' F	TCTTAACCGATGGTATACTAGCCTG
40	ScDOA10 3' R	CTCGCACGCATAGTTTTATATTTTT
45	KanMX R	CTGCAGCGAGGAGCCGTAAT
46	KanMX F	TGATTTTGATGACGAGCGTAAT
47	D10B_CRISPR1F	ATTGACACGCGCACGGATGTCTCA
48	D10B_CRISPR1R	AAACTGAGACATCCGTGCGCGTGT
49	D10B_CRISPR2F	ATTGACCATGGCACGAATTCTTGA
50	D10B_CRISPR2R	AAACTCAAGAATTCGTGCCATGGT
51	D10B_CRISPR3F	ATTGGGCATCATCAAGATCCAAGG
52	D10B_CRISPR3R	AAACCCTTGGATCTTGATGATGCC
53	D10B_CRISPR4F	ATTGCTACGCTGTCTCCCACACGG
54	D10B_CRISPR4R	AAACCCGTGTGGGAGACAGCGTAG
63	SS42	TCCCAGGATTAGAATGATTAGG
65	SS102	CACCATGTTATCACATCAATCC
78	ACT7 qPCR F	CTGGAATGGTGAAGGCTGGT
79	ACT7 qPCR R	GTGCCTAGGACGACCAACAA
92	MB498_naa20cloF	CACCATGACGACGATACGTAGATTCAG
93	MB499_naa20cloRO	ATCATACTCTAATTCATCAGGTGTGAT
100	D10AqPCR Exon 1R	GCACAAAACTCAACCGCAGA
101	D10AqPCR Exon 8F	AATCTGGACTAGACTGGTAATGCT

102	D10AqPCR Exon 8R	GGCACGCATAAAGCCGTTAG
105	D10B_T1SceenF	TTTGCCAATCACCGGAGGAA
106	D10B_T1SceenR	TTACGGCGATTGAGCCAGAG
107	D10B_T3SceenF	AGAGACTACCATGGCACGAA
108	D10B_T3SceenR	GCGTTGAAAGGCATCATCAAGA
110	ScUBC6_F	GATACTTGGAATCCTGGCTGGTCTGTCTC
111	ScUBC6_R	AAAGGGTCTTCTGTTTCATCACCTGTATTTGC
112	SQE1 Cloning F	CACCATGGAGTCACAATTATGGAATTGG
121	SQE1qPCR_E7_F	CACTCTGCGAAAGCCAGTTG
122	SQE1qPCR_E7_R	GCAAGCTTCCCTCATCTCGT
131	SQE2qPCR_E4_F	GTGGGATGACCGTTGCATTG
132	SQE2qPCR_E4_R	GGTGGAAGCTACAGGTTTTCG
135	SQE3qPCR_E5_F	TTCGCAGTTGCGATTTTCGG
136	SQE3qPCR_E5_R	TCCCTGAAGCACTCGAGATT
139	SQE4qPCR_E7_F	CATCCGCAAGCCAATGTCAG
140	SQE4qPCR_E7_R	GTCGCATTCCCTCTCTTGCT
143	SQE5qPCR_E8_F	ACTTTTTGGTTTGGCTATGAAAATG
144	SQE5qPCR_E8_R	GGCGTTGACTGGAAACAACA
147	SQE6qPCR_E8_F	TTTGGCATAGCCTCAAGC
148	SQE6qPCR_E8_R	GGCTGCATTTGCTGGAAACA
152	UPR_CNX1qPCR_F	TTTGGTGGCAAAAAGGCGG
153	UPR_CNX1qPCR_R	CCCCACTCTTTGAACTCTCCG
154	UPR_BIP1/2qPCR_F	CCACCGGCCCCAAGAG
155	UPR_BIP1/2qPCR_R	GGCGTCCACTTCGAATGTG
156	UPR_bZIP60_F	GCTGAAAACCAGTCTCTACGTT
157	UPR_bZIP60_R	AAGCAGGGAACCCAACAG
158	UPR_bZIP60(s)_R	AGCAGGGAACCCAACAGCAGACT
159	SQE1-Myc_qPCR_F	GCGGAAGGAGTTAGGCAGAT
160	SQE1-Myc_qPCR_R	ACCCGCTGTTATCAACCACTT
200	DOA10A Cloning R (Open)	GGGGACCACTTTGTACAAGAAAGCTGGGTCAGCTTCTTGTTGGATTGCAC
201	DOA10B_gCloF	GTGGAAAGCTCAAGAAATCGGTACTC
202	D10Brnai1CloF	CACCTCTCCCACACGGTTGGATTG
203	D10Brnai1CloR	GCAAAGCAAACGCCAAGAGA
204	D10Brnai2CloF	CACCCGAGAACGCTCCAGAGAGAC
205	D10Brnai2CloR	AACGCTAGTGCAAGAGGTCC
206	D10Brnai3cloF	CACCGGAGACACCCCGAGTTCTTG
207	D10Brnai3cloR	CAATCCAACCGTGTGGGAGA
208	D10Brnai4cloF	CACCCCGGCGAAGCTGTTACTACA
209	D10Brnai4cloR	TGGGAGACAGCGTAGGAGAA
210	DO10A internal 3 F	CTCTGGCAAGCTCACTTGTTC
211	ScNaa10 5' F	AATATCATGGAACAAATGATACGCT
212	ScNaa10 5' R	CTTCACGAGAACGTATCCAACTAAT
213	ScNaa10 3' F	TCCTTGCATGTGAGACAATCTAATA
214	ScNaa10 5' F	GAATAGGGTAGCCATTTTCATTTTT
215	ScNaa20 5' F	ACCCATCTTTTCTTGTAGTTTGTTG
216	ScNaa20 5' R	TCTCAATGTTCCACTATCTTTCACA
217	ScNaa20 3' F	TACATGATAATATGGCCAGACCTTT
218	ScNaa20 3' R	GAAATTATAGGGGAAACAGAGGAAA

Appendix 2 – Full length alignments of ScDoa10, AtDOA10A and AtDOA10B protein sequences

i

```
ScDOA10     1 MDVD---------------SDVNVSR------------LRDELHKVANEETDTATFND
AtDOA10A    1 MEISPADSLSISGAAASEVVSEPSVSSSSSSSSPNQASPNPFSNMDPAVSTATGSRYVDD
AtDOA10B    1 MEISPAEDKLVG------------------------------SG---EAVTTEEVSDI

ScDOA10    32 DAPSGATCRICRGEATEDNPLFHPCKCRGSIKYMHESCLLEWVASKNIDISKPGADVKCD
AtDOA10A   61 DEDEEDVCRICRNPGDADNPLRYPCACSGSIKFVHQDCLLQWLNHSN-------ARQCE
AtDOA10B   26 NNKAVDICRICQSPEEPDNPLRHPCACRGSIKYIHSDCILFLWINRRK-------RNHCE

ScDOA10    92 ICHYPIQFKTIYAENMPEKIPFSLLLSKSILTFFEKARLALTIGLAAVLYIIGVPLVWNM
AtDOA10A  113 VCKHPFSFSPVYADNAPSRLPFQEFVVGIAMKACHVLQFFLRLSFVLSVWLLTIPFITFW
AtDOA10B   78 ICKRSYSIVPVYSENAPERLPWHEFLMGLLVRALRFMNL-------ILPWILMVPFNAYC

ScDOA10   152 FGKLYTMMLDGSSPYPGDFLKSLIYGYDQSATPELTTRAIFYQLLQNHSFTSLQFIMIVI
AtDOA10A  173 IWRLAF-----VRT----------HGEAQR---------LFLSHI------STTVILTDC
AtDOA10B  131 FSF---------RP----------WGRESE---------FVNQTV---------FELSLR

ScDOA10   212 LHIALYFQYDMIVREDVFSKMVFHKIGPRLSPKDLKSRLKERFPMMDDRMVEYLAREMRA
AtDOA10A  203 LHGFLLSA----------SIVFI----FLGAT--------------SLRDYF-RHLRE
AtDOA10B  154 FPGLFYTA----------QIVSSATEMVVQME--------------TIRVLL-RR---

ScDOA10   272 HDENRQEQGHDRLNMPAAAADNNNNVINPRNDNVPPQDPNDHRNFENLRHVDELDHDEAT
AtDOA10A  232 LG--GQEERDDDVDR--------------NGA--------------------------
AtDOA10B  184 ------------------------------------------------------------

ScDOA10   332 EEHENNDSDNSLPSGDDSSRILP-GSSSDNEEDEEAEGQQQQQQPEEEADYRDHIEPNPI
AtDOA10A  248 -------RAARRPAGQANRNLAGEGNGEDAGDQGAAVGQIARRNPENVLA---RLD-IQA
AtDOA10B  184 -----------HPE--FLRRMIILENGLKDRDVTGIVLLLA--N---------HLQ-ILC

ScDOA10   391 DMWANRRAQNEFDDLIAAQQNAINRPNAPVFIPPPAQNRAGNVDQDEQDFGAAVGVPPAQ
AtDOA10A  297 ARLEAQV-------------------EQMFD--GLDDADGAEDVPFDELVGM-------
AtDOA10B  219 DWWHDQLLQLPF--------------LHIEQRGPLALAFVPRNTPLHQFGAI-------

ScDOA10   451 ANPDDQGQGPLV--INLKLKLLNVIAYFIIAVVFTAI------YLAISYLFPTFIGFG--
AtDOA10A  328 -------QGPVFHLVENAFTVLASNMIFLGVVIFVPFTLGRIILYHVSWLFAAARGPAVA
AtDOA10B  257 -------RRVFSLLSDNTFAVLAINIYWSFFRVLLPFSIGRVVLVLLRCLPH--------

ScDOA10   501 -----------LLKIYFGIF------------KVILRGLCHLYYLSGAHI-----AYNG
AtDOA10A  381 ASLHLTDTGLSLENITLKSALTAVSNLTSEGQGNGLLGQLTEMMKVNGSELNGANNTLSV
AtDOA10B  302 -------GWIAE------------------------------NASE----------

ScDOA10   532 LTKLVPKVDVAMSW-------ISDHLIHDITYLYNGYTENTMKHSIFIRALPALTTYLT
AtDOA10A  441 ATDLLKGSTVGASKLSDITTLAVGYMFIVFLVFLYLGIIA-------------LIRYAK
AtDOA10B  311 ------------------MAAGDMVIRSVLLACLG---------------------

ScDOA10   584 SVSIVCASSNLVSRGYGRENGM--SNPTRRLIFQILFALKCTFKVFTLFFIELAGFPILA
AtDOA10A  487 ------GEPLTVGRFYGIASIVEAVPSLLRQFLAAMRHLMTMIKVAFLLVIELGVFPLMC
AtDOA10B  328 -------G--------------VFTMSRDTYLTSVRTFLPSVKDTFILSFKLGVLPWLL

ScDOA10   642 GVMLDFSLFCPILASNSRMLWVPSICAIWPPFSLFVYWTIGTLYMYWFAKYIGMIRKNII
AtDOA10A  541 GWWLDVCTVRM---FGKTMSHRVQFLSISPLASSLVHWVVGIMYMLQISIFVSLLR-GVL
AtDOA10B  366 GCWLHFCTFPI---LGKTASHTVEVLSDYPLMA-DKHWLMGTLYLVSALSCMELIQ-KIV

ScDOA10   702 RPGVLFFIRSPEDPNIKILHDSLIHPMSIQLSRICLSMFIYAIFIVLGFGFHTR------
AtDOA10A  597 RPGVLYFLRDPADPNYNPFRDLIDDPVHKHARRVLLSVAVYGSLIVMLVFLPVKLAIRMA
AtDOA10B  421 QKRALWYILDVAEPNYKVTK--------LHLGPILLAFALHGTMVVIVLHLPIKTISLIS
```

```
ScDOA10    756   -IFFPFMLKSNLISVPEAYKPTSIISWKFNTILLTLYFTKRI--LESSSYVKPLLERYWK
AtDOA10A   657   PSIFPLDIS---VSDPFTEIPADMILFQIC---I-PFIIEH---FRLRTTIKSLLRCW--
AtDOA10B   473   QSFFPLQFG---VYEDE--FVFGLLVAYMC---L-VIFGPRWLANLIRPSIRPIVHKW--

ScDOA10    813   TIFKLCSRKLRLSSFILGKDTPTERGHIVYRNLFYKYIAAKNAEWSNQELFTKPKTLEQA
AtDOA10A   705   --FTGVGWALGLIDFLLPRPEDNIGQDN--------------------GNGEPGRQNR-
AtDOA10B   522   --VITISSLLKLSDFLLGEPRKHRANHN--------------------M-------R-

ScDOA10    873   EELFGQVRDVHAYFVPDGVLMRVPSSDIVSRNYVQTMFVPVTKDDKLLKPLDLERIKERN
AtDOA10A   741   -------AQVLQVGGPDRAMAALPVADDPNRS--------------------RLRAGN
AtDOA10B   550   -------LRCLVFGIAEGSMVSLHGS----------------------------------

ScDOA10    933   KRAAGEFGYLDEQNTEYDQYYIVYVPPDFRLRYMTLLGIVWLFASILMLGVTFISQALIN
AtDOA10A   772   VNTGEEYEDDDEQSDSDRYN-------FVVRIILLLLVAWVT--LLLFNSALIVVPVSL
AtDOA10B   569   -QSDTTCEK--DTNEQRDKR-------FMVRIGVMLILASLS--MFLVSTTFMALPILV

ScDOA10    993   FVCSFGFLPVVKLLLGER-NKVYVA-------WKELSDISYSYLNIYYVCVGSVCLSKIA
AtDOA10A   822   GRALFSATPILPITHGIKCNDLYAFVIGTYAFWTTISGARYAI----------------
AtDOA10B   616   GRAFFHSISFFMLSFGLKHDDLCAFWIGFCILRGIYIITCFVY----------------

ScDOA10    1045  KDILHFTEGQNTLDEHAVDENEVEEVEHDIPERDINNAPVNNINNVEEGQGIFMAIFNSI
AtDOA10A   865   ---EHVKSKRTSV-----------------------------------------ILNQI
AtDOA10B   659   ---DHFITGRTDL-----------------------------------------LLNHV

ScDOA10    1105  FDSMLVKYNLMVFIAIMIAVIRTMVSWVVLTDGILACYNYLTIRVFGNSSYTIGNSKWFK
AtDOA10A   880   W---------KWCGIVFKSSVLLAIWVFIIPVLIGLLFELLV--IVPMRVPVDESPVFL
AtDOA10B   674   L---------ME----IRNVLLFSIWISVIPGVLGLLIDLMI--IIPSQVPLGESPVYN

ScDOA10    1165  -YD----ESLLFVVWIISSMVNFGTGY-----KSLKLFFRNRNTSKLNFLKTMALELFKQ
AtDOA10A   928   LYQDWALGLIFLKIWTRLVMIDHMLPIVDDSWRAKFERVREDGFSRLQGLWV-----LR-
AtDOA10B   718   LLHDWLIGVVVLHIWIFLTMITRINCFATVAWREKLQRIRSVTINRLPFTWL-----IR-

ScDOA10    1215  GFLHMVIYVLPTIILSLVFLRD---VSTKQIIDISHGSRSHTLSLNESFPTWTRMQDIYF
AtDOA10A   982   ------EIVFPIV-MKLLTALCVPYVLARGVFP----MLGYPLVVNSAVY--------RF
AtDOA10B   772   ------DVIGSTI-VSLLFTLCVPYVVVNSLFP----ILGFSSAVNLTVQ--------RF

ScDOA10    1272  GLLIALESFTFFFQATVLFIQWFKSTVQNVKDEVYTKGRALENLPDES------------
AtDOA10A   1023  -AWIGCLSVSLFCFCAKRCHVWFRNLHNSIRDDRYLIGRRLHNFGEAALANQNQNQSSED
AtDOA10B   813   -IWPAILALIPIWFSVKLIRDLILYLHKLEFDNRYKVGERLVDFTEDLE----------

ScDOA10          --------------------------
AtDOA10A   1082  AGDGVLIGREGDVDNGLRLRRAIQQEA
AtDOA10B         --------------------------
```

**ii**

		ScDoa10
	AtDOA10A	22.87%
AtDOA10B	31.68%	19.23%

Appendix 2– Amino acid sequence alignments of full length ScDoa10, AtDOA10A and AtDOA10B
(i) Sequence alignments of amino acids. Black shading signifies identical amino acids and grey shading shows similar amino acids. (ii) Percentage identity values between proteins.
Alignments were generated using EMBL-EBI Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/) (Scievers et al. Mcwilliam et al.) with shading added using BoxShade 3.21 (https://embnet.vital-it.ch/software/BOX_form.html).

Appendix 3 – Topology predictions for ScDoa10, AtDOA10A and AtDOA10B
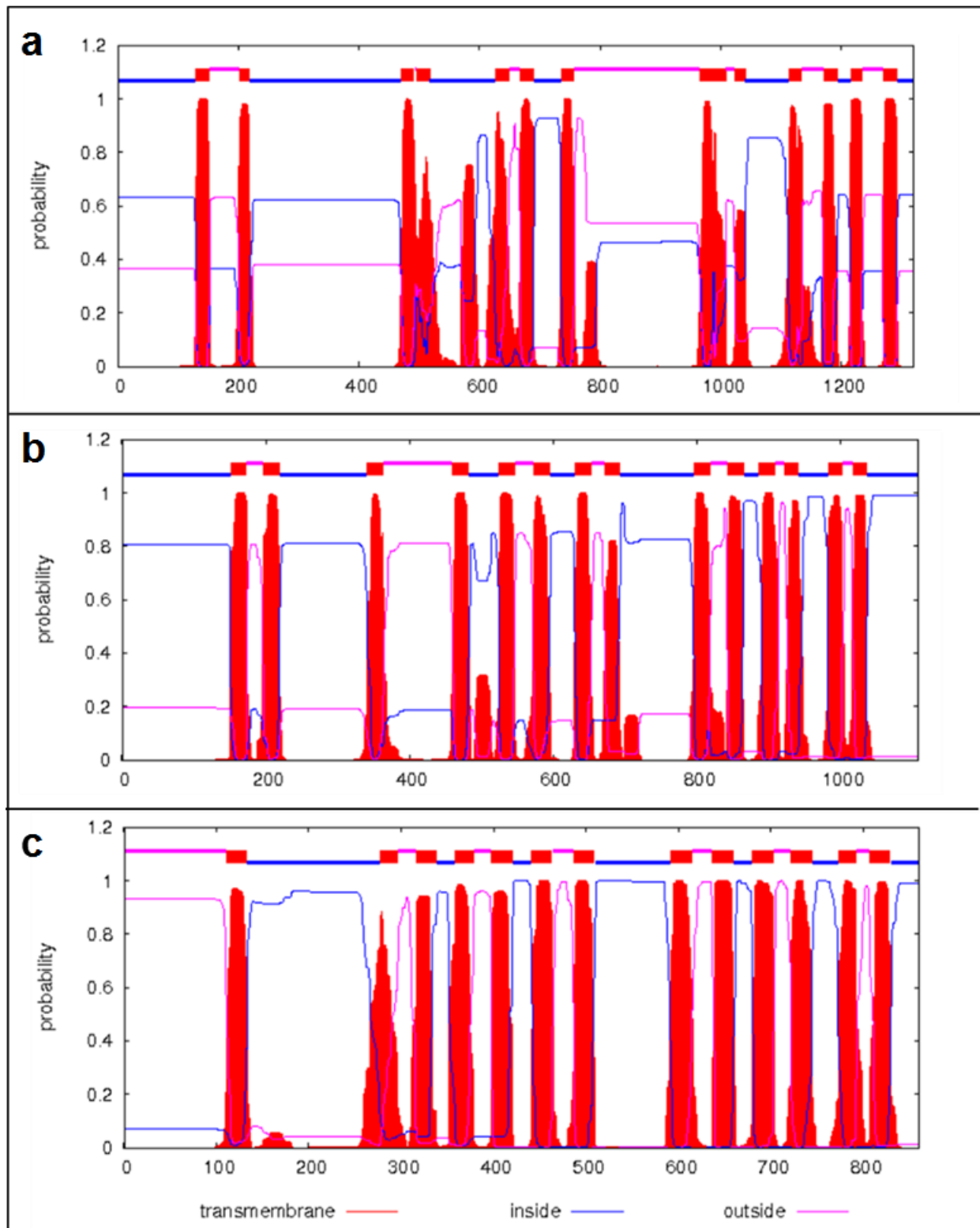
Appendix 3 – TMHMM 2.0 Outputs
Probability plots showing likelihood of residues being within a transmembrane helix for (a) ScDoa10, (b) AtDOA10A and (c) AtDOA10B as predicted by TMHMM 2.0 (http://www.cbs.dtu.dk/services/TMHMM/) (Krogh et al., 2001).

Appendix 4 – UPR-induced increases in AtCNX1 were not affected by *Atdoa10* mutations
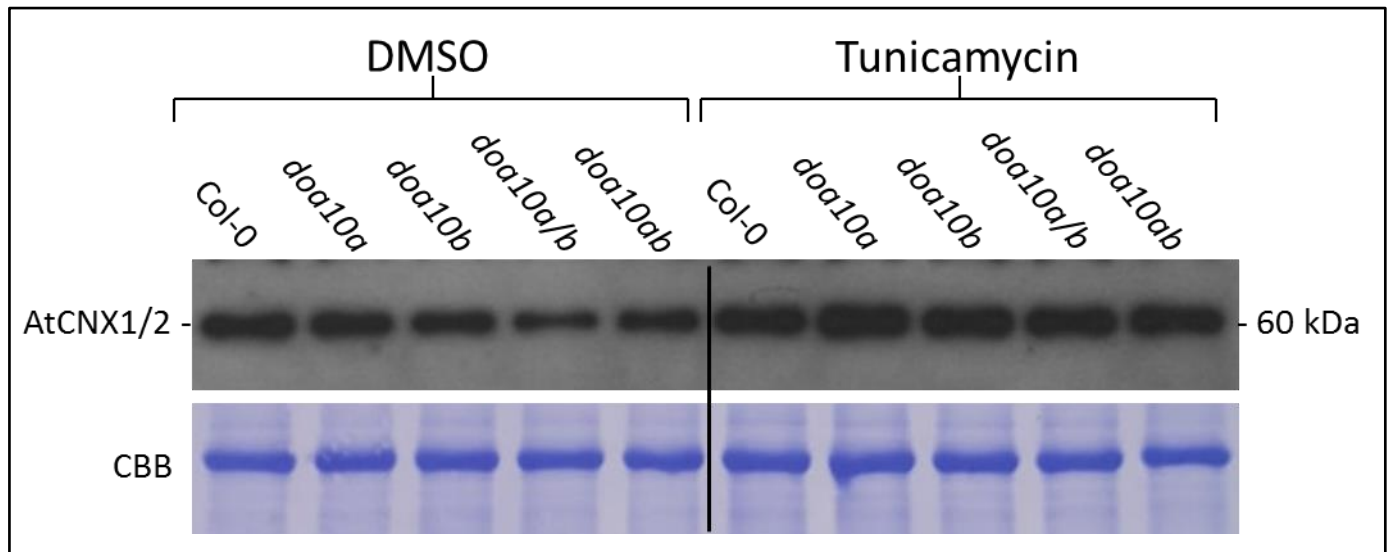
Appendix 4 – The impact of *Atdoa10* mutations on the accumulation of AtCNX1 during the UPR
Levels of AtCNX1, as detected by Western blot, were elevated by induction of the UPR through treatment of seedlings with 5 µg ml^{-1} TM but were not affected by mutations of *AtDOA10A* or *AtDOA10B*. *doa10a/b* is Atdoa10a/b is RNAi line 4-2(2), *doa10ab* is CRISPR line 3(11)-2. Proteins were detected using anti-AtCNX1/2 and antibody (see 2.4.1). CBB = Coomassie Brilliant Blue.

Appendix 5 – The GABI_588A06 T-DNA insertion lies within exon 5 of *AtDOA10A*
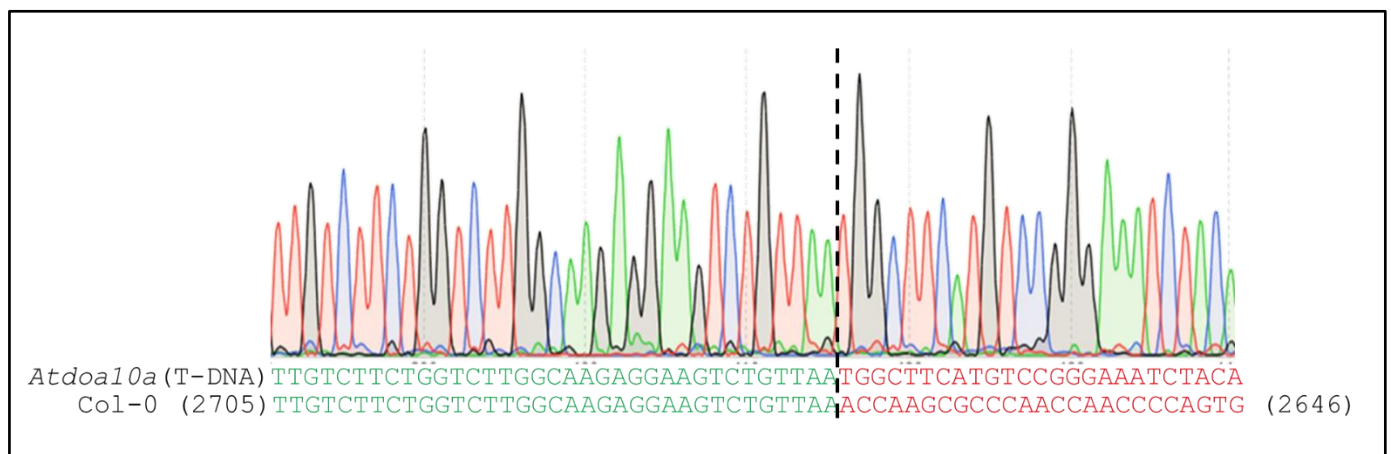
Appendix 5 – Chromatogram of *AtDOA10A* sequence in GABI_588A06
The sequence of *AtDOA10A* in GABI_588A06 (*Atdoa10a* (T-DNA)) contains a T-DNA at a position of 2670 bp (within exon 5). Sequencing in the 3' to 5' direction from exon 6 is consistent with Col-0 (green letters) until position 2670 where the sequences no longer match (red letters).

Appendix 6 – Expression of *AtDOA10B* increased from the T2 to T3 generation for RNAi Construct 4 – Line 9 plants
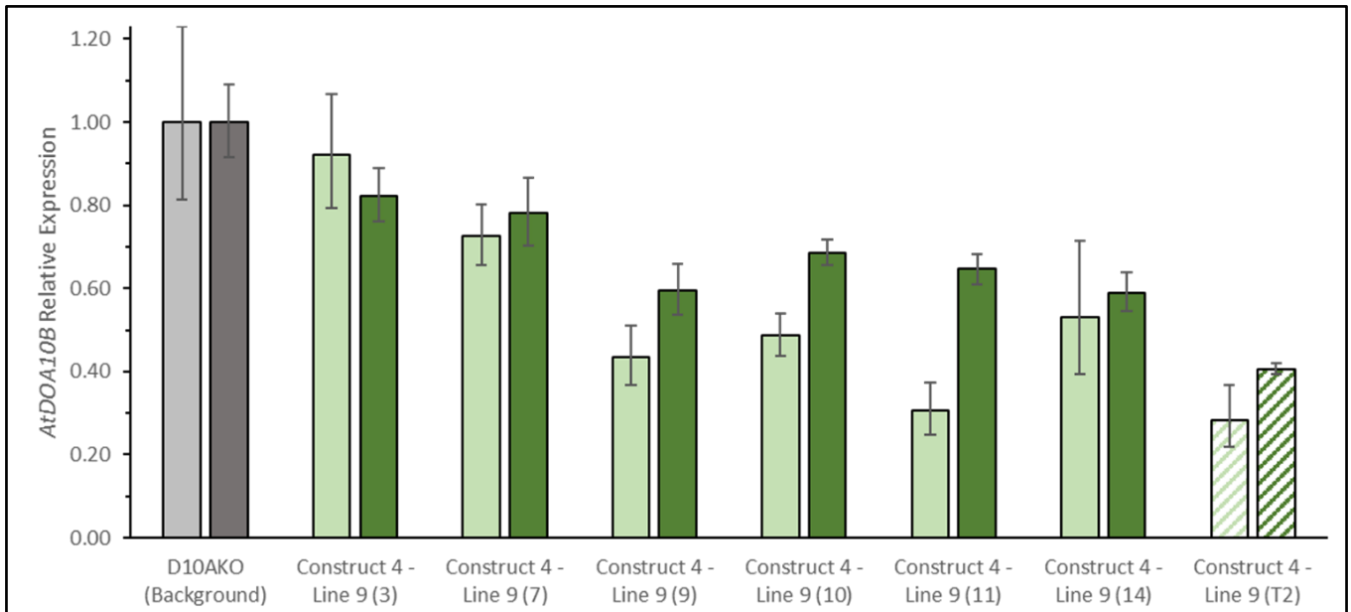
Appendix 6 – Further screening of T3 *AtDOA10B* RNAi lines

Expression levels, as determined by qRT-pCR, of six homozygous T3 lines from Construct 4 – Line 9 h T2 parents (solid green bars) were all greater than that of a heterozygous T2 population (dashed green bars). Pale bars correspond to exon 1 amplicons, dark bars correspond to exon 7 amplicons. Values of relative expression are averages of two to three technical replicates ± standard error.

Appendix 7 – An AtDOA10A-AtSQE1 interaction could not be identified using the yeast mating-based Split-Ubiquitin System
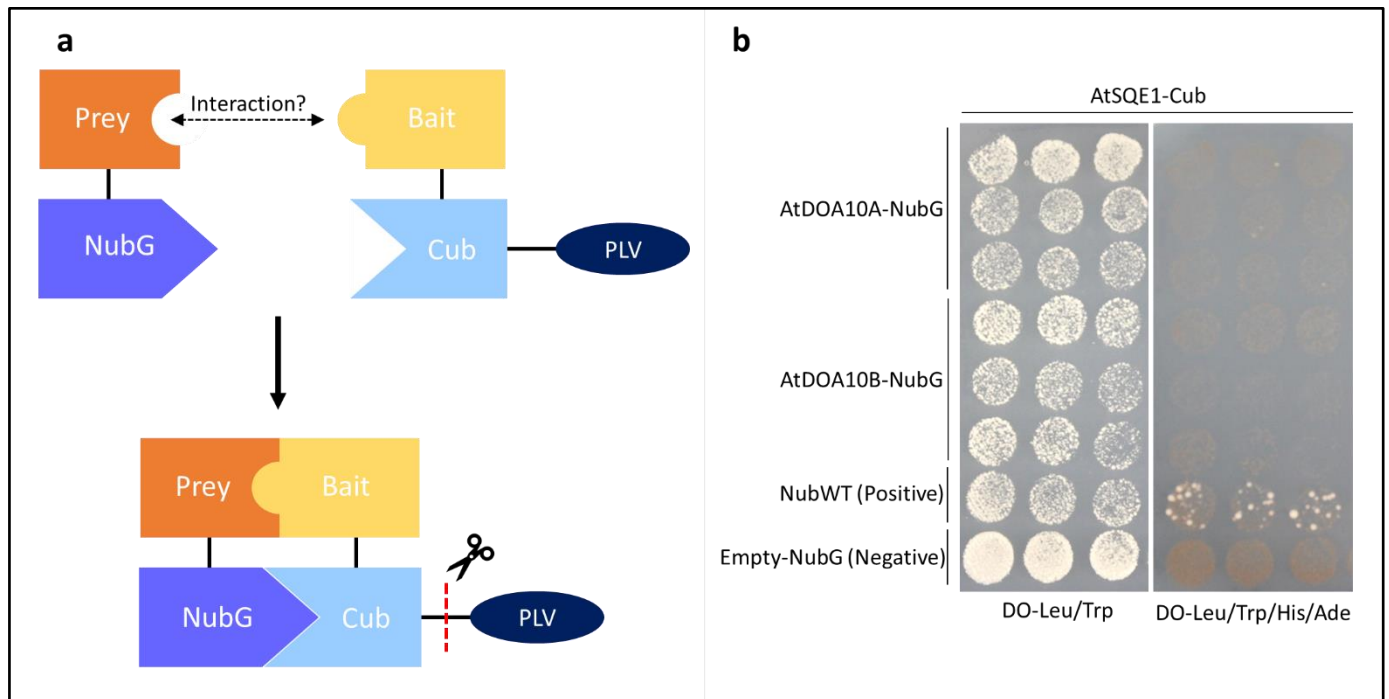
Appendix 7 – Further screening of T3 *AtDOA10B* RNAi lines
(a) Schematic diagram of the yeast mating-based Split-Ubiquitin System (Grefen et al., 2009). An interaction between Prey and Bait proteins allows the recombination of the N-terminal (Nub) and C-terminal (Cub) regions of ubiquitin, which are bound to the Pey and Bait respectively. The Nub bears an I13G mutation to prevent spontaneous recombination with Cub and is therefore denoted NubG. Reconstitution of complete ubiquitin (NubG+Cub) leads to cleavage of PLV (protein A-LexA-VP16) by constitutive deubiquitinases, allowing LexA to drive the expression of reporter genes such as *ADE2* and *HIS3*. (b) Yeast growth was observed on DO-Leu/Trp media, confirming that both the *TRP1*-containing Prey vector and *LEU2*-encoding Bait vector were being expressed but no growth for any combination of Prey (AtDOA10A/AtDOA10B) and Bait (AtSQE1) was observed on DO-Leu/Trp/His/Ade, thereby failing to confirm any interaction. Minimal growth of the positive control NubWT (non-mutated Nub that spontaneously recombines with Cub) was also observed with AtSQE1-Cub.