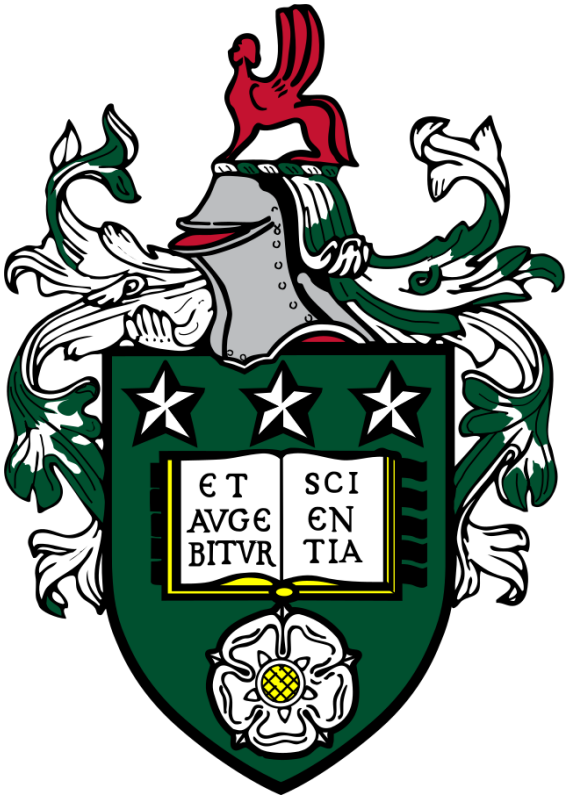


WHIRLY Protein Functions In Plant Development

Rachel Emma Taylor

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

The University of Leeds
Centre for Plant Sciences, Faculty of Biological Sciences

April 2022

Acknowledgements

I would like to thank Professor Christine H. Foyer and Dr. Christopher E. West for their supervision throughout my PhD and who took time to teach me new techniques and gave me advice to aid the completion of my studies. I would also like to thank Dr. Jacques Rouster of Limagrain, France for additional support in my industrial Collaborative Awards in Science and Engineering (iCASE) partnership.

I am thankful to my iCASE studentship from the Biotechnology and Biological Sciences Research Council as part of UK Research and Innovation for funding my project for four years at the University of Leeds as well as the three month paid extension to complete my work in the light of the university closure in 2020.

A deep expression of thanks goes to those who I shared a laboratory space with and who taught me new techniques throughout my PhD research, especially to Dr. Wanda Waterworth who has helped guide my germination and ageing experiments with her expertise. I am truly thankful to Rosalind Latham who has provided much stimulating conversation during long experiments and a great support to me as my close friend.

A special thanks to everyone I have met in the Centre of Plant Sciences over the past four years, who are always friendly faces both in and outside work. I am especially thankful to my close friends, Dr. Richa Yeshvekar and her partner Jonathan Woodward, for their support, love and excellent company. I will forever treasure the many laughs and great conversations we had throughout my PhD. I am grateful for the many Yorkshire countryside walks over the past few years with my friend Emily Aidoo-Baidoe where I have delighted in nature, alongside brilliant company, throughout challenging moments.

My deepest appreciations go to my parents, Nicolette and Stuart Taylor, as well as to my brother Joshua Taylor, who are always there for me with love, support and encouragement. Without my family I would not have been able to complete my studies.

Finally, I would like to express my love and gratitude to Robbie Phillips who has cared for me throughout my thesis writing with excellent food, love, laughs and outstanding company.

Abstract

All plants have at least two WHIRLY (WHY) proteins that are localised to organelles and the nucleus. Arabidopsis has an extra WHY3 protein. WHY proteins have multiple functions in plant growth and defence, including photosynthesis and stress tolerance. The studies reported here were designed to explore the roles of WHY1 and WHY3 in seed viability, germination, and later development of Arabidopsis including flowering timing. The Arabidopsis *Atwhy1*, *Atwhy3* and *Atwhy1why3* mutant lines were first analysed and then this data was used to inform studies in wheat with RNAi *Tawhy* knockdown lines. Germination was similar in all the mutant lines compared to the wild type Arabidopsis. However seed ageing studies revealed a significant reduction in seed viability in the *Atwhy1why3* seeds. Furthermore, analysis of RNA-seq data revealed significant differences in the transcriptome profiles of the *Atwhy1why3* seeds compared to WT. In particular, transcripts encoding heat shock response and drought response related proteins were changed in abundance. Other transcripts indicate changes in phytohormone signalling particularly involving abscisic acid (ABA). Furthermore, at the later stages of plant development the *Atwhy1why3* mutants had reduced biomass, fewer leaves and lower chlorophyll contents than the WT. Wheat mutants lacking WHY1 had a delay in flowering timing compared to the wild type. These data show that WHY proteins have important functions in seeds and in plant development. The findings reported in this thesis thus have wider implications for future food security in crops, particularly in relation to environmental change.

CONTENTS

List of Figures

List of Tables

Table of Abbreviations

Abbreviation	Meaning
ABA	abscisic acid
ANOVA	analysis of variance
cDNA	complementary DNA
CIPK	calcineurin B-like-interacting protein kinase
DEG	differentially expressed gene
DNA	deoxyribonucleic acid
DSB	double strand break
ERE	elicitor response element
ETR	electron transport rate
GA	gibberellic acid
GFP	green fluorescent protein
GO terms	gene ontology terms
HA	hemagglutinin
HSP	heat shock protein
kDa	kilodaltons
LEA	late embryogenesis abundant protein
MADS-box	*MCM1/AGAMOUS/DEFICIENS/SRF* gene family
MIKC MADS-box	MADS-box with M, I, K, C domains
MMBIR	microhomology-mediated break-induced replication
MMEJ	microhomology-mediated end joining
mRNA	messenger ribonucleic acid
NADPH	reduced nicotinamide adenine dinucleotide phosphate
NEP	nuclear-encoded RNA polymerase
PBF	potato nuclear binding factor
PCR	polymerase chain reaction
PEP	plastid-encoded RNA polymerase
Pol1B	DNA polymerase 1B
PR	pathogenesis related
PSI/ PSII	photosystem I/II
qPCR	quantitative PCR
RNA	ribonucleic acid
RNAi	RNA interference
ROS	reactive oxygen species
RuBisCO	ribulose-1,5-bisphosphate carboxylase-oxygenase
SA	salicylic acid
SPAD	soil plant analysis development
TDNA	transfer RNA
UV	ultraviolet
WHY	WHIRLY
WT	wild type

1 INTRODUCTION

1.1 THE DISCOVERY OF WHIRLY

Plants have pathogenesis-related (*PR*) genes to deal with a wide range of microbial and fungal pathogens as well as insect and herbivory attack (Sels *et al.*, 2008). As sessile organisms these are often the first line of defence and as such are highly important to the survival of a plant (Tamaoki *et al.*, 2013). The genes of sub-group *PR-10* encode small (15-18 kDa) intracellular proteins. The regulation of *PR-10a* was characterised in *Solanum tuberosum* (potato) during infection of the oomycete pathogen, *Phytophthora infestans* (Despres *et al.*, 1995). Desveaux *et al.* (2000) characterised the accumulation of the *PR-10a* mRNA in *Solanum tuberosum* (potato) via UV-crosslinking of two potato nuclear binding factors (PBF1 & 2) with high affinity to the elicitor response element (ERE). In these studies, UV-crosslinking to the ERE identified a 24 kDa protein (p24) as the DNA binding component PBF-2 which was confirmed through antibodies specific to p24 which bound to the large peptide of PBF-2 (Desveaux *et al*., 2000). In PBF-2, four p24 molecules were shown to interact in a helix-loop-helix motif to form a quaternary protein structure resembling a garden windmill, or 'whirligig' [Figure 1], hence the interesting WHIRLY (WHY) protein name (Desveaux *et al.*, 2002).

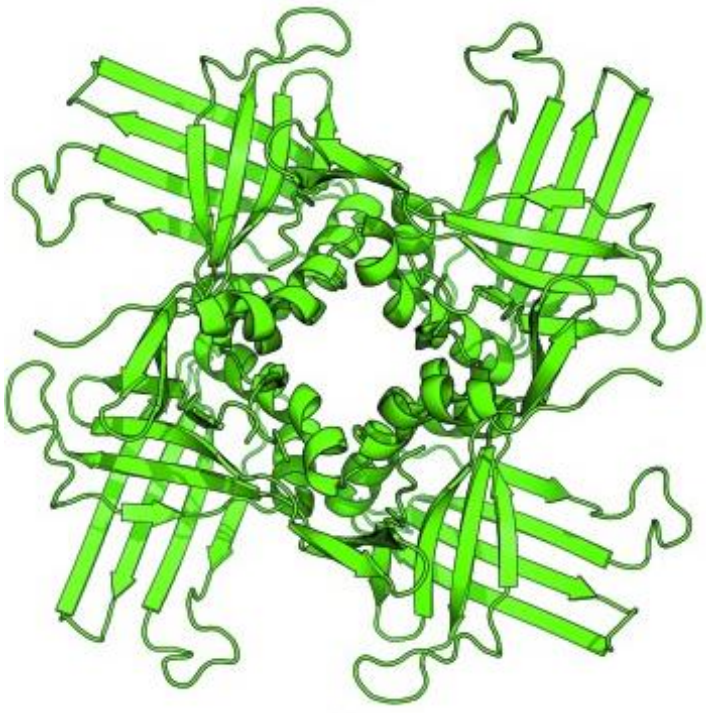


Figure 1: Cartoon representation of the crystal structure of AtWHY1 generated through crystallographic fourfold symmetry of the asymmetric unit tetramers (Cappadocia et al., 2013) using the same DNA crystallographic structure method (Cappadocia *et al.*, 2010).

1.1.1 WHIRLY Family

All plant species to-date appear to have at least 2 *WHY* genes, a dual-targeted (nuclear and plastid) *WHY1* and a mitochondria-localised *WHY2* (Krause *et al*., 2005). A neighbour-joining method of phylogenetic analysis highlighted that most WHY sequences cluster into WHY1 and WHY2 homologous groups (Krause *et al*., 2005). An additional plastid-localised gene called *WHY3* (AT2G02740), has only been identified in Arabidopsis (Krause *et al*., 2005).

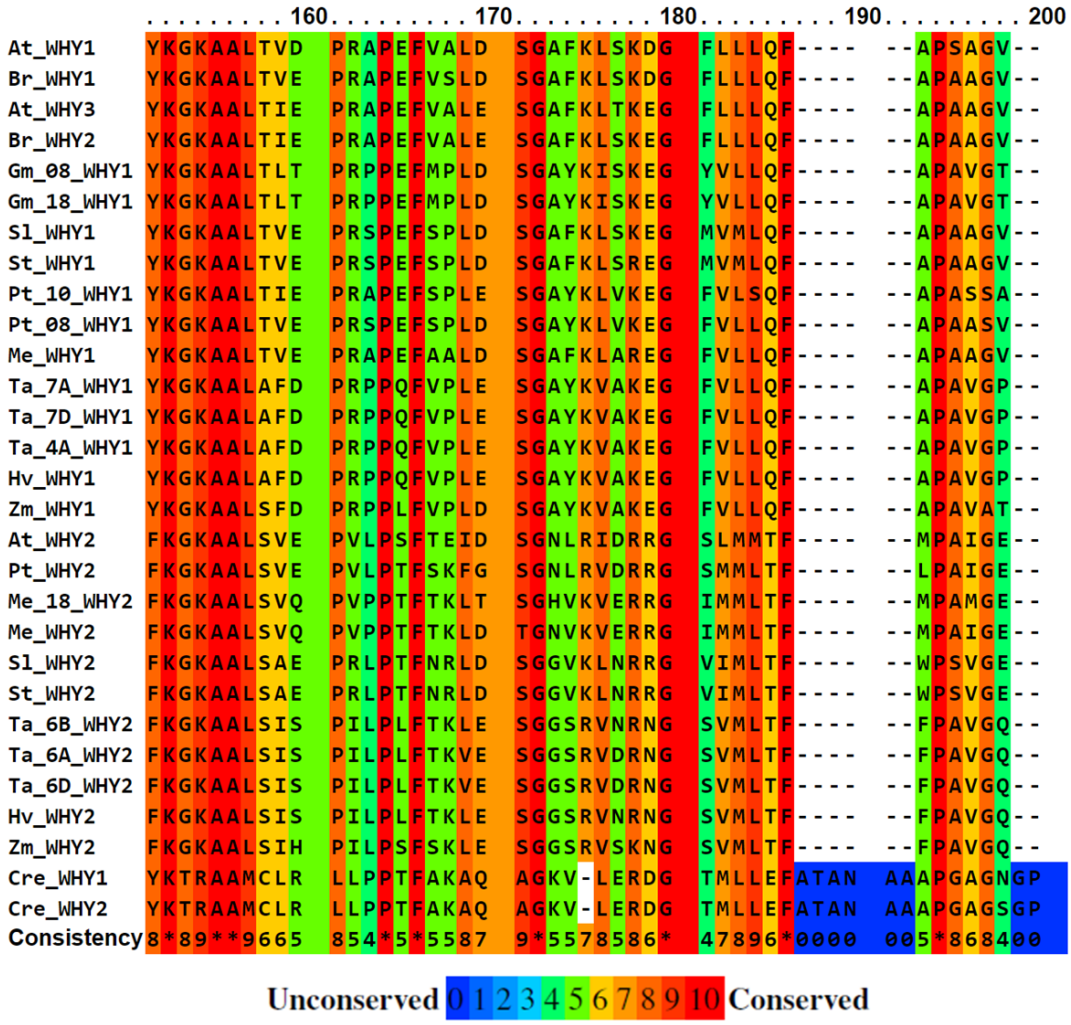


Figure 2: Multiple sequence alignment of WHIRLY protein sequences in a range of species of interest compiled through a Profile ALIgNmEnt (PRALINE) toolkit with homology extended alignment. The conserved KGKAAL DNA binding domain in WHIRLY proteins (Simossis and Heringa, 2005; Bawono and Heringa, 2014). Species are as follows: *Arabidopsis thaliana* (At), *Brassica rapa* (Br), *Glycine max* (Gm), *Hordeum vulgare* (Hv), *Manihot esculenta* (Me), *Populus trichocarpa* (Pt), *Oryza sativa* (Os), *Sorghum bicolor* (Sb), *Solanum lycospericum* (Sl), *Solanum tuberosum* (St), *Triticum aestivum* (Ta), *Zea mays* (Zm), *Chlamydomonas reinhardtii* (Cre) .The scoring scheme works from 0 for the least conserved alignment position, up to 10 for the most conserved alignment position. The colour-coded assignments are scored as conservation of alignment position from unconserved (blue) to highly conserved (red).

Throughout the plant kingdom WHY protein sequences have been found to have a degree of similarity (58% on average) to the StWhy1 protein sequence (Krause *et al*., 2009). Furthermore, the KGKAAL sequence is conserved in all WHY proteins and mutations in this domain were found to prevent all DNA binding but did not affect protein tetramerization (Krause *et al*., 2005). This sequence can be seen at a high level of conservation under a multiple protein sequence analysis of the best characterised species used in WHIRLY research [Figure 2]. The conservation of the KGKAAL domain suggests that all WHY proteins form the same structural arrangement as that observed previously in p24 crystals (Desveaux *et al*., 2002). Crystallographic analysis of the *Arabidopsis thaliana* homologue, AtWHY2 confirmed the structural arrangement and this was applied to other members of the WHIRLY family [Figure 1] (Cappadocia *et al*., 2010). This DNA binding domain allows WHY proteins to bind to ssDNA molecules of differing nucleotide sequences (Grabowski *et al*., 2008). This suggests that WHY could be functional in a range of growth and defence processes, and gives a clearer image of how it may function in both nuclear and plastid processes. The degree to which each WHY protein functions in these organelles has not been wholly demonstrated.

A comparison of the amino acid sequences of the potato p24 of Arabidopsis WHY homologs showed the highest similarity (68%) to AtWHY1 protein, followed by AtWHY2 (42%) and then AtWHY3 (30%) (Krause *et al*., 2005). It is surprising that the AtWHY1 and AtWHY3 protein sequences have differing percentage similarities to the p24 sequence because a previous study showed that the plastid-localised AtWHY1 and AtWHY3 proteins have 77% sequence similarity to each other (Desveaux *et al*., 2004). Furthermore, the prediction programs (PredictNLS) used in the later homology calculation did not show nuclear localisation signals in any of the three AtWHY proteins (Krause *et al*., 2005), despite other studies showing clear nuclear localisation of all WHY3 proteins (Krause *et al*., 2005; Xiong *et al*., 2009; Huang *et al*., 2020). This suggested that the proteins are nuclear localised without a nuclear localisation signal.

Furthermore, there is potential for a degree of functional redundancy between plastid-localised proteins in some species and an analysis of WHY protein sequences across several species found a higher degree of homology between WHY groups in monocots than with other species (Krause *et al*., 2005) More recently, genes homologous to WHY1 were found in unicellular green algae *Chlamydomonas reinhardtii* and *Klebsormidium flaccidum*, as well as common liverwort *Marchantia polymorpha* as pTAC1 (found in the proteome of transcriptionally active chromosomes) (Krupinska *et al*., 2014 and

Kobayashi *et al.*, 2015). A search of protein sequence databases found WHY1 and WHY2 in *C. reinhardtii* [Figure 2], however the KGKAAL domain was not conserved in this species, perhaps suggesting a later evolution of the angiosperm WHIRLY form which has been widely reported. Furthermore, there was no homologous genes found in cyanobacteria *Synechococcus elongates* or chloroplastic *C. paradoxa* and *C. merolae* (Kobayashi *et al.*, 2015).

Together, these results suggest that duplication of the original *WHY* gene may have occurred as an early eukaryotic component of chloroplast nucleoids (Krause *et al.*, 2005). In dicot species, the WHY proteins are more distantly related to each other compared to those in monocots; this may allow for separate species-specific protein functions, although this remains to be fully elucidated (Desveaux *et al.*, 2005). This further indicates a split in WHY function during early angiosperm evolution (Krause *et al.*, 2005).

1.2 DUAL LOCALISATION

The nuclear-encoded, chloroplast-targeted WHY1 protein was the first plant protein to be identified in the nucleus and plastids of the same cell (Krause and Krupinska, 2009). Assays of AtWHY1-GFP fusion and StWHY1-GFP fusion proteins identified sole localisation to plastids but not to the nucleus, this was potentially due to the high molecular mass of the fusion protein (Krause *et al*., 2005, Grabowski *et al*., 2008). Krause *et al.* (2005) used PredictNLS to determine that no nuclear localisation occurred in any of the Arabidopsis WHY proteins, however iPSORT and TargetP programmes, forecasted that the N-terminal of AtWHY2 could represent a putative mitochondrial import sequence, whereas the N-terminal of AtWHY1 and AtWHY3 had high probability as chloroplast transit peptides.

This raised questions whether the protein was indeed dual-localised *in vivo*, however, an immunological analysis of *Hordeum vulgare* (barley) HvWHY1 showed dual-localisation in plastid and nuclear compartments of the same cell (Grabowski *et al*., 2008). Furthermore, a recombinant hemagglutinin (HA)-tagged WHY1 protein that was expressed only in the chloroplasts was translocated to the nuclei of tobacco leaves in response to pathogen attack leading to the expression of *PR* genes (Isemer *et al.*, 2012). These observations suggest that stimuli, such as pathogen infection, were required to activate the translocation of the WHY1 from plastids to nuclei (Krause and Krupinska, 2009).

Chloroplasts from 7 week-old Arabidopsis plants with plastid and nuclear WHY1-HA tags were used in a western blot with anti-HA antibody. This revealed that WHY1 and an additional band of slightly higher molecular weight were detectable in nuclear protein extracts; this band disappeared when treated with phosphatase suggesting that WHY1 may undergo phosphorylation (Ren *et al.,* 2017). A cDNA expression library was created from five and seven week old leaves to identify a possible kinase responsible for the phosphorylation of WHY1 during this time point at which WHY1 had is first expression peak during plant development (Miao *et al.,* 2013; Ren *et al.,* 2017). A yeast two-hybrid screen was performed to look at partners interacting with a WHY1 fused to the downstream of a plastid transit peptide. The WHY1 protein which was used was truncated in the C terminus to remove the activation domain. This screen gave eleven positive colonies which had confirmed interactions with WHY1 and this identified Calcineurin B-Like-Interacting Protein Kinase 14 (CIPK14) (Ren *et al.*, 2017). A biomolecular fluorescence complementation assay was used to confirm the AtCIPK14 and AtWHY1 interaction *in vivo* with co-transformation of protoplasts from the *CIPK14-cmyc-GFPn173* and *WHY1-HA-GFPc155* constructs which resulted in fluorescence in both the nucleus and cytosol (Ren *et al.*, 2017). The mature *AtWHY1* precursor has an N-terminal chloroplast transit peptide for import into chloroplasts, when this is removed the fluorescence for the protein is found only in the nucleus (Krause *et al.*, 2005). Furthermore, western blot analysis showed that overexpression of At*CIPK14* enhanced the abundance of *AtWHY1* in a nuclear isoform whilst a plastid isoform was decreased (Ren *et al*., 2017). In another study, 95% of plants overexpressing *AtCIPK14* had a stay-green phenotype with partial defects in biogenesis of the large ribosomal subunit and decreased plastid isoform of *AtWHY1*; the other 5% had variegated pale-green phenotype and decreased whole-plastid ribosomal RNA (Guan *et al*., 2018). There is clearly a requirement for balance in the amount of WHY1 protein for its localisation in nucleus or plastid. Furthermore, WHIRLY proteins can be localised in various intracellular compartments depending on the stage of development or exposure to stress, as illustrated in Figure 3.

Moreover, there was an increased abundance of higher molecular weight phosphorylated AtWHY1 in transgenic plants overexpressing AtCIPK14, together with a decrease in the abundance of AtWHY1 in the chloroplasts (Guan *et al*., 2018). This suggested that *AtCIPK14* has a regulatory role through reversible phosphorylation of AtWHY1 in the cytosol that prevents it from entering the plastid, whilst allowing re-entry back into the

nucleus in retrograde signalling. WHY1 phosphorylation and organelle distributions regulated by CIPK14 may act as the switch between the plastid and the nucleus through retrograde signalling.

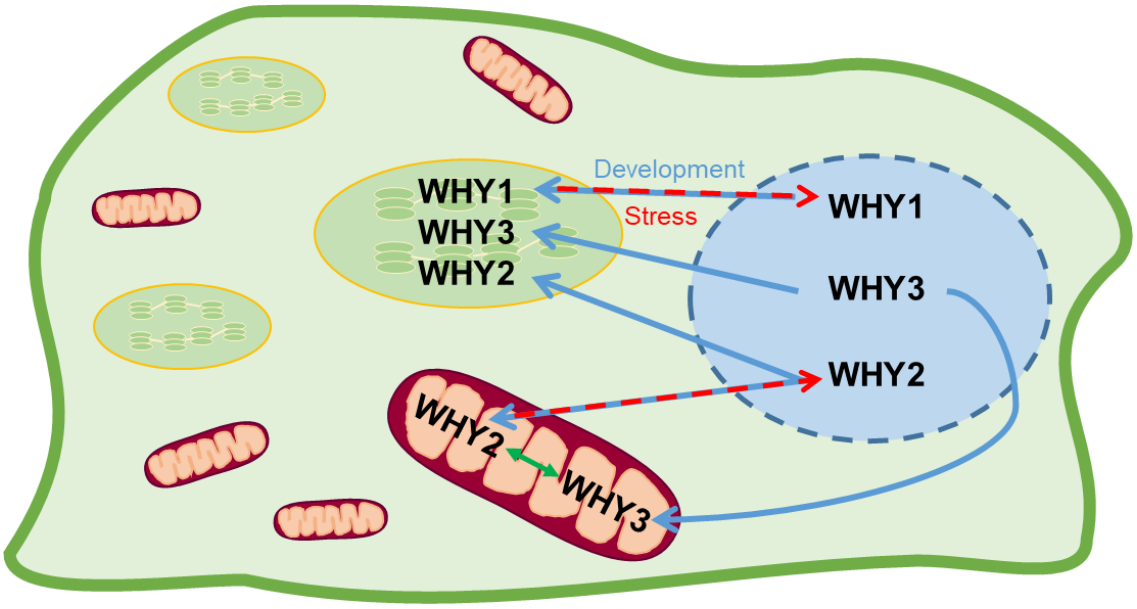


Figure 3: Schematic model of the intracellular localisation of the WHIRLY proteins. The WHY proteins are targeted to organelles and are found in the nucleus (blue) at various stages of plant development. During growth WHY1 and WHY3 are targeted to the chloroplasts (green & yellow) while WHY2 is targeted to the mitochondria (maroon & orange). WHY3 and WHY2 have redundant functions (green double arrow) in the mitochondria. WHY2 has also been found in the chloroplasts (blue arrow). This model infers that WHY1 and WHY2 can relocate to the nucleus in response to environmental stress (red dashed arrow).

1.2.1 Retrograde Signalling

Normal production of reactive oxygen species (ROS) initiates plastid-nucleus retrograde signalling and modified nuclear and chloroplast gene expression patterns to adjust photosynthetic activity due to stresses (Lepage *et al.*, 2013). WHY1 is dual localised to the plastid and nucleus which gives it an ideal position for signal transduction events, such as retrograde signals, to increase defence responses under pathogenic attack (Isemer *et al.*, 2011, Tada *et al.*, 2008). It has been theorised that WHY proteins have a role in plastid to the nucleus retrograde signalling which may explain its wide range in functionality (Foyer *et al.*, 2014).

Immunofluorescence under confocal laser scanning microscopy was used to detect WHY1 in the cytoplasm and the nuclei [Figure 4A]. Further immunogold labelling confirmed the presence of AtWHY1:HA in the nucleus where it seems to associate with

the chromatin [Figure 4D-F]. Previous studies have speculated that it is likely that WHY1 is small enough to enter through the nucleus passively, which is why no fluorescence was detected in the nucleus when a large GFP fusion protein was attached that required a nuclear import localisation signal, (Krause *et al*., 2005). However, it is likely that direct transfer of WHY1 from the plastids to the nuclei through contact sites or stromules is possible although this function this remains to be fully elucidated (Foyer *et al*., 2014; Hanson and Hines, 2018).

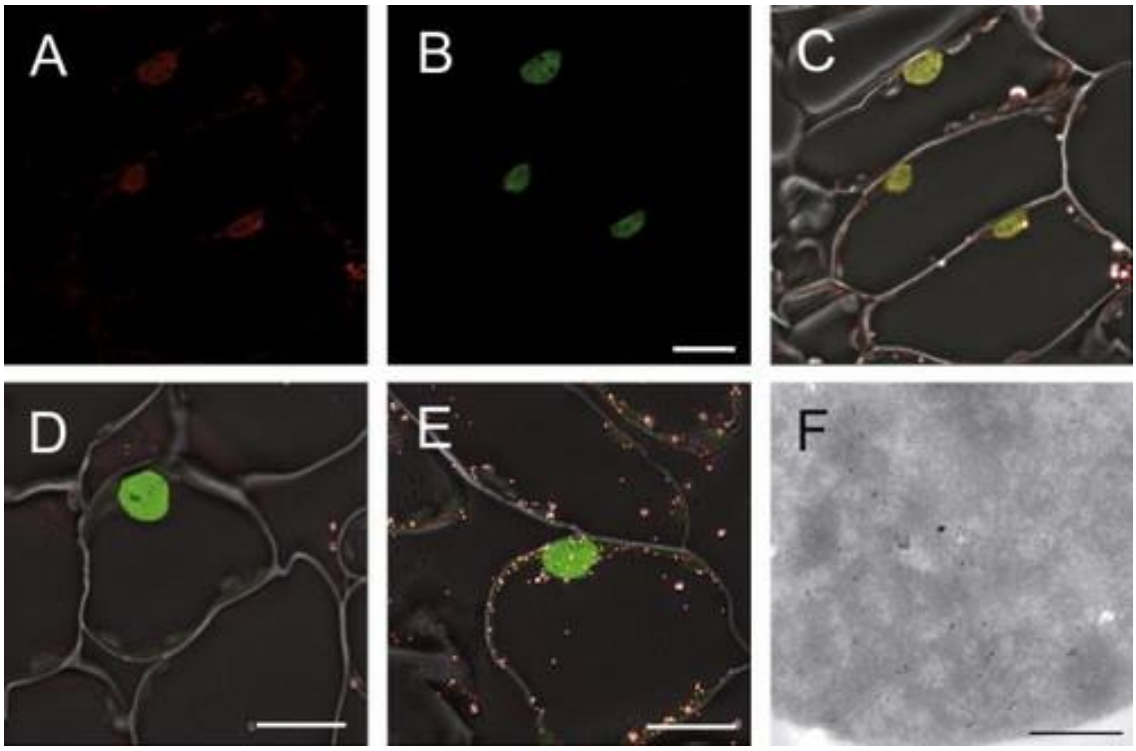


Figure 4: A-C) Immunocytological detection of AtWHY1:HA in leaves of transplastomic tobacco. A) Immunofluorescence detection of HA-tag directed primary antibody and a secondary antibody coupled to Alexa Fluor 568. B) Nuclei counterstained by TO-PRO-1 iodide. C) Overlay of fluorescence of Alexa Fluor 568 and phase contrast. Scale bar: 10 µm. (D and E) Sections with immunogold labelling and silver enhancement as seen by confocal reflection microscopy. Overlays of phase contrast, TO-PRO-1 iodide fluorescence (nuclei) and reflection (silver-enhanced gold particles) were presented. D) Control section with primary antibody without treatment. E) Section treated with primary and secondary antibody. Scale bar: 5 µm. F) Immunogold labelling of AtWHY1:HA in sections detected by transmission electron microscopy, secondary antibody was coupled to 15 nm-gold particles. Scale bar: 500 nm (Isemer et al., 2011).

The 24-oligomers of WHY1 have been isolated from chloroplasts and shown to have the ideal proximity and size to span the distance between the thylakoid and nucleoid [Figure 5] (Cappadocia *et al*., 2011). The close association of WHY1 with the thylakoid membrane raises the possibility that there may be a role in redox function (Foyer *et al*.,

2014, Tada *et al*., 2008). Furthermore, when WHY1 distribution between the nucleus and chloroplasts was altered, senescence and cellular redox homeostasis was also altered (Lin *et al.*, 2019). Once in the chloroplast, in order for WHY1 to travel back to the nucleus, a pathway must be activated possibly through the production of redox (reduction-oxidation reaction) signals such as thioredoxin [Figure 5]. Thioredoxins are a small class of redox proteins which are widely used as integrators of cellular function including linking metabolism and photosynthetic electron transport to gene transcription, translation and protein synthesis and degradation (Foyer and Noctor, 2009). Reversible thiol-disulphide exchange mechanisms may be involved in the formation of the WHY1 multimeric complexes. If this is the case, then the chloroplast thioredoxins were likely regulators of the process. In addition, this process may also involve reactive oxygen species (ROS) that are produced by the photosynthetic electron transport chain. ROS are considered to provide information concerning the redox state of the chloroplast to the nucleus through chloroplast to nucleus retrograde signalling pathways (Foyer *et al.*, 2014, Spetea *et al.*, 2014).

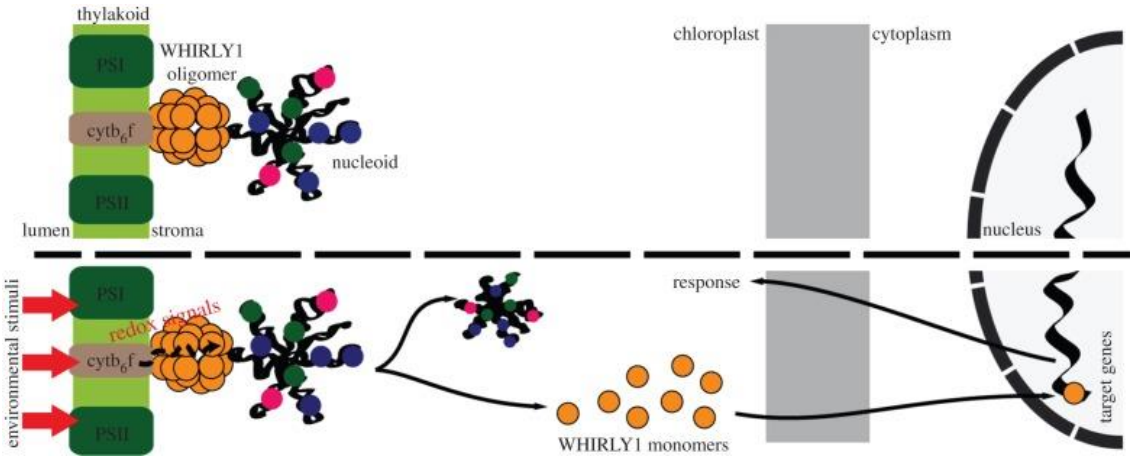


Figure 5: A theoretical model of WHY-dependent perception and transduction of photosynthetic redox signals from the thylakoid electron transport chain. WHY1 monomers form 24-oligomers which forms a bridge between the thylakoid and nucleoid for effective transfer of electrons along the thylakoid electron transport system during photosynthesis. A change in environmental stimuli which alters the photosynthetic redox state and the monomerisation of WHY1 is induced in retrograde signalling from plastid to nucleus (Figure taken from: Foyer *et al.*, 2014).

Plastid-to-nucleus retrograde signalling can be initiated through the production of high levels of ROS in the chloroplast which in turn co-ordinates the expression of the plastid-encoded photosynthetic genes with the expression of nuclear genes which can adjust photosynthetic activity (Escoubas *et* al., 1995; Lepage *et* al., 2013; Huang *et al.*, 2017).

Changes in either the rate of electron transport or the path of electron flow were suggested to regulate the partitioning of WHY1 between the chloroplasts and nuclei through redox-dependent changes in the oligomerisation of the protein (Foyer *et al.*, 2012). This environmental-stimuli driven gene regulation could be regulated though the redox states of the plastoquinone pool and cytochrome b_6f complex [Figure 5] which have been previously implicated in short-term control of the electron transport, gene expression and long-term changes in membrane protein content (Foyer *et al*., 2012). Redox-modulated destabilisation of the larger WHY1 complexes may lead to increased availability of the monomers in the stroma. These could be directly transferred to the nuclei to ensure rapid retrograde signalling and increased expression of nuclear-encoded genes. Furthermore, the leaves of barley RNAi WHY1-knockdown lines (W1-1, W1-7 and W1-9), as described above, had a much greater abundance of plastome-encoded transcripts encoding the thylakoid NADPH complex, the chloroplast RNA polymerase, the cytochrome b_6f complex and chloroplast ribosomes than the wild type (Comadira *et al*., 2015). This thylakoid electron transport system could be activated by developmental triggers such as senescence, and /or environmental changes, for example through abiotic stresses such as light, heat or drought, or biotic factors such as the presence of pathogens. This matches with the wide-reaching functionality of the WHIRLY protein in many different stress factors.

1.3 WHIRLY FUNCTIONS

1.3.1 Species Differences

The single knock-out *Atwhy1* and *Atwhy3* mutant plants were reported to be phenotypically similar to wild type plants and there was largely no phenotypic difference between the *Atwhy1why3* double knock-out mutant and wild type. Presumably these phenotypes were recorded under standard growth conditions, however he reportin of the exact growth conditions was not clear (Maréchal *et al.*, 2009). Comparatively in a controlled growh chamber growth chamber (16 h light, 24 °C/ 8 h dark, 19 °C), there were stark differences in phenotype with *Zmwhy1* mutants which were white in colour and died at the fourth leaf stage compared to the green wild type plants (Prikryl *et al.*, 2008). Furthermore, under the same growth conditions as maize, RNAi WHY1 knockdown barley lines were phenotypically similar to the wild type with similar redox metabolite levels to the wild type, with comparable photosynthesis and senescence rates but the leaves had significantly more chlorophyll and less sucrose (Comadira *et al.*, 2015).

Additionally, in *C. reinhardtii* vegetative cells CreWHY fluorescence was faint in chloroplast nucleoids, but in early zygotes it was intense (Kobayashi *et al.*, 2015). These major differences in reported WHY1 phenotypes means that WHY1 must act in a different manner in different species, this could be to do with the type of photosynthesis (C3 or C4) or where they diverged in evolution - flowering monocots and dicots or non-flowering chloroplastic organisms. The true extent to which members of the WHIRLY differ needs much further investigation.

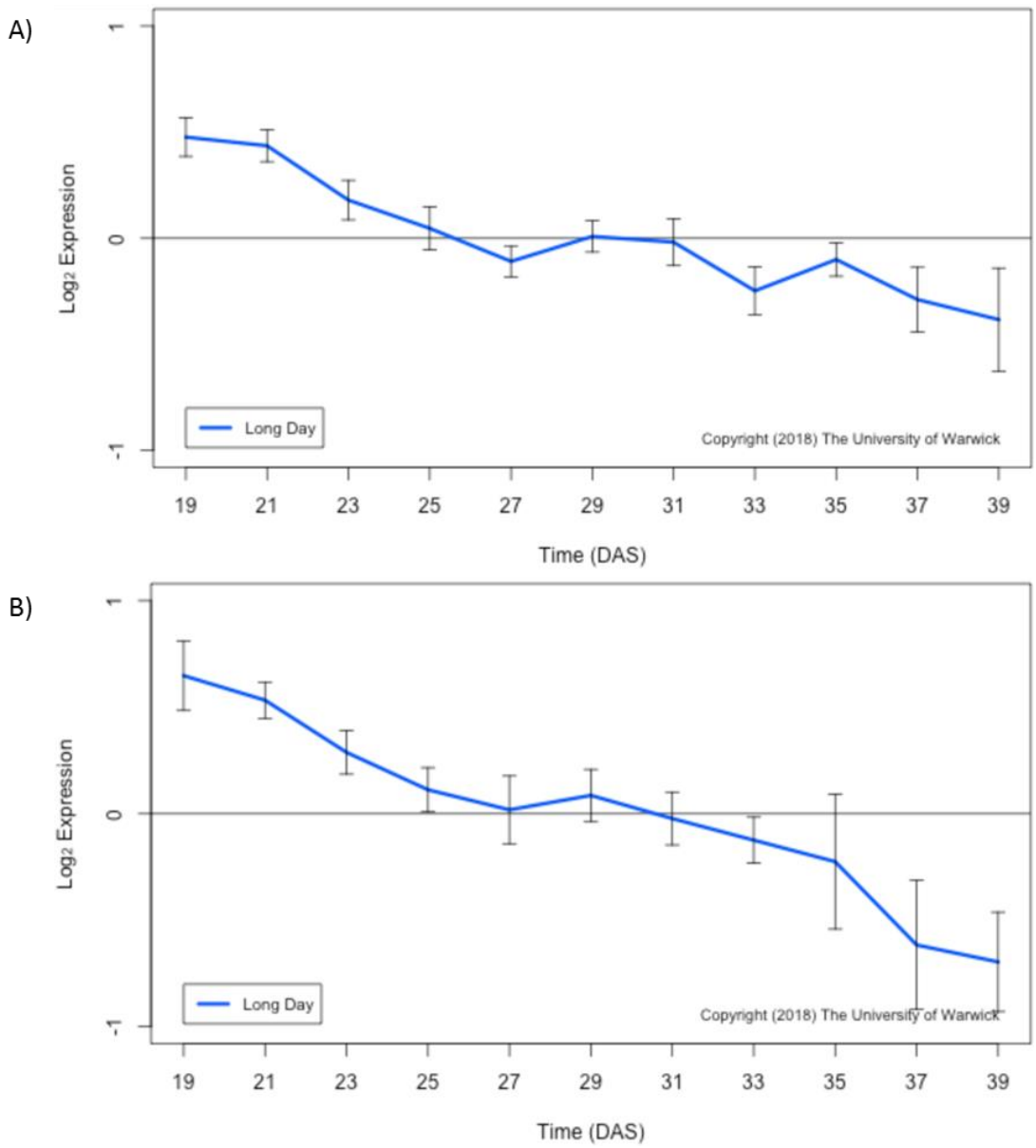


Figure 6: Relative transcript expression levels of A) *AtWHY1* and B) *AtWHY3* in long-day experiments throughout time (days after sowing). All genes have a mean expression of zero regardless of absolute expression. All lines represent the mean of the biological replicates which are the morning and afternoon results combined. Error bars indicate standard error. Figures provided by Emily Breeze from published data (Breeze *et al.*, 2011).

Breeze *et al.* (2011) showed that relative transcript levels of both *AtWHY1* and *AtWHY3* were downregulated throughout growth [Figure 6]. Both WHY1 and WHY3 proteins have been implicated in early senescent phenotypes (which will be discussed in more detail in section 1.3.5 Nuclear Gene Expression) and were differentially expressed and gave significant differences in an ANOVA (p values after multiple testing corrections: *AtWHY1*, 0.003; *AtWHY1*, 0.014) during leaf senescence. Additionally, the Whirly transcripton factor family was significantly downregulated ($p<0.05$) between 21 and 35 days after sowing compared to the entire dataset of transcription factor families (Breeze *et al.*, 2011). These results are surprising as they suggest that another factor may be involved in their connection to the senescence pathway. However, this data was compiled using relative transcript levels which are only one part of the process (Breeze *et al.*, 2011). Post-transcriptional regulatory factors may have a more important role in the cellular activity changes, both in growth and defence, caused by the expression of *WHY* genes.

1.3.2 Flowering Gene Interactions

There was significantly reduced seed yields in barley W1-1 with severe WHY1 knockdown. In W1-7 which had slightly higher WHY1 expression than W1-1, although still significantly knocked-down compared to WT, fewer fertile tillers were exhibited although seed yield per fertile tiller was similar to wild type plants (Comadira *et al,*. 2015). A homologue to an *Oryza sativa* type II MIKC MADS box gene, was also found with increased abundance in W1-7 leaves relative to the wild type (Comadira *et al,*. 2015). Expression of a type II MIKC MADS box homologue in wheat, also involved in flowering time and seed development, AGAMOUS-like33 (TaAGL33), was found to repress flowering and cell elongation. This occurred by down-regulation of a group of genes homologous to *Arabidopsis* FLOWERING PROMOTING FACTOR1 (Winfield *et al*., 2009; Greenup *et al.*, 2011). The MIKC MADS box homologue is also key regulator of flowering developmental processes in barley, including seed development, which may explain the reduced numbers of fertile tillers observed in WHY1-deficient barley lines which overexpressed this floral inhibitor. However, lower levels of AtWHY1 were found in floral tissue compared to the rosettes (Yoo *et al*., 2007) which contradicts the reported phenotype in barley. Furthermore, CreWHY fluorescence in stroma regions did not overlap chloroplast nucleoids (Kobayashi *et al.*, 2015) which matches with the phenotypes previously reported in Arabidopsis. Despite this reported interaction with flowering genes there have been no reports of a change in seed yield in plants overexpressing the WHIRLY protein.

1.3.3 DNA Repair Pathways

Some individuals of the *Atwhy1why3* mutant line were smaller than the wild type and 4.6% of mutant plants had at least one variegated leaf sector where there was no variegation in WT plants (Maréchal *et al.*, 2009). This finding indicated that the removal of both plastid-localised WHY proteins led to interference with chloroplast development and function, which is evidenced by the minor amount of plants which had variegated leaf sectors (Maréchal *et al*., 2009). Moreover, the variegated *Atwhy1why3* mutants proved similar to the ivory phenotype exhibited in *Zmwhy1* mutants, which is consistent with a role in chlorophyll production and photosynthesis (Prikryl *et al*., 2008). Furthermore, the *Atwhy1why3* double mutants showed higher levels of plastid genome instability compared to the wild type; the absence of both proteins lead to DNA rearrangements (Xiong *et al*., 2009, Maréchal *et al*., 2009). After self-crossing the variegated *Atwhy1why3* mutants to produce progeny, the plastid damage in variegated sectors was found to be irreversible, as the re-introduction of either or both WHY1 and WHY3 could not rescue the phenotype (Maréchal *et al*., 2009).

There may be a redundant role between AtWHY1 and AtWHY3 as the complete ivory phenotype was only observed in ZmWHY1 single knock-out mutants where WHY3 does not exist (Prikryl *et al.*, 2008). Although WHY3 functions are largely unknown, it has been previously identified in the plastid-encoded RNA polymerase (PEP) protein complex that is involved in plastid gene expression in a network with WHY1 (Lepage *et al*., 2013). Plastid genes are transcribed by two types of RNA polymerase; a bacterial type, PEP, and a nuclear-encoded RNA polymerase (NEP) (Dietz and Pfannschmidt, 2011). The PEP enzyme is composed of plastid encoded core subunits and additional protein factors including polymerase associated proteins and sigma factors which are required for promoter recognition (Dietz and Pfannschmidt, 2011).

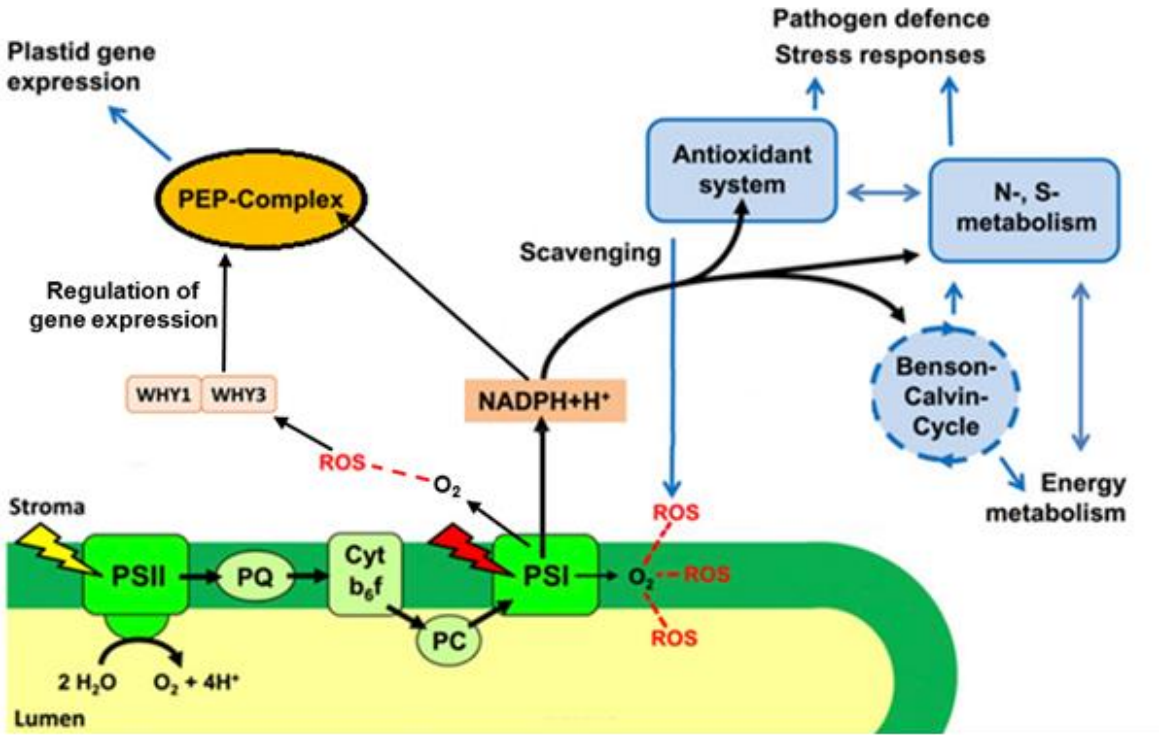


Figure 7: The interaction of the photosynthetic electron transport chain and chloroplast RNA PEP with the addition of the putative interaction of WHY1 and WHY3 with the PEP complex via ROS. The intrinsic and extrinsic protein complexes of the photosynthetic electron transport chain and associated redox mediators (green). Separate light-dependent charges in the PSII and PSI reaction centres are indicated by yellow and red flashes. The main electron flow from water to NADP and the subsequent cellular processes (blue) are represented by thick black arrows. Thin black arrows indicate the flow of a minor proportion of electrons used for regulatory redox reactions. The potential interactions of regulators on distinct cellular processes are marked by blue arrows. Adapted from Dietz and Pfannschmidt (2011).

The DNA polymerase (Pol1B), is one of two type-1 chloroplast PEPs which were proposed to perform chloroplast DNA replication; Pol1B is also implicated in repair of double strand break (DSB) sites suggesting a role in DNA repair, in which WHY proteins also have a role (Parent *et al*., 2011). Thus it is possible that a genetic interaction exists between *Pol1B* and *WHY1* or *WHY3* [Figure 7]. To investigate this further, an *Atwhy1why3pol1b-1* triple mutant was produced, with defective AtWHY1, AtWHY3 and AtPol1B proteins which exhibited a severe growth defect and a yellow variegated leaf phenotype in all of the mutant plants which is even more severe than the other single and double mutants (Lepage *et al*., 2013, Parent *et al*., 2011).

The yellow variegated sections of the *Atwhy1why3pol1b-1* triple mutant had reduced photosynthetic electron transport efficiency and increased replication errors (Parent *et al*., 2011). White variegation was previously reported in *Atwhy1why3* mutant phenotypes,

although the distinct yellow-colouration and increase in error-prone rearrangements in the repair pathway of *Atwhy1why3pol1b-1* suggests a separate phenotypic change (Maréchal *et al*., 2009, Lepage *et al*., 2013). An increased accumulation of ROS was observed in the *Atwhy1why3pol1b-1* plants and ROS can have deleterious effects if improperly scavenged, including cell death (Redza-Dutordoir and Averill-Bates, 2016), therefore the cause of the variegation phenotype may be an increase in dead cells rather than a lack of chlorophyll. Furthermore, the mutant and wild-type plants were grown under low-light conditions to address the redox imbalance that was likely causing the variegated phenotype; a decreased electron flow across the photosynthetic electron transport chain rescued the phenotype, thus elevated ROS levels played a role in the phenotypic change (Lepage *et al.*, 2013). However, it is unclear whether these growth conditions affected the other reported phenotypes of the plants. Furthermore, *Atwhy1why3* and *Atwhy1why3pol1b-1* plants were hypersensitive to paraquat-caused redox imbalances and ROS production localised to the chloroplasts was found to be the cause of the plastid DNA rearrangements and thus variegation phenotype (Lepage *et al.*, 2013). It is proposed that the AtWHY1 and AtWHY3 proteins are putative regulators of the PEP complex via ROS activity and this in turn alters plastid gene expression (Parent *et al*., 2011). These results suggest that WHY proteins stabilise the plastid genome by limiting illegitimate recombination events which ensures that mutations in genes encoding PEPs do not occur (Cappadocia *et al.*, 2010, Lepage *et al*., 2013).

According to Chevigny *et al.*, (2020), the chloroplast genome is subject to significantly fewer rearrangements compared to the nuclear genome because apart from the large inverted repeat there are not many other repeat sequences in chloroplast DNA. Microhomology-mediated end joining (MMEJ), which is a DSB repair pathway, was responsible for chloroplast DNA rearrangements in *Atwhy1why3* mutants (Maréchal *et al*., 2009). MMEJ U-turn rearrangements were also reported to be repressed by WHY proteins in plastids (Zampini *et al.*, 2015). More recent studies have found that DNA polymerases blocked the MMEJ pathway in the presence of AtWHY2 (García-Medel *et al.*, 2019). Despite this interaction with DNA repair mechanisms there have been no studies on seed ageing and whirly. DNA repair is a significant determinant of seed longevity which is important in the preservation of crops (Waterworth *et al*., 2019). The only study on whirly in germination implicated WHY2 and did not consider the effects of WHY1 and WHY3. In this study a *Atwhy2* mutant with altered mitochondrial structure had reduced seed germination (Golin *et al*., 2012). However there have been no studies

into yield or germination changes in *Atwhy1*, *Atwhy3* or *Atwhy1why3* mutants, which will be explored in Chapter 3: Germination.

1.3.4 Plastid Gene Expression

WHY1 was found as a soluble protein in both the stroma and thylakoid fractions of *Zea mays* (maize) leaves (Prikryl *et al*., 2008) and later also in barley chloroplasts (Melonek *et al*., 2010). Furthermore, ZmWHY1 was associated with the thylakoid membrane which upon DNAse treatment is released into the stroma, suggesting an interaction of the protein with the plastid nucleoid (Prikryl *et al*., 2008). Contrastingly, chloroplast development was impaired in *Zmwhy1* mutants which have severely reduced levels of plastid RNA (Prikryl *et al*., 2008). An electron microscopy analysis of nucleoids in RNAi-mediated *Hvwhy1* knock-down plants showed reduced nucleoid compactness suggesting a role of WHY1 in plastid genome stability (Krupinska *et al*., 2014).

1.3.5 Nuclear Gene Expression

The WHY1 protein was first discovered as a *trans*-acting factor of *PR* gene expression (Desveaux *et al*., 2005). RNA from the previously described HA-tagged overexpressing AtWHY1 plants were analysed in quantitative real-time PCR and showed that relative transcript levels of *PR1* and *PR2* genes in the nucleus were enhanced by factors of 700 and 70, respectively compared to WT (Isemer *et al*., 2011), these results suggest that AtWHY1 regulates the expression of *PR* genes in the nucleus.

A gel shift assay with a labelled telomeric repeat oligonucleotide showed that AtWHY1 specifically bound to single-stranded, but not double-stranded telomeric DNA sequences (Yoo *et al*., 2007). Two TDNA *AtWHY1* deficient lines did not exhibit growth or developmental defects but did show a significant increase in telomere tracts and telomerase activity in consecutive generations (Yoo *et al*., 2007). Transgenic overexpressing plants showed the opposite phenotype, shortened telomere tracts and decreased telomerase activity (Yoo *et al*., 2007). Therefore, there is a clear role of AtWHY1 as a telomere end-binding protein to regulate telomere-length homeostasis in the nucleus. Maintaining telomere lengths has widely been reported as important in ageing processes.

An early senescence phenotype was observed in *Atwhy1* mutant rosette leaves and a semi-quantitative reverse transcription-PCR showed increased expression of senescence-associated proteins, WRKY53 and SAG-12 (Miao *et al*., 2013). Furthermore, immunodetection of HA-tagged WHY1 and a GUS reporter construct of the *WRKY53*

promoter in tobacco found that WHY1 acted as a direct repressor of *WRKY53* gene expression (Miao *et al.*, 2013). Furthermore, phosphorylation of AtWHY1 led to enhanced expression of *AtWRKY53*, which encodes a key transcription factor that regulates leaf senescence and plant aging (Ren *et al*., 2017). Expression of AtWRKY53 is induced during leaf senescence and cell death, thus AtWHY1 has a role as an upstream suppressor of At*WRKY53* during leaf-senescence (Miao *et al*., 2013). The expression of *WRKY53* regulates several plant ageing and leaf senescence genes, thus nuclear-localised WHY1 indirectly regulates senescence in the plant through WRKY53. Interestingly, WHY1 has also been implicated in the regulation of other WRKY family members (Miao *et al*., 2013). Intriguingly, immunogold labelling of nuclear-localised HvWHY1 was found in heterochromatin (Grabowski *et al.*, 2008). This suggests that as in the plastid, WHY1 may have a function in chromatin compaction in the nucleus. Moreover, dark-induced senescence of detached leaves was not affected in leaves lacking the HvWHY1 protein (Comadira *et al.*, 2015). However, when the plants were grown at high light intensity, senescence was delayed in RNAi W1 plants compared to the wild type suggesting that WHY1 does not regulate age- and dark-dependent senescence, but it is vital in photo-oxidative stress response (Comadira *et al*., 2015).

1.4 FUTURE OF WHIRLY STUDIES

An increasing number of publications establish the importance of the WHY proteins in plant development and stress tolerance [Table 1] which is an increasing area of research interest considering the effects of climate change on crops. The knowledge of WHY proteins, in particular WHY1 has significantly changed recently and it is becoming more apparent as a regulator of development under heightened abiotic and biotic stress conditions [Table 1] which will be discussed in further chapters of this thesis. However, a more comprehensive analysis of WHY functions in Arabidopsis is required before the functions of WHY can be beneficially transferred to the more commercially valuable crops of *Triticum aestivum* (wheat) and *Zea mays* (maize).

Table 1: Recent developments in the functions of WHIRLY proteins in several species of interest

Key Findings:	**Species**	**Protein**	**Reference**
SlWHY1 was expressed more widely than *SlWHY2*. Drought and salt stress enhanced levels of *SlWHY1* and *SlWHY2* transcripts.	*Solanum lycopersicum*	WHY1 WHY2	Akbudak and Filiz, 2019
OsWHY1 and OsWHY2 have the highest coverage of proteins bound to OsPAL2;3, an allelopathy promoter. OsWHYs negatively regulate OsPAL2;3.	*Oryza sativa*	WHY1 WHY2	Fang *et al.*, 2019
The MMEJ pathway was blocked by DNA polymerases in the presence of AtWHY2 (and other ssDNA-binding proteins) in templates of single-stranded regions longer than 12 nts.	*Arabidopsis thaliana*	WHY1 WHY2 WHY3	García-Medel *et al.*, 2019
The levels of *AtWHY* transcripts were decreased in shoots up to 12h after phytotoxic citral treatment to roots. The AtWHYs had a strong affinity for citral isomer binding and low *in silico* molecular docking.	*Arabidopsis thaliana*	WHY1 WHY2 WHY3	Graña *et al.*, 20
SlWHY1 expression was increased by chilling. SlWHY1 acts as a positive regulator of *SlpsbA*, which enhanced chloroplast D1 synthesis. *SlAMY3-L*, a starch-degrading enzyme, and inhibitor of *SlISA2* starch synthesis-related enzyme was also regulated by SlWHY1.	*Solanum lycospericum*	WHY1	Zhuang *et al.*, 2019
One of the two putative Sorghum WHY TFs that is crucial for pollen development is orthologous to AtWHY2.	*Sorghum bicolor*	WHY2	Dhaka *et al.*, 2020
Plastid genome instability and increased accumulation of ROS were observed in *Atreca1why1why3* mutants, which had leaf growth defects, white variegated sectors, higher accumulations of plastid DNA rearrangements, and reduced fertility.	*Arabidopsis thaliana*	WHY1 WHY3	Duan *et al.*, 2020
Seed germination was reduced in *Atwhy2* mutants that had an altered mitochondrial structure, disordered nucleoids and increased *AtWHY3* levels compared to the wild type. WHY3 was dual targeted to the chloroplasts and mitochondria in protein transport *in organello* experiments.	*Arabidopsis thaliana*	WHY2 WHY3	Golin *et al.*, 2020
WHY2 was localized in the mitochondria, plastids and nucleus during leaf ageing. The chloroplasts of pericarp cells of AtWHY2 OE lines had increased starch granule numbers and jasmonate-associated gene expression linked to early senescence. The opposite phenotype was observed in the *Atwhy2* mutants.	*Arabidopsis thaliana*	WHY2	Huang *et al.*, 2020

The *Atwhy1* mutants showed early senescence with an early peak in salicylic acid (SA) levels. This was prevented by nuclear WHY1 (nWHY1) expression. Plastid WHY1 (pWHY1) expression enhanced SA levels. The levels of stress-related transcripts were changed in *pWHY1* lines. In contrast, transcripts associated with plant development and early growth were changed in the nWHY1 lines.	*Arabidopsis thaliana*	WHY1	Lin *et al*., 2020
SlWHY2 RNAi lines showed a severe wilting phenotype under drought with decreased fresh weight, chlorophyll contents and photosynthesis, as well as decreased expression of mitochondrial DNA repair and recombination genes. ROS accumulation was increased in the *SlWHY2* RNAi lines compared to the wild type.	*Solanum lycospericum*	WHY2	Meng *et al*., 2020
Weak interactions were observed between CsWHY1 and the highly expressed Irregular Vasculature Patterning (CsIVP). CsIVP targets developmental regulators and functions in downy mildew resistance.	*Cucumis sativus*	WHY1	Yan *et al*., 2020a
The expression of MeWHYs and MeCIPK23 was significantly increased 10-20 days of drought. Plants lacking in any or all MeWHYs and/ or MeCIPK23 were more sensitive to drought stress.	*Manihot esculenta*	WHY1 WHY2 WHY3	Yan *et al.*, 2020b
SlWHY1 OE lines showed reduced wilting under heat stress, with increased levels of *SlHSP21.5A* transcripts, greater membrane stability and higher soluble sugar contents. ROS levels were decreased relative to the wild type. RNAi lines lacking SlWHY1 showed the opposite phenotype.	*Solanum lycospericum*	WHY1	Zhuang *et al.*, 2020a
SlWHY1 was a positive regulator of RuBisCO expression under cold stress directly binding to the promoter of the rbcS gene that encodes the small subunit.	*Solanum lycospericum*	WHY1	Zhuang *et al.*, 2020b
WHY2 was shown to be a major regulator of the root apical meristem developmental network.	*Arabidopsis thaliana*	WHY2	McCoy *et al*., 2021
Recruitment of the WHY1 and 3 proteins, and AtRNH1C to the same genomic site promoted homologous recombination repair. These proteins maintain chloroplast genome integrity through AtRecA1 interaction. Deletion of *WHY1*, *3* or *AtRNH1C* suppressed RNApol binding. In contrast, WHY1 and 3 promoted recruitment of PEP RNApol.	*Arabidopsis thaliana*	WHY1 WHY3	Wang *et al.*, 2021

2 RESEARCH AIMS

Previously published research led to the hypothesis that WHY1 and WHY3 proteins work together to integrate environmental signals including biotic and abiotic stresses to inform plant growth from seed viability to yield. The aim of the work described in this report was to elucidate the function of WHIRLY proteins throughout plant development and under heat stress. This project provided a much-needed insight into the fundamental role of WHIRLY1 and its interaction with WHIRLY3 in *Arabidopsis thaliana*. These findings were then used to inform the research of WHIRLY in *Triticum aestivum* (common bread wheat) which had not been previously explored. For ease, the project aims were split into four sections:

Germination:

1. Explore the roles of WHY1 and WHY3 in seed longevity and germination using Arabidopsis *Atwhy1*, *Atwhy3* and *Atwhy1why3* mutant lines.
2. Explore the potential roles of WHY1 and WHY3 in germination using Arabidopsis *Atwhy1*, *Atwhy3* and *Atwhy1why3* mutant lines that have been subject to accelerated ageing.

Gene Interactions:

1. Identify targets of WHY1 and WHY3 in germination through RNA-seq analysis of *Atwhy1why3* seeds and wild type seeds.
2. Identify mechanisms for ageing hypersensitivity in *Atwhy1why3* 7 days aged mutants.

Arabidopsis Development:

1. To characterise the effect of WHY1 and WHY3 throughout Arabidopsis growth
2. To compare photosynthesis in *Atwhy1*, *Atwhy3*, and *Atwhy1why3* mutants relative to the wild type.
3. To determine the responses of the *Atwhy1*, *Atwhy3*, and *Atwhy1why3* mutants and the wild type plants to heat stress.

Wheat Development:

1. To confirm transformations of TaWHY knockdown plants and select lines to carry forward.
2. To characterise the early and late phenotypic growth changes of TaWHY knockdown plants compared to null segregants and WT.
3. To determine the effects of TaWHY deficiency on plant fertility and seed yield.

3 MATERIALS AND METHODS

3.1 MATERIALS

3.1.1 *Arabidopsis thaliana*

The *Arabidopsis thaliana* lines used in the following studies originated from the Salk Institute Genomic Analysis Laboratory (La Jolla, CA, USA) and the Seattle TILLING Project (Till *et al*., 2003). The Col-0 wild type (WT) accession and the three *Atwhy* mutants: *Atwhy1*, *Atwhy3* and *Atwhy1why3*, were provided by the Normand Brisson group (Department of Biochemistry, Université de Montréal, Montréal, Canada).

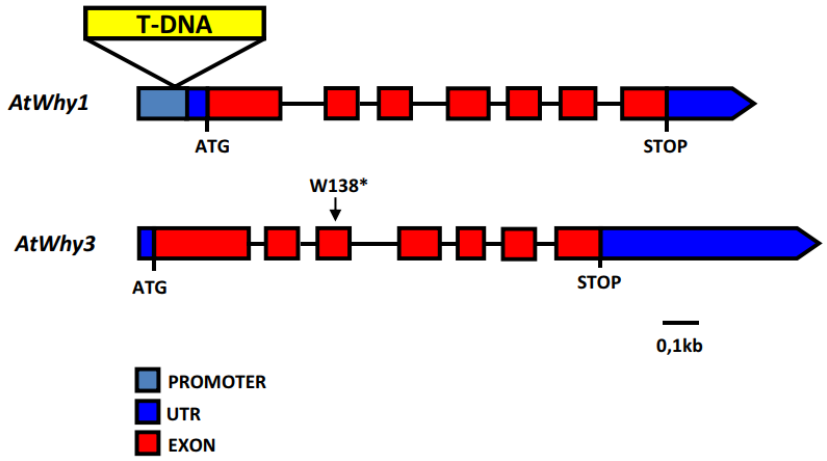


Figure 8: The physical maps of the *AtWhy1* (AT1G14410) and *AtWhy3* (AT2G02740) genes and the positions of the changes to make the knock-out mutants. The TDNA position in the knock-out *AtWhy1* line is indicated and knock-out *AtWhy3* line is a TILLING line with a mutation that changes the TGG to a TGA stop codon at position W138* inside the *AtWhy3* gene Adapted from: (Maréchal *et al.*, 2009).

The *Atwhy1* mutant line (SALK_099937) was produced as a result of a T-DNA insertion in the *AtWhy1* gene, mapped at nucleotide -102 relative to the initial ATG as characterised in Figure 8. The *Atwhy3* mutants were produced as a Targeting Induced Local Lesions IN Genomes (TILLING) line from the Seattle TILLING Project (Till *et al*., 2003) where a premature stop codon occurred in the *AtWhy3* gene as a result of a single nucleotide polymorphism from TGG to TGA as characterised in Figure 8 (Maréchal *et al.*, 2009).

The double knockout mutant *Atwhy1why3* line was previously reported (Maréchal *et al*., 2009). The lines were presumably produced by crossing the above lines followed by selection for the double mutations, although this not clear from the described information (Maréchal *et al.*, 2009), the genotypes of these lines were checked using PCR (10: Appendix).

3.1.2 *Triticum aestivum*

The *Triticum aestivum* (wheat) RNAi *TaWhy1* knock-down plants used in the following studies were produced in Biogemma (now Limagrain) (Centre de Recherche, Chappes, Puy-de-Dôme, France) and provided to this project by Jacques Rouster.

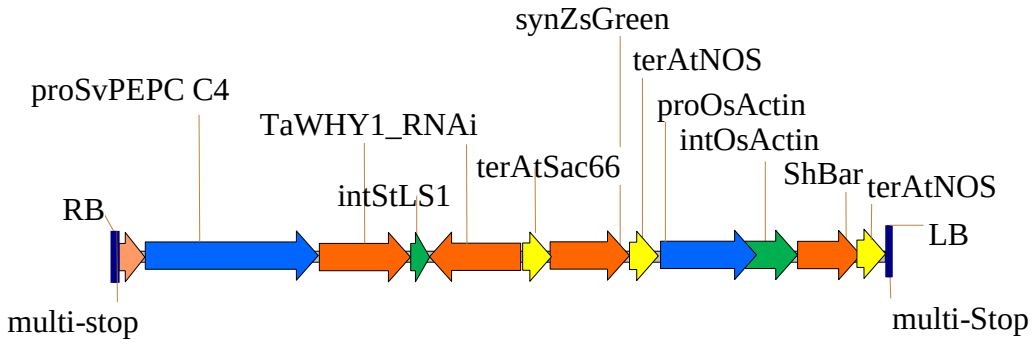


Figure 9: A map of the sequence of the T10715 RNAi TaWHY1 down-regulated wheat plants as follows: right border (RB), *Sorghum vulgare* C4 PEPC promoter (proSvPEPC C4), wheat WHY1 RNAi knockdown (TaWHY1_RNAi), intron of StLS1 (intStLS1), Arabidopsis polygalacturonase Sac66 terminator (terAtSac66), synthesised ZsGreen fluorescence (synZsGreen), Arabidopsis nitric-oxide synthase terminator (terAtNOS), *Oryza sativa* actin promoter (proOsActin), *Oryza sativa* actin intron (intOsActin), Bar gene for Basta glufosinate tolerance (ShBar), Arabidopsis nitric-oxide synthase terminator (terAtNOS), left border (LB) (Biogemma, 2019).

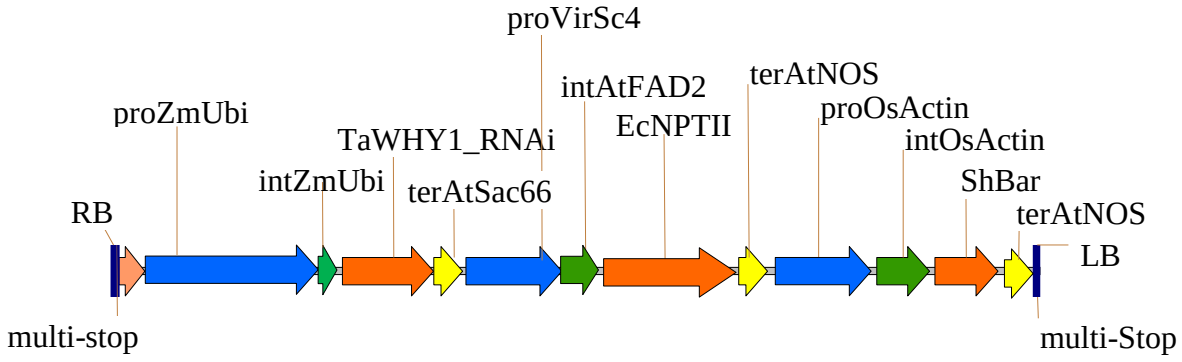


Figure 10: A map of the sequence of the T10716 RNAi TaWHY1 down-regulated wheat plants as follows: right border (RB), *Zea mays* Ubiquitin promoter (proZmUbi), intron of ZmUbi (intZmUbi) wheat WHY1 RNAi knockdown (TaWHY1_RNAi), Arabidopsis polygalacturonase Sac66 terminator (terAtSac66), promoter of VirSc4 (proVirSc4) intron of Arabidopsis FAD2 gene (intAtFAD2), *E. coli* NPTII gene (EcNPTII), Arabidopsis nitric-oxide synthase terminator (terAtNOS), *Oryza sativa* actin promoter (proOsActin), *Oryza sativa* actin intron (intOsActin), Bar gene for Basta glufosinate tolerance (ShBar), Arabidopsis nitric-oxide synthase terminator (terAtNOS), left border (LB) (Biogemma, 2019).

The RNAi lines with endogenous TaWhy1 knock-down comprised of two transformation experiments from RNAi hairpin constructs, these were labelled T10715 and T10716. The T10715 transformation [Figure 9] used a sorghum PEPC_C4 promoter and produced 5 lines, whilst the T10716 transformation used a maize Ubi1 promoter [Figure 10] which produced 6 lines. These lines were genotyped using qPCR to determine the level of knockdown of each individual line which is detailed in Chapter 7: Wheat.

3.2 METHODS

3.2.1 Surface sterilisation of Arabidopsis seeds

The *Arabidopsis thaliana* seeds upon arrival to Leeds were surface sterilised before being used for further experiments to prevent contamination in growth chambers. The seeds were aliquoted into 1.5 mL Eppendorf tubes and weighed before being placed in tube racks inside a 50 L desiccator which was exposed to chlorine gas formed by mixing commercial bleach (100 mL) and 100 % Hydrochloric Acid (3 mL) for 15 H. Seeds were then used for further experiments and to produce more seed stock for further experiments.

3.2.2 Growth Conditions

Seeds were sown onto Petri dishes containing 0.5% w/v Murashige and Skoog (MS) (Murashige and Skoog, 1962) basal salts, 1% w/w agar, pH 5.7 under a flow hood to maintain sterile conditions which were sealed with micro pore tape before being placed into a 4 °C chamber for 48 hours for cold stratification to ensure synchronised germination. Thereafter, the plates were placed in a controlled environment chamber with a 20 ºC /16 ºC day/ night temperature regime in a 16 h light and 8 h dark photoperiod (65-95 μmol m^{-2} s^{-1} irradiance) and 60% relative humidity for seven days . They were then transferred to compost (Sinclair Professional potting compost, Sinclair Pro, Cheshire) on p15 cell trays. Each cell contained one seedling and each tray contained 15 cells. All 4 genotypes were equally distributed in a random order in each tray.

These trays were placed in a controlled environment chamber with a 20 ºC /16 ºC day/night temperature regime, with a 16 h light and 8 h dark photoperiod (250 μmol m^{-2} s^{-1} irradiance) and 60% relative humidity. 1 cm of water was added to the bottom of trays containing cells with perforated drainage holes twice a week or as needed when the soil was dry. The plants were watered to fill trays 0.5 cm full during silique filling for seed production until senescence occurred. Seeds were collected and then further generations of plants for future experiments were grown as above unless otherwise specified.

3.2.2.1 Seed Production

Seeds were produced according to the Arabidopsis Biological Resource Centre (ABRC), Ohio State University (ABRC, 2013). The inflorescences were transferred to a labelled bag before siliques browned. Seeds were then sifted through a sieve to remove debris and placed in labelled Eppendorf tubes for storage in a controlled low humidity (20-40%) room, no surface sterilisation was performed on these seeds.

3.2.3 PCR genotyping

Primers were designed to genotype Arabidopsis mutant plants compared to the WT, sequences and band sizes are shown in Table 2.

Table 2: Primer Sequences and expected base pair sizes

Primer Name	Primer Sequence (5'-3')	Predicted Band Size (bp)
AtCat2 (Housekeeping)	Left: CGAGGTATGACCAGGTTCGT Right: GATGCTTGGTCTCACGTTCA	565
AtAct1 (Housekeeping)	Left: CCTGACAATTTCTCGCTCGG Right: CCTCGGTCAGCAGTATAGGG	404
AtWhy1 Left	TCGAATGACCCACGTAAAATC	WT *AtWhy1*: 1100
AtWhy1 Right	TGACCAACAAACTGTTGATGG	
LBB.3	ATTTTGCCGATTTCGGAAC From: (Signal.salk.edu, 2005)	With *AtWhy1* Right, *Atwhy1*: 750
AtWhy3 Left	TCGTCTTCTTCGCAAAACGGTAG	With *AtWhy3* Left, WT WHY3: 535 With *AtWhy3* Specific Left, *AtWhy3:* 925
AtWhy3 Specific Left	TGCAAAAGTTGGAGGGCTTTCTT	
AtWhy3 Right	CCTGCAACATTTCAAATCAAACA	

Leaf samples (100 mg fresh weight) were frozen in liquid nitrogen and then placed in a -70 °C freezer until analysis. The GenElute™ Plant Genomic DNA Miniprep Kit Protocol G2N70 (Sigma-Aldrich, USA) was used to extract and analyse the DNA.

Table 3: PCR cycling conditions for identification of mutant and wild-type *Arabidopsis* plants.

Step	Temperature (°C)	Time (seconds)
1	94	40
2	94	10
3	57	15
4	72	30
5	Go to step 2 for 37 x	
6	72	5 minutes
7	4	∞

The eluted DNA sample was quantified for purity on a Nanodrop (ND-1000 Spectrophotometer, Labtech International, UK) via ratio of absorbance at 260 nm and 280 nm. The ratio of $A_{260/280}$ between 1.8 and 2.0 indicates DNA as pure. The samples were then placed at -20 °C for storage. The housekeeping gene *AtCat2* (AT1G58030.1) was used in a PCR assay as an endogenous control to ensure that the DNA extraction was efficient for PCR conditions outlined in Table 3. The primers described in Table 2 were further used under the same conditions to identify the genotypes of the mutants.

3.2.4 Germination Assays

3.2.4.1 Seed Viability and Vigour Assay

Seeds of comparable age and storage conditions were plated on blue blotter germination paper in 90 mm Petri dishes with 7 mL of dH_2O and cold stratified at 4 °C for 48 h to remove dormancy. All plates were removed from cold stratification at the same time and placed in 20 °C growth chamber with 60% RH and a 16 h light and 8 h dark photoperiod (65-95 μmol m^{-2} s^{-1} irradiance). Seeds were scored for germination by the emergence of the radicle from the endosperm under a light microscope at the same time at daily intervals. The final germination percentage (viability) and time to germinate (vigour) across the different lines compared to WT was recorded.

3.2.4.2 Seed Ageing Assay

Aliquots of each of the seeds were counted and placed inside open 1.5 mL microcentrifuge tubes on a tube-rack placed above a salt solution (100 mL dH_2O, 50g KCl) to produce 80% humidity inside a sealed air-tight plastic storage container which was further sealed with parafilm and placed inside a 35°C oven to simulate ageing. Seeds were kept in this arrangement for 7 days and 14 days to measure seed viability and vigour loss respectively. The germination of seeds was measured as described in 3.2.4.1.

3.2.5 RNA-seq

This experiment was carried out to determine the differences between unaged and aged seeds to determine if both WHY1 and WHY3 had a role in DNA repair in aged seeds. Therefore, this experiment compared the WT and *Atwhy1why3* double mutant.

Seed samples were split into 2 categories; aged for 7 days (as described in 3.2.4.2) and unaged straight from the plant. The seeds of each genotype were from the same plant or tray of plants grown in exactly the same conditions at the same time. There were 3 aliquots of seeds per genotype and treatment. The treatments were as follows; unaged seeds with 0 h imbibition, unaged seeds with 6 h imbibition, aged 7-day seeds with 6 h imbibition.

3.2.5.1 RNA Isolation

The Promega SV Total RNA Isolation System kit (Madison, WI, USA) was used to extract total RNA by centrifugation based on Kobs (1998). Small batches of seeds were prepared so that the RNA could be extracted quickly and to ensure that all samples were kept cold. The seed aliquots were weighed and age treated (as described in 3.2.5). The Promega protocol was followed for the RNA isolation which included DNAse treatment after wash steps.

3.2.5.2 Yield Purity

The yield of total RNA obtained from the kit must be determined before sequencing. The nanodrop was used for spectrophotometric quantification at 260 nm. At this wavelength, 1 absorbance unit was equal to 40µg of single-stranded RNA per mL and anything between 1.7 – 2.1 A260/S280 ratio was accepted as sufficient purity.

3.2.5.3 Sequencing

The total RNA samples were sent to Novogene (Cambridge, UK) for paired end reads using the Illumina NovaSeq 6000 Sequencing System. Here the samples went through extra quality control steps, library construction, library quality control, sequencing, data

quality control and bioinformatic analysis. Novogene then sent the paired and unpaired read data which was analysed using the Galaxy software (UseGalaxy.eu), produced by the Freiburg Galaxy Team (Giardine *et al*., 2005). The details of the analysis are further described in Chapter 5: Whirly gene interactions during ageing hypersensitivity

3.2.6 Arabidopsis Macroscopic Phenotyping

3.2.6.1 Rosette Measurements

For each measurement, 24 plants per genotype were grown in 4 trays. Photographs of the trays were taken every week after 2 weeks post cold stratification through the vegetative growth phase until seed production on a Nikon Digital Camera D5100. Rosette diameters were measured from the photographs using Image J (Schneider *et al*., 2012).

At later growth stages when Image J photograph analysis was difficult due to bolting, calliper measurements (Clarke Precision Digital Vernier Calipers Model: CM145, Essex, UK) were used as a secondary analysis. Statistical significance was calculated both by multiple comparison ANOVA for consistent differences between genotypes across time points and t tests for significances at each time point.

3.2.6.2 Leaf Images

For each measurement, 24 plants per genotype were grown in 4 trays. After two weeks post germination, photographs were taken every week throughout the vegetative growth phase until seed production began as follows. Leaves were removed from three plants per genotype and laid out on white paper in order of emergence. Photographs were then taken under flash exposure using a Nikon Digital Camera D5100 from a stand. Image J was used for RGB analysis for leaf area and number of leaves.

3.2.6.3 LiCOR Measurements

3.2.6.3.1 Light Saturation

The *Atwhy* mutants and WT were grown at 20 ºC /16 ºC day/ night temperature regime in an 8 h light and 16 h dark photoperiod (200 µmol m^{-2} s^{-1} irradiance) and 60 % RH. The plants were monitored until the leaves were large enough to fit into the 2 cm^2 clamp on the Li-6400 and Li-6800, this was between 4 and 6 weeks and on plants which had not bolted. A series of measurements were taken at 5 minutes under each light intensity of increasing photosynthetically active radiation (0, 50, 100, 150, 200, 300, 400, 500, 600, 900, 1200, 1500) to allow stomatal changes to take place before readings were taken under the following conditions. The conditions of the LiCORs were: 200 µmol s^{-1} flow,

42 % RH, 400 µmol mol^{-1} of CO_2 at 20 °C. An average of four plants per genotype was taken to get the photosynthetic gas exchange measurement for each genotype.

3.2.6.3.2 Temperature

Plants were grown as above and conditions were the same as for the light saturation curves, a light of 1000 µmol s^{-1} was used based on information gathered from the light saturation curves. Measurements were taken at a range of temperatures (20-40 °C), spaced evenly at 2 °C apart. Temperatures were held for 5 minutes to allow stomatal changes to take place before values were taken. An average measurement of four plants per genotype was taken to get a measurement of photosynthetic gas exchange at different temperatures for each genotype.

3.2.6.4 Seed Yield Assays

Yield is a broad term, but here it is used to describe whole plant yield including total seed weight, and wet and dry biomass of plants. The number of siliques per plant and number of seeds per silique were also counted and weighed to assay yield of *Atwhy* mutants compared to the wild-type.

3.2.6.5 Total Plant Yield

Plants were grown in the controlled chamber as described in 3.2.2. For total seed weight plants once siliques began to ripen after inflorescence, the plants were sealed into Glassine paper bags until senescence was completed. The stems were cut and removed from the pot of soil and the sealed Glassine bags were placed into a box to dry. Once dried, the siliques were sieved and a spatula was used to break them open to expose seeds. The total amount of seeds was funnelled into a 1.5 mL Eppendorf tube and weighed to give total seed yield per plant. This was repeated at least 5 times for each genotype under the same growth conditions.

For total fresh biomass the plants were grown in the controlled chamber as described in 3.2.2 and 3 plants of each genotype were removed every week, in the time period of 3 to 6 weeks post germination. Soil was washed from the plants and then the plants were patted dry with lab roll and weighed. Photographs of the plants were taken from above as described in 3.2.6.1. These plants were then placed individually into labelled Glassine bags and sealed. These bags were placed into a 100 °C oven for 48 H. After this time the bags were removed from the oven and their contents were weighed. This value was compared to the fresh weight value to obtain the fresh: dry weight ratio.

3.2.6.6 Silique Counts

Plants were grown in the controlled chamber as described above in 3.2.2and placed into a large autoclave bag to ensure that siliques do not escape. After senescence was completed the plants were removed from the walk-in chamber and the siliques were counted on each branch of each plant. This was repeated at least 5 times for each genotype and the total and averages were calculated. Additionally, on each plant 3 random siliques were removed carefully and placed into 1.5 mL Eppendorf tubes, ensuring that the seeds did not fall out. These tubes were weighed and the siliques were opened to count the seeds. This gave the seeds per silique measurement and was averaged across at least 5 plants per genotype.

3.2.6.7 Chlorophyll and Carotenoid Pigment Extraction

One of three rosette portions harvested as above was used for the analysis of leaf pigments on ice as follows:

The frozen samples (~ 100 mg each) were ground to a fine powder with a pestle in a precooled mortar (using liquid nitrogen). 1 mL of cold 100% acetone was added to each powder sample and ground again until thoroughly mixed. A further 3 mL of 100% acetone was added to the mixture, additional volumes of 100% acetone were added to bring the total volume to 5 mL. These were stored on ice in the dark until analysis. The samples, mixed using a vortex and subjected to centrifugation (Centrifuge 5804R, Eppendorf, UK) at maximum speed (14,000 x g) for 10 minutes at 10 ºC. The supernatant fractions were collected and stored on ice.1 mL of the supernatant leaf extract was transferred to a 1 cm pathlength cuvette and placed in a spectrophotometer after using a blank of 1 mL of 100% ethanol. For these assay the 6715 UV/ Vis Jenway 67 series (Staffordshire, UK) single cell holder spectrophotometer was used.

Absorbance readings were taken at 470 nm, 649 nm and 664 nm and applied to the appropriate formula [Equations 1] to determine leaf pigment content (µg/mL extract solution). This was multiplied by 5 mL to obtain the total amounts of chlorophyll a, chlorophyll b and total carotene in the 5 mL extracts. The mean of 3 replicates was used to determine pigment content per genotype (Lichtenthaler, 1987).

Equations 1: The estimate of the concentration of leaf pigments in ethanol extracts (Lichtenthaler, 1987).

$$Chlorophyll\ a\ (Ca) = 13.36A_{664} - 5.19A_{649}$$

$$Chlorophyll\ b\ (Cb) = 27.43A_{649} - 8.12A_{664}$$

$$Total\ chlorophyll\ (Ca + Cb) = 5.24A_{664} + 22.24A_{649}$$

$$Carotene = \frac{1000A_{470} - 2.13Ca - 97.64Cb}{209}$$

3.2.6.8 Tetrazolium Assays

These assays were used to determine the viability of seeds using the colourless compound 2, 3, 5-triphenyltetrazolium chloride, a cationic dye, which turns into the scarlet-red coloured formazan, which is water-insoluble. This reaction occurs by cellular dehydrogenases which catalyse the hydrogen transfer reaction from NADH-dependent reductases. This reaction occurs in all respiring tissues, and tetrazolium enters the cells of living and dead tissue, but is only catalysed into non-diffusible formazan stain in viable, living tissue. Therefore, seeds which remain unstained are non-viable and would not normally germinate. These results can be used in addition to germination studies to measure seed coat permability because the intensification of the red colouration is directly proportional to the seed permeability. These methods of quantifiable staining were adapted from Vishwanath *et al.* (2014), Porter *et al*. (1947) and Wharton (1955) and timings were chosen based on the results of these studies.

3.2.6.8.1 Seed Penetration Assay

Three 50 (± 1) mg aliquots of seeds per genotype and treatment were weighed and placed into microcentrifuge tubes with 1 % w/v tetrazolium red (Sigma-Aldrich, UK). Tubes were mixed by inverting 3 times before placing in a 30 °C heat block in the dark for 48 h. After this incubation period was completed, a Pasteur pipette was used to remove around 100 seeds to observe colour changes under a light microscope. Photographs were taken with a Nikon Digital Camera (D5100) and Image J was later used to count and measure seeds that had turned red, to create data on percentages of viable seeds that have the potential to germinate. During the observation and photography, seed tubes were kept on ice at 4 °C until the extraction of formazans in 3.2.6.8.2.

3.2.6.8.2 Formazan Seed Coat Permeability Assay

The tetrazolium red solution was removed from the above tubes using a 1 mL pipette and seeds were washed with 1 mL ddH_2O twice, with tubes kept on ice. A cut tip was used to remove the seeds from the tubes, which were transferred to a clean mortar with 1.5 mL of 95 % ethanol. A pestle was used to grind the seeds into a fine paste and then transfer to a clean 2 mL microcentrifuge tube on ice. The extracts were adjusted to give a final volume of 1.5 mL with the addition of 95% ethanol. Tubes were centrifuged at 15,000 x g for 3 minutes. Supernatant was collected into a new clean microcentrifuge tube on ice in the dark until all extracts had been performed. These final steps be completed for all samples in under an hour to ensure that the tetrazolium salts don't interact with the embryo cells after seed disruption. The supernatants were transferred into 1.5 mL spectrophotometer cuvettes and the absorbance was measured at 485 nm, with 95 % ethanol as the blank reference cuvette.

3.2.7 Wheat Macroscopic Phenotyping

3.2.7.1 Growth Conditions

The wheat *Tawhy1* RNAi knockdown seedlings in tubes of agar were initially placed in a walk-in growth chamber, as described above, for three days to acclimatise after dark transportation from Limagrain, France. Thereafter they were re-potted into compost (Sinclair Professional potting compost, Sinclair Pro, Cheshire, UK) in 3 L pots with slow-release fertiliser (Osmoscote Exact Standard 8-9, Everris, Suffolk, UK) and the lines were randomly distributed and placed into a controlled environment glasshouse. The growth conditions were as follows: 22 ºC day and night temperature regime where natural sunlight was supplemented over a 16 h light and 8 h dark photoperiod with LED lamps (Plessy, Devon, UK) at 400 µmol m^{-2} s^{-1} irradiance. Water was added into trays up to a 1cm mark every 3 days as needed.

3.2.7.1.1 Differing Growth Conditions at Limagrain

The *Tawhy* plants which were grown at glasshouses at Limagrain (Chappes, France) had a higher varying temperature from 25 °C to 40 °C, however the 16 h light and 8 h dark photoperiod with LED lamps was the same as the Leeds growth conditions (3.2.7.1).

3.2.7.2 Wheat Phenotyping

The plants were checked every 2 days for any visible changes in phenotype and photographs were taken of the full plant every 2 weeks until the tips of spikes were

present. Any plants that died were marked as such on the date that it occurred. 100 mg leaf samples of the wheat plants were taken throughout development.

The height of the transgenic wheat plants was measured using a tape measure from bottom of shoot to last emerged leaf every week from 5 weeks until senescence commenced. This experiment was repeated for early stage development from germination to mid-development. Other measurements that were taken include the number of tillers, number of leaves per tiller, date of anthesis and number of spikes. Photographs of wheat plants were taken at early development, mid development and late development at 5 week intervals.

3.2.7.2.1 Chlorophyll Pigment Measurements

The Soil Plant Analysis Development (SPAD) 502 plus chlorophyll meter (Minolta Camera Co., Osaka, Japan) provides an instantaneous measurement of chlorophyll content. The SPAD value given is the difference between the transmittance of red (650 nm) and infrared (940 nm) light through the leaf (Uddling *et al.,* 2007). Five measurements were taken on the flag leaf of each plant every week to track chlorophyll pigment throughout growth.

3.2.7.3 qPCR

Samples of 120 mg of leaf tissue were collected into a 96-well plate and a tissue lyser was used to grind it into a powder. The RNeasy QIACUBE 96 HT (Qiagen, Hilden, Germany) kit was used to isolate RNA and optimised for plant RNA extraction using the 'Isolation of total RNA from plants' protocol on the Qiagen website (Qiagen, 2006). Initial quality controls were carried out using the Agilent RNA 6000 Nano (Agilent Technologies Inc, Santa Clara, California, USA) as per the kit specifications.

3.2.7.3.1 Quality Controls

After extraction, the RNA was assayed in a UVSTAR 384- well UV plate (Greiner Bio-One, Kremsmünster, Austria), along with a blank sample to calibrate the assay. This was created using the Biomek FX pipetting robot (Beckman, Brea, California, United States) as per the instructions and then analysed on the Infinite M200PRO (Tecan Group Ltd, Männedorf, Switzerland).

R software was used to analyse the results with the concentrations of the quantity of the RNA and the 260/280 ratio. The acceptable lower limit was placed at 35 ng / μL, the upper limit was 570 ng / μL and the A260/A280- ratio was placed at a threshold greater

than 1.8. After this assay was performed, the RNA was normalised to a 35 ng /µL concentrations in a final volume of 16.2 µL with the Biomek FX pipetting robot (Beckman, Brea, California, United States). The plates were spun down in a centrifuge at 500 rcf for 30 seconds and then stored at -80 ºC for subsequent use. This included a 2 µL aliquot for control of gDNA contamination and the remaining 14.2µL was used for cDNA synthesis.

3.2.7.3.2 Control of gDNA contamination

The dilution of the RNAs and creation of the qPCR plate was carried out using the Biomek FX pipetting robot (Beckman, Brea, California, United States) according to the Biogemma protocols. These protocols were carried out during the iCASE placement in Limagrain, France. The qPCR plate was created by adding the primers, FastStart Universal SYBR Green Master Rox (Roche, Basel, Switzerland) to the qPCR plates and ran on the LightCycler480 (Roche, Basel, Switzerland).

3.2.7.3.3 Synthesis of cDNAs and Reverse Transcription

A mix was prepared of 5 µL of each primer and 90 µL of ddH_2O which contained a sufficient amount of the Primer Mix [Table 4] and Master Mix in the order described [Table 5] for all of the samples, plus an additional 10 sample volume for error. The mix was then vortexed for 10 seconds and the Biomek FX pipetting robot (Beckman, Brea, California, United States) was used to apply the mix to each well of the prepared plate, then centrifuge the plate at 500 rcf for 30 seconds.

Table 4: qPCR primers for wheat

Target	Forward Sequence	Reverse Sequence
TaWHY1	GGAGGAGCGATGAAGGTAA	CCCATGATGTGCGGTATGATG
TaAct1 (Housekeeping)	ACTGGGACGACATGGAGA	TCTTGGCATTTTCCAGCTCT

Table 5: Primer mix for qPCR

Reagent	Volume for 96 well plate reaction (µL)
ddH_2O	552
Forward Primer	24
Reverse Primer	24
Total:	600

Table 6: Master Mix for qPCR

Reagents	Volume for one reaction (μL)
Primer Mix	5
SYBR Green	12
Total for one well	17
cDNA for each well	2

The qPCR was carried on the CFX96 Real-Time PCR Detection System (Bio-Rad, Hercules, California, United States) under the qpcr_CFX_3StepAmp+Melt protocol at a Tm of 50 °C. Further information into the analysis and use of the ΔCT method was described in Chapter 6: Wheat.

4 EFFECT OF WHIRLY ON GERMINATION IN ARABIDOPSIS

4.1 INTRODUCTION

The foundation of the majority of world agriculture starts at seed germination. This initiation of plant development is important in both industrial and economical settings, as well as in natural ecosystems. Seed dormancy ends when appropriate environmental cues including temperature, light and phytohormones, such as gibberellic acid (GA) and abscisic acid (ABA), are integrated by a spatially embedded decision-making centre to establish seed germination (Topham *et al*., 2017). Germination can be defined as the imbibition of water and emergence of the radicle from the seed coat (Bewley and Black, 1994). Germination is usually recorded as the percentage of germinating seeds, defined by vigour and viability. Seed vigour is generally defined as the potential performance of viable seeds in agricultural practice and it is determined by many complex genetic and environmental interactions (Finch-Savage and Bassel, 2016). A delay in radicle emergence, which increases the time to germinate, reduces seed vigour. Seed viability is the overall rate of successful germination (Waterworth *et al.,* 2019).

The genetic determinants of viability and vigour have not yet been fully elucidated. However, reduced seed quality is associated with DNA damage caused by various processes ranging from natural ageing to chemical and UV damage, which can occur in aged seeds (Waterworth *et al.,* 2019). Seed aging is especially important when considering the collection of seed resources in agriculture. There has been extensive work to preserve germplasm reserves throughout the world for use in research to improve food and nutritional security, plant breeding, and education. The establishment of *ex situ* conservation seed vaults, most notably the Svalbard Global Seed Vault, has provided extra protection for seed resources against natural or man-made disasters (Asdal and Guarino, 2018).

In the experiments described in this chapter, seed vigour was determined as the time taken to germinate and seed viability was denoted as the final germination percentage. These parameters are important, because yield losses and reduced uniformity of crops caused by low vigour or low viability seed lots can be costly to the producer and consumers. Furthermore, low yielding seeds showed a heightened vulnerability to abiotic and biotic stresses as well as competition with weeds (Finch-Savage and Bassel, 2016; Waterworth *et al.*, 2019). WHIRLY proteins are involved in wide-ramge of abiotic and biotic stress responses (Taylor *et al*., 2022), as well as having effects on competition in some plants

(Graña *et al.*, 2020). Therefore, elucidating the function of WHIRLY in seed germination and seed ageing may reveal further functions which will be beneficial to agriculture.

4.1.1 Previous studies on WHIRLY in germination

There have been limited studies on the functions of WHIRLY proteins in germination. In the nucleus WHY1 acts as a transcriptional repressor of genes such as *WRKY53*, which is implicated in regulating senescence and post germination growth (Huang *et al.*, 2018). Another member of the WRKY family, *WRKY38*, regulates salicylic acid (SA) metabolism in *Hordeum vulgare* (Xie *et al.*, 2007). Furthermore, in Arabidopsis WHY1 was found to function as a downstream component of the SA defence pathway (Desveaux *et al.*, 2004). Such findings open up many questions about WHY1 protein interactions and their effects on germination and growth.

A recent study showed a link between the application of SA and improved germination in *Limonium bicolor* seeds under salt stress (Liu *et al.*, 2019). Moreover, exogenous treatment with SA resulted in a three-fold increase in *AtWHY1* and a four-fold increase in *AtWHY3* transcripts in Arabidopsis plants, together with a decrease in Arabidopsis *KINESIN-LIKE PROTEIN 1* (*AtKP1*) transcripts (Xiong *et al.*, 2009). AtKP1 is a member of the kinesin-14B subfamily in Arabidopsis, which is involved in microtubule-based movement, aerobic respiration and regulation of seed germination at low temperatures (Yang *et al.*, 2011). AtWHY1 and AtWHY3 are transcription factors binding to the *AtKP1*-related element-binding factor 1 (KBF1). ChIP assays showed upstream binding of both WHY proteins to the *AtKP1* promoter *in vivo*. Additionally, a two-fold decrease in *AtKP1* transcripts was reported in AtWHY1-OE and AtWHY3-OE lines (Xiong *et al.*, 2009). Interestingly, Arabidopsis *PATHOGENESIS RELATED* 1 (*AtPR1*), a SA signal marker gene was activated eight-fold in AtWHY1-OE and twenty-five-fold in AtWHY3-OE lines, whilst overexpression of both WHYs led to a fifty-fold increase in *AtPR1* (Xiong *et al.*, 2009). These results confirm that both WHY1 and WHY3 are involved in SA-signalling pathways.

Many studies have shown antagonistic relationships between SA and ABA in germination (Cao *et al.*, 2011). However, *Atwhy1* mutants and transgenic WHY1-OE lines showed reduced sensitivity to both SA and ABA during germination (Isemer *et al.*, 2012). Seedlings with WHY1 only in the nucleus displayed a similar insensitivity to ABA as the *Atwhy1* mutants. These studies suggest that the responsiveness of germination to ABA is altered by the localisation of WHY1 in the plastids or nuclei, the plastid-localised WHY1

form being responsible for enhanced responsiveness to ABA (Isemer *et al.*, 2012). While the mechanisms involved in this regulation remain to be fully explained, these results suggest that the intracellular localisation of WHY1 has a significant impact on the synthesis and responsiveness to ABA and SA (Isemer *et al.*, 2012, Lin *et al.*, 2020). No studies to date however, have considered the impact of interactions between the WHY1, WHY2 and WHY3 proteins on germination. Previous studies have found that overexpression of SlWHY2 in *Nicotiana benthamiana* enhanced drought stress tolerance during seed germination and seedling growth (Zhao *et al.*, 2018). A more recent study showed reduced seed viability in *Atwhy2* mutants. Disruption of WHY2 led to altered mitochondrial structure, disordered nucleoids and increased expression levels of AtWHY3 compared to the wild type (Golin *et al.*, 2020).

4.1.2 Role of WHIRLY in DNA Repair Mechanisms

The putative DNA binding domain, KGKAAL that is characteristic of WHY proteins is essential for DNA binding but mutations in this domain do not disrupt tetramerization or protein structure (Desveaux *et al.*, 2002). The ssDNA binding function is inherent to the role of WHY proteins in DNA repair mechanisms, as illustrated in Figure 11 (Desveaux *et al*., 2002; Grabowski *et al*., 2008). Characterisation of the *Atwhy1* and *Atwhy3* mutants has revealed their roles in DNA in repair mechanisms. These mutants are phenotypically similar to wild type plants under normal growth conditions (Maréchal *et al*., 2009). The functions of the WHY proteins in germination will be explored in this chapter, together with effects on shoot development.

The plastid genome is protected by WHY1 and WHY3 binding (Figure 11; Maréchal *et al*., 2009). The *Atwhy1why3* double mutants had reported variegation in 4.6% of plants [covered in more depth in Chapter 5], which may have occurred through plastid DNA rearrangements occurring during microhomology-mediated break-induced replication (MMBIR). This process is repressed by WHY proteins (Figure 11a; Maréchal *et al.*, 2009). The plastid interference observed in the *Atwhy1why3* mutants is more prominent when other chloroplast proteins, such as Pol1B, described in 1.3.3 (Parent *et al.*, 2011), or chloroplast localised ATP-dependent recombinase 1 (RecA1) (Zampini *et al*, 2015; Duan *et al*., 2020), are also mutated. Pol1B is a DNA polymerase required for chloroplast DNA replication. The lack of Pol1B increases replication errors. The *Atwhy1why3pol1b-1* triple mutants had severe phenotypes with variegated leaves, growth defects, lower photosynthetic rates and an accumulation of reactive oxygen species (Lepage *et al.*, 2013, Parent *et al.*, 2011).

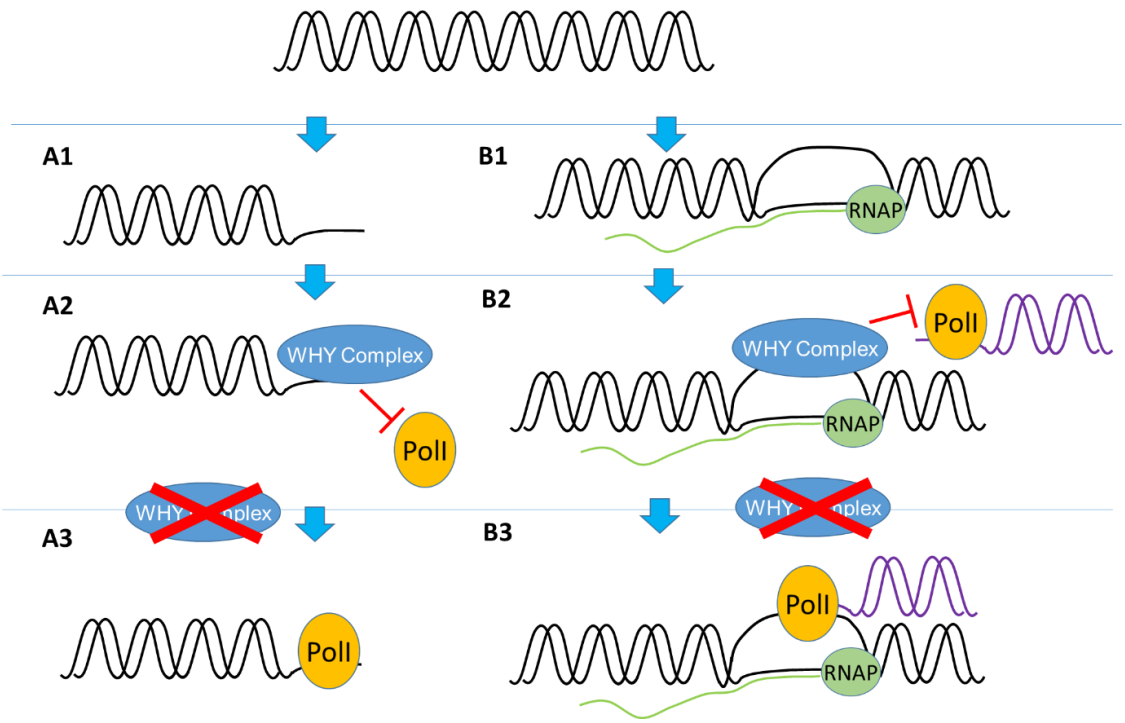


Figure 11: Model of WHIRLY involvement in DNA repair pathways. WHIRLY is involved in two pathways; A1) by blocking microhomology-mediated end joining (MMEJ) by DNA polymerases in the presence of AtWHY2 where single-stranded regions are longer than 12 nucleotides (García-Medel *et al.*, 2019) and B1) the accumulation of RNA: DNA hybrids (Pérez Di Giorgio *et al.*, 2019). The transcribed DNA strand (black) is associated with the RNA (green) and RNA polymerase (RNAP). WHY proteins protect chloroplast and mitochondrial DNA by preventing the error-prone MMBIR pathway in both chloroplasts and mitochondria (Cappadocia *et al.*, 2010). This leads to the formation of A2) error-free homologous recombination, or B2) R loops. In the absence of WHY proteins the double strand break could be repaired by MMEJ/ microhomology-mediated break-induced replication (MMBIR) by the recruitment of PolI A3) alone or B3) in conjunction with RNA, which are both error-prone pathways.

The *Atwhy1why3reca1* triple mutant also has a similarly severe phenotype with a 60-fold increase in DNA rearrangements compared to the wild type and a reduction in seed viability to 41% (Duan *et al*., 2020). The quadruple mutant *Atwhy1why3reca1pol1b* is embryo lethal (Zampini *et al*, 2015). Pérez Di Giorgio *et al.* (2019) showed that mutations in WHY1, WHY3, RecA1 and Pol1B reduced transcriptional levels in Arabidopsis plastids with increases in RNA: DNA hybrids, also known as R loops, due to DNA rearrangements. High levels of transcription can lead to increased mutational load and increases in R loops which can displace untranscribed DNA strands [Figure 11b]. Furthermore, the *Atwhy1why3rnh1c* triple mutant had a severe bleaching phenotype (Wang *et al*., 2021), suggesting increased plastid genome instability. Analysis using fluorescence microscopy and immunoprecipitation assays showed colocalization of WHY1 and WHY3 with RNase H family protein (RNH1C) and RecA1 in Arabidopsis

(Wang *et al*., 2021). These findings suggest that WHY proteins work in unison with other plastid proteins to protect against DNA damage, which could be considered to occur, for example, during germination.

4.1.3 WHIRLY Expression Levels During Germination

RNA-seq analysis was used to investigate the transcripts that were differentially expressed during germination (Narsai *et al*., 2011). The time points analysed in this study were freshly harvested from the silique (Silique Harvest), 15 day old seeds imbibed in darkness (0 S), 4°C dark stratified seeds at 1 h (1 S), 12 h (12 S), 48 h (48 S), which were then moved to continuous light for 1 h (1 L), 6 h (6 L), 12 h (12 L), 24 h (24 L), and 48 h (48 L). The 24 L time point was considered by the authors to be the time of radicle emergence from the testa, while the 48 L time point was stated as post-germination (Narsai *et al*., 2011). This RNA-seq dataset is available on an eFP-Seq browser (http://bar.utoronto.ca/eplant/). WHY1, WHY2 and WHY3 were identified amongst the differentially expressed genes [Figure 12].

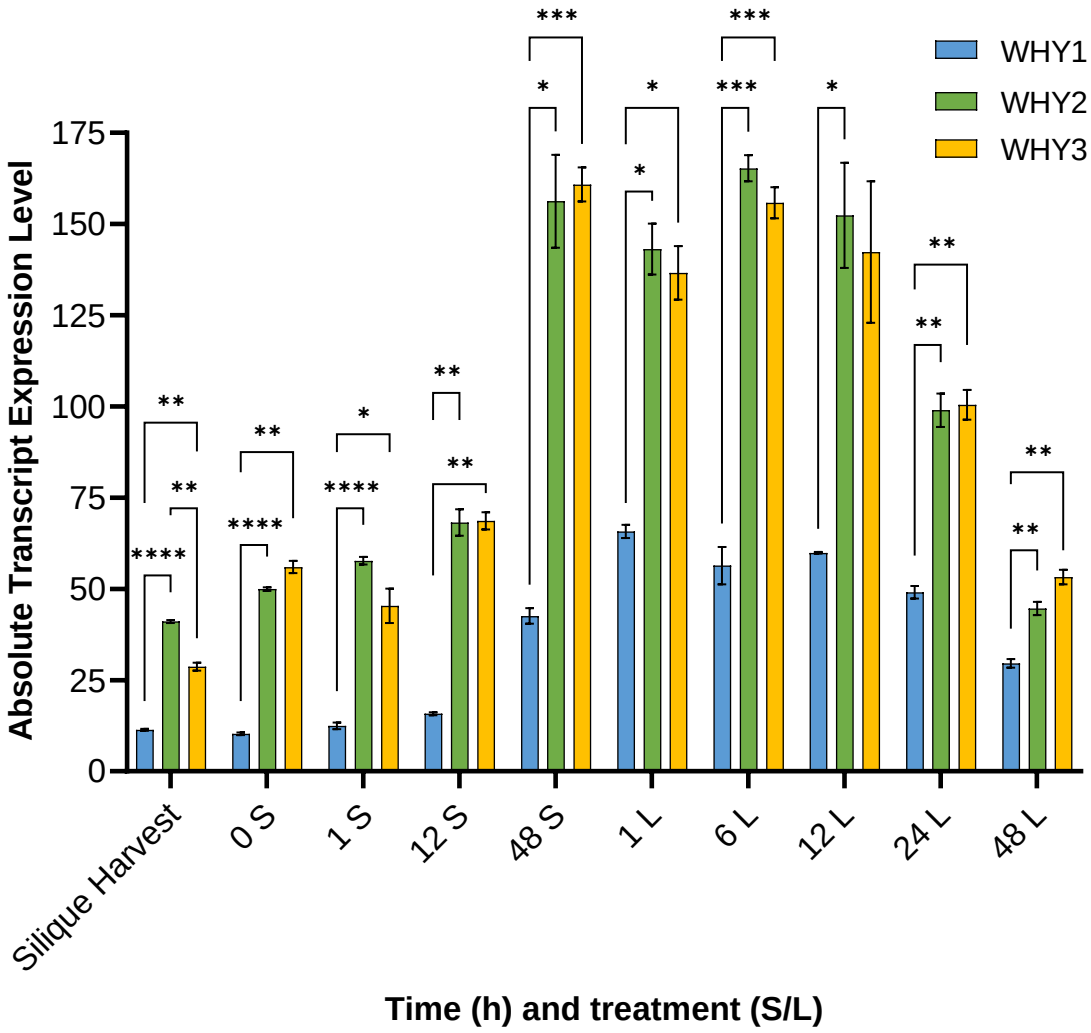


Figure 12: Mean absolute transcript expression levels of WHY1, WHY2 and WHY3 straight from harvest from the silique, during 0 to 48 hours in cold stratification treatment (S) and then from 1 to 38 hours in a light growth chamber (L). Raw germination eFP data (RNA-Seq data): WHY1 (AT1G14410, blue); WHY2 (AT1G71260, green) and WHY3 (AT2G02740, yellow) from BAR ePlant (http://bar.utoronto.ca/eplant/) on 07/04/2020, means were calculated from 3 samples and error bars are standard error of means. ANOVA and Tukey's pairwise comparison test was conducted: *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$; ****, $p < 0.0001$. Data was previously published in Narsai *et al.* (2011).

There was significantly higher absolute transcript expression levels of both WHY2 and WHY3 compared to WHY1 from silique harvest all the way to 48 hours of germination in the light (48 L) [Figure 12]. The transcript expression levels of all three WHYs increase more than two-fold from 12 S to 48 S and then maintain similar levels throughout the first 24 hours of light treatment (24 L) [Figure 12]. The highest WHY1 transcript expression levels during germination occurred after just 1 hour in the light (1 L), whilst WHY2 and WHY3 had much a higher expression after 6 hours in the light (6 L) and both WHY2 and WHY3 expression levels were significantly higher than WHY1 [Figure 12]. Although the

expression of WHY2 and WHY3 was decreased at 48 h of germination, these were still significantly larger than WHY1. These data suggest that the WHY1 and WHY3 proteins fulfil different roles in germination.

4.1.4 Objectives

There have been several studies on the interaction of WHY proteins and phytohormones in germination but there is no published work on the potential involvement of WHY proteins in DNA repair in germination. As stated above, DNA repair is a significant determinant of seed longevity, important in the preservation of germplasm reserves. Both WHY1 and WHY3 have been implicated in functioning in DNA repair mechanisms but display different expression patterns [Figure 12]. Therefore, the aims of the studies reported in this chapter are twofold:

1. Explore the roles of WHY1 and WHY3 in seed longevity and germination using Arabidopsis *Atwhy1*, *Atwhy3* and *Atwhy1why3* mutant lines.
2. Explore the potential roles of WHY1 and WHY3 in germination using Arabidopsis *Atwhy1*, *Atwhy3* and *Atwhy1why3* mutant lines that have been subject to accelerated ageing.

4.2 RESULTS

4.2.1 Both WHY1 and WHY3 are required for germination

Adverse environmental conditions such as high temperature and high relative humidity result in genome damage through oxidative modifications (Job *et al*., 2005; Waterworth *et al*., 2016; Waterworth *et al*., 2019). To investigate the role of WHY1 and WHY3 proteins in germination, seeds of the single and double mutants were placed under accelerated ageing conditions of high temperature (35 °C) and high relative humidity (80 %) for a period of 7 or 14 days (as described in Materials and Methods: 3.2.4.2). Germination was then compared in the aged and unaged seeds.

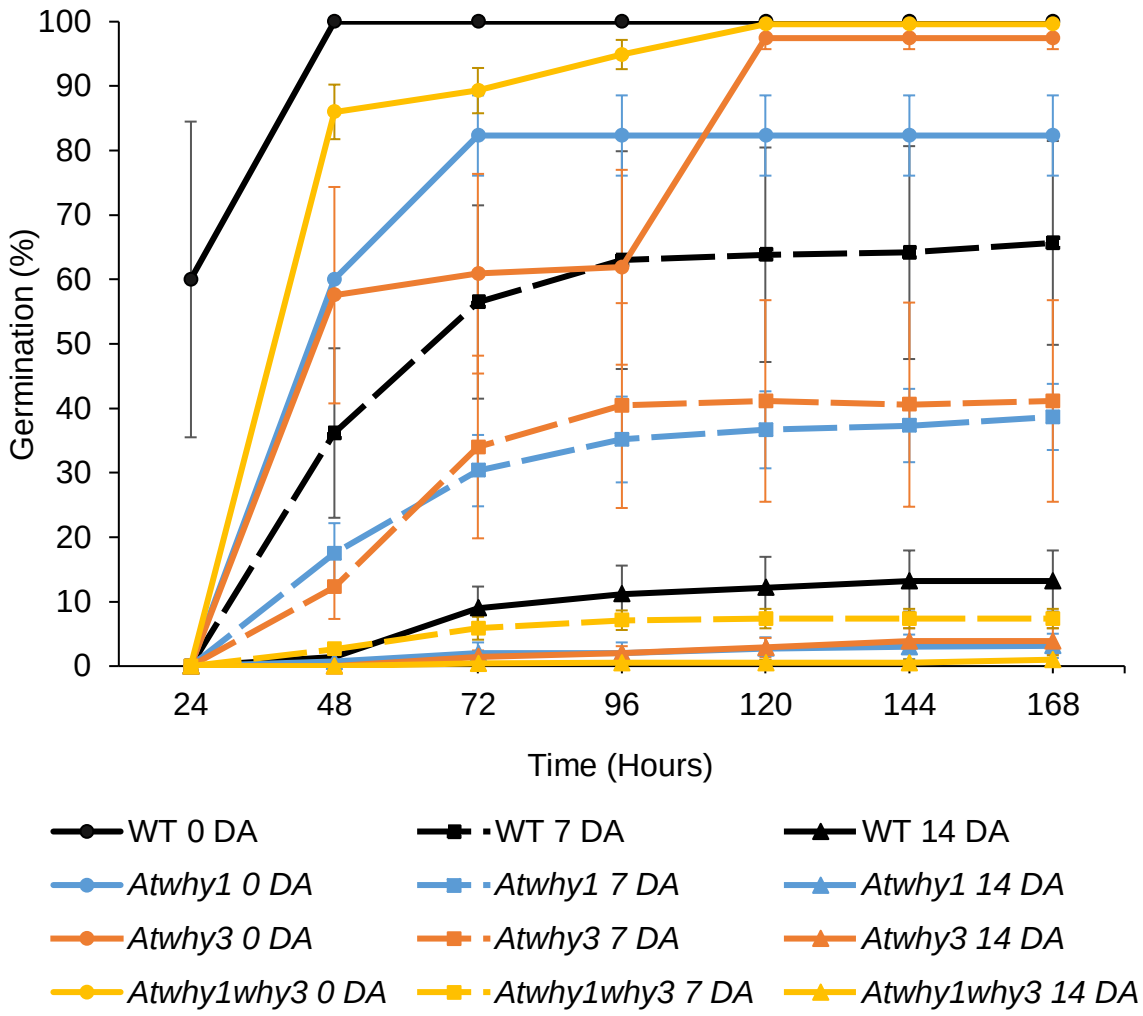


Figure 13: Mean percentage of seed germination after 48 h stratification. Each genotype is shown: unaged, 0 DA (days aged) circle unbroken line; 7 DA square broken line and 14 DA triangle dotted line, from 24 h post stratification to 168 h (7 days). Error bars are standard error of means calculated from 3 biological replicates, each with 100 seeds per replicate.

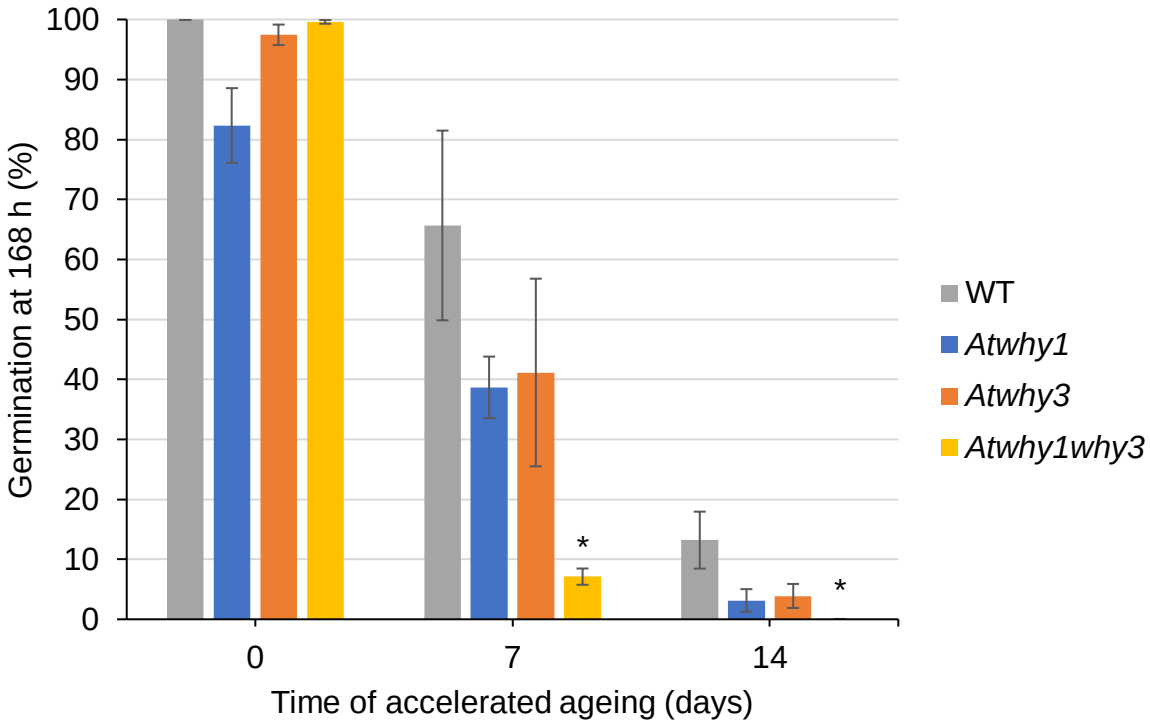


Figure 14: Mean percentage of seeds germinated at 168 h (7 days) after 48 h in cold stratification. Each treatment is shown per genotype. Error bars are standard error of means calculated from 3 biological replicates, each with 100 seeds per replicate. There was no germination in any *Atwhy1why3* seeds after 14 days ageing. T test against WT per aged section *, p <0.05.

Table 7: Two sample t test for each ageing treatment compared to high quality seeds of the same genotype.

Genotype	7 days aged seeds	14 days aged seeds
WT	0.082	5.3×10^{-5}
Atwhy1	3.4×10^{-3}	3.4×10^{-3}
Atwhy3	0.011	2.3×10^{-16}
Atwhy1why3	3.3×10^{-13}	1.2×10^{-19}

The first 24 - 48 h period is generally considered to be the time when the radicle emerges from the endosperm in Arabidopsis seeds, which are then said to have germinated. This crucial time point is important for comparisons of control and age-accelerated seeds. This time point has previously been identified as important in mutants involved in DNA damage via ageing (Waterworth *et al.*, 2016). The high-quality unaged seeds of all

genotypes germinated within the first 48 h post cold stratification (WT, 100 %; *Atwhy1*, 60 %; *Atwhy3*, 57.6 %; *Atwhy1why3*, 86 %). This rate of germination was reduced at 7 days in aged seeds of all genotypes (WT, 36.17 %; *Atwhy1*, 17.5 %; *Atwhy3*,12.29 %; *Atwhy1why3*, 2.6 %) and then again in 14 days aged seeds (WT, 1.4 %; *Atwhy1*, 0.71 %; *Atwhy3*, 0 %; *Atwhy1why3*, 0 %) [Figure 13].

Due to the severely delayed germination rate in aged seeds, the seeds were scored for viability after 7 days (168 h) post cold stratification [Figure 14]. The high-quality unaged seeds of all genotypes maintained a high vigour at this time point (WT, 100 %; *Atwhy1*, 82.3 %; *Atwhy3*, 97.4 %; *Atwhy1why3*, 99.6 %). There were no significant differences between the mutants and the wild type in the absence of accelerated ageing at 7 days post cold stratification. Germination vigour was reduced in 7 day-aged seeds of all genotypes, the effect being most pronounced in the mutant lines (WT, 65.7 %; *Atwhy1*, 38.7 %; *Atwhy3*, 47.14 %; *Atwhy1why3*, 7.38 %). Germination vigour was reduced further in the seeds aged for 14 days (WT, 13.2 %; *Atwhy1*, 3.14 %; *Atwhy3*, 3.89 %; *Atwhy1why3*, 1 %) [Figure 14]. The germination of 7 day- and 14 day-aged seeds was more irregular and delayed. Therefore, the decrease in seed vigour was greater in the aged mutant seeds compared to WT and high-quality unaged seeds [Figure 14]. The apparent decreases in end viability of germinated high quality unaged *Atwhy1* seeds (after 168 h) was not significant relative to control WT seeds. In contrast, significant decreases in the viability of *Atwhy1why3* double mutant seeds was observed after 7 ($p = 0.014$) and 14 days of ageing ($p = 0.0498$) compared to the aged wild type seeds [Figure 14]. When these data are compared against the high-quality unaged seed of each genotype, the differences become clearer [Table 7]. All mutant lines had significantly reduced viability after 7 and 14 day of ageing. Hence, WHY1 and WHY3 appear to have an additive effect because seeds lacking both proteins show a greater reduction in seed viability.

4.2.2 Viability and Permeability Assays

Higher plants have the testa (seed coat) and endosperm that protect the embryo from physical and chemical damage (Debeaujon *et al*., 2000). Many studies have been conducted on the interactions between seed coat permeability and germination, namely its role in dormancy and defence against pathogenic attack (Debeaujon *et al*., 2000; De Giorgi *et al*., 2015; Vishwanath *et al*., 2013). Furthermore, conditions during seed storage may affect permeability and the ability of the seed to germinate (Debeaujon *et al*., 2007). Therefore, it is important to gain an understanding of seed permeability during germination.

The tetrazolium penetration assay is commonly used to test seed viability (Wharton, 1955). (Vishwanath *et al*., 2013). This assay uses 2, 3, 5-triphenyltetrazolium chloride, a cationic dye, which turns from a colourless compound into the scarlet-red coloured formazan in a reaction catalysed by dehydrogenases in the endoplasmic reticulum of respiring tissues (Porter *et al*. 1947). Tetrazolium enters the cells of living and dead tissue, but it is only converted into formazan in viable, living tissues. Therefore, seeds which remain unstained are non-viable and would not normally germinate. Additionally, the intensity of the red colour is directly proportional to seed permeability (Porter *et al*. 1947; Wharton, 1955). The absence of respiration in aged seeds prevents formazan production and so aged tissue remains unstained (Verma *et al*., 2013). This assay, which is particularly useful for providing insights into the viability of aged seeds, was used to compare the viability of the three *Atwhy* mutant lines with the wild type [Figure 15]. Control high quality unaged seeds and seeds that had undergone accelerated ageing treatment for seven days were used in this analysis.

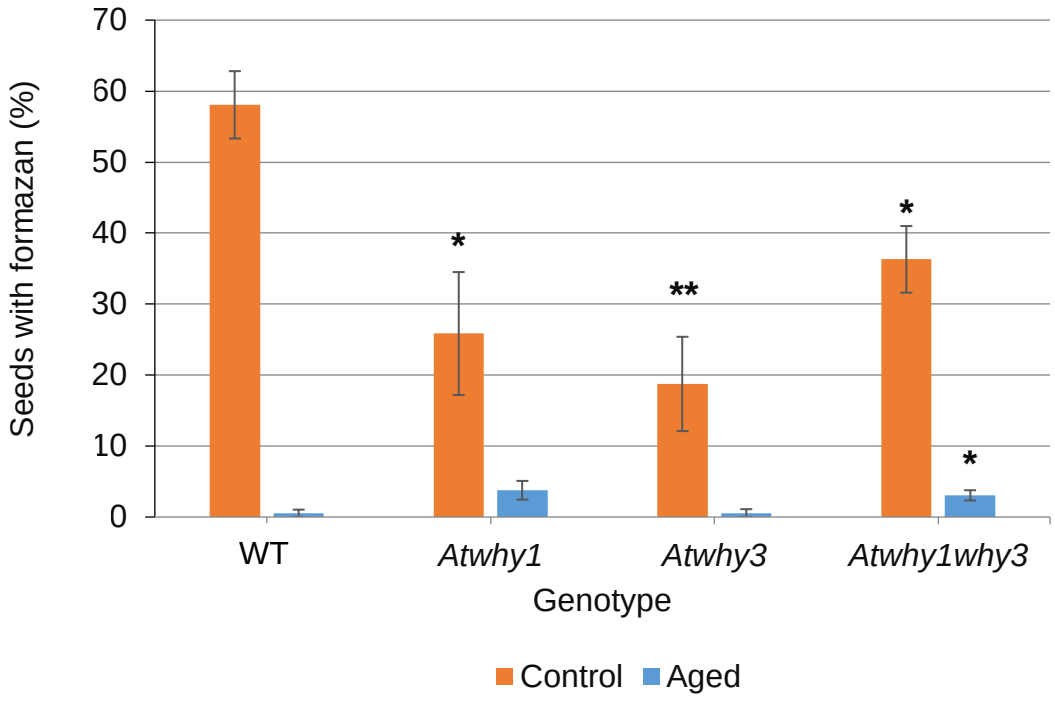


Figure 15: Percentage of seeds showing red formazan after 48 h of incubation with tetrazolium, high quality control seeds (orange) and 7-day aged seeds (blue) were compared from photographs of the stained seeds using ImageJ analysis. Error bars are standard error of means calculated from 3 biological replicates, each with 100 seeds per replicate. T test against WT of the same level of ageing: *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$.

The high quality unaged wild type seeds showed the highest percentage of formazan. The high quality unaged *Atwhy* seeds showed significant reductions in the percentage of seeds containing formazan relative to WT (WT, 58.07%; *Atwhy1*, p=0.045; *Atwhy3*, p=0.0085; *Atwhy1why3*, p=0.031) [Figure 15]. The *Atwhy3* mutants had the greatest loss of viability relative to the WT in the absence of ageing. Interestingly, the *Atwhy1why3* double mutant had more seeds with formazan than either the wild type or the single mutants (*Atwhy1*, 25.85%; *Atwhy3*, 18.75%; *Atwhy1why3*, 36.3%). They therefore had significantly more viable seeds than the other lines [Figure 15]. Surprisingly, seeds with only WHY1 or WHY3 present had less formazan, but seeds with both WHY1 and WHY3 missing showed increased levels of formazan.

Ageing resulted in a reduction in the percentage of seeds contain formazan compared to the high-quality unaged seeds of the same genotype; WT was reduced more than 100-fold from high quality to aged whilst the *Atwhy1*, *Atwhy3* and *Atwhy1qhy3* were reduced 7-, 34- and 12-fold respectively [Figure 15]. The only significant difference in percentage of seeds with formazan compared to the WT was observed in the *Atwhy1why3* double mutants, which had significantly more ($p = 0.047$) seeds with formazan than the aged wild type line [Figure 15].

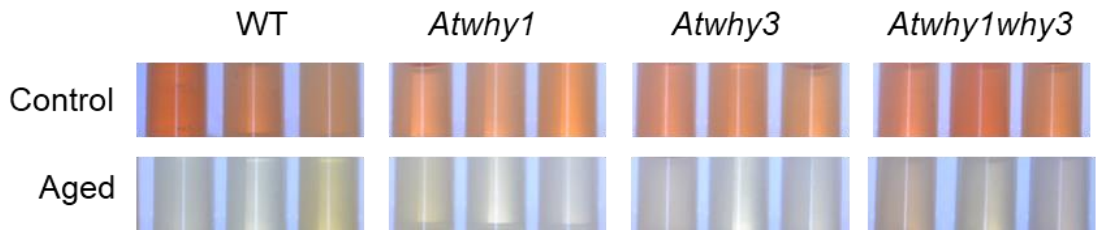


Figure 16: Colour of formazan extracts for high quality control seeds and 7-day aged seeds were compared in 1.5 mL microcentrifuge tubes, photographed after absorbance measurements were taken. These samples are from 48 h incubations with red tetrazolium.

Furthermore, it was important to rule out issues with seed permeability after ageing treatment and to ensure that this was not causing the change in germination phenotype in aged seeds as described in 4.2.1. Extracts with a higher intensity of red colouration allowed more of the tetrazolium to enter cells and converted to formazan and therefore were more permeable than lighter coloured extracts. The extracts from control seeds had visibly more formazan than extracts from aged seeds of all genotypes [Figure 16, Figure 17]. The high quality seeds showed variations in the level of red staining, whilst the aged seeds lacked red coloration and differences in shading were far less distinct [Figure 16,

Figure 17]. The red colour in seed extracts was quantified using a spectrophotometer at 484 nm [Figure 17] (Vishwanath *et al*., 2014).

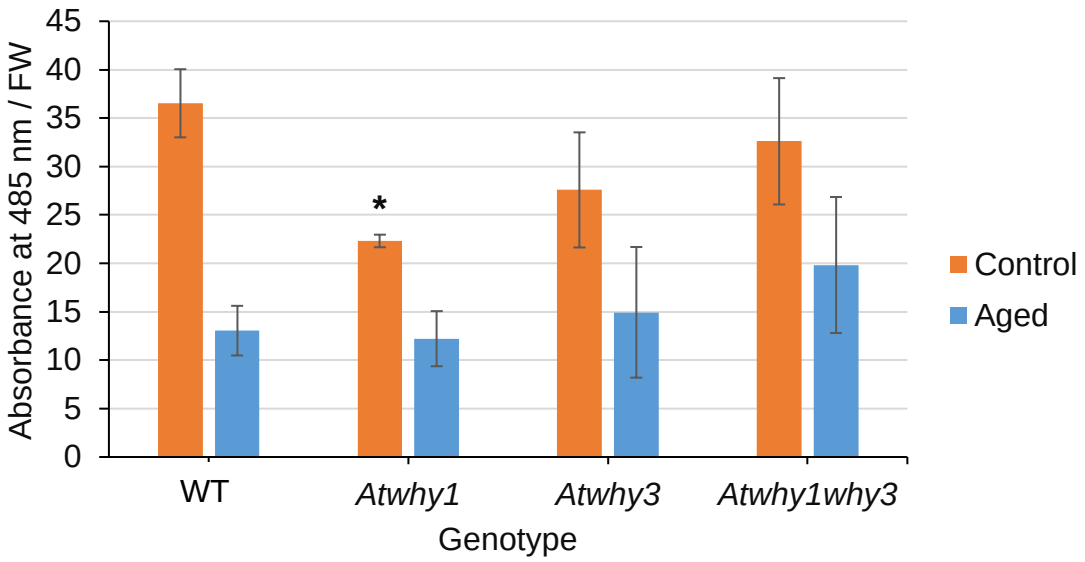


Figure 17: Seed coat permeability assay with absorbance per fresh weight (g) of the seed sample, high quality control seeds and 7-day aged seeds were compared. Absorbance was taken at 485 nm on a spectrophotometer after 48 h of incubation. Error bars are standard error of means calculated from 3 biological replicates and 3 technical replicates; a T test was performed against WT of the same level of ageing: *, p <0.05.

The high-quality control unaged seeds showed a consistently higher absorbance than aged seeds, indicting a higher number of viable seeds [Figure 17]. A significant difference in absorbance was observed between control and aged treatments for the WT (p=0.0057) and *Atwhy1* mutants (p=0.026) which showed that the ageing treatment probably reduced the permeability of the seeds. However, the absorbance values (per fresh weight) were similar in both the high-quality unaged control seeds and the aged seeds, only the *Atwhy1* high unaged quality seeds (p = 0.016) were significantly different from the wild type of the same treatment [Figure 17]. This finding suggests that WHY1 may play a role in permeability protection in high-quality unaged seeds. However, the double mutant, which is lacking in WHY1 and WHY3 had had significantly reduced germination did not show a lower absorbance than the WT in either control unaged or aged seeds, therefore these results cannot account for the change in phenotype between high quality and aged seeds.

4.3 DISCUSSION

Germination is a crucial development transitional stage in the plant life cycle. As such understanding Arabidopsis seed viability has been a topic of research for several decades. However, many interactions between proteins, transcriptional factors and phytohormones still remain unclear (Debeaujon *et al*., 2000; De Giorgi *et al*., 2015). Previous studies on the role of WHY1 in germination have concentrated on interactions with SA and the role of WHY1 localisation in the plastids and nucleus (Isemer *et al.*, 2012; Lin *et al.*, 2020). There have been no studies to date on the function of WHY1 and WHY3 in germination or seed ageing (Desveaux *et al*., 2002; Maréchal *et al*., 2009; Parent *et al.*, 2011, Lepage *et al.*, 2013, Zampini *et al*, 2015; Pérez Di Giorgio *et al.* 2019; Duan *et al*., 2020; Wang *et al*., 2021). The data presented here is important because it reveals new functions for the WHY1 and WHY3 proteins in seed viability and ageing.

Previous studies had shown that the AtWhy1 gene was expressed less than other AtWhy forms during germination [Figure 12] (Narsai *et al.*, 2011). However, the data presented here show that the viability of the single and knockout lines were similar to the wild type. Ageing for 7 days decreased seed vigour in all genotypes. This treatment delayed germination more in the mutant lines than the WT. For example, the *Atwhy* mutant seeds germinated in the 96 - 120 h range compared to 48 h to the wild type seeds [Figure 13]. There appears to be a degree of functional redundancy between WHY1 and WHY3 in protecting the seeds against ageing. However, end-viability was significantly reduced in the double mutants such that only 1 % of the *Atwhy1why3* seeds germinated after 168 h in the light after 14 days ageing. These data demonstrate that both WHY1 and WHY3 are required for full germination.

Germination was reduced by ageing for 7 days and 14 days in all genotypes, as would be predicted from previous studies on seed ageing (Waterworth *et al*., 2016). The germination of the *Atwhy1* and *Atwhy3* lines was reduced compared to high quality unaged seeds, similar to the WT. Crucially, however, the *Atwhy1why3* seeds showed the highest levels of inhibition of germination [Table 7]. These findings suggest that there is a degree of functional redundancy between WHY1 and WHY3 and that only when both forms are knocked-out does viability differ significantly from aged wild type. This data shows that both WHY1 and WHY3 are required for protection against ageing. However, it is not clear how this functional redundancy between the two proteins work and whether plant species without WHY3 would have a similar reduction in germination of aged seeds. A further germination study is required, including a wider range of the commonly studied

species that do not have the WHY3 protein, such as *Zea mays*, *Hordeum vulgare* and *Solanum lycopersicum*. These studies would help to identify the function of WHY1 in germination further and may provide an explanation for the redundancy between WHY1 and WHY3.

The *Atwhy1* and *Atwhy3* seeds had reduced germination after 7 and 14 days of ageing compared to the high-quality unaged seeds of the same genotype [Table 7]. However, the double *Atwhy1why3* mutants were hypersensitive to ageing. The aged seeds of this genotype showed the most significantly reduced viability and vigour compared to wild type seeds (Figure 13, Figure 14). Such changes are reminiscent of those observed in mutants lacking DNA repair factors, which also show significantly delays in germination (Waterworth *et al.*, 2016). Furthermore, the ageing treatment resulted in a significant decrease in the percentage of seeds containing formazan [Figure 15]. The reduced viability observed in the mutants lacking WHY1 may explain the lower germination observed in the high quality *Atwhy1* seeds [Figure 14]. Seed permeability was only significantly decreased in the high quality *Atwhy1* seeds, which suggests a potential role for WHY1 in protecting the seed coat against damage during germination this may provide some explanation for the reduction in germination of high quality *Atwhy1* seeds. However, this does not elucidate function of WHY in germination that caused the significant reduction in viability and vigour in the *Atwhy1why3* aged seeds. However, there is a complex interaction of environmental and genetic factors that determine seed longevity at the molecular level; these mechanisms especially in regards the function of WHY proteins in DNA repair are yet to be fully understood (Taylor *et al.*, 2022). These germination studies are limited in that they do not elucidate the mechanism involving WHY during germination. The RNA-seq analysis described in the next chapter serves to identify these mechanisms which underpin the ageing hypersensitivity phenotype in the *Atwhywhy3* double mutants. The identification of differentially expressed genes will further clarify the functions of WHY1 and WHY3 in germination, as well as revealing potential interactions with SA and ABA (Isemer *et al.*, 2012; Xiong *et al.*, 2009).

5 WHIRLY GENE INTERACTIONS DURING AGEING HYPERSENSITIVITY

5.1 INTRODUCTION

Following the results on seed germination reported in Chapter 4, it is important to explore the roles of WHY1 and WHY3 in germination and accelerated ageing. The identification of the mechanisms leading to the ageing hypersensitivity observed in the *Atwhywhy3* mutants will provide useful new information concerning the DNA repair process in seeds that may be useful for the conservation of germplasm reserves and long-term storage of seed stocks This is particularly important in conserving rare crop wild relatives (Waterworth *et al*., 2019).

Telomeres preserve genome integrity. Shortened telomeres can cause seeds to age and reduce in viability and so the protection of telomeres is particularly important in seed preservation (Donà *et al* 2013). Telomere monitoring has long been used as a marker for seed ageing, differences in telomere parameters have been found between fresh *Triticum aestivum* seeds and seeds that had naturally aged in storage (Bucholc and Buchowicz, 1992). WHY1 has been identified as a telomeric maintenance protein in Arabidopsis. AtWHY1 is important in single stranded DNA repair through telomere homeostasis and chromosome maintenance (Yoo *et al*., 2009). While developmental and/or growth defects have been reported in TDNA lines deficient in AtWHY1, telomeres were elongated over several generations (Yoo *et al*., 2009). This finding suggests that WHY1 may have a role in seeds ageing and seed viability which needs to be explored further. Data was presented in Chapter 4 showing that WHY1 and WHY3 play important roles in the germination of aged seeds. Thereafter, RNA-seq analysis was used to identify the differentially expressed genes (DEGs) in the different genotypes under unaged seeds and following ageing treatments.

5.1.1 Objectives

This chapter aims to identify transcripts that are changed in abundance following accelerated ageing stress and how these transcripts in seeds lacking WHY1 and/or the WHY3 proteins might influence germination. Furthermore, these experiments were carried out to determine the differences between unaged and aged seeds, and to determine whether WHY1 and/or WHY3 play a role in DNA repair during germination. The objectives of the studies reported in this chapter are two-fold:

1. Identify targets of WHY1 and WHY3 in germination through RNA-seq analysis of *Atwhy1why3* seeds and wild type seeds.
2. Identify mechanisms for ageing hypersensitivity in *Atwhy1why3* 7 days aged mutants.

5.2 RESULTS

5.2.1 Experimental Plan

The Col-0 and *Atwhy1why3* seeds used in the germination studies reported in Chapter 4 were used for the RNA-seq analysis [Table 8]. The 0-hour time point served as a control for the normalisation of expression levels prior to germination. Objective one was addressed through analysis of the WT and *Atwhy1why3* control treatments at the 0-hour time point and the 6-hour time point. Objective 2 was addressed through the analysis of WT and *Atwhy1why3* at the 6-hour time point, by comparing control high quality seed and seeds that were aged for 7 days at 35 °C at 80 % RH (as presented in 3.2.4.2). Three independent biological replicates were used for each treatment, time point and genotype to ensure reproducibility and that anomalies could be identified and eliminated. This was further ensured through the use of Log2 (fold-change) >1 data with <-1 threshold and adjusted $p<0.01$ in the analysis of differentially expressed genes.

Table 8: Summary of treatments and time points for each genotype that was sent for RNA-seq. The codes have abbreviations for the treatment; U = unaged, DA = days aged, HI = hours post imbibition.

Genotype	Treatment	Time point	Biological replicates	Code
WT	Control	0 hours dry seed	3	WTU0HI
Atwhy1why3	Control	0 hours dry seed	3	WHYU0HI
WT	Control	6 hours post imbibition	3	WTU6HI
Atwhy1why3	Control	6 hours post imbibition	3	WHYU6HI
WT	7 days ageing	6 hours post imbibition	3	WT7DA6HI
Atwhy1why3	7 days ageing	6 hours post imbibition	3	WHY7DA6HI

5.2.2 Quality Control of RNA-seq

High quality RNA was extracted from *Atwhy1why3* and wild-type seeds from controlled environments and from seeds that had been aged for 7 days at 35 °C and 80 % RH. All the seeds of the same genotype were harvested from the same plant and then treated.

Table 9: Initial quality control step after RNA extraction

Tube No.	Code	Seed weight (g)	RNA Isolation	Nanodrop A260/A280	Original calculated concentration (µg/ mL)
1	WT7DA6HI	0.1172	1st elution	2.12	26.7
2	WT7DA6HI	0.1159	2nd elution	2.16	133
3	WT7DA6HI	0.112	1st elution	2.15	68.8
4	WHY7DA6HI	0.1267	1st elution	2.13	27.3
5	WHY7DA6HI	0.1184	1st elution	2.16	27.3
6	WHY7DA6HI	0.1494	1st elution	2.1	27.8
7	WTU0HI	0.1671	2nd elution	2.19	97.6
8	WTU0HI	0.1314	1st elution	2.48	17.5
9	WTU0HI	0.151	2nd elution	2.13	87.2
10	WTU6HI	0.1401	1st elution	1.95	20.4
11	WTU6HI	0.1347	2nd elution	2.12	59.6
12	WTU6HI	0.1599	2nd elution	2.24	40.8
13	WHYU0HI	0.1249	1st elution	2.1	106
14	WHYU0HI	0.1312	2nd elution	2.17	27.2
15	WHYU0HI	0.1158	2nd elution	2.2	70
16	WHYU6HI	0.1325	1st elution	2.99	2.16
17	WHYU6HI	0.121	2nd elution	2.22	22.4
18	WHYU6HI	0.1231	1st elution	2.31	37.2

At the 260 nm wavelength, 1 absorbance unit was equal to 40µg of single-stranded RNA per mL and anything between above 1.8 A260/S280 ratio was accepted as sufficient purity [Table 9]. The total RNA samples were used for paired-end 150 bp mRNA sequencing of cDNA libraries using Illumina NovaSeq at Novogene (Cambridge, UK). Each sample produced between 60 and 120 million base pair reads out of 8600 million to 17700 million total raw reads [Table 10].

Table 10: Final quality control of RNA reads provided by Novogene. All have a percentage error of 0.03%.

Tube Number	Code	Number of reads (millions)	Raw reads (millions)	Effective (%)	GC (%)
1	WT7DA6HI	78.36	11753.72	96.32	47.55
2	WT7DA6HI	73.47	11020.61	99.18	47
3	WT7DA6HI	89.49	13423.97	99.17	47.13
4	WHY7DA6HI	80.55	12082.79	98.94	47.58
5	WHY7DA6HI	62.43	9363.85	99.41	46.83
6	WHY7DA6HI	82.04	12306.73	99.14	47.33
7	WTU0HI	57.90	8684.58	99.23	46.72
8	WTU0HI	71.72	10758.21	98.97	47.58
9	WTU0HI	73.46	11018.50	98.27	47.23
10	WTU6HI	78.13	11718.75	98.99	47.02
11	WTU6HI	61.88	9281.90	98.45	47.2
12	WTU6HI	67.48	10121.65	98.36	46.91
13	WHYU0HI	84.80	12719.49	98.61	46.68
14	WHYU0HI	117.51	17626.80	98.84	49.32
15	WHYU0HI	67.82	10172.70	98.43	46.87
16	WHYU6HI	60.99	9148.91	98.76	47
17	WHYU6HI	77.73	11660.16	98.53	47.59
18	WHYU6HI	62.07	9310.22	98.15	46.95

5.2.3 Extraction of data using Galaxy

Sequence data was provided by Novogene. These data were analysed using the Galaxy software (UseGalaxy.eu) produced by the Freiburg Galaxy Team (Giardine *et al*., 2005) (https://usegalaxy.eu/u/racheltaylor/w/rna-seq-rachel). This was used to design and execute a workflow and to uniformly analyse the data as shown in Table 11.

Table 11: RNA-seq analysis steps with Galaxy software functions.

Step No.	Function	Input	Output
1	Input (2)	The two datasets (.fq.gz files)	Short read data
2	FastQC (2)	Short read data	
3a	Trimmomatic paired-end	Input FastQC files R1 (first of pair) and R2 (second of pair)	fastqc_out_r1_paired fastqc_out_r1_unpaired fastqc_out_r2_paired fastqc_out_r2_unpaired
3b	FastQC paired	fastqc_out_r1_paired	FastQC on paired (2) FastQC on unpaired (2)
		fastqc_out_r2_paired	
	FastQC unpaired	fastqc_out_r1_unpaired	
		fastqc_out_r2_unpaired	
4a	Hisat2 paired	fastqc_out_r1_paired fastqc_out_r2_paired	HISAT2 on aligned reads (BAM) (3)
4b	Hisat2 unpaired	fastqc_out_r1_unpaired	
		fastqc_out_r2_unpaired	
5a	Samtools Sort paired	HISAT2 on aligned reads (BAM) paired	Samtools Sort (BAM) (3)
5b	Samtools Sort unpaired	HISAT2 on aligned reads (BAM) unpaired 1	
		HISAT2 on aligned reads (BAM) unpaired 2	
6	Samtools merge	All Samtools Sort files	Samtools merge on all data
7	Samtools Stats	Samtools merge	Samtools stats

Step No.	Function	Input	Output
8	Upload Genome file	Upload GTF genome file	
9	featureCounts	Samtools merge GTF genome file	featureCounts on Samtools merge
10	FastQC	Samtools merge	FastQC on RawData
11	MultiQC	FastQC on paired (2) FastQC on unpaired (2) FastQC on RawData Samtools stats featureCounts on Samtools merge	MultiQC on input dataset: Stats MultiQC on input dataset: webpage
12	Deseq	All FeatureCounts files (renamed to match treatment and genotype)	DESeq2 plots on all data (PCA plots) Normalized counts file on all data rLog-Normalized counts file on all data VST-Normalized counts file on all data

The inputs were the two paired samples (Step 1 Input Rx_1; Input Rx_2) and the genome file for the species of interest (Step 6) which produced a MultiQC analysis. The output from the workflow was data on differentially expressed gene files (Deseqs) which, once merged, produced principal component analysis (PCA) data plots [Figure 18] and normalised counts files of all 18 samples which were then used to compare for DEGs.

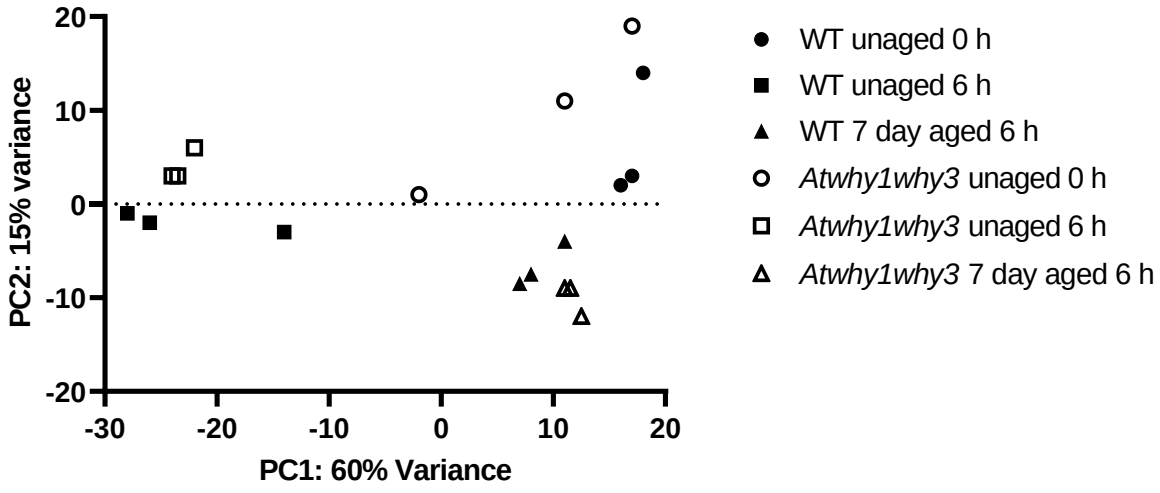


Figure 18: Principal component analysis (PCA) biplot of variability in WT plants compared to *Atwhy1why3* plants that had been untreated (unaged), treated with simulated ageing for 7 days (7 day aged) and imbibed for 0 hours or 6 hours. Each sample had three biological replicates.

PCA plots are used to interpret large datasets by adaptively grouping samples of similar variance together in a matrix, therefore samples that are very similar to each other group together (Jolliffe and Cadima, 2016). The groupings of this dataset show that some WT and WHY groups have a large spread in variance across the samples (*Atwhy1why3* unaged 0 h and WT unaged 0 h), whilst the others are grouped more closely together and even overlap in the cases of the 6 hour imbibed WHY groups (red and blue). This analysis also shows a high degree of similarity between samples of the same treatment as all the treatments are in three distinct groups [Figure 18]. Interestingly, there is some overlap between the WHYU0HI and WTU0HI samples, which implies that there are few differences in transcript levels of dry seeds at 0 hours. Also, the seeds of both genotypes exposed to the same treatment are grouped together. The transcript profiles of the unaged WT and *Atwhy1why3* seeds are grouped together, after 6-hour imbibition and after the 7-day accelerated ageing treatment. Overall, Figure 18 shows that the ageing treatment has a higher impact on transcript levels than genotype because the treatments appear in three distinct groups. Differentially expressed genes in each of these treatments will be described.

5.2.4 Differentially Expressed Genes

Differentially expressed genes (DEGs) in WT and *Atwhy1why3* mutant seeds were identified, following this the top ten transcripts that were most increased or decreased in abundance in each genotype with each treatment were identified. The treatments were: 1) dry (0 h imbibition), 2, imbibed (6 h imbibition) and 3) aged (6 h imbibition and 7 day accelerated

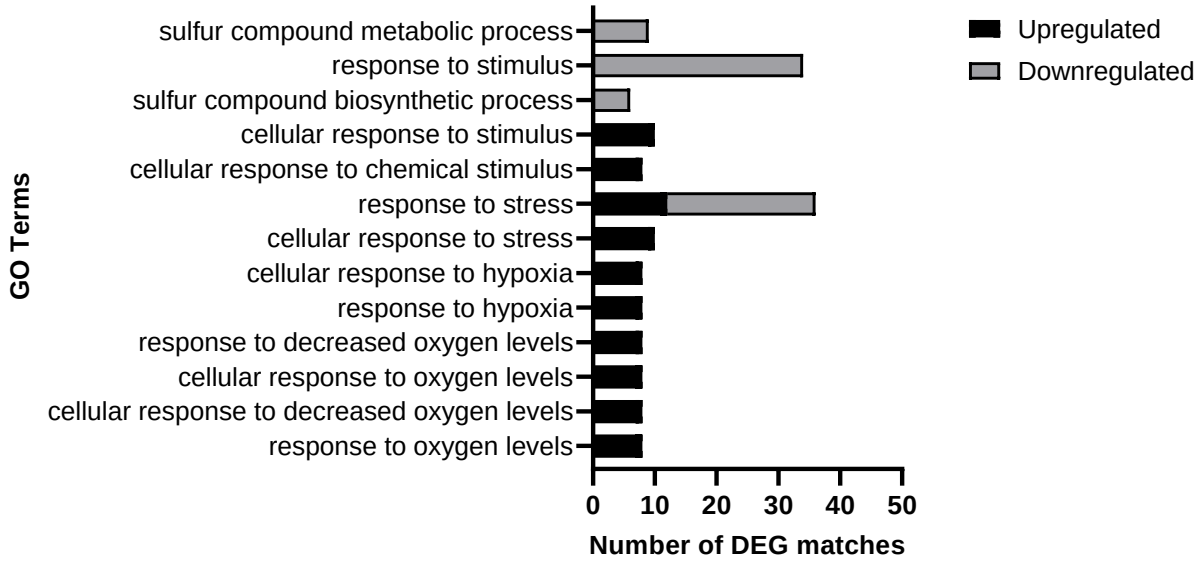
ageing). Values are presented as Log2 FC >1 and <-1 and adjusted p-value ($p<0.01$), which accounts for false discovery rates (FDR) (Benjamini and Hochberg, 2000). The TAIR (Huala *et al*., 2001) website (https://www.arabidopsis.org/tools/bulk/genes/index.jsp) was used to identify gene ontologies, and the ThaleMine tool (https://bar.utoronto.ca/thalemine/begin.do) on the Bio-Analytic Resource for Plant Biology (Provart and Zhu, 2003) was used for identification of genes that were enriched in factors with specific functions or domains as well as Gene Ontology (GO) terms. These gene lists helped to identify easy comparisons between the treatment of ageing and control as well as the effect that the *Atwhy1why3* mutants have on germination compared to WT as well as idenfiying numbers of differentially expressed genes in each list [Table 12].

Table 12: Comparisons of genotype and treatment with the number of up- and downregulated differentially expressed genes (DEGs) with a p <0.01 adjusted with Benjamin-Hochberd with the discoveries that can be made in each comparative assessment.

Gene List	**Treatment**	**Number of DEGs**			**Discovery**
		Upregulated	**Downregulated**	**Total**	
1	WT dry relative to *Atwhy1why3* dry	3	29	32	Changes in mutants without treatment
2	WT imbibed relative to *Atwhy1why3* imbibed	18	52	70	Changes in mutants that might explain poor germination
3	WT aged imbibed relative to *Atwhy1why3* aged imbibed	227	115	342	
4	WT dry relative to WT imbibed	1630	2062	3692	Changes in WT seed upon germination
5	WT dry relative to WT aged imbibed	899	373	1003	Effect of ageing on germination
6	WT imbibed relative to WT aged imbibed	1251	1092	2343	
7	*Atwhy1why3* dry relative to *Atwhy1why3* imbibed	1322	1225	2546	Changes in mutant seed upon germination
8	*Atwhy1why3* dry relative to *Atwhy1why3* aged imbibed	666	472	1138	Effect of ageing on germination in mutants
9	*Atwhy1why3* imbibed relative to *Atwhy1why3* aged imbibed	1957	1415	3372	

The lowest number of DEGs, only 32, in total were in gene list 1, which explored changes in germination in dry WT relative to dry *Atwhy1why3* mutants without treatment, in fact only 3 DEGs were upregulated [Table 12]. Additionally due to the small size of the dataset there were no significant ($p < 0.05$) gene ontology (GO) terms that could be found to link them due to Holm-Bonferroni test corrections. However, comparisons of GO enrichment of the WT and *Atwhy1why3* seeds under different treatments show commonalities in the two genotypes and differences that can help explain the accelerated ageing phenotype [Figure 19].

a)



b)

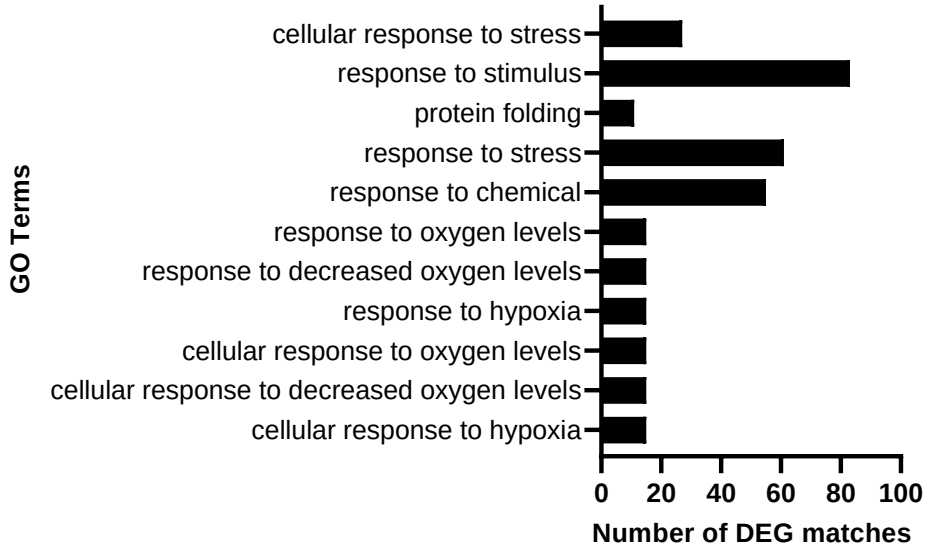


Figure 19: Gene ontology enrichment with Holm-Bonferroni test correction ($p < 0.05$) of WT seeds relative to *Atwhy1why3 seeds* after treatment for a) 6 h imbibition (gene list 2), and b) 7 day ageing and 6 h imbibition (gene list 3).

Most of the genes whose expression was changed in gene list 3 (WT aged imbibed relative to *Atwhy1why3* aged imbibed) [Figure 19a] encoded proteins that are involved in stress responses including hypoxia, anoxia and defence response. Interestingly, 24 of the decreased DEGs are involved in response to stress, and a further 35 are involved in response to stimulus. Furthermore, 18 of the decreased DEGs are involved in the sulphur metabolism, which encompasses many different functions and could be explored in future research. There were many overlaps with the pairwise comparisons of high quality WT seeds relative to high quality *Atwhy1why3* seeds and the aged WT seeds relative to aged *Atwhy1why3* which include genes upregulated in response to stress, stimulus, oxygen and hypoxia [Figure 19]. However, the main difference in the comparisons is that there are no significantly downregulated GO terms in the aged seed comparison [Figure 19b], despite the 115 downregulated genes [Table 12]. This suggests that the downregulated genes have much broader functions which will be investigated in 5.2.4.2.

The highest number of DEGs, 3692, comes from the WT dry relative to WT imbibed in gene list 4, which explored the changes in WT seed upon germination. This is useful to compare to the changes in mutant seed upon germination in gene list 7, *Atwhy1why3* dry relative to *Atwhy1why3* imbibed, which also has a high number of DEGs, 2546 [Table 12]. All of these pairwise comparisons, in addition to the 32 DEGs of the WT dry relative to the *Atwhy1why3* dry seeds, were useful to explore the changes in seeds upon germination.

5.2.4.1 Changes in seed upon germination

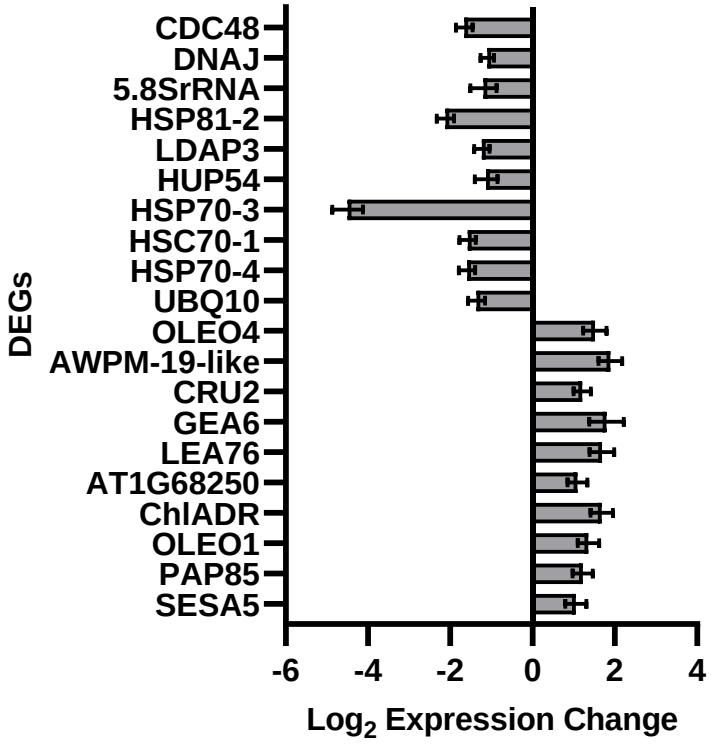


Figure 20: Expression change between dry WT seeds relative to 6 h imbibed WT seeds (Table 12: gene list 4) of the 10 most up-regulated and 10 most down-regulated genes, sorted by mean normalised counts with standard error bars, $p<0.01$.

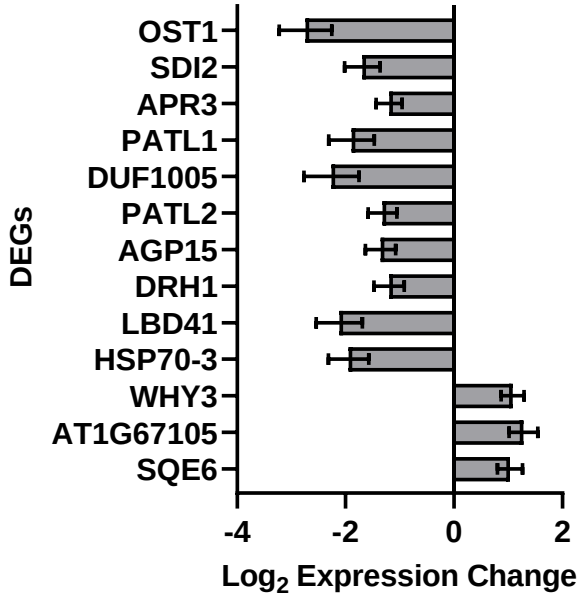


Figure 21: Expression change between dry *Atwhy1why3* seeds relative to 6 h imbibed *Atwhy1why3* seeds (Table 12: gene list 7) of the 10 most upregulated and 10 most downregulated genes, sorted by mean normalised counts with standard error bars, $p<0.01$.

Seven transcripts were commonly increased in abundance in 6 h imbibition in both genotypes [WT: Figure 20 and *Atwhy1why3*: Figure 21]. These encoded proteins that are involved in seed processes and germination: Seed Storage Albumin 5 (AT5G54740), seed lipid Oleosin 1 (AT4G25140), late embryogenesis abundant (LEA) proteins (AT2G40170, AT3G15670) and ABA-related AWPM-19-like (AT1G04560). Transcripts encoding, Chloroplast Aldehyde Reductase (CHLADR) (AT1G54870), which is involved in plastid metabolism were also more abundant after 6h of imbibition. AT1G68250 encodes a hypothetical protein. The levels of AT1G68250 transcripts were increased in both WT and *Atwhy1why3* seeds at 6h of imbibition.

Other transcripts that were found to be only in the top ten most abundant list for the WT [Figure 20] were seed and germination-related proteins such as Pap85 (AT3G22640), seed lipid OLEOSIN 3/4 (AT3G27660), and an ABA-related Cruciferin 2 (AT1G03880). The transcripts that were found to be only in the top ten most abundant list for the *Atwhy1why3* seeds [Figure 21] included ABA-related Oil Body-Associated Protein 1a (OBAP1A) (AT1G05510) and transcripts that encode LEA proteins (AT3G15670, AT3G22500) and seed lipid Oleosin 2 (AT5G40420).

There were eight genes, whose expression was downregulated in both genotypes [WT: Figure 20 and *Atwhy1why3*: Figure 21]. The transcripts (AT3G09440), that were most decreased in abundance in both lists encode Heat Shock Protein 70 (HSP70). Transcripts encoding other members of this family were also less abundant (AT3G12580 and AT5G02500) in both genotypes, as were HSP90 (AT5G56030) transcripts. The other five downregulated transcripts in both genotypes were unrelated; SA-induced Polyubiquitin 10 (UBQ10) (AT4G05320), Lipid Droplet-Associated Protein 3 (LDAP3) (AT3G05500), 5.8SrRNA (AT3G41979), Cell Division Cycle 48B (CDC48B) induced upon oilseed rape mosaic tobamovirus infection (AT3G53230). Other downregulated transcripts that were only in the top ten for the WT [Figure 20] encode a Hypoxia Response Unknown Protein 54 (AT4G27450) and a HSP40 (AT2G20560). The transcripts with increased levels, which were only seen in the top ten for *Atwhy1why3* [Figure 21] encode an Arabidopsis orthologue of human HSP70-Binding Protein 1 (HspBP-1) (AT3G09350) and ABA-Repressor 1 (AT5G64750).

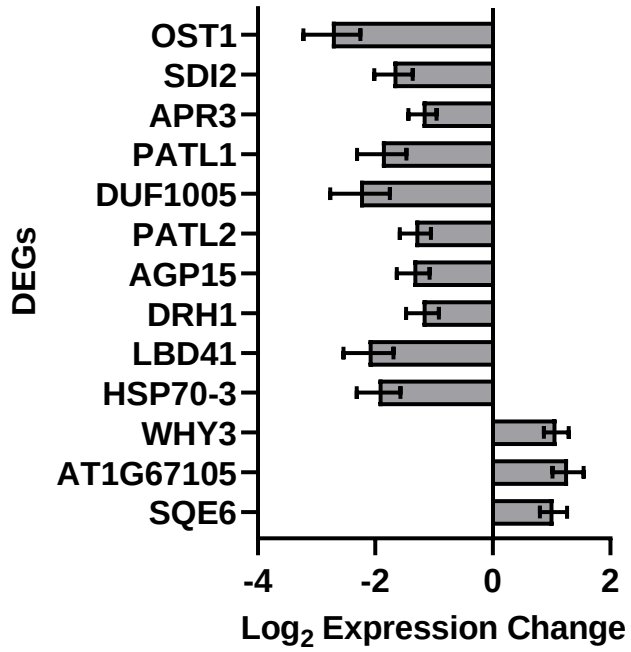


Figure 22: Expression change between dry WT seeds relative to dry *Atwhy1why3* seeds (Table 12: gene list 1) of the 3 most upregulated and 10 most downregulated genes, sorted by mean normalised counts with standard error bars, $p<0.01$.

The expression of only three genes was increased in *Atwhy1why3* dry seeds compared to the wild type [Figure 22]. These encode the proteins: WHY3 (AT2G02740), Squalene Monooxygenase 6 (AT5G24160) and encoding an unknown protein (AT1G67105). Although Why3 appears as an increased DEG in the *Atwhy1why3* mutant compared to the WT, it is a mutated form of *Atwhy3* and would not fully code for the WHY3 protein. In comparison, there were 29 transcripts that were less abundant in the *Atwhy1why3* seeds compared to the wild type [Table 12]. The ten transcripts, whose levels were most decreased encode proteins such as HSP70 (AT3G09440), which was also the most decreased DEG in the comparisons of WT [Figure 20] and *Atwhy1why3* [Figure 21] seeds. However, in a comparison of dry WT and dry *Atwhy1why3* seeds [Figure 22] the transcripts encoding Open Stomata 1 protein through calcium-independent ABA-activated protein kinase (AT4G33950) were the least abundant. The levels other DEGs that were decreased are involved in many different processes including membrane trafficking: Patellin 1 (AT1G72150), Patellin 2 (AT1G22530) and sulphur-related processes: Aps Reductase 3 (AT4G21990) and Sulphur Deficiency Induced 2 (AT1G04770). Other unrelated lower DEGS encode: Lob Domain-Containing Protein 41 (LBD41) (AT3G02550), RNA Helicase DRH1 (AT3G01540), Arabinogalactan Protein 15 (AT5G11740), and a hypothetical protein that encodes DUF1005 (AT3G19680).

5.2.4.2 Expression changes in WT compared to mutant seeds

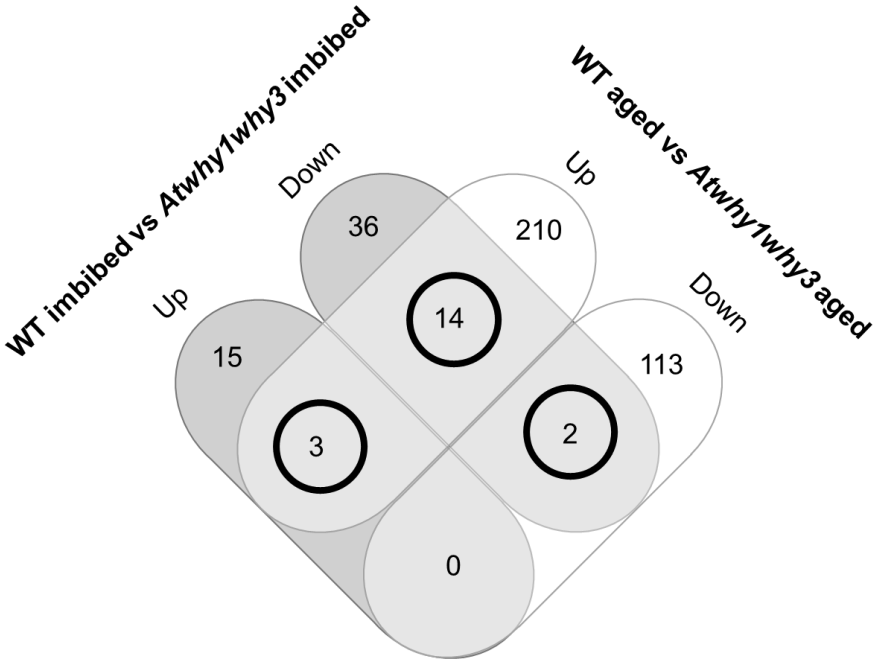


Figure 23: Number of common and unique DEGs between WT high quality imbibed seed relative to *Atwhy1why3* high quality imbibed seed (gene list 2) compared to WT aged imbibed relative to *Atwhy1why3* aged imbibed (gene list 3) split into up- and down-regulated lists. The DEGs which appeared in both gene lists were circled in black (diagram created using: http://bioinformatics.psb.ugent.be/webtools/Venn/).

In comparing high quality imbibed seeds (gene list 2) to aged imbibed seeds (gene list 3) of both genotypes it is clear that ageing produces much more DEGs compared to high quality seeds [Figure 23]. In fact, there were 210 upregulated and 113 downregulated genes in the aged comparison contrasting the 15 upregulated and 36 downregulated genes in the high quality comparison. The 3 DEGs which were increased in gene lists 2 and 3 [Figure 23] include transcripts encoding HSP40 (AT3G13310), MTO 1 Responding Down 1 (AT1G53480) and WHY1 (AT1G14410), whilst only 2 DEGs were decreased in both gene lists. These encode a protein kinase superfamily protein (AT2G30740) and Short Open Reading Frame 29 (AT3G29644). The DEGs which were downregulated in gene list 2 but upregulated in gene list 3 are those that are upregulated in seeds that have undergone accelerated ageing and downregulated in imbibed seeds, these DEGs can give us an insight into the differences in high quality and aged seed germination. These DEGs encompassed several GO terms: auxin-activated signalling pathway, cellular oxidant detoxification, cellular response to hypoxia, cellular response to oxidative stress, endoplasmic reticulum unfolded protein response,

response to abscisic acid, response to auxin, response to cadmium ion, response to microbial phytotoxins, response to molecule of bacterial origin and response to water deprivation.

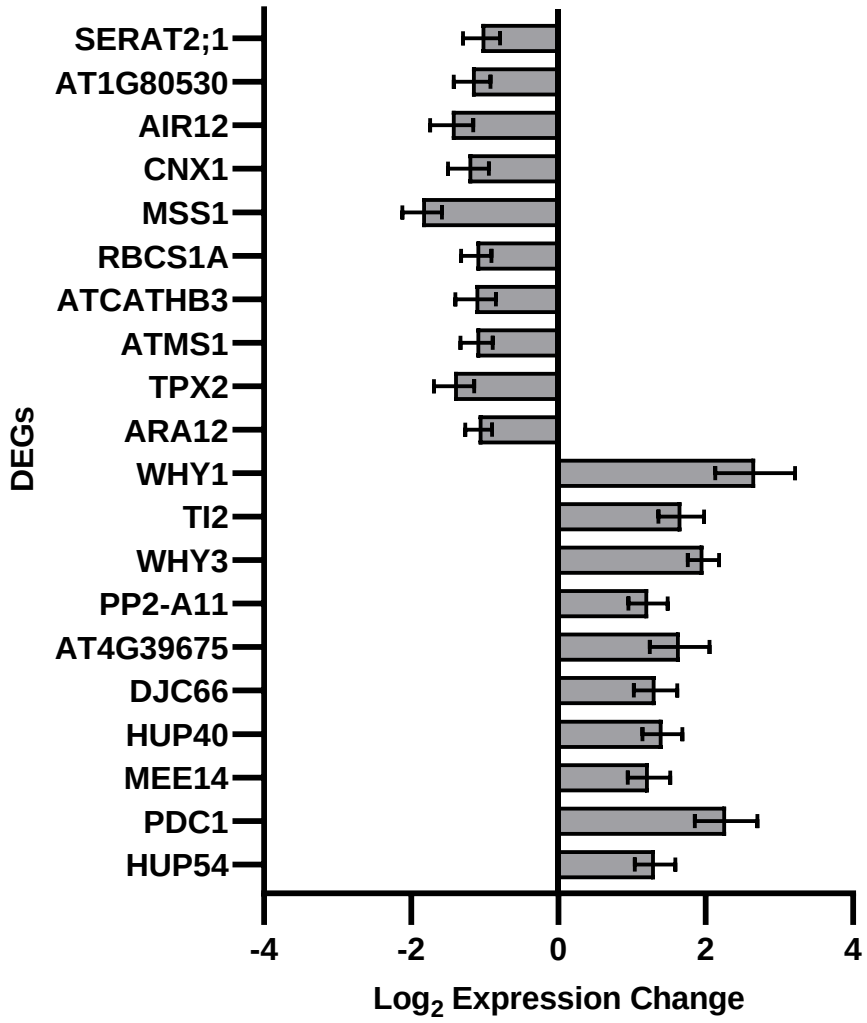


Figure 24: Expression change between 6 h imbibed WT seeds relative to 6 h imbibed *Atwhy1why3* seeds (Table 12: gene list 2) of the 10 most upregulated and 10 most downregulated genes, sorted by mean normalised counts with standard error bars, $p<0.01$.

Unsurprisingly, two of the most abundant transcripts in the comparison of imbibed WT and *Atwhy1why3* seeds [Figure 24] are WHY1 (AT1G14410), which was the most highly upregulated DEG in all comparisons and WHY3 (AT2G02740). Interestingly, HSP40 (AT3G13310) and Hypoxia Response Unknown Protein 54 (AT4G27450) were in the top ten decreased DEGs for dry relative to imbibed WT [Figure 20]. They were also in the top ten increased DEGs in imbibed WT and *Atwhy1why3* seeds [Figure 24]. Additional increased DEGs were Pyruvate Decarboxylase 1 (PDC1) (AT4G33070), CCG-Binding Protein 1 (AT2G15890), Hypoxia Response Unknown Protein 40 (AT4G24110), Phloem Protein 2-A11

(AT1G63090), Trypsin Inhibitor Protein 2 (AT2G43520) and a hypothetical protein (AT4G39675).

The ten least abundant transcripts in a comparison of imbibed WT and *Atwhy1why3* seeds [Figure 24] were involved in sugar mucilage in seeds including: Sugar Transport Protein 13 (AT5G26340), Subtilisin-Like Serine Protease 1.7 (AT5G67360), as well as other unrelated DEGs: Thioredoxin-Dependent Peroxidase 2 (AT1G65970), ATCATHB3 (AT4G01610), RuBisCO Small Chain 1A (AT1G67090), Calnexin 1 (AT5G61790), 5-Methyltetrahydropteroyltriglutamate Homocysteine Methyltransferase 1 (AT5G17920), Auxin-Induced In Root Cultures 12 (AT3G07390), Serine Acetyltransferase 1 (AT1G55920) and an unlabelled Major facilitator superfamily protein (AT1G80530).

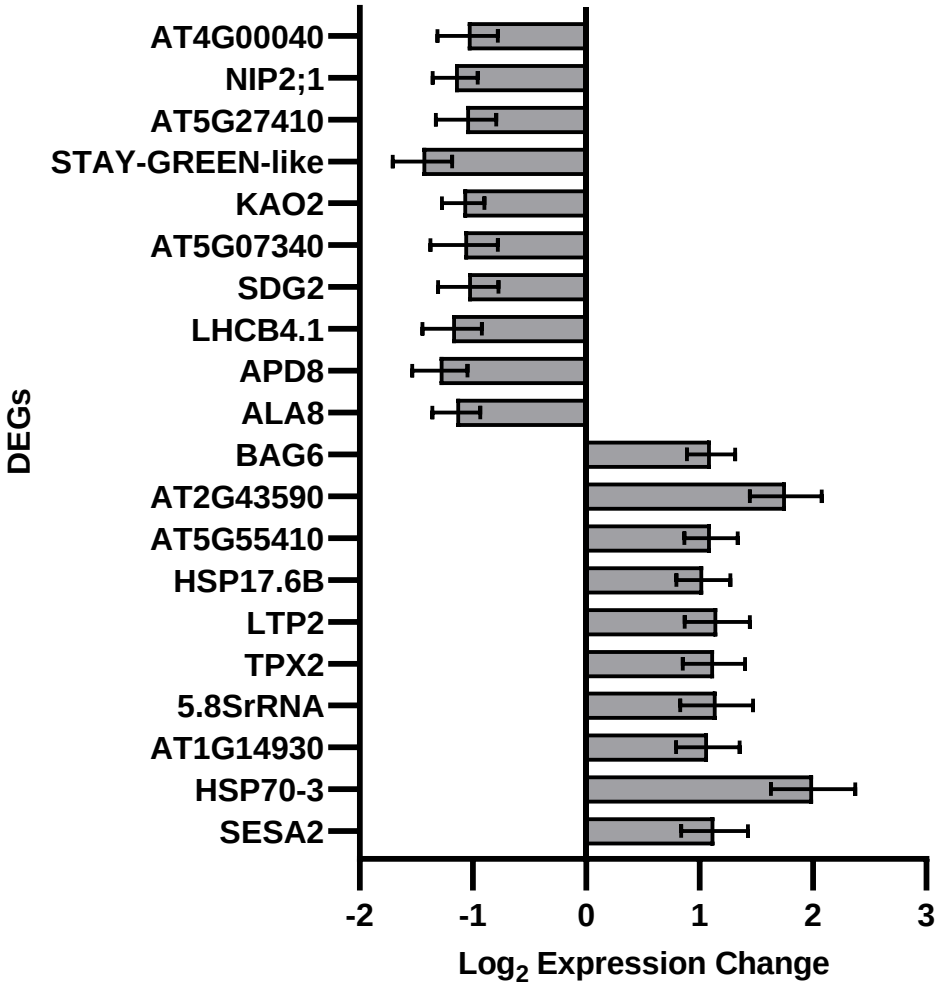


Figure 25: Expression change between 7 day aged 6 h imbibed WT seeds relative to 7 day aged 6 h imbibed *Atwhy1why3* seeds (Table 12: gene list 3) of the 10 most upregulated and 10 most downregulated genes, sorted by mean normalised counts with standard error bars, $p<0.01$.

The ten most upregulated DEGs in the aged WT and *Atwhy1why3* seed comparison [Figure 25] were mainly involved in seed storage and seed lipids: Seed Storage Albumin 2 (SESA2) (AT4G27150), a polyketide cyclase/dehydrase and lipid transport superfamily protein (AT1G14930), Lipid Transfer Protein 2 (LP2) (AT2G38530), and a bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein (AT5G55410). Three of the enhanced DEGs were also involved in heat stress: HSP70 (AT3G09440), HSP17.6B (AT2G29500) and BCL-2-Associated Athanogene 6 (ATBAG6) (AT2G46240). The other top ten increased DEGs are involved in different processes. These include transcripts encoding Thioredoxin-Dependent Peroxidase 2 (AT1G65970), 5.8SrRNA (AT3G41979), a chitinase family protein (AT2G43590).

The top ten decreased DEGs in the aged WT and *Atwhy1why3* seed comparison [Figure 25] were: Aminophospholipid Atpase 8 (AT3G27870), Protein Phosphatase 2C Clade D 8 (APD8) (AT5G06750), Light Harvesting Complex Photosystem II (AT5G01530), Set Domain Protein 2 (AT4G15180), a calreticulin family protein (AT5G07340), Ent-Kaurenoic Acid Hydroxylase 2 (AT2G32440), Stay-Green-like protein (AT4G11911), D-amino acid aminotransferase-like PLP-dependent enzymes superfamily protein (AT5G27410), aquaporin NIP2.1 (AT2G34390), and a chalcone and stilbene synthase family protein (AT4G00040).

5.2.4.3 Effect of ageing on germination

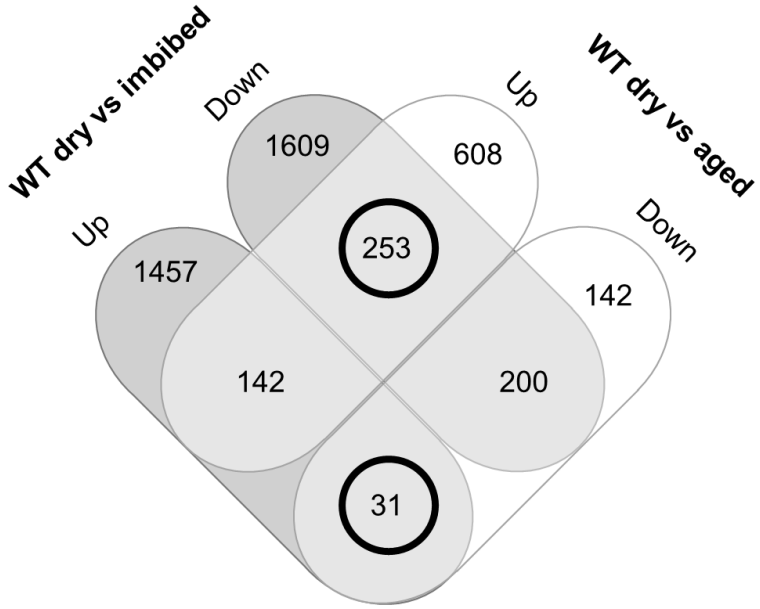


Figure 26: Number of common and unique DEGs between WT dry relative to imbibed seed (gene list 4) relative to WT dry relative to aged imbibed (gene list 5). The DEGs which were opposingly regulated were circled in black (diagram created using: *http://bioinformatics.psb.ugent.be/webtools/Venn/*).

In order to determine the effects that ageing had on germination and to identify changes in gene expression WT dry relative to imbibed seed and WT dry relative to aged imbibed seeds, gene lists 4 and 5 were compared [Figure 26]. Transcripts that were least abundant in one case but increased in the other could reflect changes that effect germination due to ageing in WT seed. The 253 DEGs which were downregulated in gene list 4 but increased in gene list 5 were divided into three GO terms: establishment of sister chromatid cohesion, negative regulation of auxin metabolic process and negative regulation of hormone metabolic process. Additionally, 4 DEGs (AT1G27360, AT1G27370, AT1G69170 and AT5G43270) encode microRNA 156 and 157 targets and Squamosa-promoter Binding Protein (SBP)-like proteins. However, these transcripts did not appear in the top 10 most changed DEG lists. The 31 DEGs which were more abundant in gene list 4 but decreased in gene list 5 [Figure 26] were also less abundant in the ageing comparison. However, there were no significant ($p<0.05$) GO term changes identified in the Holm-Bonferroni test corrections.

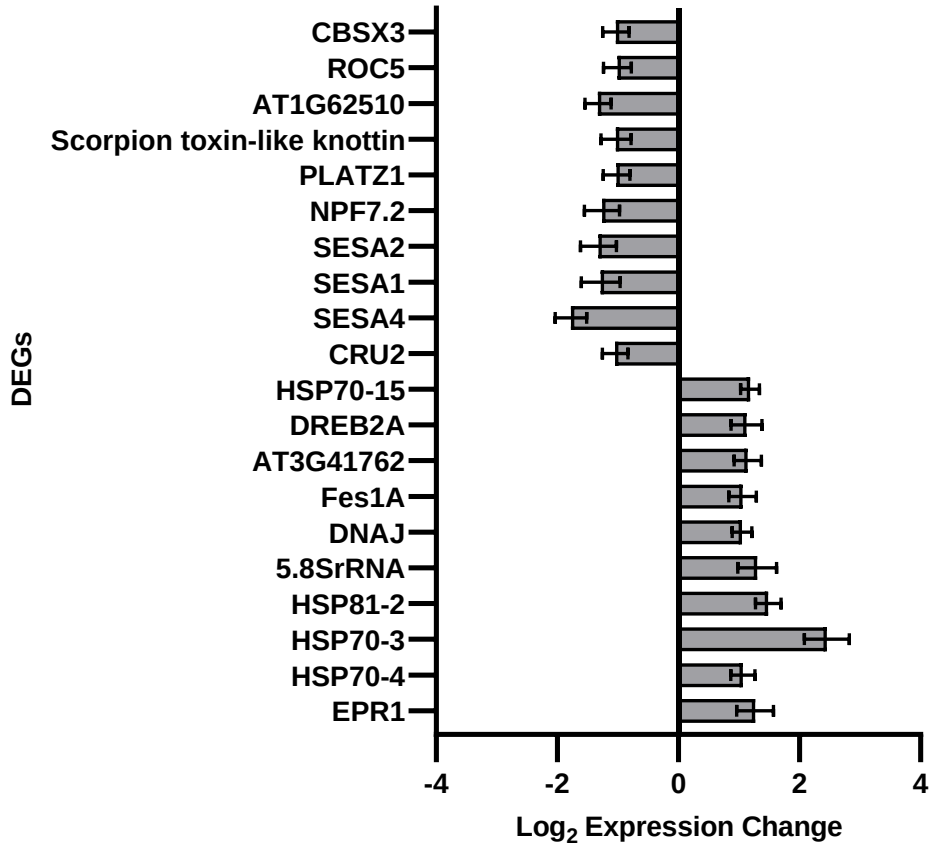


Figure 27: Expression change between dry WT seeds relative to 7 day aged 6 h imbibed WT seeds (Table 12: gene list 5) of the 10 most upregulated and 10 most downregulated genes, sorted by mean normalised counts with standard error bars, $p<0.01$.

The transcripts that were most increased in the comparison of dry WT seed and seed that had undergone accelerated ageing and been imbibed for 6 h [Figure 27] encode HSP70 (AT3G09440). Other transcripts that also encode HSP70 were also increased (AT3G12580 and AT1G79920) as were HSP BP-1 (AT3G09350) transcripts. HSP70 was the most increased DEG in the aged WT and *Atwhy1why3* seed comparison [Figure 25]. The levels of transcripts encoding other HSPs such as HSP90 (AT5G56030), HSP40 (AT2G20560) were also more abundant in this comparison [Figure 27]. Other DEGs encode proteins involved in different processes such as seed germination: Extensin Proline-Rich 1 (AT2G27380), drought stress tolerance in Dehydration-Responsive Element Binding Protein 2 (AT5G05410), protein translation 5.8SrRNA (AT3G41979). There is also an unknown hypothetical protein (AT3G41762).

The majority of the top ten DEGs that were decreased are involved in seed processes and germination. These include SESA4 (AT4G27170), SESA1 (AT4G27140), SESA2 (AT4G27150), ABA-Induced Expression 1 (AT1G21000) and an ABA-related Cruciferin 2 (AT1G03880). The other decreased transcripts encode proteins that are involved in a range of processes: Nitrate Transporter 1.8 (AT4G21680), a scorpion toxin-like knottin superfamily protein (AT1G47540), a protein expressed in the root cortex (AT1G62510) and Cyclophilin 1 (AT4G34870), CBS Domain Containing Protein 3 (AT5G10860) which modulates development by regulating the thioredoxin system.

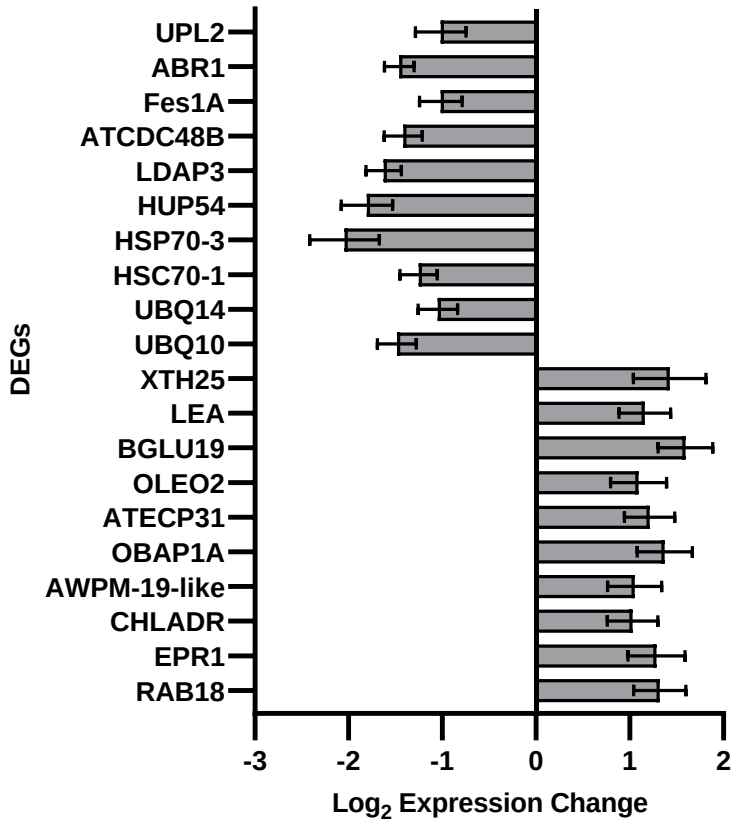


Figure 28: Expression change between 6 h imbibed WT seeds relative to 7 day aged 6 h imbibed WT seeds (Table 12: gene list 6) of the 10 most upregulated and 10 most downregulated genes, sorted by mean normalised counts with standard error bars, $p<0.01$.

The ten DEGs showing the greatest increases in the WT imbibed relative to WT imbibed and aged comparison [Figure 28] mainly encode drought-stress related proteins: drought-induced Responsive To ABA 18 (AT5G66400), ABA-related AWPM-19-like (AT1G04560), OBAP1A (AT1G05510) which was modulated in response to ABA as well as other-stress related proteins: Dehydrin LEA (AT2G21490), Beta Glucosidase 19 (BGLU19) (AT3G21370) involved in salt stress. In addition, two DEGs showing the greatest increases are involved in seed germination: Extensin Proline-Rich 1 (AT2G27380), LEA Protein ECP31 (AT3G22500), seed lipids: OLEOSIN 2 (AT5G40420), cell-wall related Xyloglucan Endotransglucosylase/Hydrolase 25 (XTH25) (AT5G57550) and CHLADR (AT1G54870). The latter was also increased in Figure 21.

The DEGs showing the greatest decreases include HSP70 (AT3G09440) [Figure 28] as was also found in comparisons between dry and 6 h imbibed WT seeds [Figure 20], dry and 6 h imbibed *Atwhy1why3* seeds [Figure 21] and dry WT seeds and dry *Atwhy1why3* seeds [Figure 22]. Additional DEGs showing the large decreases encode heat-shock related proteins such as

Heat Shock Cognate Protein 70-1 (AT5G02500) and HspBP-1 (AT3G09350) [Figure 28]. Other DEGs, which encode stress-related proteins were Hypoxia Response Unknown Protein 5 (AT4G27450) and ABA Repressor1 (AT5G64750). UBQ10 (AT4G05320) were also lower in the dry relative to imbibed comparisons for both WT and *Atwhy1why3* [Figure 20, Figure 21]. The polyubiquitin, UBQ14 (AT4G02890) was also lower as was as Ubiquitin-Protein Ligase 2 (AT1G70320). Other DEGs which were also lower in dry relative to imbibed comparisons for both WT and *Atwhy1why3* [Figure 20, Figure 21] include: LDAP3 (AT3G05500) CDC48B (AT3G53230).

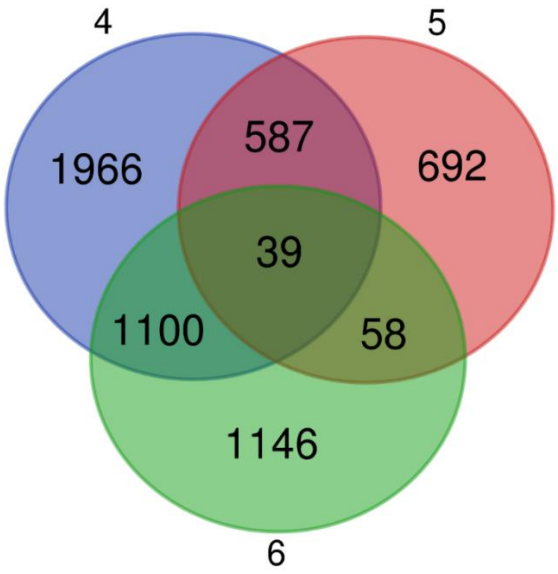


Figure 29: Comparison of genes that overlap from 3 pairwise comparisons of the WT seed: 4) WT dry relative to imbibed, 5) WT dry relative to aged imbibed, 6) WT imbibed relative to aged imbibed. (diagram created using: http://bioinformatics.psb.ugent.be/webtools/Venn/).

There were 39 DEGs common to all three pairwise comparisons, however the only DEG which was upregulated in all 3 gene lists encodes PLANT U-BOX 19 (AT1G60190), which is involved in salt inhibition of germination. However, there were 6 DEGs which were lower in all three comparisons. These encode a range of proteins: NPR1-Like Protein 3 (AT5G45110), HSP40 (AT3G13310), Glutathione S-Transferase TAU 1 (AT2G29490), an alpha/beta-Hydrolases superfamily protein (AT1G74300), a C2H2 type zinc finger transcription factor family (AT1G49900) and a putative methyltransferase family protein (AT1G79915). Additionally, there are only 2 DEGs higher in all of the ageing pairwise comparisons (gene lists 3, 6 and 9. These are APD8 (AT5G06750) and bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein (AT5G55460) [Figure 29].

5.2.4.4 Expression changes in mutant seeds

In order to determine the effects of ageing on germination in the *Atwhy1why3* mutants, gene list 7 and 8 were compared. This will allow identification of the changes in *Atwhy1why3* dry relative to imbibed seed and *Atwhy1why3* dry relative to aged imbibed seed [Figure 30]. The DEGs that were lower in one and higher in the other indicate genes that effect germination due to ageing in *Atwhy1why3* seed.

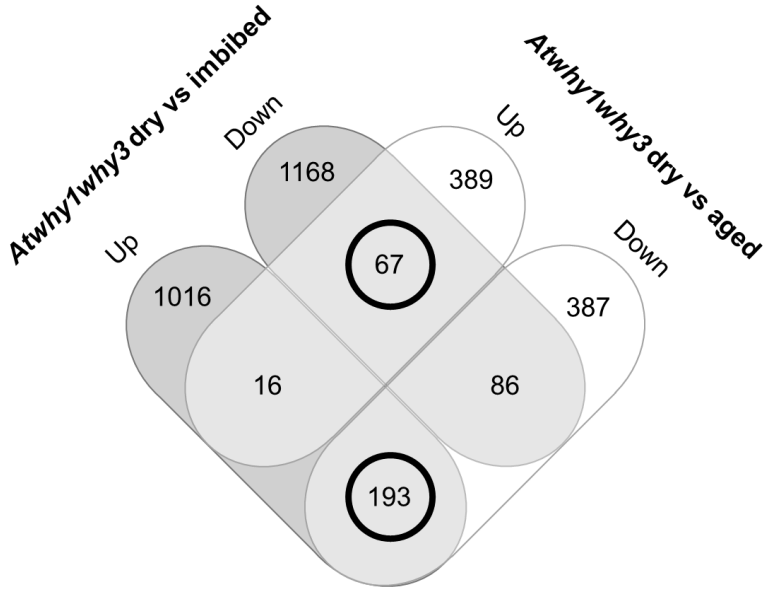


Figure 30: Number of common and unique DEGs between *Atwhy1why3* dry relative to imbibed seed (gene list 7) relative to *Atwhy1why3* dry relative to aged imbibed (gene list 8). The DEGs which were opposingly regulated were circled in black (diagram created using: http://bioinformatics.psb.ugent.be/webtools/Venn/).

The 67 DEGs that were lower in gene list 7 but higher in gene list 8, are the ones that are most abundant in the aged *Atwhy1why3* seeds. However, there were no significant ($p<0.05$) GO terms through Holm-Bonferroni test corrections. Whilst the 193 DEGs that were lower in gene list 8 and increased in gene list 7 were the ones that are decreased as a result of ageing in the *Atwhy1why3* seeds. These were categorised into 18 GO terms with numbers of genes in each group in brackets: response to abscisic acid (29), response to alcohol (29), response to lipid (31), response to water (23), response to inorganic substance (34), response to oxygen-containing compound (45), response to acid chemical (23), response to water deprivation (22), response to hormone (41), response to chemical (60), response to endogenous stimulus (41), response to organic substance (45), response to stimulus (89), response to abiotic stimulus (44), water transport (6), fluid transport (6), aging (10), cellular response to abscisic acid stimulus (11). Additionally, protein domains were found to be enriched in major intrinsic proteins, such as aquaporin transporters and pectinase inhibitor domains. The DEGS that are most upregulated

in this study encode the following proteins: MAPKKK18 (AT1G05100) involved in drought stress tolerance, Ribonuclease 1 (AT2G02990) involved in wound-induced signalling, Senescence-Associated Gene (SAG) 13 (AT2G29350) required for defence against ROS and fungal pathogens, DUF679 (AT3G21520) involved in ER membrane fission and the tonoplast during leaf senescence, Heavy Metal Associated Protein 31 (AT3G56240), Methyl Esterase 16 (AT4G16690), DUF679 Domain Membrane Protein (AT4G18425), SAG29 (AT5G13170), Glutamine Synthetase 2 (AT5G35630) and Aleurain-Like Protease (AT5G60360) which encodes a senescence-associated thiol protease.

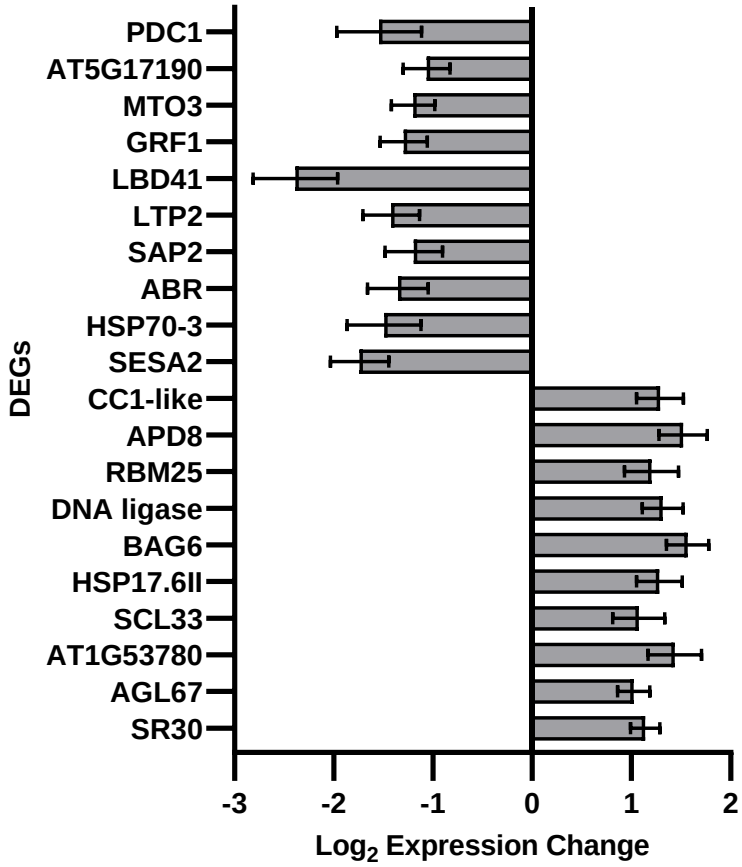


Figure 31: Expression change between dry *Atwhy1why3* seeds relative to 7 day aged 6 h imbibed *Atwhy1why3* seeds (Table 12: gene list 8) of the 10 most upregulated and 10 most downregulated genes, sorted by mean normalised counts with standard error bars, $p<0.01$.

The top ten increased DEGs in the comparison between dry *Atwhy1why3* seed relative to imbibed *Atwhy1why3* that had undergone accelerated ageing [Figure 31] encode the following proteins: Serine-Arginine Protein 30 (AT1G09140), SC35-LIKE Splicing Factor 33 (AT1G55310), Agamous-Like 67 (AT1G77950), which modulates seed germination under high temperatures, HSP17.6II (AT5G12020), a 26S proteasome regulatory complex ATPase

(AT1G53780), a DNA ligase (AT1G75860) and ATBAG6 (AT2G46240), which was also higher in the aged WT and *Atwhy1why3* seed comparison [Figure 24]. Other decreased DEGs in Figure 31 are RNA-Binding Protein 25 (AT1G60200), splicing factor, CC1-like protein (AT5G09880) and APD8 (AT5G06750), which was also lower in the aged WT and *Atwhy1why3* seed comparison [Figure 24].

The downregulated DEGs in dry *Atwhy1why3* seed and imbibed *Atwhy1why3* encoded a range of different proteins: SESA2 (AT4G27150), HSP70 (AT3G09440), ABA-Response Protein (AT3G02480), Secreted Aspartic Protease 2 (AT1G03220), LP2 (AT2G38530) involved in lipid transfer, LBD41 (AT3G02550), General Regulatory Factor 1 (AT4G09000) involved in stress response and cell cycle, Methionine Adenosyltransferase 4 (AT3G17390), B-cell receptor-associated-like protein (AT5G17190) and PDC1 (AT4G33070).

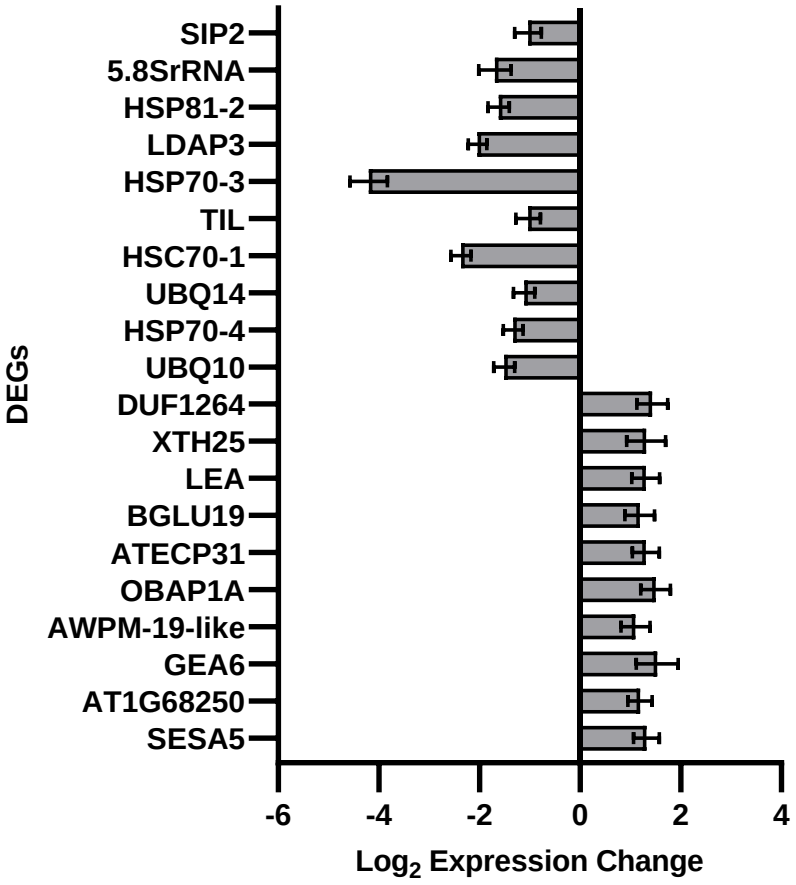


Figure 32: Expression change between 6 h imbibed *Atwhy1why3* seeds relative to 7 day aged 6 h imbibed *Atwhy1why3* seeds (Table 12: gene list 9) of the 10 most upregulated and 10 most downregulated genes, sorted by mean normalised counts with standard error bars, $p<0.01$.

The top ten most abundant DEGS for 6 h imbibed relative to 6 h imbibed 7 day aged *Atwhy1why3* seeds [Figure 32] encode mainly germination-specific proteins: LEA6 (AT2G40170), LEA Protein ECP31 (AT3G22500), Dehydrin LEA (AT2G21490), SESA5 (AT5G54740), ABA-related AWPM-19-like (AT1G04560), OBAP1A (AT1G05510), and other unrelated proteins: BGLU19 (AT3G21370), XTH25 (AT5G57550) and histone acetyltransferase (DUF1264) (AT5G45690), as well as a hypothetical protein (AT1G68250).

Once again, the lowest DEG was HSP70 (AT3G09440). Other members of the HSP70 family were also decreased (AT3G12580, AT5G02500 and AT5G58070), as were transcripts encoding HSP90 (AT5G56030). The other top ten lowest DEGs for 6 h imbibed relative to 6 h imbibed 7 day aged Atwhy1why3 seeds [Figure 32] encode UBQ10 (AT4G05320), UBQ14 (AT4G02890), LDAP3 (AT3G05500), 5.8SrRNA (AT3G41979), Seed Imbibition 2 (AT3G57520).

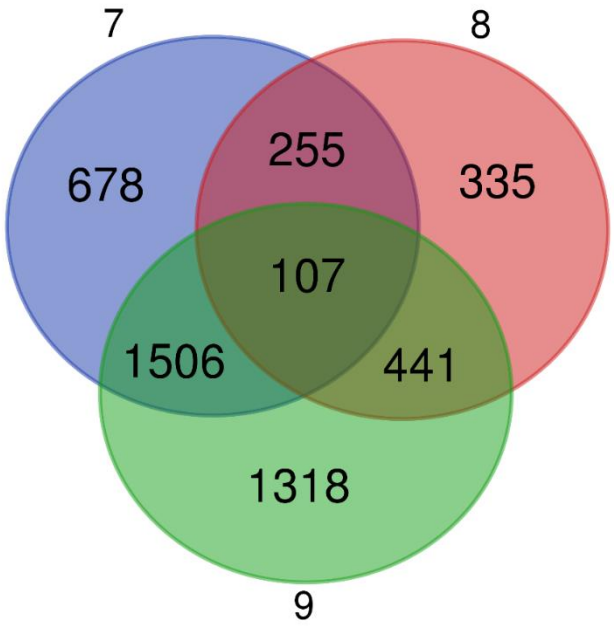


Figure 33: Comparison of DEGs that overlap from 3 pairwise comparisons of the mutant: 7) *Atwhy1why3* dry relative to imbibed, 8) *Atwhy1why3* dry relative to aged imbibed, 9) *Atwhy1why3* imbibed relative to aged imbibed (diagram created using: http://bioinformatics.psb.ugent.be/webtools/Venn/).

There are 107 DEGs which are present in all three gene lists. Of these, there are only three DEGs that are more abundant in gene list 7 but less abundant in lists 8 and 9. These encode Glutamine Synthetase 2 (AT5G35630), which has been previously identified as an ageing-related protein, Sucrose-Proton Symporter 2 (AT1G22710) and Thioglucoside Glucohydrolase 1 (AT5G26000). There are no transcripts that are simultaneously lower in list 7 but higher in

lists 8 and 9. There are 66 DEGs, which are lower in list 7 but higher in list 8. However, there were no significant ($p<0.05$) GO terms in the Holm-Bonferroni test corrections.

There are 10 DEGs that were more abundant in gene lists 4 and 7 but less abundant in gene lists 5 and 8. As a reminder, gene lists 4 and 7 compare dry and imbibed seed in WT and *Atwhy1why3* respectively, whilst gene lists 5 and 8 compare dry to aged imbibed seed in WT and *Atwhy1why3* respectively. Therefore, it is interesting to identify the DEGs that were more abundant in dry and imbibed seeds but less abundant in dry and ageing seed. These encoded Embryo Defective 2171 (AT3G04400), WES1 Auxin-responsive GH3 family protein (AT4G27260), an F-box family protein (AT2G26850), Nitrilase 2 (AT3G44300), Sucrose-Proton Symporter 2 (AT1G22710), a pectin lyase-like superfamily protein (AT3G07850), SINE1 (AT1G54385) which plays a role in the nucleus positioning guard cells, Methyl Esterase 16 (AT4G16690), (AT1G73260) and Ribonuclease 1 (AT2G02990). However, there were no significant GO terms, or protein domain enrichments in the group.

There were 11 DEGs that were less abundant in WT dry relative to imbibed and *Atwhy1why3* dry relative to imbibed and more abundant in WT dry relative to aged imbibed and *Atwhy1why3* dry relative to aged imbibed, which may provide some insights into the genes that participate in the ageing of the mutants. None of the transcripts encode proteins associated with ageing. They encode the following proteins: a zinc metalloproteinase-like protein (AT5G35688), a protein kinase superfamily protein (AT2G39110), an intron maturase, type II family protein (ATMG00520), squamosa promoter-like 11 (AT1G27360), pfkB-like carbohydrate kinase family protein (AT4G10260), an actin-binding formin homology 2 family protein (AT1G42980), an ankyrin repeat family protein (AT4G03440), Urease Accessory Protein D; (AT2G35035), Glycine-Rich Protein (AT1G04800), a hypothetical protein (AT1G48780) and an antisense transcript (AT1G29357). There were no significant GO terms, or protein domain enrichments in the group.

5.3 DISCUSSION

The gene ontology analyses and an analysis of the top ten upregulated DEGs showed a high degree of these genes code for proteins involved in response to stress and stimulus, this suggests that WHY proteins are important in stress responses as has been consistently reported in the literature [Table 1]. A novel interesting consistent finding was the appearance of HSP70 (AT3G09440) in almost all pairwise comparisons. Overall, HSP70 (AT3G09440) appeared in every pairwise comparison apart from WT relative to *Atwhy1why3* imbibed, WT relative to *Atwhy1why3* imbibed [Figure 34]. This transcript had the lowest levels of all identified DEGs. It also appeared in the RNA-seq data for WT relative to *Atwhy1why3* imbibed. However, it was not significant in this situation (Benjamin-Hochberd adjusted p = 0.97).

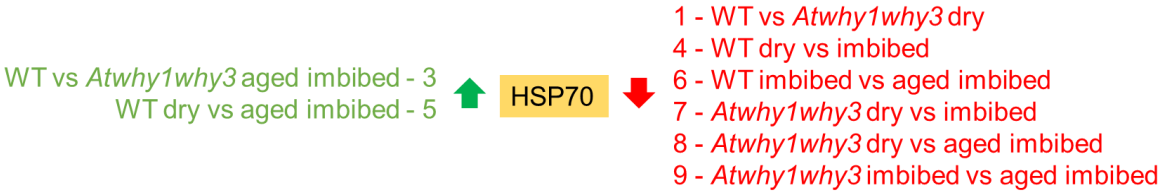


Figure 34: Comparison of up- (green) and down- (red) regulated HSP70 as a DEG in gene lists of RNA-seq data, vs is in place of 'relative to'.

In addition to HSP70, other HSPs or heat-stress related proteins were also differentially expressed in the pairwise comparisons. HSP90 (AT5G56030) is the least abundant transcript in WT dry relative to imbibed [Figure 20], *Atwhy1why3* dry relative to imbibed [Figure 21] and *Atwhy1why3* imbibed relative to aged imbibed [Figure 32], while the transcripts encoding HspBP-1 and heat shock cognate protein 70 appear in WT imbibed relative to aged imbibed [Figure 28]. The most abundant transcripts encoding heat shock proteins include HspBP-1 in WT relative to *Atwhy1why3* aged imbibed [Figure 25], WT dry relative to aged imbibed [Figure 27] and *Atwhy1why3* dry relative to imbibed [Figure 21] and HSP90 in WT dry relative to aged imbibed [Figure 27]. Of the HSP transcripts with the lowest levels HSP40 appears in WT relative to *Atwhy1why3* imbibed [Figure 23], WT relative to *Atwhy1why3* aged imbibed and WT relative to *Atwhy1why3* aged imbibed [Figure 25] and WT dry relative to aged imbibed [Figure 27], HSP17.6B appears in WT relative to *Atwhy1why3* aged imbibed [Figure 25], HSP17.6II in *Atwhy1why3* dry relative to aged imbibed [Figure 31] and heat-stress related proteins ATBAG6 in WT relative to *Atwhy1why3* aged imbibed [Figure 25] and *Atwhy1why3* dry relative to aged imbibed [Figure 31], together with Agamous-Like 67 in *Atwhy1why3* dry relative to aged imbibed [Figure 31]. Interestingly, in *Solanum lycopersicum* (tomato) plants exposed to heat stress, overexpression of SlWHY1 was found to reduce wilting and had

increased levels of SlHSP21.5A transcripts (Zhuang *et al*., 2020). Furthermore, these overexpression lines were found to have greater membrane stability and higher soluble sugar content, and decreased ROS levels compared to the WT, whilst the opposite phenotype was observed in RNAi lines with inhibited *SlWHY1* (Zhuang *et al*., 2020). The decreased expression of SlHSP21.5A transcript levels in these RNAi lines matches with the RNA-seq data obtained in this seed experiment. Clearly the interaction of HSPs and WHY proteins needs to be explored further.

To date, there are no reported studies on the role of HSP70 in seed ageing. However, the large number of heat shock proteins that appear in top ten DEGs for this treatment suggests that a there is a major function for heat shock proteins in seed ageing that should be explored further in aged seeds. Interestingly, transcripts encoding the rice HSP called OsHSP18.2 were significantly increased in during seed ageing. This protein was implicated in seed vigour and longevity. It also improves germination and seedling establishment under abiotic stress (Kaur *et al.*, 2015). Lines overexpressing *OsHSP18.2* showed improved germination and seedling establishment under stress situations. It is possible that overexpression of HSP70 might serve a similar function in Arabidopsis. Overexpression of the sunflower HaHSFA9 heat stress transcription factor in transgenic *Nicotiana tabacum* seeds increased resistance to the adverse effects of accelerated ageing (Prieto-Dapena *et al.*, 2006).

Another interesting finding arising from the RNA-seq dataset concerns the presence of LEA proteins, which have long been associated with ageing and stress responses. LEA proteins accumulate during seed maturation and in response to abiotic stresses as well as dehydration stress (Hundertmark *et al*., 2011). LEA proteins and HSPs are abundant in seeds that have undergone ageing. Like HSPs, LEA proteins are important in seed desiccation tolerance and longevity (Kaur *et al.*, 2015). A number of transcripts encoding LEA proteins were more abundant: LEA6 in WT dry relative to imbibed [Figure 20], *Atwhy1why3* dry relative to imbibed [Figure 21] and *Atwhy1why3* imbibed relative to aged imbibed [Figure 32] LEA76 in WT dry relative to imbibed [Figure 20] and *Atwhy1why3* dry relative to imbibed [Figure 21]; LEA protein ECP31 in *Atwhy1why3* dry relative to imbibed [Figure 21] and *Atwhy1why3* imbibed relative to aged imbibed [Figure 32]; and DEHYDRIN LEA in WT imbibed relative to aged imbibed WT imbibed relative to aged imbibed [Figure 28] and *Atwhy1why3* imbibed relative to aged imbibed [Figure 32]. Dehydrins and LEA proteins have roles in seed longevity and the dry dormant stage, as well as in germination and salt stress tolerance (Prieto-Dapena *et*

al., 2006). Furthermore, LEA protein transcripts were more abundant during dehydration in *Sporobolus stapfianus*, a C4 grass (Yobi *et al*., 2017).

The data presented in this chapter suggest that HSP and LEA proteins work together with WHY1 and WHY3 to protect germination under conditions of accelerated ageing stress. Other PhD students at the University of Leeds have also linked WHY1 with LEA5 via the binding partner of chloroplast DEA (D/H)-box RNA helicase 22 (Razak, 2019). It will be interesting to determine whether WHY1 and WHY3 bind to the promoters of genes encoding HSP and LEA protein in the nucleus to regulate their expression. How these proteins work together to regulate different processes associated with seed germination and ageing remains to be fully elucidated.

Perhaps unsurprisingly, most of the top ten DEGs for the RNA-seq analysis of WT and *Atwhy1why3* seeds under any treatment were related to seed storage and seed lipid processes. Additionally, 4 common DEGs (AT1G27360, AT1G27370, AT1G69170 and AT5G43270) in the WT pairwise comparisons [Figure 26] were later identified as microRNA targets for example miR156 and miR157 are Squamosa-promoter Binding Protein (SBP)-like proteins (Rhoades *et al*., 2002). These SBP-like proteins are involved in plant growth and development and are highly conserved in all green plants as they function in a range of processes from flowering time to shoot architecture (Preston and Hileman, 2013). These are less abundant in WT dry relative to imbibed but more abundant in WT dry relative to aged imbibed.

Due to the nature of this RNA-seq, DEGs were only identified that had an impact on the immediate initial 6 hours of growth from the point of imbibition compared to dry seeds. An expanded dataset with longer timeframes could show some interesting interactionsof genes with WHY1 and WHY3 throughout germination and processes that are initiated or blocked by the WHY proteins could be more thoroughly analysed. The 6 hour imbition was chosen as previous RNA-seq experiments in this lab found this to be a good time point for when germination processes are normally initiated and transcript levels are increased (Waterworth *et al*., 2016). However, perhaps in this should have been further adapted for the *Atwhy1why3* mutants which were slower to germinate than the WT. The high quality *Atwhy1why3* seeds did not have any germinating seeds until 48 hours after cold stratification and 7 day aged *Atwhy1why3* seeds had only 2.6% germination at 48 hours [Figure 13]. Future experiments should look at more time points such as 12 h, 24 h and 48 h in addition to the initial 6 h imbition to fully analyse the roles of WHY1 and WHY3 in germination.

6 WHIRLY FUNCTIONS IN ARABIDOPSIS DEVELOPMENT

6.1 INTRODUCTION

Arabidopsis is widely used as a model organism for studying the large-scale determination of gene function in plants. Analysis of mutants throughout the growth and development of Arabidopsis can provide key insights into potential targets for commercial crop breeding programmes (Boyes *et al*., 2001). Many different studies have recorded the developmental changes of rosettes, individual leaf area, leaf colour, plant biomass and photosynthetic ability to determine functional growth changes in Arabidopsis caused by mutations in the genome (Lièvre *et al*., 2016; Vanhaeren *et al*., 2015; Tanaka *et al*., 2013). Measuring specific changes of different genotypes during the growth phases of Arabidopsis acts as a detailed mechanism for understanding how genes effect plant morphology during optimal conditions as well as under abiotic or biotic stress (Boyes *et al*., 2001). This can lead to the detection of novel phenotypes and thus determine the function of genes which can then be applied to other plants.

Much of what we know about WHIRLY is through analysis of *Atwhy1*, *Atwhy3* and *Atwhy1why3* mutants in Arabidopsis, however there has been little consistent insights into the function of these proteins during plant development and stress response (Taylor *et al*., 2022). Single knock-out *Atwhy1* and *Atwhy3* mutant plants were phenotypically similar to wild-type plants (Maréchal *et al.*, 2009). However, the growth conditions for these plants in the study of Maréchal *et al.* (2009) were not clear. WHIRLY is very environment-sensitive, thus it is important that the reported phenotypes above are under stated standard growth conditions, therefore this needs to be explored further. Similarly, RNAi WHY1 knockdown barley lines were phenotypically similar to the WT with similar redox metabolite levels to the WT, with comparable photosynthesis and senescence rates but the leaves had significantly more chlorophyll and less sucrose (Comadira *et al.*, 2015).

Comparatively, there were stark differences in phenotype between the single Arabidopsis mutants and the *Zmwhy1-1* mutants, which were white in colour and died at the fourth leaf stage, compared to the green wild-type maize plants (Prikryl *et al.*, 2008). However, the removal of both *Atwhy1* and *Atwhy3* plastid localised WHY proteins in Arabidopsis in a double knockout *Atwhy1why3* mutant line presented differences in the WT. Some *Atwhy1why3* individuals were visibly smaller than the WT and 4.6% of plants had at least one variegated leaf sector, compared to no variegation in the WT (Maréchal *et al*., 2009). The small percentage of cases where double mutants had variegation also had a strong interference with chloroplast

development and function (Maréchal *et al.*, 2009), which is similar to the white coloured leaves in *Zmwhy1* mutants (Prikryl *et al.*, 2008). Furthermore, the *Atwhy1why3* double mutants showed higher levels of plastid genome instability compared to the wild type; the absence of both proteins led to DNA rearrangements (Xiong *et al.*, 2009, Maréchal *et al.*, 2009). Two lines of F_0 *Atwhy1why3* plants with variegation were crossed and progeny was formed. The variegated male parents produced no variegated F_1 progeny whilst variegated plants used as female produced variegation in 46% of the F_1 progeny (Maréchal *et al.*, 2009). Furthermore, the plastid damage in variegated sectors of the crossed F_1 progeny was found to be irreversible. These results suggested that the variegation was maternally inherited and the re-introduction of either or both WHY1 and WHY3 could not rescue the phenotype.

Further studies investigating this phenotype suggested that the previously reported plastid interference in *Atwhy1why3* mutants was more prominent when other chloroplast localised proteins, DNA polymerase 1B (Pol1B) (Parent *et al.*, 2011) or ATP-dependent recombinase 1 (RecA1) (Duan *et al.*, 2020), were also mutated. The *Atwhy1why3pol1b-1* triple mutant exhibited a significant growth defect and a yellow variegated leaf phenotype in all of the mutant plants, which is much more severe than the previously reported single and double mutant phenotypes (Lepage *et al.*, 2013, Parent *et al.*, 2011). The yellow variegated sections of the *Atwhy1why3pol1b-1* triple mutant had reduced photosynthetic electron transport efficiency and increased replication errors compared to wild-type (Parent *et al.*, 2011). Additionally, the *Atreca1why1why3* triple mutants had severe growth defects in leaves and white variegated sectors, as well as a higher accumulation of plastid DNA rearrangements compared to the wild-type plants (Duan *et al.*, 2020). Furthermore, the *Atreca1why1why3* mutants had lower fertility and produced siliques with several white ovules and abnormal shapes compared to the wild type (Duan *et al.*, 2020). This is an important area of study because there has been no previously reported effect of any of the *Atwhy* mutants on embryo lethality and yield. These results show that there is a possibility that the interaction between WHY1, WHY3 and REC1A in Arabidopsis maintains embryo development, possibly due to the reported increase in DNA repair errors which may also have caused the leaf deformations (Parent *et al.*, 2011). Therefore, it is also important to identify any phenotypic defects in leaf development in all *Atwhy* mutants, including single mutants, where it has previously only been reported in a small percentage of *Atwhy1why3* mutants (Maréchal *et al.*, 2009).

In addition to affecting plant growth and development, chloroplasts are also important sensors of environmental change as it is the site of many photosynthetic functions (Foyer *et al.*, 2014).

WHIRLY proteins have been implicated in photosynthesis; WHIRLY1 was found to be closely associated to thylakoid membranes through immunogold labelling studies with *Hordeum vulgare* (Grabowski *et al.*, 2008). This localisation at the boundary between thylakoids and nucleoids, makes WHY1 appropriately placed to function in the photosynthetic electron transport chain (Foyer *et al*., 2014). As such, a previous study found that the loss of WHY1 led to higher photochemical quantum yield of photosystem I (PSI), higher electron transport rate (ETR) and a lower thermal energy dissipation of chlorophyll a in non-photochemical quenching (NPQ) compared to the wild type (WT) (Huang *et al*., 2017). Mutants with loss of WHY, both *Atwhy1* and *Atwhy1why3* mutants had decreased chloroplast NAD(P)H dehydrogenase-like complex (NDH) activity and the accumulation of the NDH- PSI super-complex and several genes encoding this complex were also upregulated in these mutants (Huang *et al*., 2017). However, in these experiments, seven-week-old plants were used, which may well have undergone other physical changes such as bolting which is a sign of the plant shifting from growth to senescence (Hinckley and Brusslan, 2020), therefore it is important to obtain data on the effect of WHY1 before these developmental changes occur.

Interestingly, overexpression of plastid localised WHY1 (*oepWHY1*) in Arabidopsis was found to disrupt leaf senescence through decreased H_2O_2 levels, whilst *Atwhy1* mutants were found to have increased H_2O_2 levels 37 days post germination which coincided with early leaf senescence (Lin *et al*., 2019). The level of H_2O_2 in the mutant were restored to WT levels with additional expression of *WHY1* driven by the native promoter. These data show that WHY1 is essential for the homeostasis of H_2O_2 levels in plants and this can account for the altered leaf senescence phenotype and potentially provide an explanation for the variegation phenotype. Variegation is lack of chlorophyll, although in small plants like Arabidopsis, it could easily be mistaken for cell death. Additionally, Ren *et al.* (2022) recently discovered that leaf senescence was accelerated in a mutant line with knockdown of microRNA840 which targeted overlapping 3'UTRs of WHY3 and the pentatricopeptide repeat (PPR) family protein. There has been no previously reported effect of WHY3 on senescence and therefore the phenotype of *Atwhy3* mutants must be fully explored throughout growth. It is essential that WHY3 is also explored in Arabidopsis as WHY1 and WHY3 appear to share a lot of dual functions. Furthermore, it is important to use plants that have not already started to undergo senescence as there is much evidence for the changing function of WHY1 and WHY3 during senescence, which is explored in more detail later in this chapter.

Additionally, WHY proteins have been implicated in transcripts responsive to phytohormones which are involved in leaf senescence. *Atwhy* mutants had increased expression levels of *ERF13* (a coupling element 1 binding protein) which confers ABA hypersensitivity and *ERF109* which is induced by singlet oxygen and increased in response to Jasmonate (JA) (Huang *et al*., 2017). Exogenous treatment of JA has been found to induce leaf senescence and concentrations of JA are higher in senescent leaves compared to non-senescent ones (Hu *et al*., 2017). Furthermore, ERF109 enhances high light defence response in young leaves whilst in older leaves it induces senescence and chlorosis (Huang *et al*., 2017). Therefore, elevated expression of these genes might lead to ABA hypersensitivity, resistance to high light levels and early senescence. In contrast to normal light conditions, there was lower ETR (electron transport rate) under high light conditions (800 μmol m^{-2} s^{-1}) in both *Atwhy1* and *Atwhy1why3* mutant lines compared to WT. Additionally, WHY1 was found to physically interact with Light-harvesting protein complex I (LHCA1) which functions as a light receptor and delivers excitation energy to photosystem I (Huang *et al.*, 2017). This suggests that WHY1 has a role in maintaining ETR levels under high light or even light stressed conditions which may have led to an early senescent phenotype, although this needs further exploration (Huang *et al*., 2017).

One of the leaf senescence genes that WHY1 potentially regulated based on ChIP data in *Hordeum vulgare* (barley) was *Hvs40*, a leaf senescence marker which is expressed in response to biotic and abiotic stresses (Janack *et al.*, 2016). Drought stress induced the establishment of euchromatic histone modifications in H3K9me2 levels in HvS*40* in WT, however RNAi WHY1 knockdown plants there was no change in H3K9me2 levels under drought stress or senescence. Furthermore, the WHY1 knockdown RNAi plants had delayed senescence during drought (Janack *et al.*, 2016). This data suggested that the function of WHY1 may be related to the establishment of epigenetic markers during stress and development or even that epigenetic markers are resulted from gene expression changes during stress. This suggests that there is a potential for HvWHY1 to be an activator of senescence under both biotic and abiotic stresses. Furthermore, RNAi-mediated *Hvwhy1* lines of a partial knockdown had decelerated leaf senescence in primary foliage (Comadira *et al.*, 2015, Kucharewicz *et al.*, 2017). This data suggests that the leaf senescence phenotype may be due to interactions with even more protein families than previously thought and therefore senescence in *Atwhy* mutants is important to thoroughly define before further action is taken.

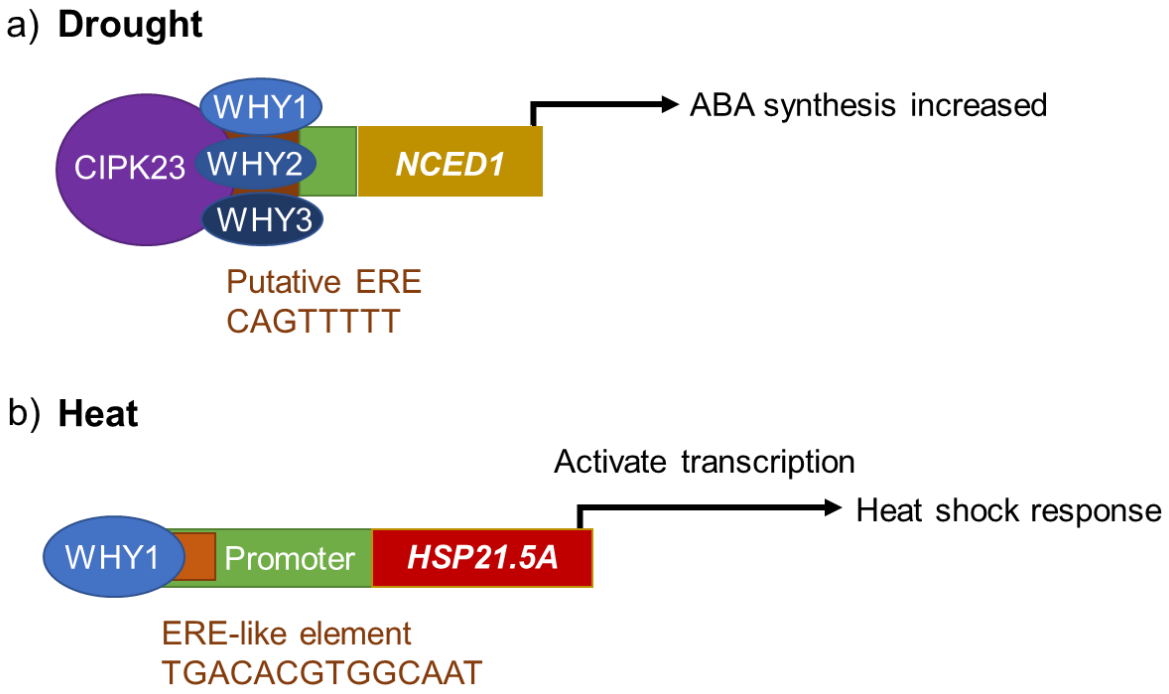


Figure 35: a) WHY proteins bind to an upstream region in the NCED1 gene promoter and to the CIPK23 promoter of *Manihot esculenta* (cassava) in response to drought stress to activate abscisic acid biosynthesis (Yan *et al.*, 2020); b) WHY1 binds to the ERE-like element in the promoter of the HSP21.5A gene of *Solanum lycospericum* (tomato) in response to heat stress to activate the heat shock response (Zhuang *et al.*, 2020a).

Many studies have reported increases in the levels of WHY transcripts in plants exposed to environmental stresses including high salt levels, drought (Akbudak and Filiz, 2019; Yan *et al.*, 2020), heat (Zhuang *et al.*, 2020a), oxidative stress (Tunc-Ozdemir *et al*, 2009) and infection with the fungus, *Botrytis cinera* (Akbudak and Filiz, 2019). Overexpression of tomato SlWHY1 in *Nicotiana tabacum* (tobacco) led to enhanced drought tolerance (Zhao *et al.*, 2018). Under droughted conditions the expression of *MeWHY1*, *MeWHY2* and *MeWHY3* in *Manihot esculenta* (cassava) were significantly up regulated as well as increased *MeCIPK23* transcript levels, which has been shown to have a strong protein interaction with MeWHYs (Yan *et al.*, 2020; Janack *et al.*, 2016). It is unclear whether the discussed MeWHY3 is similar to the *Arabidopsis* AtWHY3 protein as it is only briefly mentioned and never mentioned separately of MeWHY1 and MeWHY2 (Yan *et al.*, 2020).

The expression of cassava *9-cis-epoxycarotenoid dioxygenase* MeNCED genes, which encode the key enzymes controlling ABA biosynthesis displayed a dramatic decrease in RNAi silenced *Mewhy* plants (Yan *et al.*, 2020). Both MeCIPK23 and the three MeWHYs were found to bind to the promoter element of MeNCED1 to activate its transcription which resulted in increased ABA levels and an increased drought stress response [Figure 35]. Furthermore RT-qPCR data

showed that both *SlWhy1* and *SlWhy2* genes were upregulated under salt and drought stress in tomato (Akbudak and Filiz, 2019) [Figure 35]. Moreover, the dual localisation of WHY1 in the plastid and nucleus gives it an ideal position for transducing information, such as retrograde signals, to increase defence responses (Isemer *et al.*, 2011, Tada *et al.*, 2008). Under droughted conditions, *SlWHY2* RNAi plants had an increased wilting phenotype as well as a decrease in the major measurable photosynthetic traits such as lower fresh weight, chlorophyll content, quantum yield of photosystem I and maximal photochemical efficiency of PSII when compared to the wild type (Meng *et al.*, 2020). These lines also had a high accumulation of ROS compared to the wild-type and a decrease in the expression levels of seven DNA repair and recombination genes in the mitochondria (Meng *et al.*, 2020); all these indicate major stress on the plant and therefore the phenotypes could also be pleotropic effects arising from mitochondrial dysfunction. This in-depth study showed that whilst there is some effect of WHY1 on droughting stress, the majority of the function may come from the mitochondrial-based WHY2 protein. However, an interaction between the two proteins cannot be ruled out and further investigation into the role of WHY1 in droughting stress is needed.

There has been very little research on the interaction of WHY1 and WHY3 under heat stress or heat shock. This is surprising as the 24-mer assembly of both WHY1 and WHY3 proteins resembles bacterial heat shock proteins which form non-whirly spherical shapes. The WHY proteins 24-mer structure resembles bacterial heat shock proteins (HSPs) (Cappadocia *et al*., 2012). It is possible that WHY proteins function in a heat shock or temperature regulatory function. It was recently discovered that SlWHY1 bound to and regulated SlHSP21.5A (heat shock protein) expression. This potentially promotes thermal tolerance observed in lines with increased WHY1 [Figure 35]. Overexpression lines of SlWHY1 had upregulated levels of SlHSP21.5A and RNAi lines displayed downregulated levels compared to WT (Zhuang *et al*., 2020a). Additionally reduced membrane stability and soluble sugar content and increased ROS accumulation was found in the RNAi lines under heat stress compared with the WT, suggesting a wide-ranging role of WHY1 in heat stress response. Furthermore, immunoblot analysis showed SlWHY1 protein levels were increased in seedlings after 12 hours of heat treatment (Zhuang *et al*., 2020a). However, there have not been any studies on the effect of WHY3 on heat shock or whether there is any interaction of both WHY1 and WHY3 under heat shock.

6.1.1 Objectives

This chapter aims to identify the functions of the WHY1 and WHY3 proteins in later development of Arabidopsis plants. As discussed above, accumulating evidence demonstrates that WHY proteins function as transcriptional activators in the nucleus to co-ordinate plant development and defence and the function of WHY proteins may change throughout growth. There have been little consistent insights into the function of the WHY proteins throughout vegetative growth. Therefore, phenotypes of the *Atwhy1*, *Atwhy3* and *Atwhy1why3* mutants were characterised in post germination growth. Moreover, it is clear from previous studies that there is a function of WHY1 in stress response and that WHY proteins have a function in photosynthetic capacity which can impact growth and development. However the two have not previously been linked for studies, thus changes in the photosynthetic capacity of the *Atwhy* mutants were compared under optimal and heat stressed conditions, in order to better understand any interaction between WHY1 and WHY3 in photosynthesis.

Therefore, the objectives of this chapter were three-fold:

1. To characterise the effect of WHY1 and WHY3 throughout Arabidopsis growth
2. To compare photosynthesis in *Atwhy1*, *Atwhy3*, and *Atwhy1why3* mutants relative to the wild type.
3. To determine the responses of the *Atwhy1*, *Atwhy3*, and *Atwhy1why3* mutants and the wild type plants to heat stress.

6.2 RESULTS

The rosette phenotypes of the *Atwhy* mutants and the wild type plants were observed post germination under optimal growth conditions to determine whether the loss of WHY1 and/ or WHY3 functions altered rosette growth, senescence and seed production. Furthermore, the Li-6400 and Li-6800 Portable Photosynthesis Systems (LiCORs) were used to determine whether the loss of WHY1 and/ or WHY3 functions had effects on the photosynthetic capability of plants. A further study was conducted on the responses of photosynthesis to heat stress and rosette phenotypes were also assessed in plants that were temporarily heat stressed to determine if WHY1 and/ or WHY3 had an effect on recovery from short term heat stress. These short-term experiments induced HSPs as part of the heat shock response (Jagadish *et al.*, 2021); RNA-seq experiments of Arabidopsis seeds [Chapter 5] found a link between WHY1, WHY3 and HSPs. The following results address the effects of WHY1 and WHY3 proteins on Arabidopsis development post germination.

6.2.1 WHIRLY Maintains Rosette Phenotype Throughout Growth

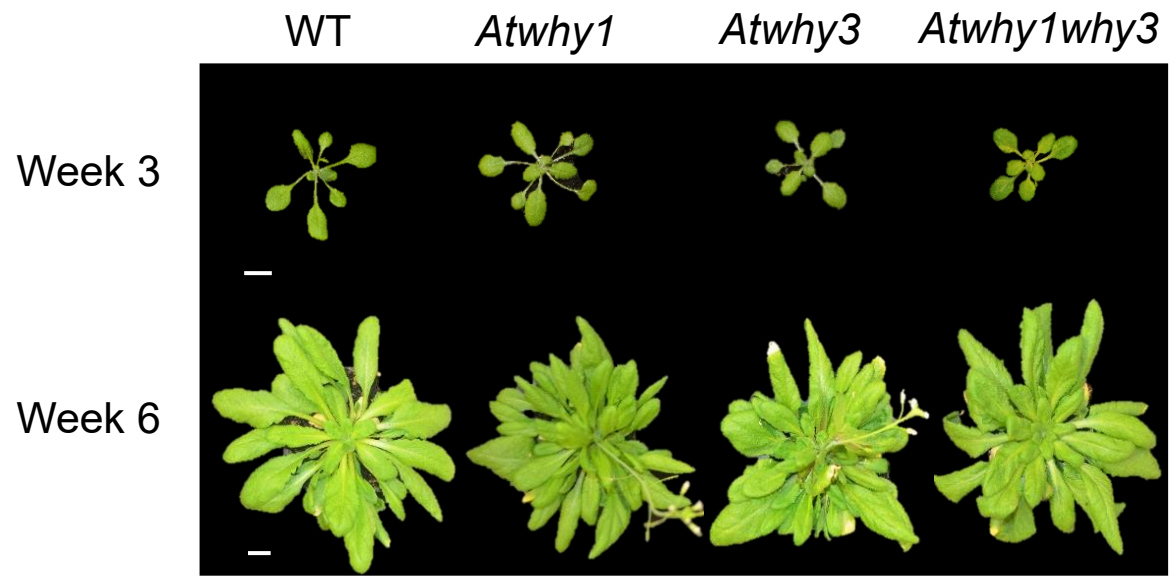


Figure 36: Rosette phenotypes of the *Atwhy* mutants compared to WT at three- and six-weeks post germination. Photographs were taken with a Nikon Digital Camera D5100. Scale bars: 10 mm

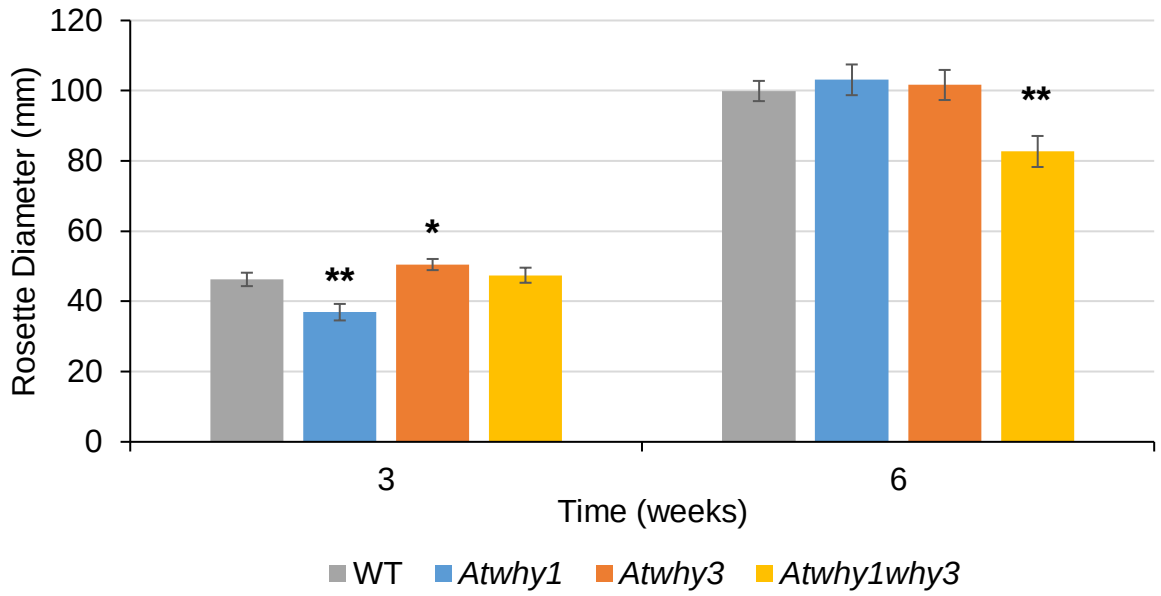


Figure 37: Rosette diameters throughout growth at three- and six-weeks post germination. Means were calculated from three biological replicates. Error bars were standard error of means and students t tests were conducted to compare each mutant to the WT at each time point: *, $p < 0.05$; **, $p < 0.01$

All three mutants and the WT were grown under the same conditions and rosette phenotype was compared. There were no visible differences between the *Atwhy1*, *Atwhy3* and *Atwhy1why3* mutants and the WT after 3 or 6 weeks of growth [Figure 36]. However, when the rosette diameters were quantified using Image J (Schneider *et al.*, 2012) [Figure 37] the *Atwhy1* mutants were significantly smaller ($p = 0.0042$) than the WT plants and *Atwhy3* mutants were significantly larger ($p = 0.0472$) in diameter compared to the WT plants after three weeks. There were no significant differences with either of the single mutants compared to the WT at week 6. However, the *Atwhy1why3* mutants were significantly smaller ($p = 0.0011$) in diameter compared to the WT at 6 weeks.

The same batch of plants were used to determine leaf number and total leaf area as a measurement of plant growth, photographs were taken of the leaves at each week of growth [Figure 40] which were then analysed using ImageJ RGB analysis to determine leaf area [Figure 38] and counted to determine leaf number [Figure 39]. The leaf photographs were also used to show leaf greening in the three *Atwhy* mutants compared to the wild type which were further analysed for chlorophyll pigment content [Figure 41].

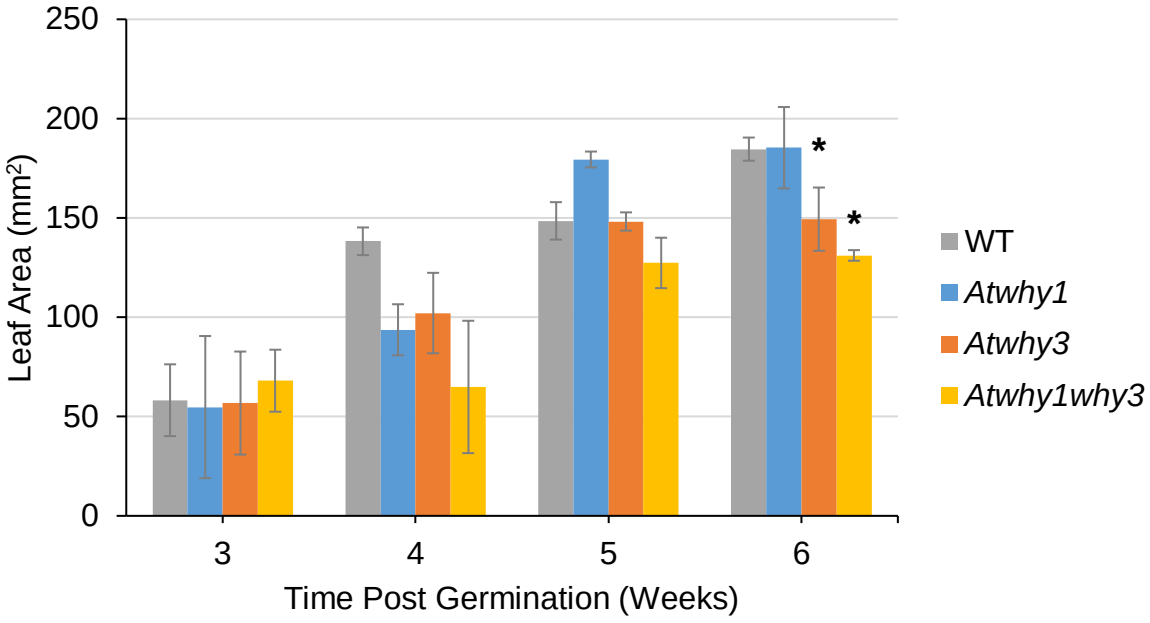


Figure 38: Whole rosette leaf measurements of average leaf area of the three *Atwhy* mutants compared to WT throughout growth. Means were calculated from three biological replicates. Error bars were standard error of means and students t tests were conducted to compare each mutant to the WT at each time point: *, $p<0.05$

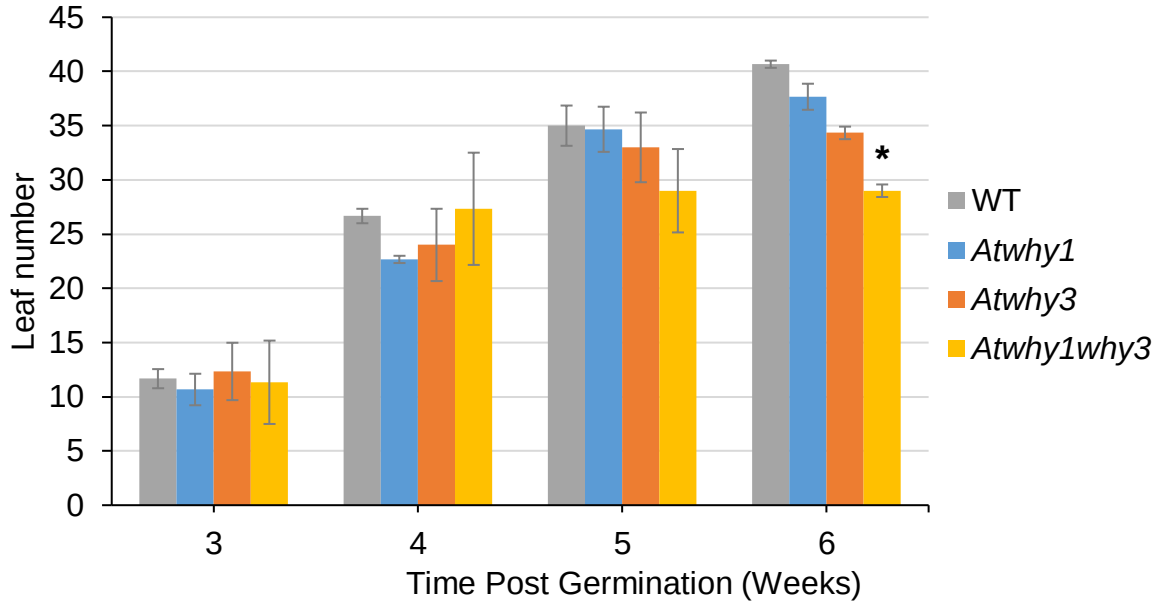


Figure 39: Numbers of leaves on whole rosettes of the three *Atwhy* mutants compared to WT throughout growth. Means were calculated from three biological replicates. Error bars were standard error of means and students t tests were conducted to compare each mutant to the WT at each time point: *, $p<0.05$

Leaf area [Figure 38] and leaf number [Figure 39] increased throughout growth in all *Atwhy* mutants and WT plants. Despite the significant changes in rosette diameter at week 3 [Figure

37], there were no significant changes in leaf area [Figure 38] or leaf number [Figure 39] in early rosette development (weeks 3 to 5). However, in later development both the *Atwhy3* and *Atwhy1why3* mutants had a significantly smaller average leaf area [Figure 38]. This suggests that the lack of WHY3 in both single and double mutant has a role in later leaf development, interestingly however, the number of leaves were not significantly reduced in the *Atwhy3* mutants compared to the WT, whereas there were significantly fewer leaves in the double mutant at the 6 week time point compared to the WT [Figure 39], which matches with the smaller diameter of the *Atwhy1why3* mutant plants [Figure 37]. The leaf images on Figure 40 also match with this data. The leaves of all genotypes were of similar colour, size and shape in week 3 and 4. However by weeks 5 and 6 the leaves of *Atwhy3* and *Atwhy1why3* mutants were paler and showed signs of senescence [Figure 40]. In the next section chlorophyll pigment content was analysed to ascertain whether WHY3 was involved in the greening process of *Arabidopsis* plants.

6.2.2 WHIRLY Maintains Leaf Colour and Chlorophyll Pigment

Figure 40: Whole rosette leaf photographs of the three *Atwhy* mutants compared to WT throughout growth from 3 to 6 weeks post germination. Scale bars: 10 mm

Leaf senescence, as judged by the yellow colouration of the leaves, began to develop at 5 weeks post germination in all genotypes, and almost half of leaves in all *Atwhy* mutants had yellowing colouration by week 6, in contrast fewer yellow leaves were observed on the WT leaves [Figure 40]. There were no signs of variegation observed on the leaves of any of the *Atwhy* mutants or WT plants. These leaves were used for chlorophyll extractions which showed no significant differences in chlorophyll pigment in any of the *Atwhy* mutants compared to WT in week 3 or 4 [Figure 41], just as there was no visible differences in any of the *Atwhy* mutant leaves compared to the WT at this early growth stage [Figure 40]. However, differences in both colouration [Figure 40] and chlorophyll pigmentation [Figure 41] became more prominent in later growth stages as stated above. This was further analysed by comparing chlorophyll a and b ratios of each genotype at each growth stage [Figure 42].

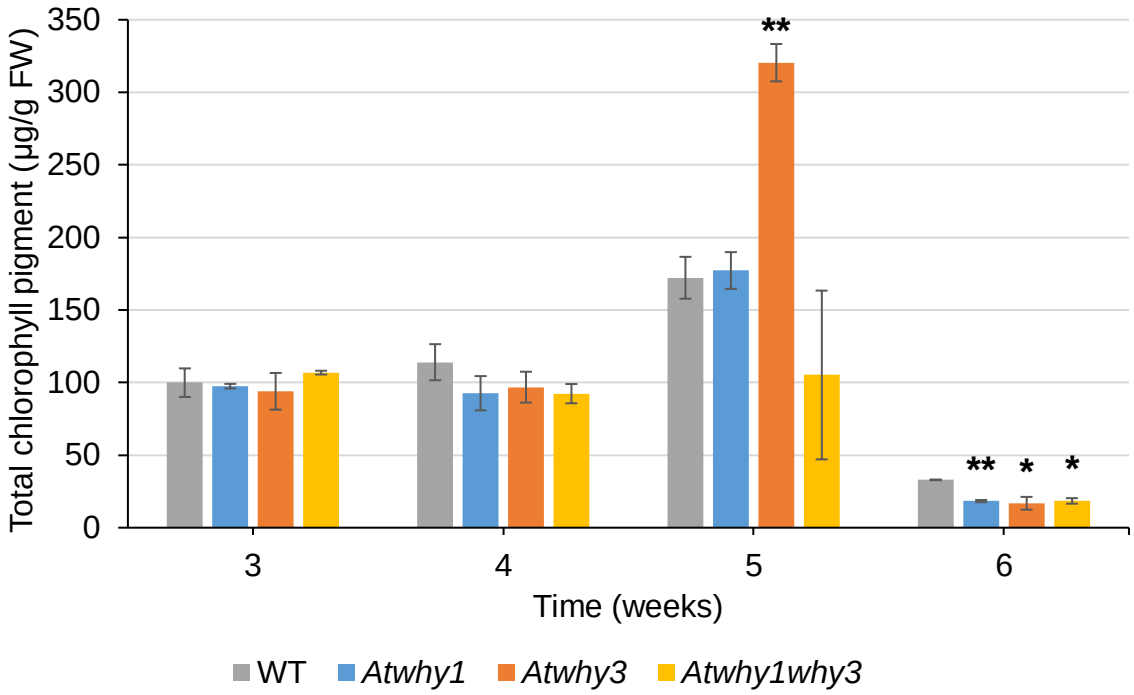


Figure 41: Chlorophyll a and b pigment content in leaves harvested at 6 weeks of growth from *Atwhy* mutants compared to WT. Means were calculated from three biological replicates. Error bars were standard error of means and students t tests were conducted to compare each mutant to the WT at each time point: *, $p < 0.05$; **, $p < 0.01$.

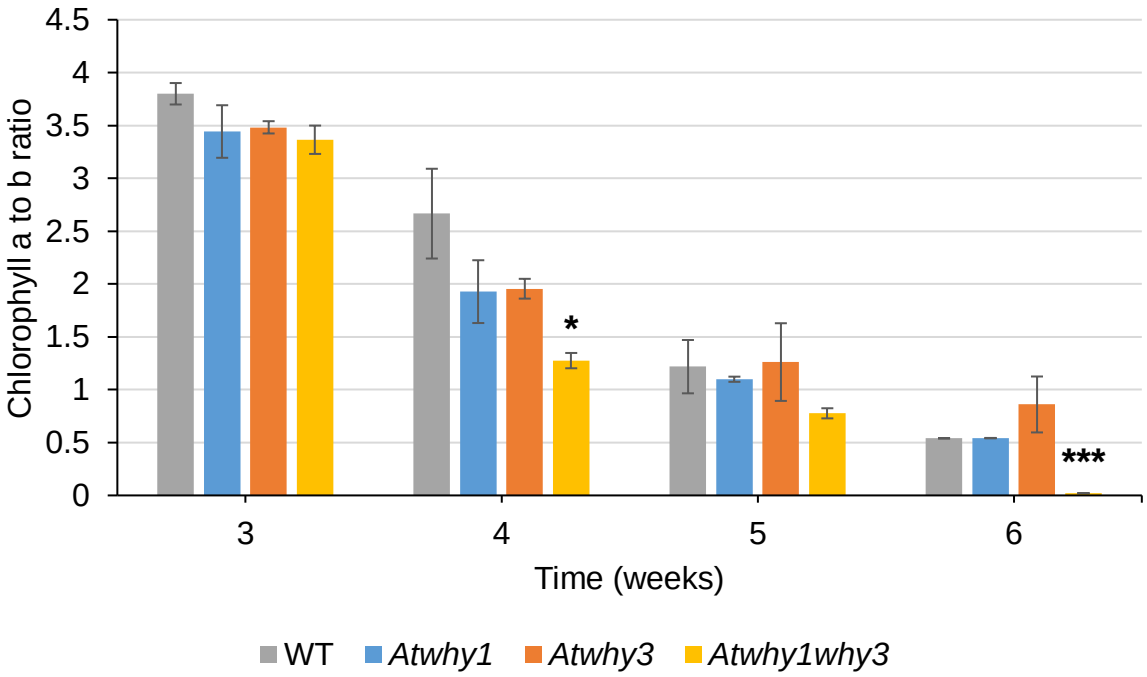


Figure 42: Comparison of Chlorophyll a mg/g FW to Chlorophyll b mg/g FW ratio per week in each genotype. Ratios are a mean of three biological replicates. Error bars were standard error of means and students t tests were conducted to compare each mutant to the WT at each time point: *, $p < 0.05$; ***, $p < 0.001$.

Levels of chlorophyll a and b increased in the leaves of the single *Atwhy1* mutants and the WT between weeks 3 and week 5, from around 100 μg/g FW to just over 170 μg/g FW [Figure 41]. Similar increases in leaf chlorophyll were observed in the *Atwhy3* mutants. By week 5 there was a significant increase, with doubled levels compared to the other *Atwhy* mutants and the WT from the other. There was significantly more chlorophyll a and b in the *Atwhy3* mutants compared to the WT [Figure 41] which also exhibited early senescence compared to the WT and the other mutants [Figure 40]. At 6 weeks post germination, both *Atwhy1* and *Atwhy1why3* mutants had significantly less chlorophyll a and b compared to the WT [Figure 41]. Chlorophyll levels of all genotypes decreased by week 6, the same time when senescence was visible in all leaves in Figure 40.

Chlorophyll a to b ratios decreased with the ageing of rosettes. However, the *Atwhy1why3* mutants had significantly lower chlorophyll a/b ratios in week 4 and week 6 compared to the WT [Figure 42]. Interestingly, chlorophyll a/b ratios were decreased in ratio by around 7-fold between week 3 and week 6 in WT, by around 6-fold in Atwhy1 mutants and by around 4-fold in *Atwhy3* mutants. Contrastingly, *Atwhy1why3* mutants were reduced by almost 160-fold from 3.36 in week 3 to 0.02 in week 6, this was the largest ratio difference in any of the genotypes [Figure 42].

6.2.3 Loss of WHIRLY Reduced Biomass

All *Atwhy* mutants and WT increased in both fresh [Figure 43] and dry [Figure 44] biomass throughout growth. The *Atwhy* mutants had lower biomass in both fresh and dry weights after 6 weeks of growth when senescence had begun based on leaf phenotype data analysed in 6.2. 2.. However, only *Atwhy1why3* was significantly greater than WT in fresh biomass in weeks 3 [Figure 43]. By week 6, all of the plants without WHY3, both the *Atwhy3* single and the *Atwhy1why3* mutants, had significantly reduced fresh biomass compared to the WT and *Atwhy1* was also reduced in weight, but this was not significantly different [Figure 43]. All three of the *Atwhy* mutants were significantly reduced in dry biomass at week 3 compared to the WT [Figure 44]. However, none of the mutants were significantly different to the WT in week 4 or 5, yet by week 6 all of the *Atwhy* mutants were again significantly reduced relative to the WT in dry biomass [Figure 44].

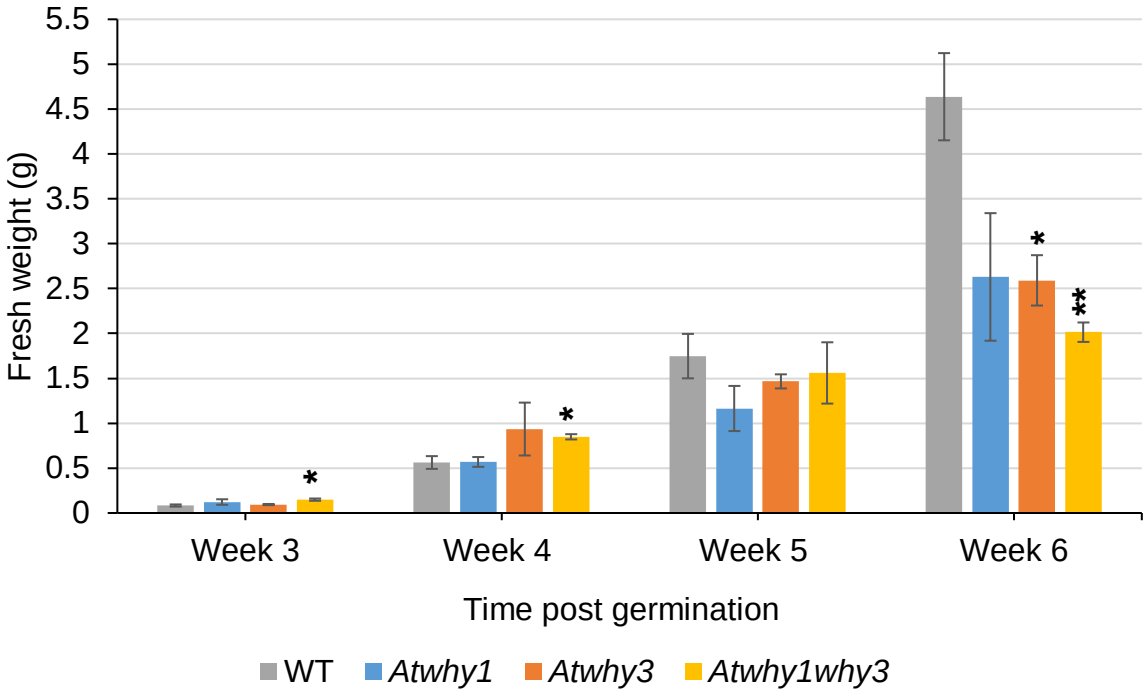


Figure 43: Fresh weight of WT and *Atwhy* mutants throughout development. Means were calculated from three biological replicates. Error bars were standard error of means and students t tests were conducted to compare each mutant to the WT at each time point: *, $p < 0.05$; **, $p < 0.01$.

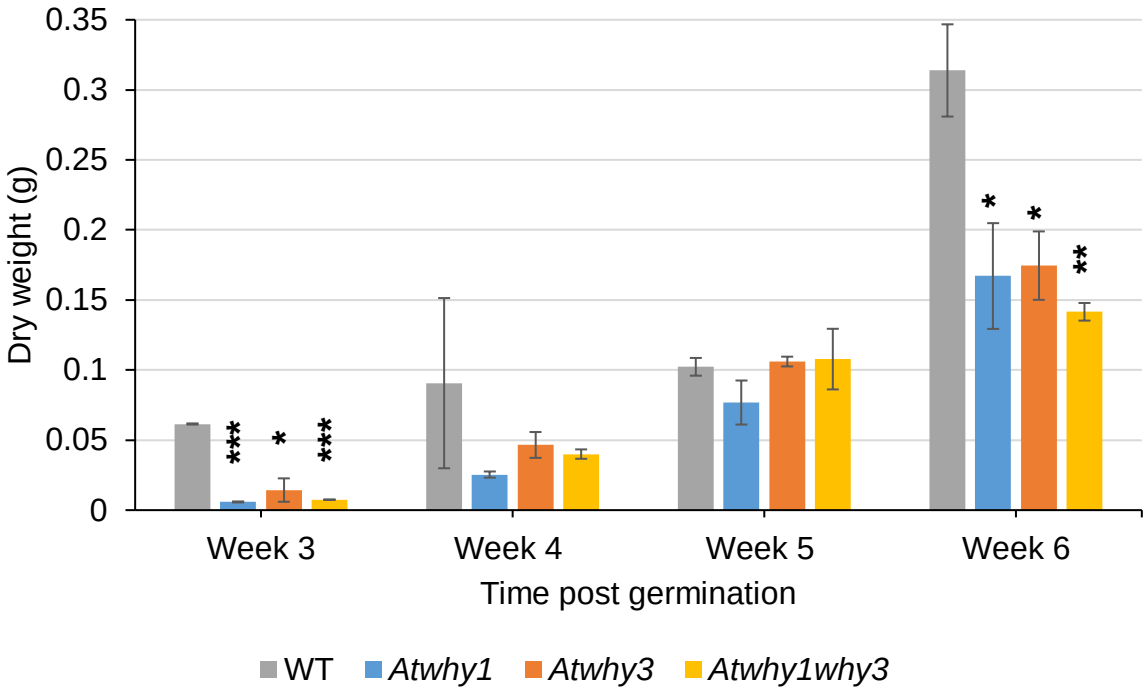


Figure 44: Dry weight of WT and *Atwhy* mutants throughout development. Means were calculated from three biological replicates. Error bars were standard error of means and students t tests were conducted to compare each mutant to the WT at each time point: *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$

6.2.4 Loss of WHIRLY Delayed Bolting

The plant bolting and seed set was observed for differences between the three *Atwhy* mutants compared to the WT to determine whether significant differences seen in rosette measurements extended beyond the phenotypes seen in week 6.

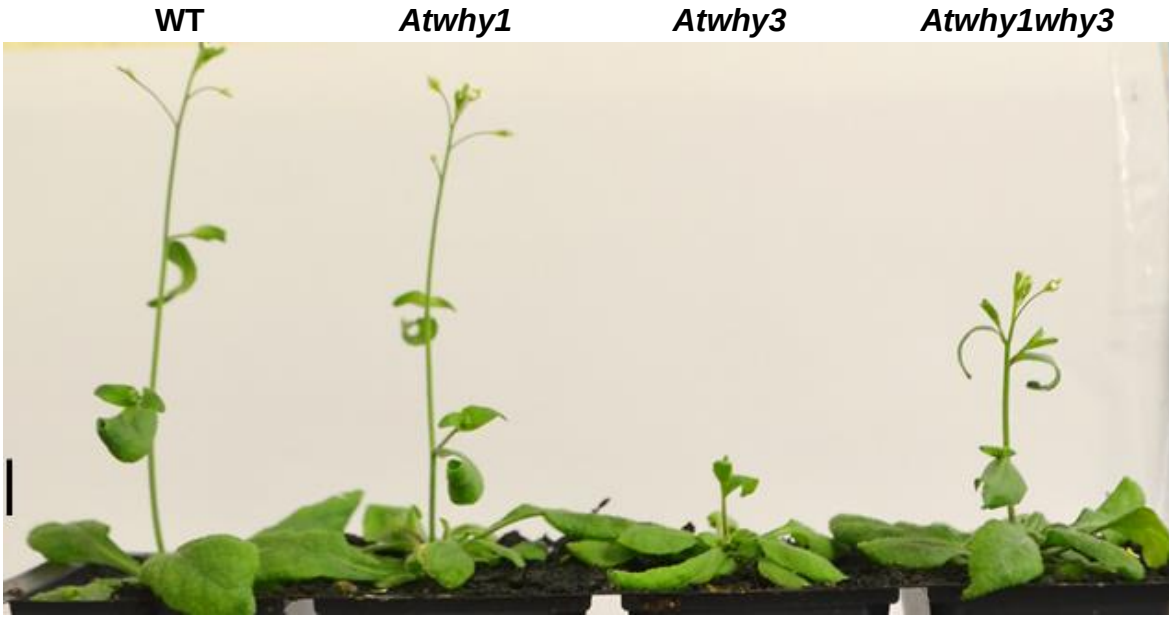


Figure 45: Vertical rosette photograph depicting bolting phenotypes of the three *Atwhy* mutants compared to WT at 6 weeks post germination. Scale bar on left of image: 10 mm.

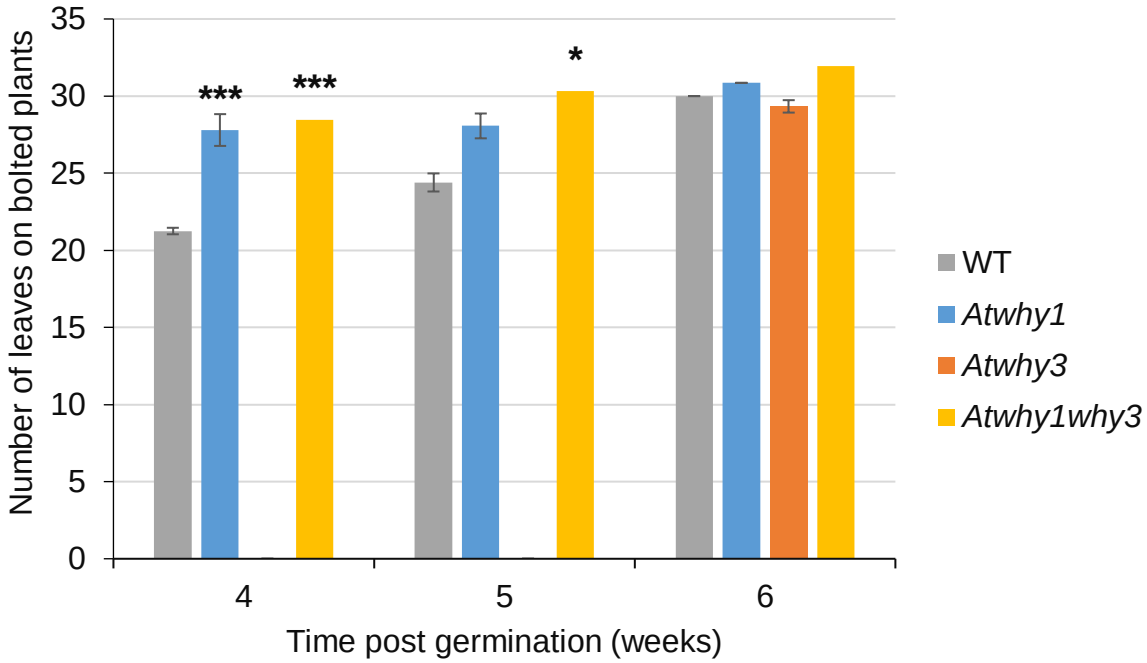


Figure 46: Number of leaves on bolted plants of the three *Atwhy* mutants compared to WT throughout growth. Means were calculated from three biological replicates. Error bars were standard error of means and students t tests were conducted to compare each mutant to the WT at each time point: *, $p < 0.05$; ***, $p < 0.001$

Plants of all genotypes had bolted by 6 weeks post germination [Figure 45]. However, visible inspection suggested that that there was a delay in bolting in the *Atwhy3* mutants compared to all the other lines. The number of leaves present on rosettes at same time as the appearance of the first flowering stem acted as a measure for developmental changes in bolted plants between the different genotypes [Figure 46]. Analysis of this parameter revealed that there was a delay in flowering in the *Atwhy3* mutants compared to all other lines [Figure 46]; there were no bolted *Atwhy3* plants before week 6. However, after a 2 week delay, they had caught up to the levels of bolting and leaf development of the other genotypes, and at this time point there was no significant difference in the bolting phenotype of any of the plants. It is important to note that more plants of every genotype bolted between weeks 5 and 6 which accounts for the increase in number of leaves on bolted plants week on week. This data further suggested that plants which bolted later on in development had more leaves than those which bolted earlier in development. The extreme bolting delay present in the *Atwhy3* mutants was not seen in the *Atwhy1* and *Atwhy1why3* plants which had significantly more leaves on bolted plants compared to WT at 4 weeks post germination [Figure 46] and *Atwhy1why3* mutants had significantly more leaves on bolted plants compared to the WT in week 5 too [Figure 46]. Even though these significant differences between the plants lacking WHY1 and the WT did not occur by week 6, it was important to determine if these changes had an effect on the seed development or seed yield of these plants. Therefore, the number of siliques per plant were counted as well as the number of seeds per silique [Figure 47].

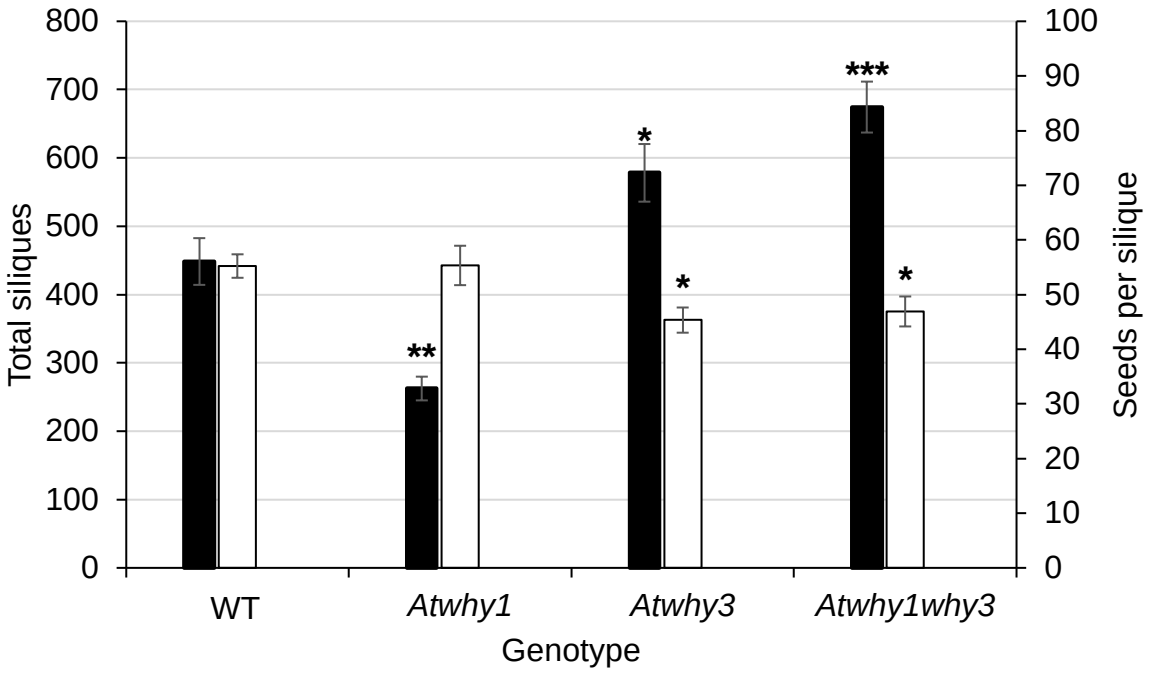


Figure 47: Total siliques per plant (black) and number of seeds per silique (white) in the three *Atwhy* mutants compared to the WT. Means were calculated from three biological replicates. Error bars were standard error of means and students t tests were conducted to compare each mutant to the WT at each time point: *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$

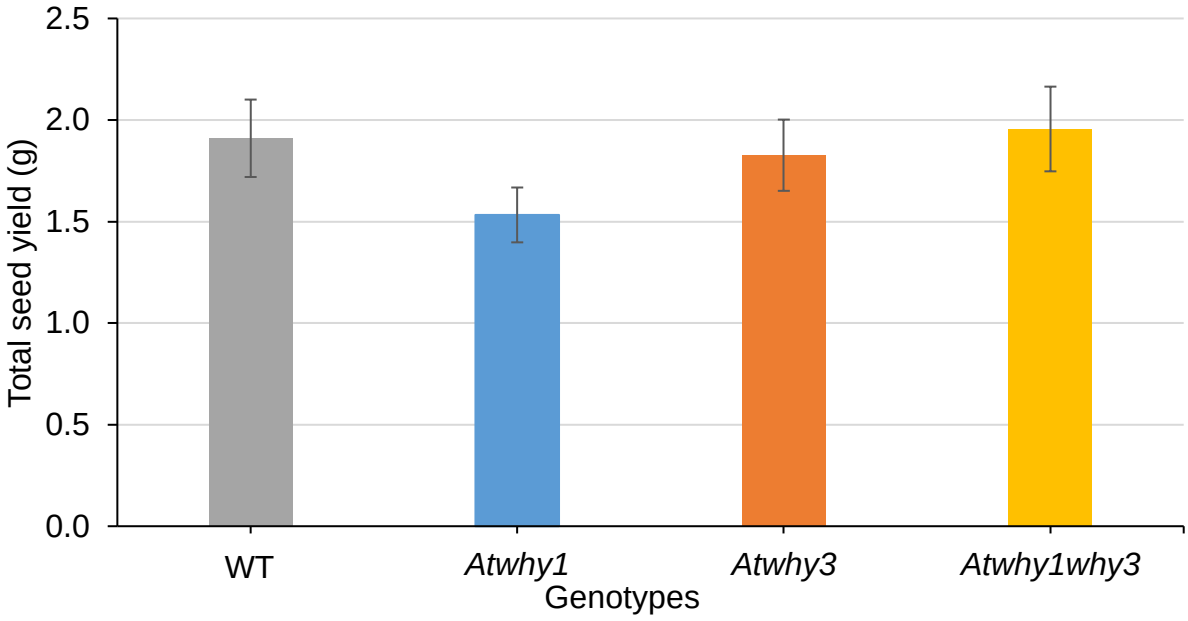


Figure 48: Total seed yield of the three *Atwhy* mutants compared to WT plants. Means were calculated from thirty biological replicates across six separate growth trays. Error bars were standard error of means and students t tests were conducted to compare each mutant to the WT at each time point: $p > 0.05$

All genotypes produced siliques and seeds. The *Atwhy3* and *Atwhy1why3* mutants had significantly more total siliques per plant than the WT, whilst the *Atwhy1* plants had significantly fewer siliques per plant than the WT [Figure 47]. There were no significant differences in the numbers of seeds per silique in the *Atwhy1* mutants compared to the WT and there were significantly fewer seeds per silique in the *Atwhy3* and *Atwhy1why3* mutants [Figure 47]. Therefore, in the plants lacking WHY3 there were significantly more siliques compared to the WT but less seeds per silique whilst in the *Atwhy1* single mutants there are fewer siliques compared to the WT but no change in the seeds per silique. In order to understand this further the total seed yields were calculated for all of the *Atwhy* mutants compared to the WT [Figure 48]. This found that only *Atwhy1* plants had a reduction in total seed yield compared to the WT, although this was not significantly different [Figure 48].

6.2.5 WHIRLY and Photosynthetic Capacity

In order to discover whether the change in rosette or bolting phenotype was caused by a change in photosynthetic capacity during the middle growth stages (week 4), a LiCOR was used to determine any changes in photosynthetic gas exchange in each of the *Atwhy* mutant plants compared to the WT. These data were recorded using plants that had not yet bolted.

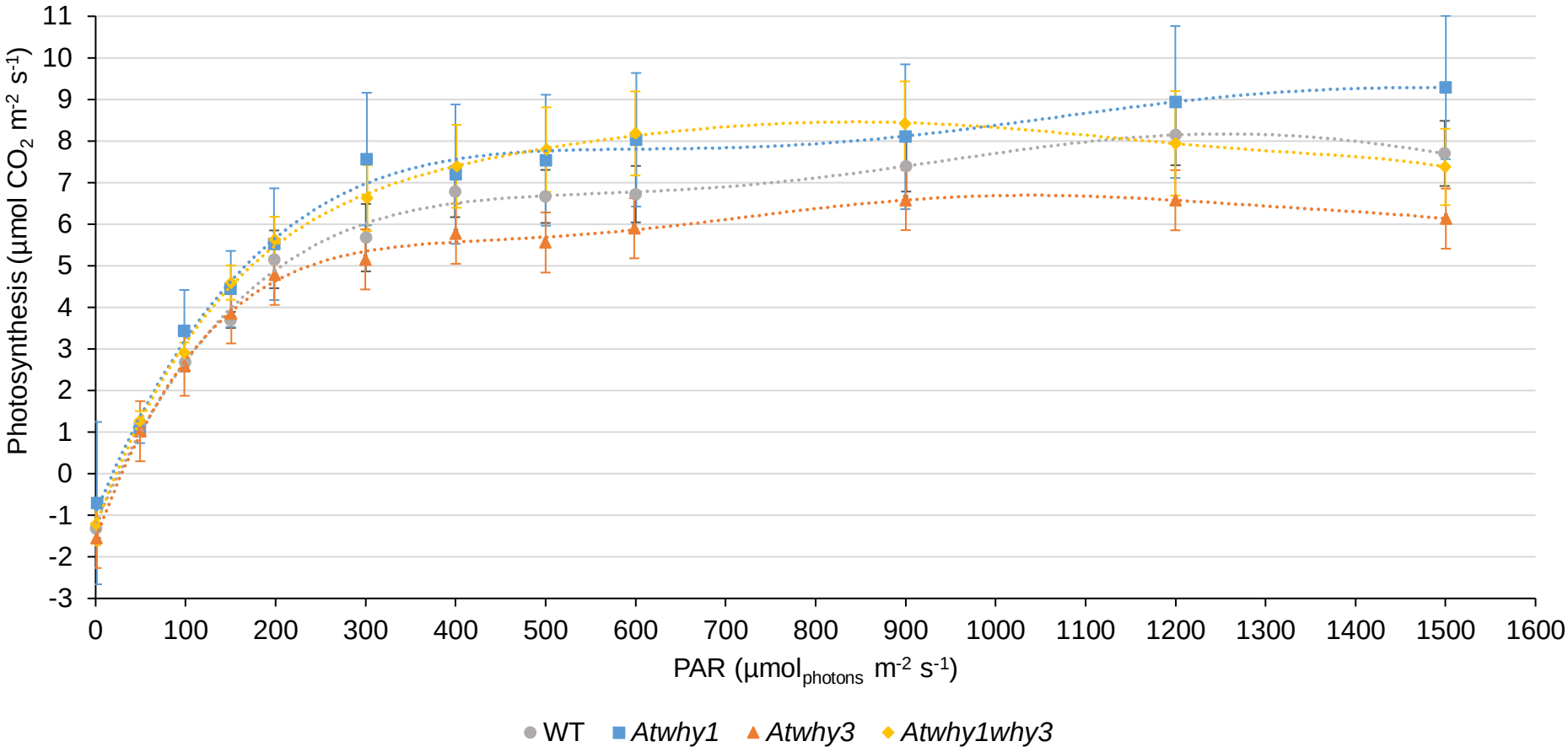


Figure 49: Light response curves for photosynthesis in the three *Atwhy* mutants compared to WT at increasing levels of photosynthetically active radiation (PAR). Means were calculated from four biological replicates. Plants were four weeks old when measured. Error bars were standard error of means. Trendline for each genotype shows the polynormial trend to the order of 4, calculated based on information from Lobo *et al.* (2013): WT ($y = 2E\text{-}14x^5 - 8E\text{-}11x^4 + 1E\text{-}07x^3 - 0.0001x^2 + 0.0499x - 1.203$), *Atwhy1* ($y = 1E\text{-}14x^5 - 6E\text{-}11x^4 + 1E\text{-}07x^3 - 0.0001x^2 + 0.0524x - 0.9456$), *Atwhy3* ($y = -2E\text{-}17x^6 + 1E\text{-}13x^5 - 3E\text{-}10x^4 + 4E\text{-}07x^3 - 0.0002x^2 + 0.0603x - 1.5832$), and *Atwhy1why3* ($y = -2E\text{-}17x^6 + 1E\text{-}13x^5 - 3E\text{-}10x^4 + 3E\text{-}07x^3 - 0.0002x^2 + 0.0584x - 1.2492$).

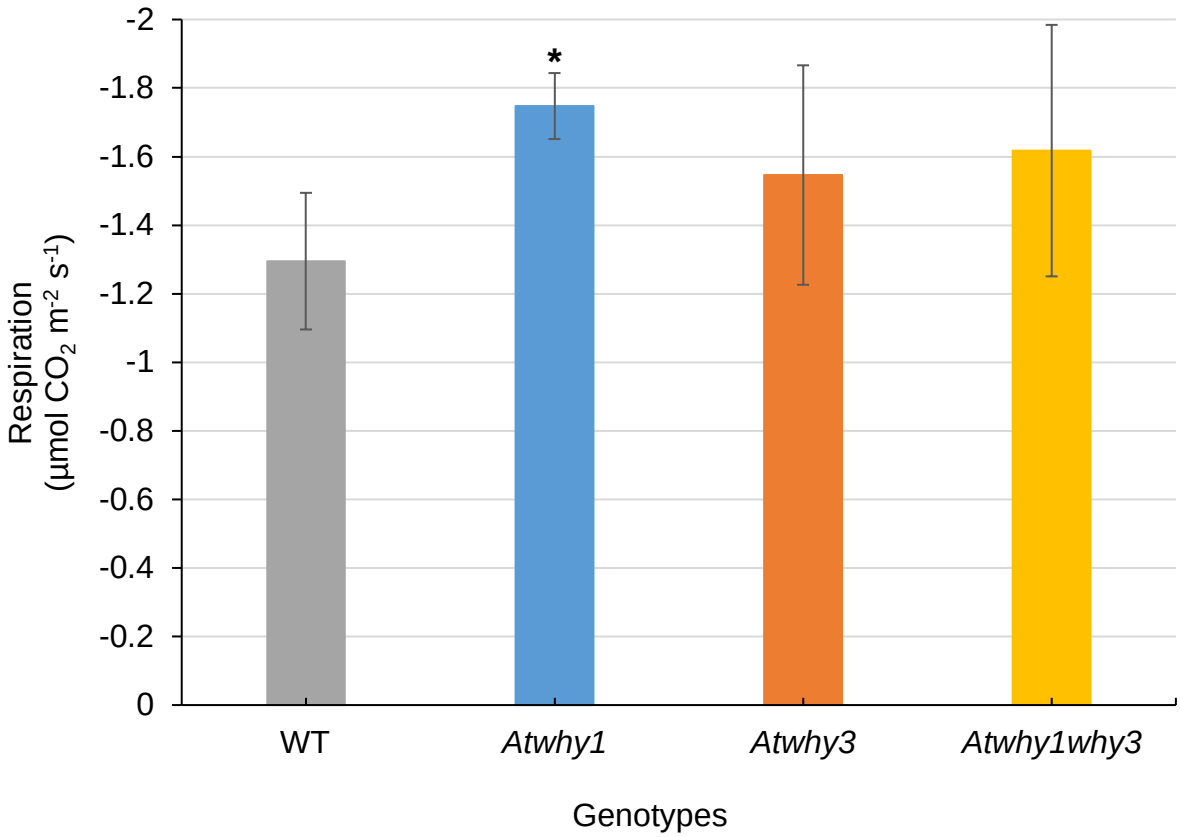


Figure 50: Respiration of plants grown at 8-hour photoperiod at 20 °C. Means were calculated from four biological replicates. Plants were four weeks old when measured. Error bars were standard error of means and students t tests were conducted to compare each mutant to the WT at each time point: *, p <0.05

All of the lines followed the same photosynthetic curve as light intensity increased up to 600 PAR [Figure 49]. These trendlines were created based on net photosynthetic light-response curves (Lobo *et al*. 2013) and show that leading up to saturating light intensity (900 PAR) the *Atwhy1why3* plants had the highest photosynthetic rate, however above saturating light intensity the photosynthetic rate reduced lower than the WT and *Atwhy1* lines [Figure 49]. Contrastingly, the *Atwhy1* lines also increased after WT saturating light intensity. Whereas the *Atwhy3* mutant maintained a photosynthetic rate between 6 and 7 µmol $CO_2 m^{-2} s^{-1}$. Generally, CO_2 assimilation rates were increased in plants lacking WHY1 but were decreased in single *Atwhy3* mutants. None of these values were statistically significantly different from the WT. However, all mutants had higher respiration rates, taken in the dark, compared to the WT [Figure 50]. The single *Atwhy1* mutants had a statistically significantly higher respiration rate compared to WT [Figure 50]. As the only statistically significant difference was observed in dark respiration rates, the stomatal conductance [Figure 49] and non-photochemical quenching (NPQ) [Figure 52] were also compared.

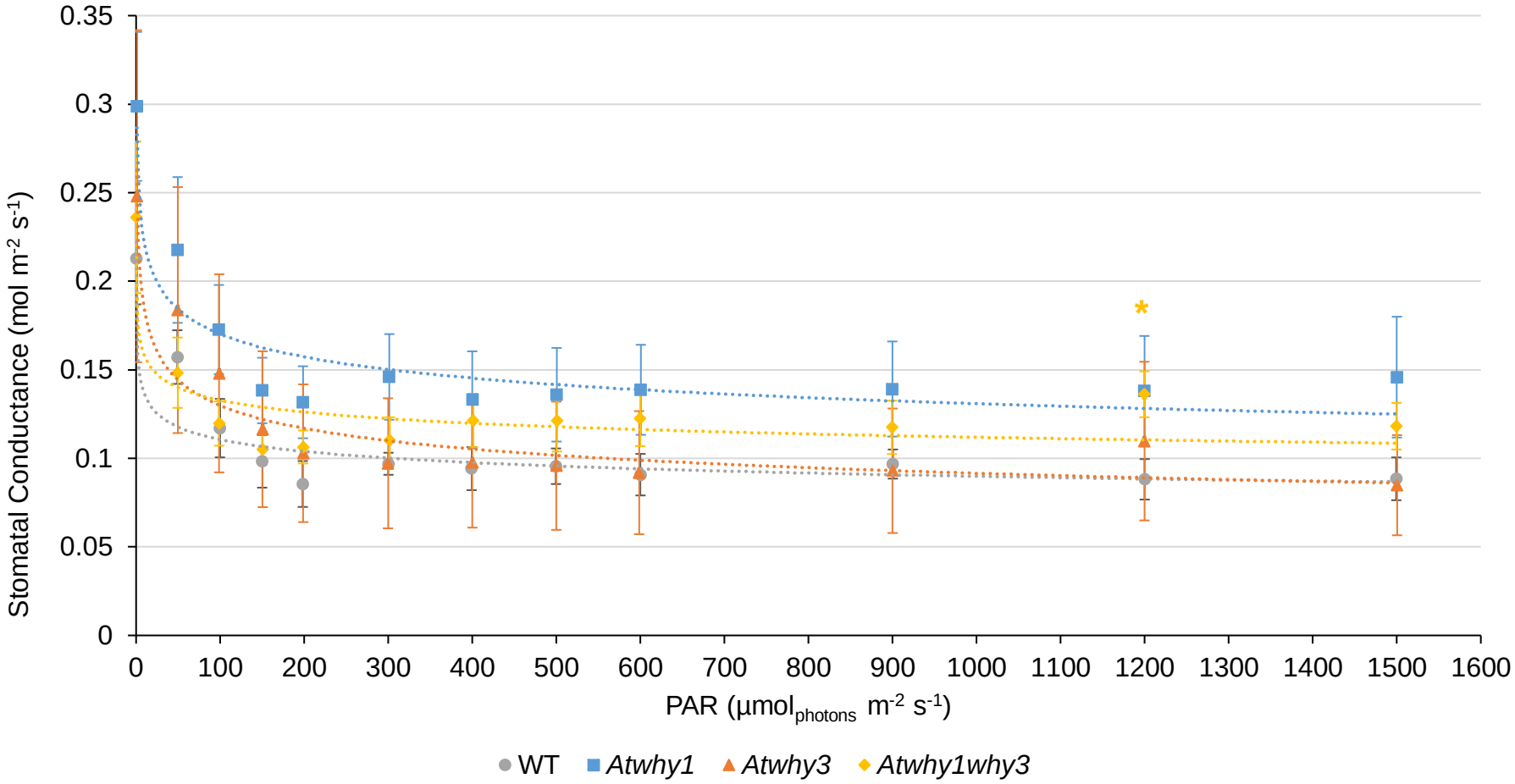


Figure 51: Stomatal conductance of *Atwhy* mutants and WT plants at increasing levels of photosynthetically active radiation (PAR). Means were calculated from four biological replicates. Plants were four weeks old when measured. Error bars were standard error of means and students t tests were conducted to compare each mutant to the WT: *, $p < 0.05$. Trendline for each genotype shows the power of the values based on information from Lobo *et al*. (2013): WT ($y = 0.1674x^{-0.09}$), *Atwhy1* ($y = 0.2881x^{-0.114}$), *Atwhy3* ($y = 0.2613x^{-0.152}$), and *Atwhy1why3* ($y = 0.1873x^{-0.075}$).

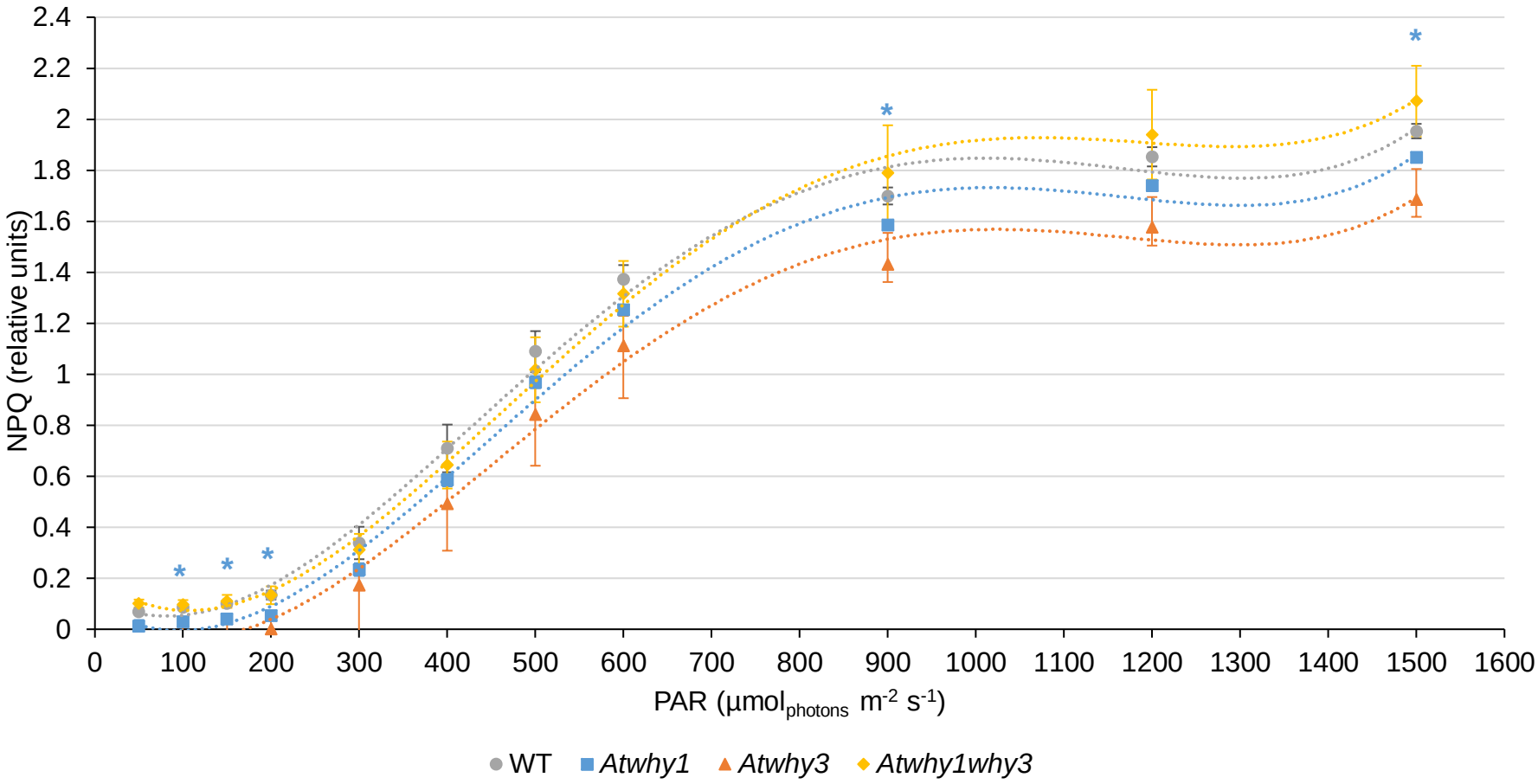


Figure 52: Non-photochemical quenching (NPQ) of *Atwhy* mutants and WT plants at increasing levels of photosynthetically active radiation (PAR). Means were calculated from four biological replicates. Plants were four weeks old when measured. Error bars were standard error of means and students t tests were conducted to compare each mutant to the WT at each PAR level: *, $p < 0.05$. Trendline for each genotype shows the polynomial trend to the order 4 based on information from Lobo *et al*. (2013): WT ($y = 4E\text{-}12x^4 - 1E\text{-}08x^3 + 1E\text{-}05x^2 - 0.0019x + 0.1236$), *Atwhy1* ($y = 4E\text{-}12x^4 - 1E\text{-}08x^3 + 1E\text{-}05x^2 - 0.0022x + 0.0956$), *Atwhy3* ($y = 4E\text{-}12x^4 - 1E\text{-}08x^3 + 1E\text{-}05x^2 - 0.0022x + 0.0666$), *Atwhy1why3* ($y = 4E\text{-}12x^4 - 1E\text{-}08x^3 + 1E\text{-}05x^2 - 0.0026x + 0.2047$).

All *Atwhy* mutant lines followed the same pattern of stomatal conductance [Figure 51] and NPQ levels [Figure 52] as the WT. Stomatal conductance at saturating light intensity was increased in plants lacking WHY1 but no different in single *Atwhy3* mutants compared to the WT plants [Figure 51]. The only significant difference in the data was found at 1200 PAR where *Atwhy1why3* had significantly more stomatal conductance compared to the WT [Figure 51]. The NPQ level was increased in *Atwhy1why3* mutants but decreased in both single mutants compared to WT [Figure 52]. Interestingly, *Atwhy1* was repeatedly significantly lower in NPQ compared to WT, from 100 to 200 PAR, again at 900 PAR and then again at 1500 PAR [Figure 52]. Both *Atwhy3* and *Atwhy1why3* are more on the peripheral of the WT values, but these mutants have a lot of variation in NPQ levels. Taken together there is not a clear impact for WHY on photosynthetic effiency in changing light, and thus no consistent insight into stress on *Atwhy* plants. In order to understand whether *Atwhy* plants are under stress during growth temperature must also be considered. The LiCOR was used to maintain leaf temperature during the photosynthesis measurements to identify if temperature change or heat stress had an impact on WHIRLY function in photosynthesis and therefore the growth of *Atwhy* mutant plants.

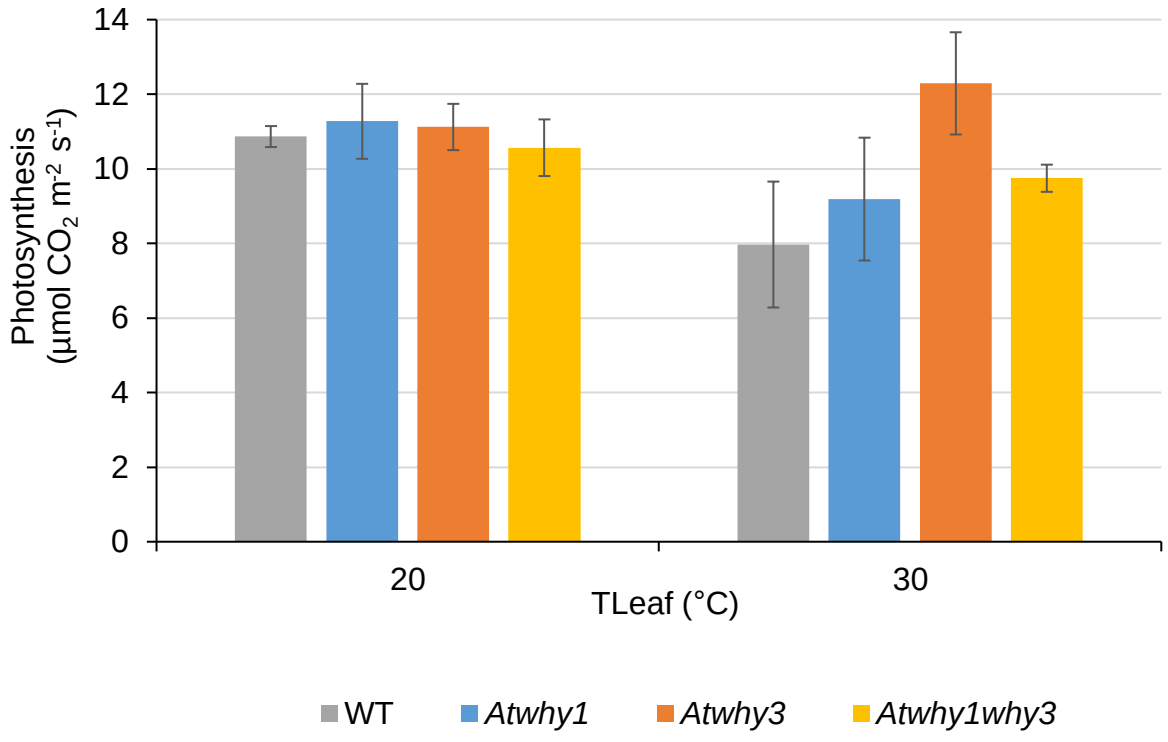


Figure 53: Comparison of photosynthesis rates at standard growth temperature (20 °C) and high temperature (30 °C), measured by the temperature of the leaf (TLeaf) of the three *Atwhy* mutants compared to WT at 900 PAR. Plants were four weeks old when measured. Means were calculated from four biological replicates. Error bars were standard error of means.

At standard growth temperature (20 °C), photosynthetic rates were similar in all three of the *Atwhy* mutants compared to the WT plants, however, when the temperature of the leaf was raised to 30 °C some changes in photosynthesis began to become visible. The WT CO_2 assimilation rates were constant at around the 10 µmol CO_2 m^{-2} s^{-1} mark at 20 °C with very little variation, however this became more variable after the leaf temperature was increased and the photosynthetic ability of the WT was reduced by around 2 µmol CO_2 m^{-2} s^{-1} and with a wider range of variation [Figure 53]. The *Atwhy1* and *Atwhy1why3* mutants had similar photosynthesis levels to that of the standard leaf temperature but there was more variation in the plants measured at 30 °C compared to 20 °C. [Figure 53]. There was no significant difference in the photosynthetic results of any of the *Atwhy* mutants and WT between 20 °C and 30 °C, however there was a relative increase in *Atwhy3* photosynthesis at 30 °C compared to 20 °C [Figure 53].

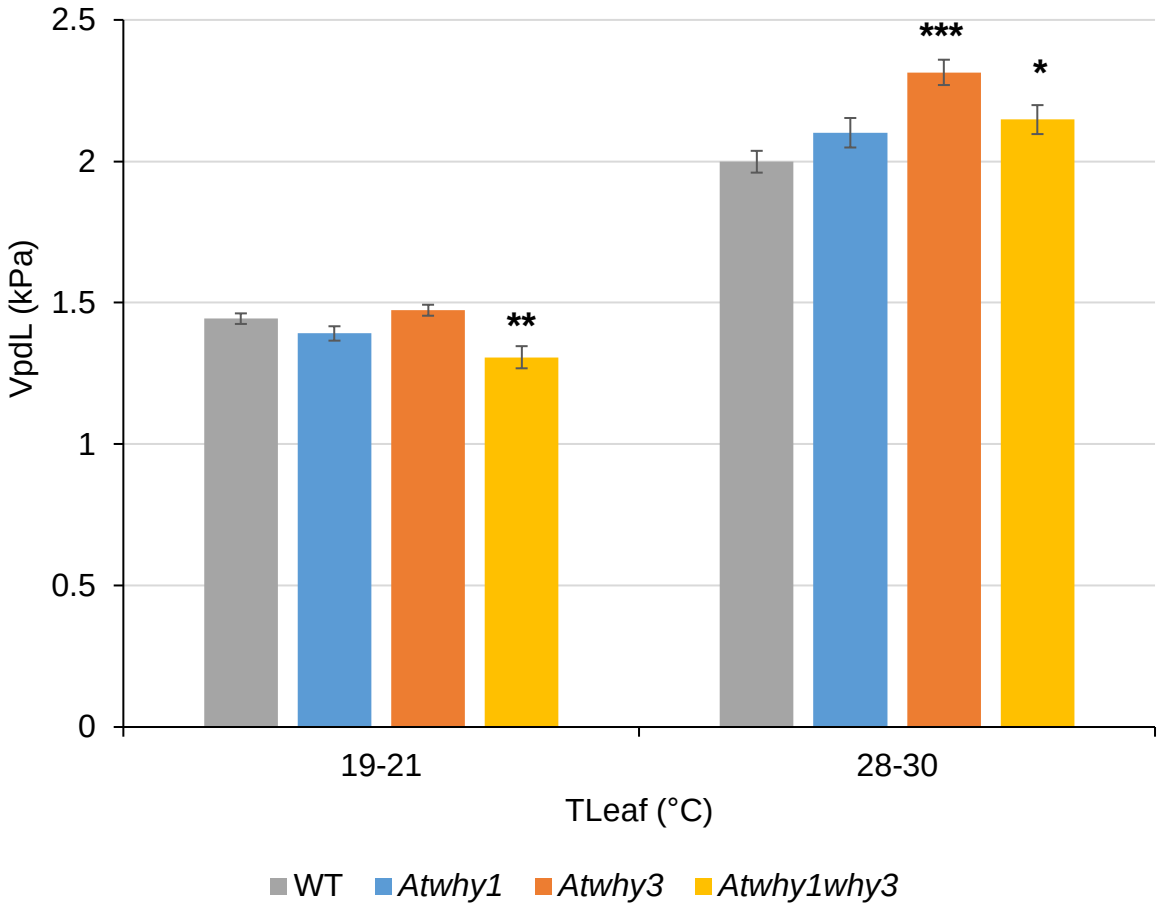


Figure 54: A comparison of leaf vapour pressure deficits (VpdL) of the three *Atwhy* mutants compared to WT under standard (19-21 °C) and high (28-30 °C) leaf temperatures at 900 PAR. Plants were four weeks old when measured. Means were calculated from four biological replicates. Error bars were standard error of means and students t tests were conducted to compare each mutant to the WT at each time point: *, $p < 0.05$, **, $p < 0.01$, ***, $p < 0.001$

In order to explore the cause of this shift further, the plant response to water vapour was explored as this gives an idea of stomatal conducatance efficiency [Figure 54]. The vapour pressure deficit response showed the plant response to changing water vapour at different leaf temperatures. The data shows that the WT plants maintained VpdL even at very high temperatures, however all of the mutants had a much lower VpdL at lower temperatures compared to WT and a much higher VpdL than WT at higher temperatures [Figure 54]. Furthermore, the VpdL of *Atwhy1why3* was significantly lower than the WT at 19-21 °C [Figure 54] and significantly higher at 28-30 °C where *Atwhy3* plants also had a significantly higher VpdL compared to the WT [Figure 54]. These data show that plants lacking both WHY 1 and WHY3 had shifted photosynthetic rates caused by temperature increase.

6.3 DISCUSSION

Previous studies of *Atwhy1*, *Atwhy3* and *Atwhy1why3* mutants provided little consistent insight into the phenotypic effects of the lack of the functional *WHY* genes on vegetative and reproductive growth. The single knock-out *Atwhy1* and *Atwhy3* mutants were found to be phenotypically similar to that of the WT. However, 4.6 % of *Atwhy1why3* double mutants were reported to have variegated sectors (Maréchal *et al.*, 2009). It has also been reported that the *Atwhy1* and *Atwhy1why3* lines had lower photosynthetic electron transport rates (ETR) under high light (800 μmol m^{-2} s^{-1} irradiance) than the WT (Huang *et al* 2017). The lower ETR could explain the greater susceptibility of these mutants to high light. To provide detailed phenotypes of the *Atwhy1*, *Atwhy3* and *Atwhy1why3* mutants, plants were vigorously re-evaluated under more specific conditions than previous studies. Plants were grown under conditions of 20 ºC/ 16 ºC day/ night temperature regime, with a 16h photoperiod (250 μmol m^{-2} s^{-1} irradiance) and 60% relative humidity. Previous results did not publish specific growth conditions and this is where the wide variety in published phenotypes, including variagtion in the *Atwhy1why3* mutants may have originated (Maréchal *et al.*, 2009). The rosette phenotypes of all the mutant lines were very similar to the wild type throughout vegetative growth [Figure 36, Figure 37]. While some small differences were observed at some individual time points, there were no consistent significant differences between the mutants and WT during the earlier vegetative growth stages. However, in later development both single and double mutants lacking the WHY3 protein had significantly smaller leaf area than the WT [Figure 38]. Moreover, senescence, judged by the yellowing of the leaves, became visible by week 5 in all the mutant lines, whereas the WT remained green. These changes were most prominent in the leaves of *Atwhy3* and *Atwhy1why3* mutants at week 6 [Figure 40]. Unlike a previous report (Maréchal *et al.*, 2009), no variegation was seen on the leaves of the *Atwhy1why3* mutants under the growth conditions used in these experiments [Figure 40]. However, symptoms of senescence [Figure 40] and loss of chlorophyll pigment [Figure 41] were evident in the *Atwhy1why3* mutants before the other lines.

Chlorophyll levels were lower in the leaves of all genotypes at week 6 [Figure *41*] suggesting the onset of leaf senescence. The leaves of the *Atwhy1* and *Atwhy1why3* mutants had significantly less chlorophyll than the WT at this stage [Figure *41*] in line with the visual appearance of senescence [Figure 40]. There was a large reduction in the chlorophyll a/b ratios in all of the mutants relative to the WT [Figure 42]. Furthermore, all the *Atwhy* mutants had ratios decreased by 6-fold from week 3 to week 6, with all final ratios below 1, whilst the WT

ratios were only halved from week 3 to week and was above 1. The concentration of chlorophyll a is 2-3 times higher than the concentration of chlorophyll b in healthy plants and it is normal for this ratio to decrease from around 3.4 in early development to around 1.5 in later development (Nath *et al.*, 2013).

Biomass accumulation was similar in the *Atwhy1*, *Atwhy3* and *Atwhy1why3* mutants to the wild type in terms both of fresh weight [Figure 43] and dry weight [Figure 44]. However, multiple measurements are required to determine plant growth phenotypes because measurements of leaf area or biomass can have a non-linear relationship during vegetative growth (Weraduwage *et al.*, 2015). In contrast to the *Atwhy1* or *Atwhy1why3* mutants which flowered at the same time as the wildtype, a significant delay in bolting was observed in the *Atwhy3* mutants, which had not produced any flowering stems until week 6 [Figure 45, Figure 46]. There were significantly fewer siliques per plant but more seeds per silique in the *Atwhy3* and *Atwhy1why3* mutants [Figure 47]. The overall seed yield was similar in all genotypes [Figure 48]. These findings would suggest that WHY1 and WHY3 fulfil roles in both in leaf development and leaf senescence which may have an impact on other developmental changes such as bolting timing and silique development.

Significant differences between the *Atwhy1why3* mutants and the other lines were observed in a number of the measured parameters. While effects seem to be minimal during early rosette development, they become more apparent as the plants move from vegetative to reproductive development. However, it is difficult to determine what is causing the phenotypes, perhaps an early developmental change such as rosette diameter, number of leaves and chlorophyll a to b ratios have an added compound effect on timings of other developmental changes on the plant.

The photosynthetic parameters measured in this study revealed very few differences between the lines. There were no significant differences in carbon dioxide assimilation rates or stomatal conductance values [Figure 51]. However, higher respiration rates were observed in the *Atwhy1* mutants compared to WT [Figure 50]. The *Atwhy1* plants had significantly lower levels of thermal energy dissipation than the WT at a range of low and high light intensities [Figure 52]. Additionally, poor induction of NPQ can indicate lower resistance to abiotic stresses such as salt, heat, drought and chilling (Zhao *et al.*, 2017). However, the data presented here show that photosynthesis of the *Atwhy1*, *Atwhy3* and *Atwhy1why3* mutants was no more sensitive to heat stress than the wild type [Figure 53]. However, the plants lacking WHY3 had significantly higher leaf vapour pressure deficit (VpdL) values than the WT following exposure to higher

temperatures [Figure 54]. This observation could suggest that plant physiology is more affected in these mutants as high VpdL could cause plants to close their stomata to minimise water loss (Grossiord *et al*., 2020). The VpdL of the *Atwhy1why3* leaves was significantly lower than the WT when measured at 20 °C at 1.31 kPa but significantly higher at 30 °C, and at 2.51 kPa it was almost twice the pressure than at optimum growth temperatures [Figure 54]. These findings may indicate that that WHY3 and its interactions with WHY1 could be important in heat stress responses. Furthermore, data presented in the Germination and RNA-seq Chapters provided evidence of interactions between WHY1, WHY3 and heat shock proteins in seeds and germination. Previous studies with 4-week-old tomato plants showed that the expression of the SlWHY1 protein was increased after 12 hours of heat treatment (Zhuang *et al.*, 2020a). While the responses to heat shock may vary during vegetative development in different species, no consistent changes in the responses of photosynthesis to heat shock response were observed in 4-week-old plants in the present study.

The data presented in this chapter have highlighted the modest roles that WHY1 and WHY3 play in the growth and development of Arabidopsis plants. The *Atwhy1why3* double mutants show the strongest growth and senescence phenotypes, suggesting that the WHY1 and WHY3 proteins have overlapping or redundant functions during plant development.

7 WHIRLY FUNCTIONS IN WHEAT

7.1 INTRODUCTION

Wheat is the one of most widely grown staple crops in the world, and is grown on 135 million farms yearly (Erenstein, 2021). This number is predicted to decrease to 130 million farms by 2030, partly due to an upsurge in pests and diseases because of global temperature increases and unpredictable weather patterns (Erenstein, 2021). Recent technological developments have overcome many of the barriers to progress in wheat genomics research caused by its highly repetitive and large polyploid genome. The sequencing of the *Triticum aestivum* (common bread wheat) genome has enabled the identification of gene families that could help improve the yields of wheat crops under stress (IWGSC, Appels *et al.*, 2018; Walkowiak *et al.*, 2020).

The WHIRLY (WHY) family of proteins have been implicated in disease resistance pathways in many crop species and thus they could be suitable candidate genes for improving disease resistance in wheat. Gene re-annotation studies suggested that a WHY transcription factor could be involved in disease defence in the wheat wild relative, *Dasypyrum villosum*. The role of this gene in disease resistance was studied in relation to the powdery mildew infection caused by *Blumeria graminis f. sp. tritici* (Bgt) in wheat (He *et al*., 2016). The levels of *DvWHY* transcripts were enhanced in both the resistant Yangmai18 and susceptible Yangmai9 varieties, which had been inoculated by Bgt YZ01 isolates, reaching a peak at 12 hours post infection. Furthermore, barley stripe mosaic virus-induced gene silencing of *DvWHY* lead to enhanced development of *Bgt* hyphae and conidiophores. He *et al*. (2016) showed that DvWHY conferred resistance to powdery mildew, a process mediated by the *Pm21* resistance gene. However, it is potentially difficult to generate knock-out, knock-down or overexpression lines that have a significantly different expression of WHY from wild-type in *T. aestivum* because of the presence of multiple copies of WHY genes localised on different chromosomes due to hexaploidy nature of this wheat. Hexaploid wheat is allopolyploid and the genome consists of two copies of the consists of the three genomes A, B and D in the form AABBDD which evolved from the spontaneous hybridization event between tetraploid, *Triticum turgidum* L (AABB), with the diploid, *Aegilops tauschii* (DD), as such the genome is large (16 Gb) and around 85% of the genome has repetitive elements (Travella *et al*., 2006; Walkowiak *et al*., 2020). As such, many genes have three of more copies and there is a high degree of functional redundancy in them and as suvh it is difficult to do reverse genetics experiments in bread wheat (Mochida *et al.*, 2003).

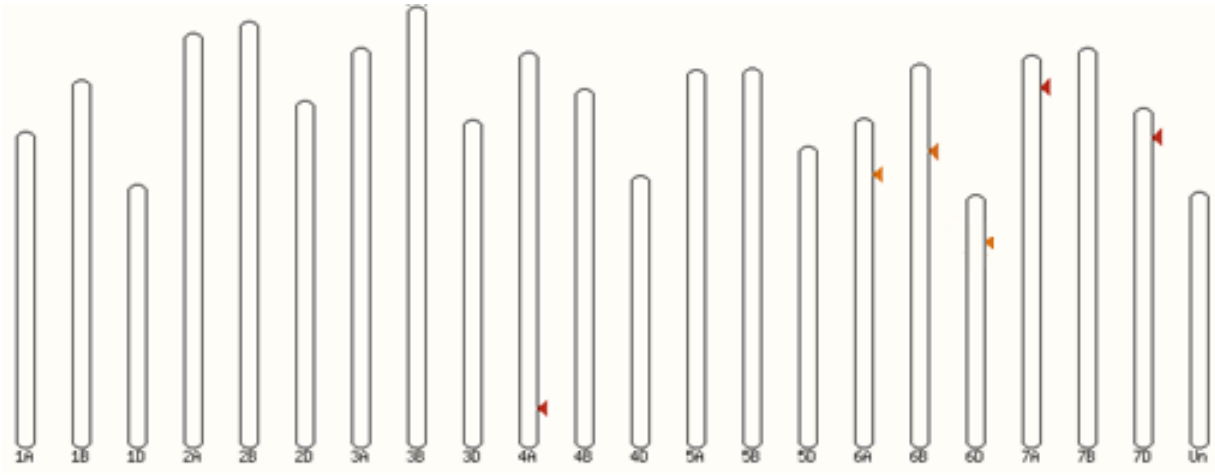


Figure 55: High-scoring segment pair distribution on *Triticum aestivum* genome from ensemblPlants (https://plants.ensembl.org/). The sequences with the highest sequence similarity to the input are in red (4A, 7A, 7D) and the sequences with lower sequence similarity to the input are in orange (6A, 6B, 6D).

WHIRLY proteins were first identified in *T. aestivum* by protein sequence alignments of seven p24 homologs with the conserved KGKAAL domain (Desveaux *et al.*, 2002). The same group later identified two WHY family members (WHY1 and WHY2 homologues: Desveaux *et al.*, 2005). Krause *et al.*, (2009) showed that TaWhy1 has a 48.5 % sequence similarity to StWhy1, whilst TaWhy2 has a 43.3 % sequence similarity to StWhy2. As a monocotyledonous hexaploid species, it is likely that there are multiple copies of the two WHY genes normally found in diploid species. Romeuf *et al.* (2010) established that wheat had more than three-times the number of WHY genes than would was predicted in comparisons to *Oryza sativa* (rice). A TBLASTN search for WHY1 was performed with the ZmWHY1 protein sequence against the *T. aestivum* genome. This study suggested that there were multiple copies of *WHY1* in wheat, which are in located on chromosomes 4A (TraesCS4A02G395700), 6A (TraesCS6A02G131700), 6B (TraesCS6B02G160100), 6D (TraesCS6D02G121500), 7A (TraesCS7A02G095300) and 7D (TraesCS7D02G091500) [Figure 55].

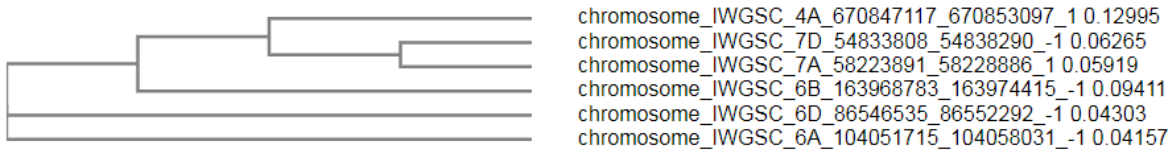


Figure 56: Phylogenetic tree of a multiple sequence alignments of *Triticum aestivum* Whirly sequences using Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/)

All the genes had the conserved KGKAAL domain when translated to proteins. It is likely that some of these genes may encode the mitochondria localised WHY2. To differentiate between

the putative *Why1* and *Why2* genes, a BLASTN analysis of the *T. aestivum* genome was performed using ZmWhy2. This study identified TraesCS6A02G131700 (Chromosome 6A) and TraesCS6B02G160100 (Chromosome 6B) and TraesCS6D02G121500 (Chromosome 6D) as likely WHY2 homologues. Thus, it is likely that the WHY1 homologues are TraesCS4A02G395700 (chromosome 4A), TraesCS7A02G095300 (chromosome 7A) and TraesCS7D02G091500 (chromosome 7D). Based on synteny, where there is conservation of blocks of order within chromosomes (Myers, 2008), it would be predicted that *WHY1* homologues occur on chromosome 7B that match those on chromosome 6. Therefore, it appears that a *WHY* gene is missing from Chromosome 7B and that there is a spurious *WHY* gene on chromosome 4A. However, there was a one-way translocation from Ta7B to Ta4A during evolution which could involve this *WHY* gene (Ling *et al.*, 2018). Multiple sequence alignments of these gDNA sequences showed that the gene on chromosome 4A is the least similar to the others and that the genes on chromosome 6 are near-identical. Perhaps the gene on chromosome 4A has a similar function to *ATWHY3* as this is the most recent divergent gene in the species [Figure 56]. However, the presence of a *WHY3* gene in wheat has not been confirmed. Additionally, there are no published studies on the functions of the WHY1 protein in wheat development. These aspects should be characterised further in order to better understand how modification of *WHY* expression may benefit wheat crops.

There have been numerous studies on WHY protein functions in *Hordeum vulgare* (barley) and *Zea mays* (maize), which are also in the Poaceae family along with *Triticum aestivum*. The *Zmwhy1-1* mutants exhibited an ivory phenotype, and a pale green phenotype was observed in *Zmwhy1-2* mutants. The *Zmwhy1-1* mutants, which had a transposon insertion 35-bp downstream of the predicted start codon, had higher levels of ZmWHY1 on an immunoblot analysis compared to *Zmwhy1-2*, which had an insertion 38-bp upstream of the predicted start codon (Prikryl *et al.* 2008). However, the authors did not account for photosynthetic enzyme complexes in this immunoblot which may have over-represented the total protein in the leaf extracts. The pale green and albino phenotypes were linked to impaired chloroplast development in the mutants, which had severely reduced levels of ribosomal RNA (rRNA) and were thus unable to support photosynthesis (Prikryl *et al*. 2008). A similar explanation was given for the reported variegation phenotype in the *Atwhy1why3* mutants (Maréchal *et al.* 2009). Furthermore, *Hvwhy1* knockdown barley leaves had a similar phenotype to the wild type, as had the *Atwhy1* and *Atwhy3* mutants (Melonek *et al*., 2010; Maréchal *et al*. 2009). The leaves of the *Hvwhy1* knockdown plants were slower to green and establish photosynthesis.

However, once green, the *Hvwhy1* leaves had significantly more chlorophyll and less sucrose than the wild type, with similar redox metabolite levels and comparable rates of photosynthesis (Grabowski *et al*., 2008; Comadira *et al.*, 2015). The *Hvwhy1* lines had similar senescence phenotypes to the WT at low light intensity and had similar patterns of dark induced senescence (Kucharewicz *et al*., 2017). However, high light intensity-induced senescence was delayed in the *Hvwhy1* lines compared to the wild type.

Partial redundancy between AtWHY1 and AtWHY3 may explain the differences in phenotypes observed in Arabidopsis and other species. However, the ivory phenotype has only been reported in ZmWHY1 single knock-out mutants, in which WHY3 does not exist (Prikryl *et al.*, 2008). It is not clear why such a range of phenotypes has been reported in different species or why such a marked difference in phenotypes exists between the maize and barley mutants because they are both monocotyledonous plants. However, maize is a C4 species while barley, like Arabidopsis, is a C3 species. Perhaps, WHY proteins operate rather differently in C4 leaves than their C3 counterparts. Wheat also operates C3 photosynthesis and thus WHY protein functions may result in a different phenotype than is observed in barley.

The leaves *Hvwhy1* knockdown lines had a much greater abundance of plastome-encoded transcripts than the WT (Comadira *et al*., 2015). These encode subunits of the the thylakoid NADPH complex, the chloroplast RNA polymerase, the cytochrome b6f complex and chloroplast ribosomes (Comadira *et al.*, 2015). These data show that WHY1 functions are important for the operation of photosynthesis. However, the regulation of the intracellular localisation of the WHY1 proteins may be different in in barley and maize and this may have a profound impact on the phenotype. WHY1 functions have not been previously characterised in wheat. Hence, the data reported in this chapter are novel and may have significant agricultural and economic implications for the wheat industry

7.1.1 Objectives

The aim of the studies reported in this chapter was to characterise the functions of the WHY1 protein in *Triticum aestivum* (wheat) in the Fielder background. As discussed above, current evidence suggests that there are multiple copies of the *WHY1* gene in *T. aestivum*. However, the RNAi knockdown approach used by Biogemma should achieve knockdown of one or more forms of the WHY1 protein (detailed in Materials and Methods). All wheat plants were grown at the glasshouses in Limagrain, France during the iCASE placement for the project.

The objectives of the studies reported in this chapter are three-fold:

1. To confirm transformations of TaWHY knockdown plants and select lines to carry forward.
2. To characterise the early and late phenotypic growth changes of TaWHY knockdown plants compared to null segregants and WT.
3. To determine the effects of TaWHY deficiency on plant fertility and seed yield.

7.2 RESULTS

7.2.1 Genotyping of RNA interference (RNAi) knockdown lines

A series of RNA interference (RNAi) knockdown lines were produced by Biogemma (see 3.1.2). Two transformation experiments (T10715 and T10716) were performed on soft white spring wheat (*Triticum aestivum* cv. Fielder) to produce RNAi knockdown lines of the TaWHY1 protein. Five RNAi lines were selected from the T10715 transformation experiment [T15012 (512), T15022 (522), T15031 (531), T15042 (542), T15051 (551)] and six RNAi lines [T16011 (611), T16022 (622), T16023 (623), T16031 (631), T16042 (642), T16051 (651), and T16062 (662)] were selected from the T10716 transformation experiment. Null segregants were selected for each transformation event as internal controls for genotyping and phenotyping assays. Only one null segregant (NS) line per transformation experiment [T15054 NS (554 NS) and T16023 NS (623 NS)] was used in the following experiments

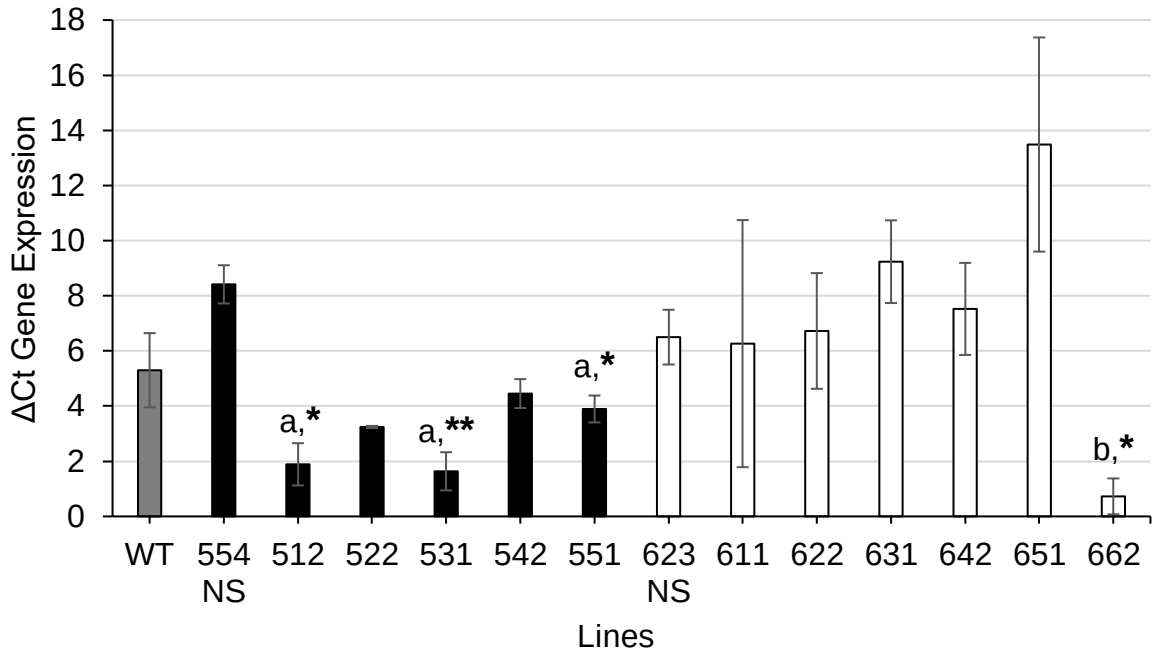


Figure 57: Relative expression levels in TaWHY1 RNAi knock-down plants from the two transformation experiments: T10715 (black) and T10716 (white) and their relative null segregant (NS) lines and the Fielder wild type (WT, grey). WHIRLY transcripts are expressed relative to the ACTIN housekeeping gene. Each line has is an average of two technical replicates and three biological replicates, geometric mean was given, and error bars were standard error of the mean. Students t test was calculated and compared each line to the NS of the same transformation, 554 NS (a) and 623 NS (b) respectively: *, $p<0.05$; **, $p<0.01$; t tests using the WT gave $p > 0.05$.

The ΔCt method was used to calculate the relative fold gene expression (Livak and Schmittgen, 2001). The expression of the actin housekeeping gene was compared to Tawhy1 expression relative to cycle threshold [Figure 57]. The geometric mean was used for normalisation (Vandesompele *et al.*, 2002). There were no significant differences between the WT and the NS lines. The NS lines were therefore used for the statistical analysis of ΔCT gene expression for each of the respective transformation experiments. There were variations in relative TaWHY1 expression between both T10715 and T17016 transformation experiments [Figure 57]. The T10715 lines showed reduced TaWHY1 expression compared to the null segregant, however only three lines had significantly lower ΔCt values compared to the 554 NS (8.42 ± 0.69). These are 512 (1.89 ± 0.77, p = 0.026), 531 (1.64 ± 0.69, p = 0.0097) and 551 (3.90 ± 0.49, p = 0.042) [Figure 57]. Interestingly, most of the T10716 lines appeared to have much higher TaWHY1 expression compared to the 623 NS (6.50 ± 0.99) apart from 662 (0.73 ± 0.65, p = 0.013), which had significantly reduced expression compared to the NS [Figure 57]. The data show that Tawhy1 transcripts are lower in RNAi lines 512, 522, 531, 542, 551 and in 662. Therefore, the four lines (512, 531, 551 and 662) with significantly lower levels of Tawhy1 transcripts were carried forward for further analysis experiments. These lines will henceforward be referred to as *Tawhy* lines, together with their relative NS lines (554 NS and 623 NS) and the Fielder WT.

7.2.2 Effects of loss of WHY1 functions on early plant development

The growth of the *Tawhy* RNAi lines was monitored throughout vegetative and reproductive development to monitor any phenotypic differences between the knockdown lines, and the NS controls. The morphological measurements included a germination study [Figure 58], and an analysis of early growth parameters including height [Figure 59], number of leaves [Figure 60] and colouration of leaves [Figure 61] . These measurements were repeated at later development stages [Figure 62, Figure 63 and Figure 64].

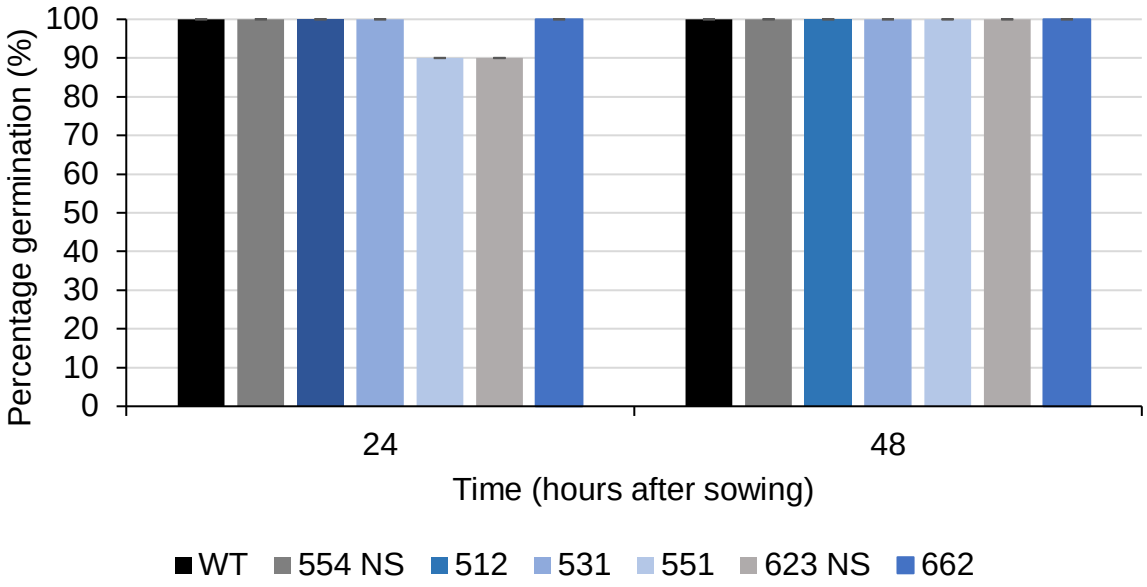


Figure 58: Percentage germination of RNAi down-regulated *Tawhy* wheat lines compared to relative NS, values are a mean of ten biological replicates split across two separate 0.5% MS plates, error bars are standard error of means.

All of the *Tawhy* seeds showed similar germination to the WT [Figure 58]. Additionally, there was little variation between the lines. These data suggest that there was no effect of the kock-down on TaWHY1 on the germination of wheat seeds under standard conditions.

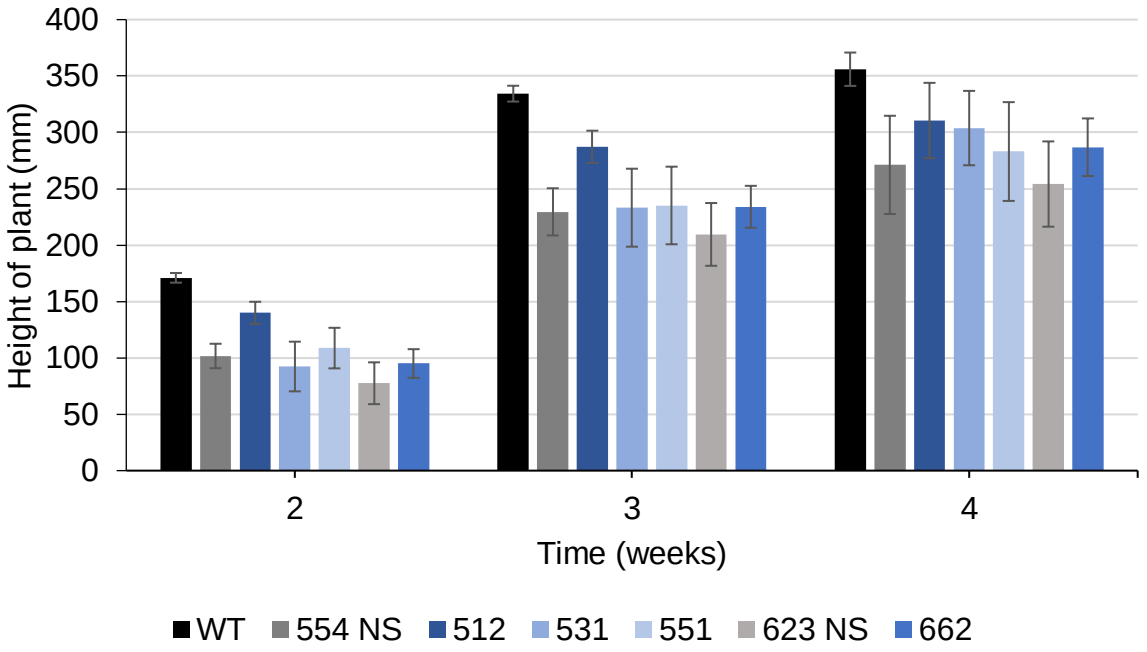


Figure 59: The length of the tallest tiller (from soil to tip) in the *Tawhy* lines compared to WT and NS controls in early development. Values are a mean of six biological replicates. Error bars are calculated from standard error of means of 6 plants per line and t tests conducted with the corresponding NS gave $p > 0.05$

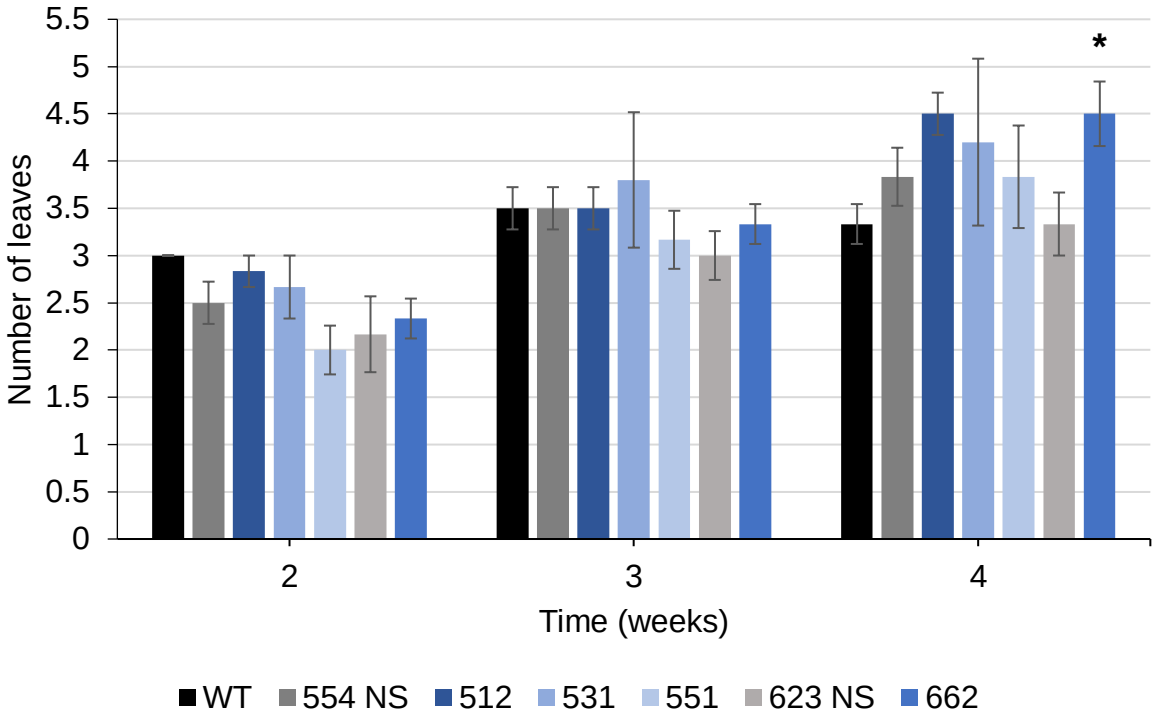


Figure 60: The relative number of leaves on the *Tawhy* plants compared to WT and NS controls. Values are a mean of six biological replicates. Error bars are calculated from standard error of means of 6 plants per line and t tests were conducted against the corresponding NS: *, $p < 0.05$.

The tillers were significantly shorter in the transgenic lines and the NS lines than the wild type in early development (2 and 3 weeks), however when compared to the corresponding NS there were no significant differences in the height of the plant at any time point [Figure 59]. Therefore, RNAi-mediated knockdown of WHY does not affect the height of the tillers in early development. The only statistically significant difference in leaf number was observed in 662 at 4 weeks (4.5 ± 0.21 leaves, $p = 0.035$) compared to 623 NS (2.17 ± 0.40) [Figure 60]. However, as there was no consistent effect on development in height or number of leaves these data suggest that that the knockdown in WHY1 function did not have an effect on leaf development.

Chlorophyll pigment analysis was carried out using a Soil Plant Analysis Development (SPAD) 502 plus chlorophyll meter. The SPAD value given is the difference between the transmittance of red (650 nm) and infrared (940 nm) light through the leaf (Uddling *et al.*, 2007). Five measurements were taken on the flag leaf of each plant.

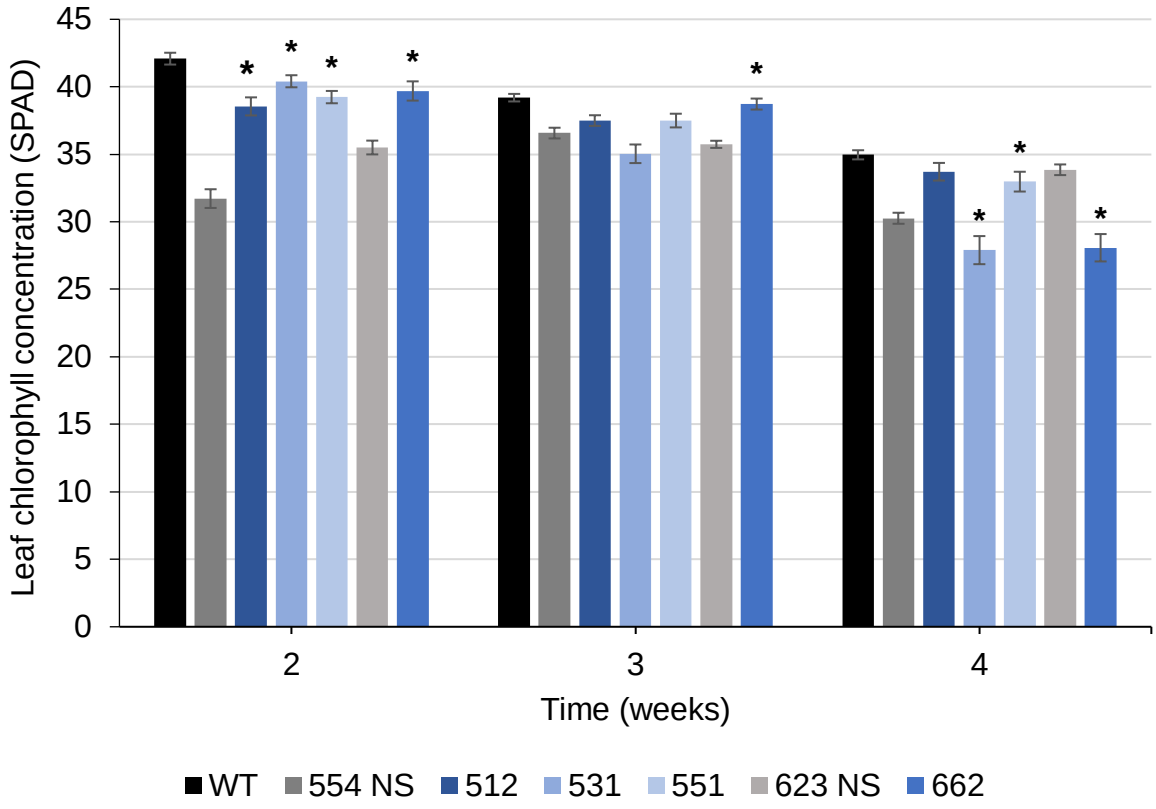


Figure 61: Relative SPAD measurements in *Tawhy*, WT and NS leaves. Values are the means of six biological replicates. Error bars are calculated from standard error of means of 6 plants per line and ANOVA t tests were conducted against the corresponding NS line: *, $p < 0.05$

The SPAD measurements decreased in the flag leaves of all lines with time. Variations in the chlorophyll concentrations were observed between lines and their corresponding control: 554 NS and 623 NS [Figure 61]. However, there were no consistent differences in the levels of chlorophyll between the TaWHY1 lines and controls. These data suggest that loss of WHY1 functions did not consistently affect photosynthesis in the flag leaves during early development [Figure 57]

7.2.3 Effects of loss of WHY1 functions on later plant development

The growth of the *Tawhy* lines and controls was monitored from 4 weeks until anthesis to determine later developmental effects on plants. Plant height was measured in all the lines from five to ten weeks post germination. This parameter increased in all of the lines with time.

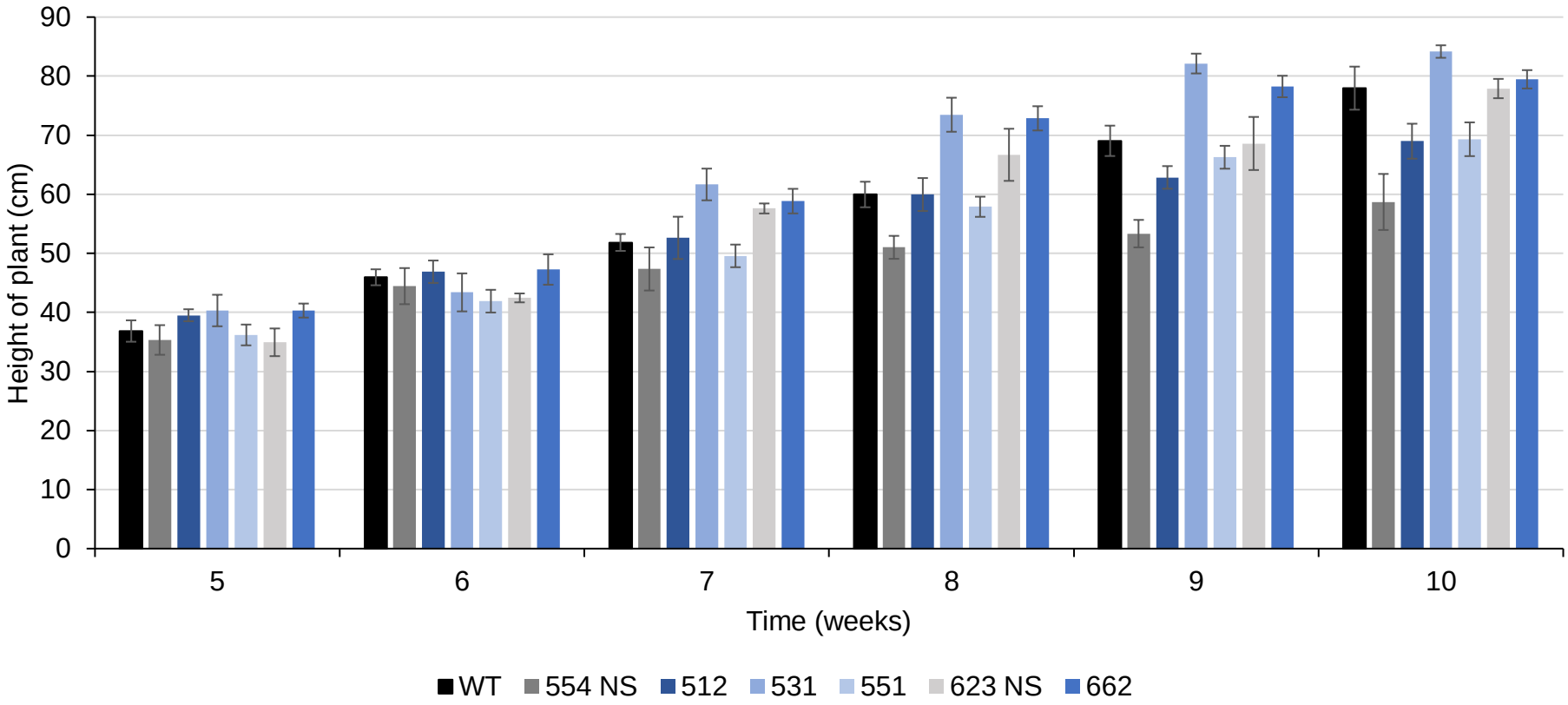


Figure 62: The relative heights of the plants from top of the tallest tiller to the base of the stem of the *Tawhy*, WT and NS lines in late development. Values are a mean of six biological replicates. Error bars are calculated from standard error of means of 6 plants per line and and t tests conducted with the corresponding NS gave $p > 0.05$.

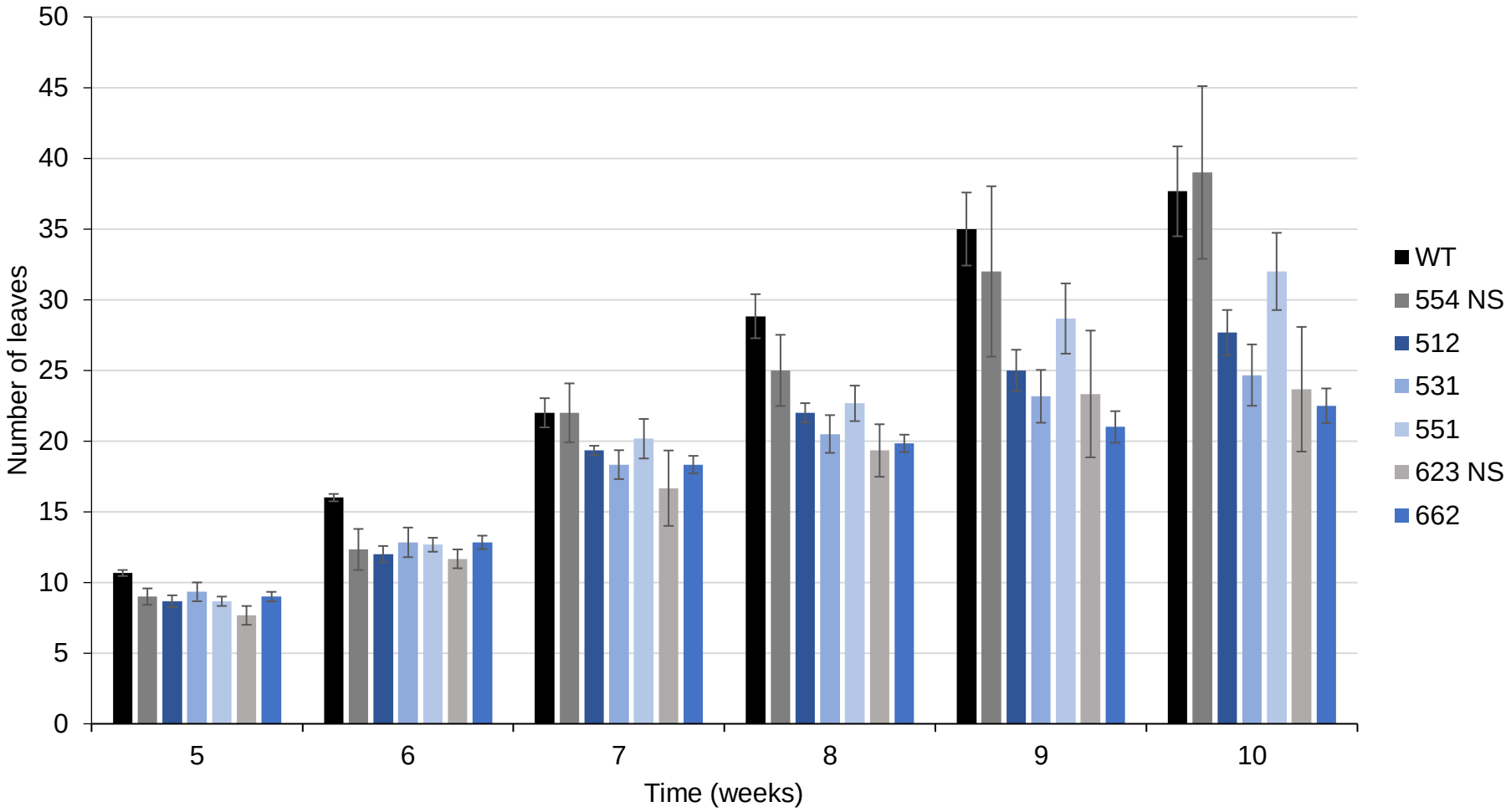


Figure 63: The relative number of leaves in the *Tawhy*, WT and NS lines in late development. Values are a mean of six biological replicates. Error bars are calculated from standard error of means of 6 plants per line and t tests conducted with the corresponding NS gave $p > 0.05$.

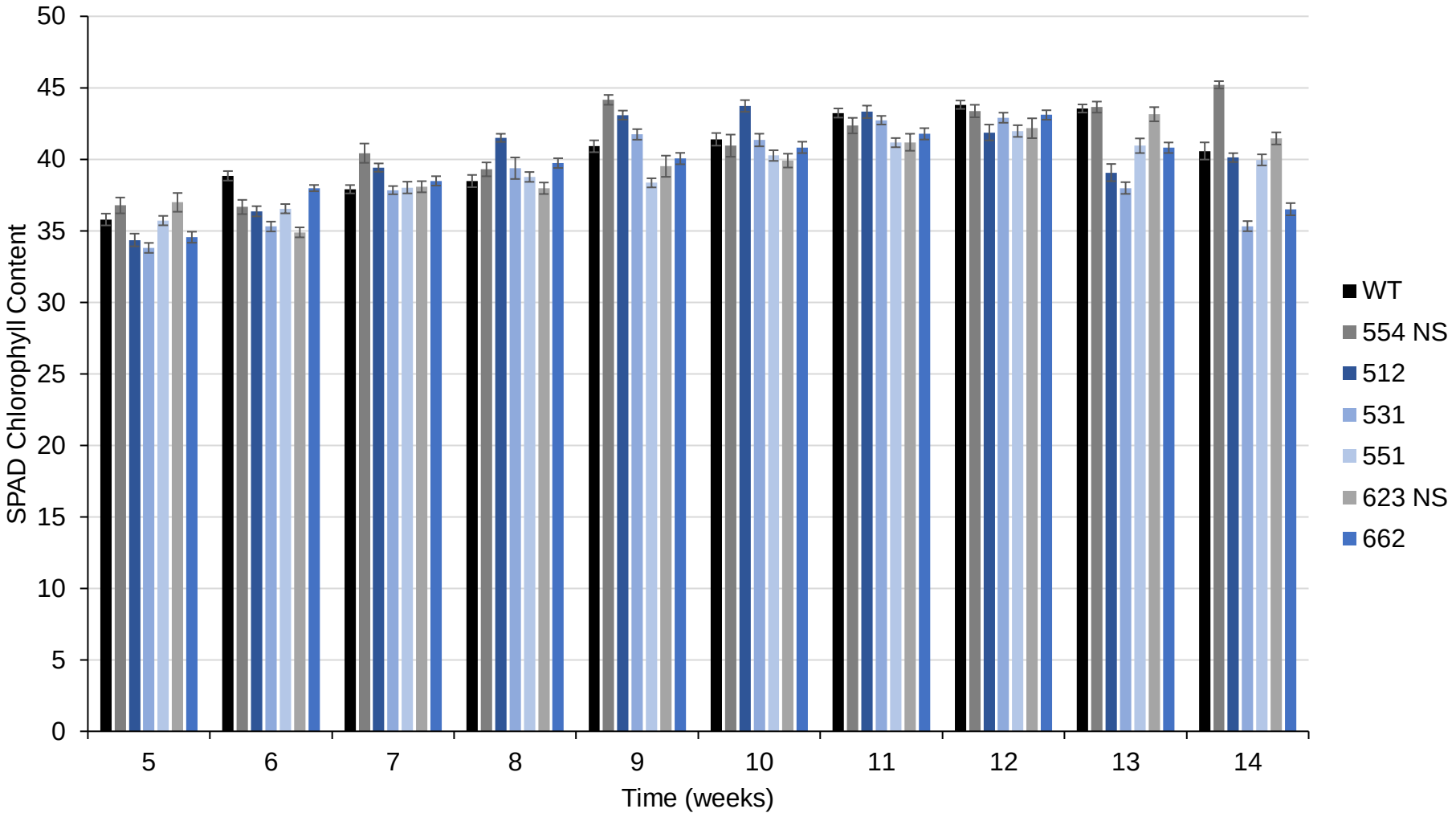


Figure 64: Relative SPAD measurements of the *Tawhy* and control lines during later development. Values are a mean of six biological replicates. Error bars are calculated from standard error of means of 6 plants per line and t tests conducted with the corresponding NS gave $p > 0.05$.

Plant height was similar in all of the lines during early vegetative growth (five and six), however none of the lines were significantly different to the corresponding NS at any time point [Figure 62]. This phenotype was again observed in the number of leaves and the SPAD measurements of each of the lines compared to the NS [Figure 63, Figure 64]. Taken together, these data suggest that loss of WHY1 functions results in smaller plants with fewer leaves.

7.2.4 The effects of WHY1 on seed production and flowering

The *Tawhy* RNAi lines were monitored for changes in fertility. Tiller numbers were monitored [Figure 65], as well as the time leading up to anthesis, the extrusion of flowers from the spikelet [Figure 66]. Additionally, seed yields were measured through the production of ears [Figure 67, Figure 68] and thousand grain weight per line [Figure 69].

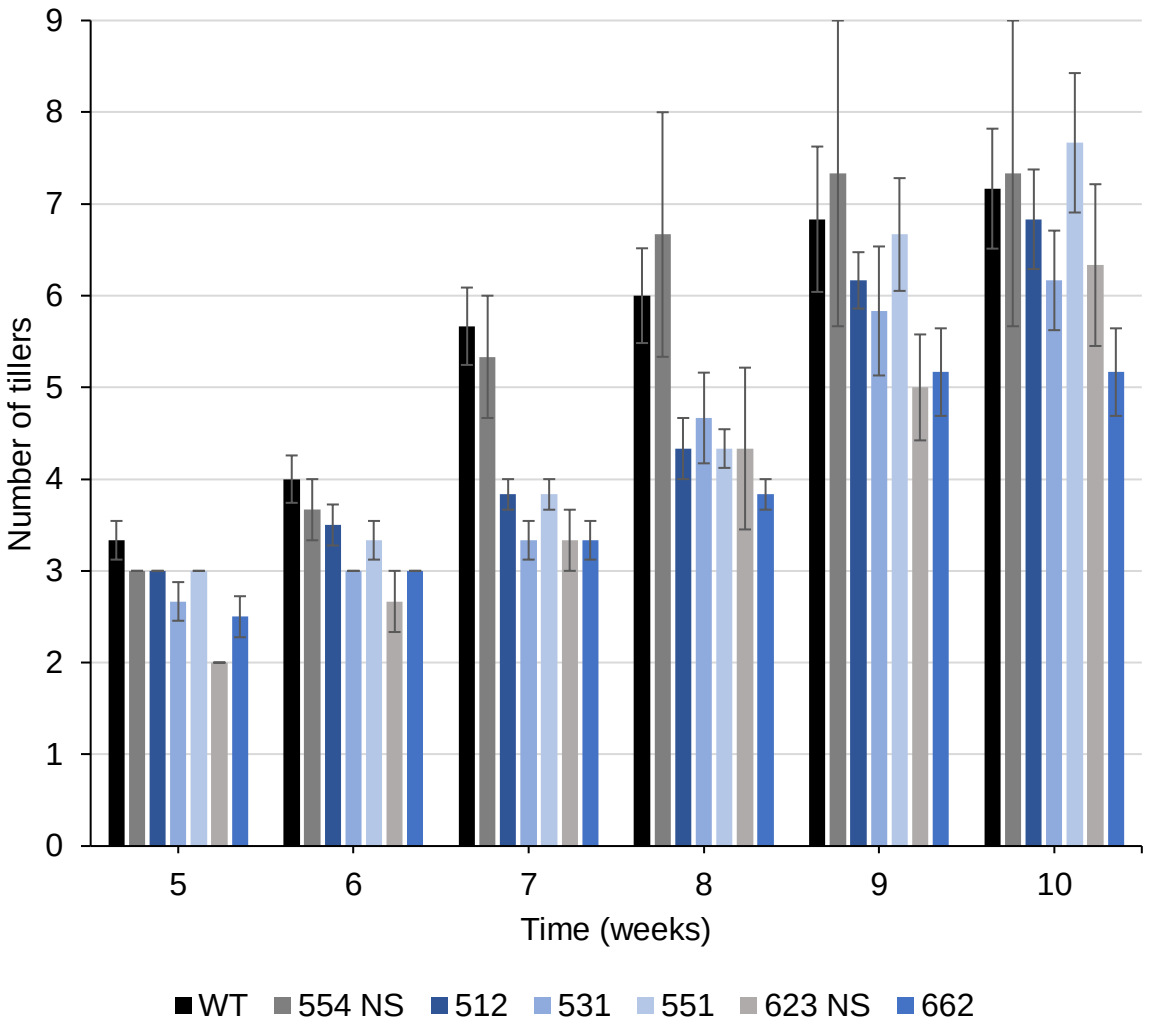


Figure 65: Relative numbers of tillers in *Tawhy* lines and controls. Error bars are calculated from standard error of means of 6 plants per line and t tests conducted with the corresponding NS gave $p > 0.05$.

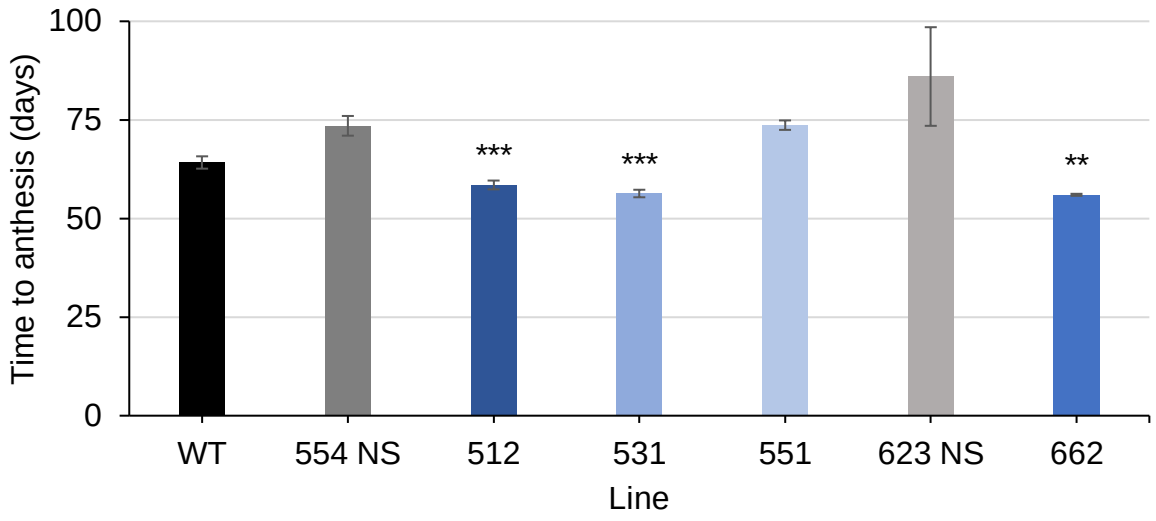


Figure 66: The number of days to anthesis in the different *Tawhy* lines. Values are a mean of six biological replicates. Error bars are calculated from standard error of means of 6 plants per line and t tests were conducted against the respective NS: **, p <0.01; ***, p < 0.001.

None of the lines had significantly more or fewer tillers than NS controls [Figure 65]. A significant difference was found between all groups (p value of 9.6×10^{-7}). All the lines, apart from line 551 which had a p value of 0.948, had significantly shorter time to anthesis compared to the NS control in an ANOVA comparison [Figure 66]. This data suggest that there is an effect of WHY1 on flowering time.

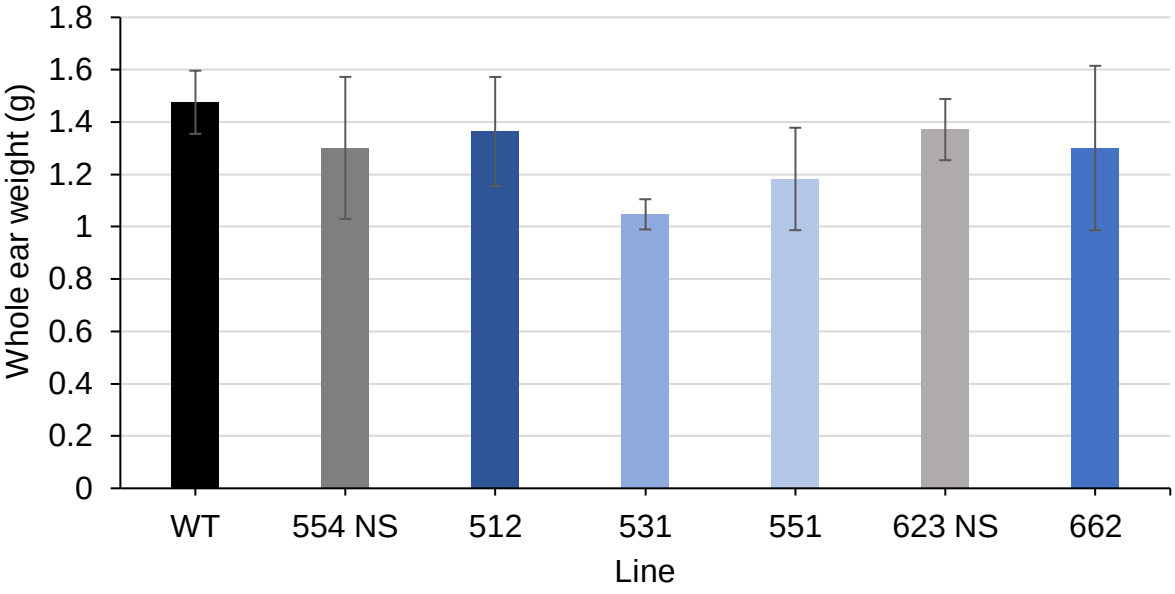


Figure 67: A comparison of whole ear weights (g) between the *Tawhy* lines Error bars are calculated from standard error of means of 6 plants per line and t tests conducted with the corresponding NS gave p > 0.05.

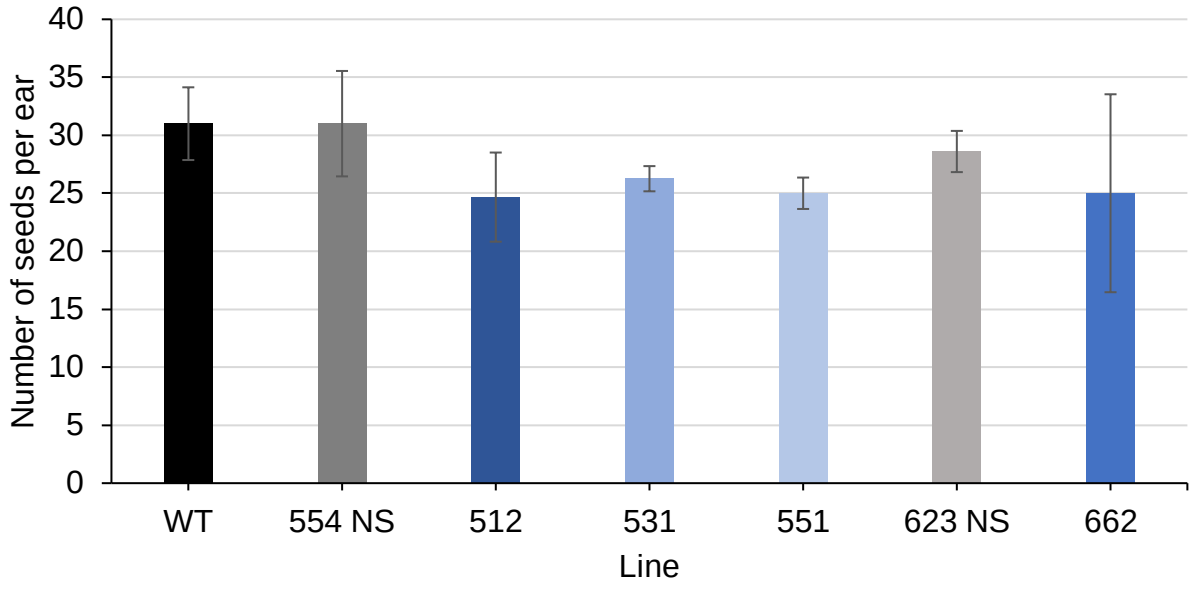


Figure 68: A comparison of total seeds per ear in the *Tawhy* lines. Error bars are calculated from standard error of means of 6 plants per line and t tests conducted with the corresponding NS gave $p > 0.05$.

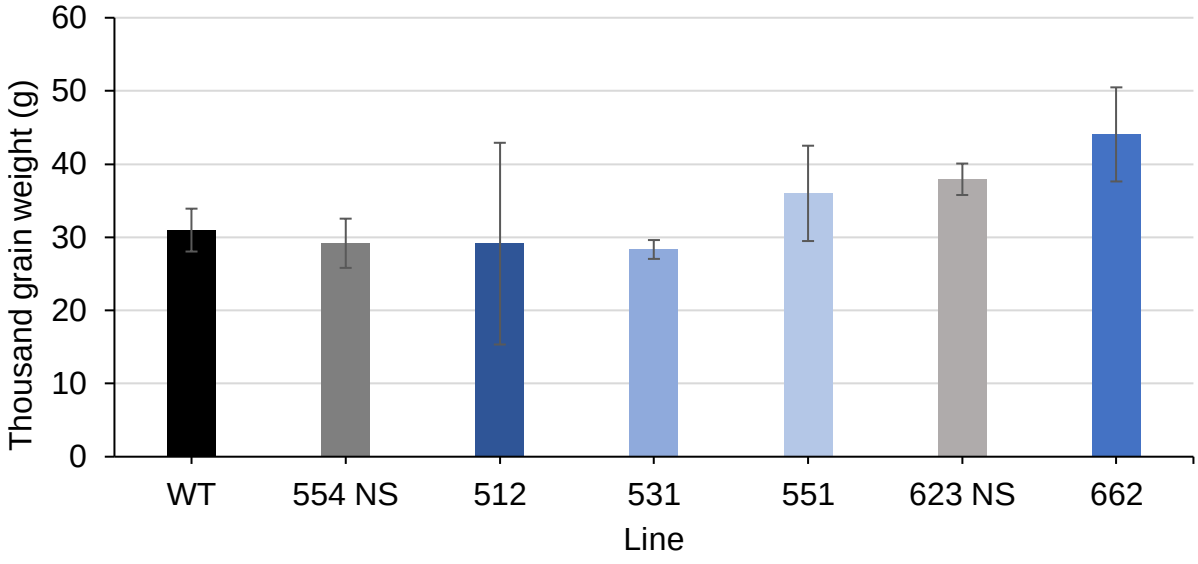


Figure 69: Comparison of seed yields expressed as weight per thousand seeds or thousand grain weight (g) in the *Tawhy* lines. Error bars are calculated from standard error of means of 6 plants per line and t tests were conducted the corresponding NS: *, $p < 0.05$;

Interestingly, despite the change in time to anthesis compared to the NS, there were no significant differences between the lines and the seed yield measurments [Figure 67, Figure 68, Figure 69]. These results suggest that loss of WHY1 functions influence anthesis by speeding up the flowering time [Figure 66], however this did not have an effect on final yields.

7.3 DISCUSSION

There are no reports in the literature concerning the functions of the WHY1 protein in wheat development. The hexaploid nature of *T. aestivum* means that there are probably multiple copies of the Why genes (Romeuf *et al*. 2010), adding complexity to the analysis of WHY1 protein functions. BLAST results reveal the presence of multiple copies of each of the Why genes [Figure 56]. Hence, reverse genetics approaches to study WHY protein functions in wheat are challenging. The RNAi knockdown lines created by Limagrain and analysed here provide an ideal tool for the analysis of WHY1 protein functions in wheat. The qPCR analysis reported in this chapter confirmed that four RNAi knockdown lines (512, 531, 551 and 662) have significantly decreases in *WHY1* transcripts compared to the NS controls (554 NS and 623 NS) and the Fielder WT [Figure 57]. The phenotypes of these lines were therefore characterised throughout development.

Germination was similar in the *Tawhy* lines and the NS and WT controls [Figure 58]. There was a minor delay in germination of both 551 and 623 NS in the first 24 hours after cold stratification however these lines had 100% germination by the 48h time point. There were no significant differences in the height of any of the lines at either early [Figure 59] or late development [Figure 62]. Similarly, there were no significant differences in the number of leaves in any of the lines in early [Figure 60] or late development [Figure 63]. Therefore, loss of Why1 functions had no apparent effect on early plant development. These data are consistent with observations in the barley *Hvwhy1* knockdown lines, which had similar photosynthesis rates and senescence characteristics to the WT (Comadira *et al.*, 2015). These observations are also consistent with previous observations in Arabidopsis [Chapter 4: Arabidopsis Figure 3, Figure 5]. Only the *Atwhy1why3* double mutants showed phenotypic changes relative to the wild type. A more marked phenotype might become evident in wheat if more of the WHY proteins were knocked down simultaneously.

There were few differences in the chlorophyll contents of any of the lines during early plant development. Line 662 had higher pigment level than the NS control at weeks 2 and 3 [Figure 61]. However, line 662 line had significantly lower chlorophyll content than the NS control at week 4 and thereafter [Figure 64]. No consistent phenotypes related to pigment contents were observed in any of the lines. In contrast, the leaves of barley *Hvwhy1* knockdown lines had significantly more chlorophyll than controls (Comadira *et al.*, 2015; Grabowski *et al.*, 2008). In contrast to maize, which is a C4 plant, Arabidopsis, barley and wheat are C3 plants. Hence,

it may be expected that these species would show similar phenotypes in response to loss of WHY1 functions. The data presented in this thesis would support this conclusion.

There was a significant delay in the appearance of the flowering stem in the *Atwhy3* mutants compared to the WT [Chapter 4: Arabidopsis Figure 11, Figure 12]. A delay in flowering was also seen in most of the *Tawhy* lines, which reached anthesis faster than the WT and the NS control lines. The lines 512 and 531 flowered 15 and 17 days earlier respectively at 58.5 and 56.3 days compared to the 554 NS line which flowered after 73.5 days. Furthermore, the most significant shift in flowering time was in the 662 line which flowered 30 days earlier than the 623 NS at 86 and 56 days respectively [Figure 66] The exception was line 551, which had longer time to anthesis than the WT [Figure 66]. Taken together, these findings suggest that WHY has a function in the control of flowering, a trait that could be useful in future crop breeding strategies as this is a consistent phenotype seen in both Arabidopsis and wheat. It would be useful to expand these studies to other widely-cultivated crops such as *Zea mays*, *Oryza sativa* or even *Solanum lycospericum* as it has been recently studied a lot in relation to WHY functions.

Despite the early flowering phenotype, there was no significant difference in whole ear weights [Figure 67], the number of seeds per ear [Figure 68] or the thousand grain weight [Figure 69] compared to the corresponding NS. This study is limited in that no measurements of senescence phenotypes were made. This was due to limited time at Limagrain, France for the iCASE placement, this may have uncovered the reason why there was an earlier time to anthesis but not an increase in grain yield, for example development could have been shifted earlier and the plants may also have senesced earlier after flowering early and thus yields would not be increased. It also would have been useful to repeat these experiments in the controlled glasshouse conditions at Leeds to provide further information on the growth and photosynthesis phenotypes.

These results that the *Tawhy* lines are a useful tool for the elucidation of WHY protein functions in wheat. Further studies are required to explore the interactions between WHY1 and other WHY proteins in wheat and how they may act to regulate flowering. The early flowering phenotype was observed in both wheat and Arabidopsis plants lacking a functional WHY1. More confirmation experiments are required to indicate the decreased gene expression throughout growth and interactions this may have with development, furthermore, miRNA studies are needed to confirm the RNAi knockdown targeted all copies, if this is not the case

this may explain the lack of a consistent phenotype throughout development in the RNAi lines compared to their NS. Further research is required to understand the mechanisms involved and how the different WHY proteins contribute to the control of plant growth and development. For example, it may be good to make a triple-knockdown RNAi plant of all of the WHYs found on chromosome 6 [Figure 56] which may give more severe phenotypes than the ones discussed in this chapter.

8 DISCUSSION

8.1 RESEARCH AIMS

This project was designed in order to elucidate the function of WHIRLY during plant development. The findings presented in this thesis provide new information on the roles of WHIRLY on seed viability in Arabidopsis and on development in wheat. For ease, the project aims were split into four sections:

Germination:

1. Explore the roles of WHY1 and WHY3 in seed longevity and germination using Arabidopsis *Atwhy1*, *Atwhy3* and *Atwhy1why3* mutant lines.
2. Explore the potential roles of WHY1 and WHY3 in germination using Arabidopsis *Atwhy1*, *Atwhy3* and *Atwhy1why3* mutant lines that have been subject to accelerated ageing.

Gene Interactions:

1. Identify targets of WHY1 and WHY3 in germination through RNA-seq analysis of *Atwhy1why3* seeds and wild type seeds.
2. Identify mechanisms for ageing hypersensitivity in *Atwhy1why3* 7 days aged mutants.

Arabidopsis Development:

1. To characterise the effect of WHY1 and WHY3 throughout Arabidopsis growth
2. To compare photosynthesis in *Atwhy1*, *Atwhy3*, and *Atwhy1why3* mutants relative to the wild type.
3. To determine the responses of the *Atwhy1*, *Atwhy3*, and *Atwhy1why3* mutants and the wild type plants to heat stress.

Wheat Development:

1. To confirm transformations of TaWHY knockdown plants and select lines to carry forward.
2. To characterise the early and late phenotypic growth changes of TaWHY knockdown plants compared to null segregants and WT.
3. To determine the effects of TaWHY deficiency on plant fertility and seed yield.

8.2 KEY FINDINGS

The germination results suggested that there was no difference in the viability of high-quality seeds in the *Atwhy1*, *Atwhy3* and *Atwhy1why3* mutants compared to the WT (4.2.1). However, after accelerated ageing, both viability and seed vigour were reduced in all mutant seeds compared to the WT; the most significant reduction was seen in the *Atwhy1why3* mutants. This data suggested that both WHY1 and WHY3 were required for recovery from ageing in seeds [Figure 70].

This phenotype was further explored in Chapter 5 where the mechanisms that underpin ageing hypersensitivity in the *Atwhy1why3* mutants were explored through analysis of the change in transcripts. Upon imbibition of high quality *Atwhy1why3* and WT seeds, the most upregulated genes were those that encoded for seed storage, seed lipid and late embryogenesis abundant (LEA) proteins which were not upregulated in gene lists with aged *Atwhy1why3* seeds. However, the transcript that appeared most commonly in the lists of top ten differentially expressed genes were those that encoded heat shock proteins (HSPs), namely, HSP70 which was downregulated in most pairwise comparisons apart from two comparisons where WT aged and imbibed seeds were compared. This transcript had the first and second lowest levels of all identified DEGs which appeared in the WT dry seeds relative to WT imbibed seeds comparisons and the WHY aged imbibed seeds relative to WHY imbibed seed comparisons respectively. There were only two comparisons that gave upregulation for genes that encoded HSP70, these were WT aged imbibed relative to WT dry seed and WT aged imbibed relative to WHY aged imbibed seeds, therefore these genes are upregulated in *Atwhy1why3* mutants which have been aged but the opposite is seen in WT aged seeds, where HSP70 is downregulated. Furthermore, transcripts that encoded LEA proteins were more abundant in comparisons of *Atwhy1why3* aged seeds. This data suggested that HSPs and LEA proteins are important in the germination of *Atwhy1why3* seeds under accelerated ageing

The photosynthetic capacity phenotypes of *Atwhy1*, *Atwhy3* and *Atwhy1why3* plants were recorded at optimal, and heat stressed conditions in 6.2.5. The *Atwhy3* and *Atwhy1why3* mutants had significantly higher leaf vapour pressure deficits (VpdL) compared to the WT when 4-week-old plants were exposed to higher temperatures. Furthermore, the *Atwhy1why3* double mutant had a significant reduction in VpdL at optimal growth temperatures. This data suggests that the WHY3 protein has an impact on photosynthetic function at higher temperatures, however this phenotype requires further exploration in regard to the switch of function at optimal and stressed conditions and the length of heat stress.

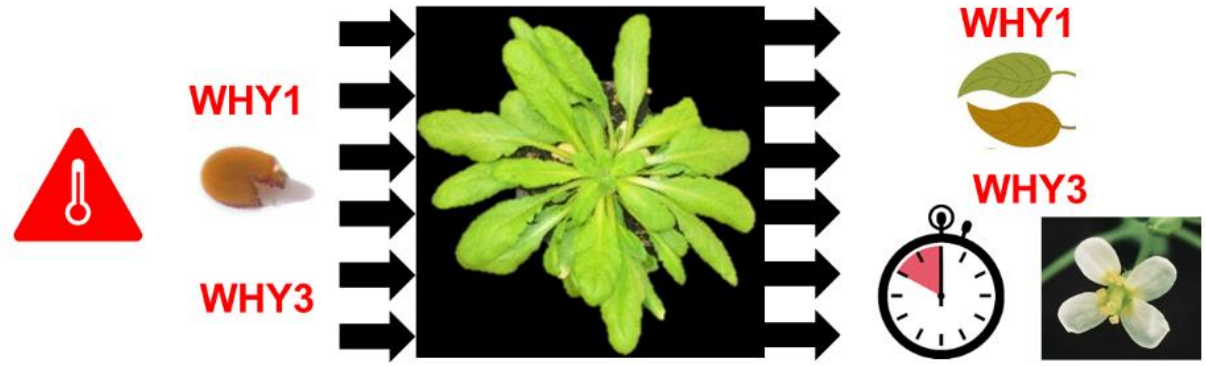


Figure 70: Schematic diagram illustrating the thesis conclusions in relation to the research of WHY1 and WHY3 throughout Arabidopsis development. Both WHY1 and WHY3 may be required for germination in seeds that have undergone accelerated ageing which has an interaction with HSP proteins as identified through RNA-seq data. WHY1 may have a role in the later development of leaves and WHY3 likely has a role in flowering timing.

There was no consistent significant difference found in the early vegetative growth of the *Atwhy* mutants compared to the WT. However, in later development growth differences became more prominent. Plants lacking the WHY3 protein had significantly smaller leaf area, reduced fresh weight and more siliques per plant and fewer seeds per silique compared to the WT [Figure 70]. Furthermore, all *Atwhy* mutants had significantly reduced chlorophyll content and visible yellowing of leaves compared to the WT after six weeks of growth. The most consistent significant differences were seen in the *Atwhy1why3* mutants after six weeks where plants had significantly reduced rosette diameter, fewer leaves, reduced chlorophyll a to b ratios, reduced biomass, and a delay in time to bolting compared to the WT plants [Figure 70]. This data suggests that both WHY1 and WHY3 are important in later development of leaves and rosettes which may have a knock-on causal effect on plant fertility and seed yield. Therefore, in analysis of *Tawhy* lines in the wheat chapter was split into three sections; early development, late development and plant fertility.

In the wheat chapter, four RNAi lines were found to be significantly knocked down in Tawhy compared to their relative null segregants. These were used for analysis of the functions of TaWHY in wheat development. There were no consistent significant differences in any of the *Tawhy* lines compared to their NS or the WT in early or late vegetative development. However, there was a flowering developmental phenotype. Most of the *Tawhy* lines, except line 551, had a longer time to anthesis compared to the WT and their relative NS lines. These results suggest that WHY has an effect on flowering timing in wheat, however the growth phenotypes of wheat need further exploration with a range of *Tawhy* mutants as there is a large potential multiple

copies of genes, as identified in Figure 56, and due to the hexaploidy nature of *Triticum aestivum* there is a possibility for functional redundancy in Why genes (Romeuf *et al*. 2010; Mochida *et al.*, 2003). Therefore, use of wider range of *Tawhy* mutants from more of these genes and crosses of them could produce more significant phenotypes and elucidate the function of TaWHY in flowering development.

8.3 INTERPRETATION OF RESULTS

Plants lacking both WHY1 and WHY3 were significantly more sensitive to accelerated ageing during germination. Both single mutants and the double mutant were significantly reduced in germination viability and vigour after ageing treatment. No previous studies have found a link between WHY proteins and ageing in germination, however high quality *Atwhy2* seeds were found to have reduced seed viability after four days, potentially because of the high levels of mitochondrial activity during germination (Golin *et al.*, 2020). Previous studies have found that *Atwhy1* mutants had reduced sensitivity to ABA and SA, whilst transgenic lines overexpressing AtWhy1 were found to be hypersensitive to ABA and overexpressing only the nuclear form of AtWhy1 led to the same level of ABA insensitivity as *Atwhy1* mutants (Isemer *et al.*, 2012). The exogenous application of ABA on high quality seeds is known to inhibit their germination (Garciarrubio *et al*., 1997), therefore plastid localised WHY1 was implicated in the responsiveness of germinating seeds to ABA even when applied exogenously (Isemer *et al.*, 2012). The *Atwhy1* mutants were found to have a shift in vigour under normal growth conditions and an increase in germination by up to 40% was observed in the presence of high concentrations of SA and ABA (Isemer *et al.*, 2012). There was a slight reduction in the end viability of high quality *Atwhy1* seed germination compared to all other lines (4.2.1), however this difference was not significantly different to the WT. Interestingly, ABA has also been found to prevent degradation of seed storage proteins (Garciarrubio *et al*., 1997), the reduced ABA sensitivity of *Atwhy* mutants may lead to increased storage protein remobilisation which may explain the significant reduction of viability and vigour in all aged *Atwhy* mutants. The *Atwhy1why3* seeds showed the most significant inhibition in germination. This new function shows that there may be an interaction of WHY1 and WHY3 proteins in protection of seeds during the ageing process, or in the recovery of ageing-induced cellular damage. Furthermore, the RNA-seq analysis identified many seed storage related genes had upregulated transcripts in high quality seeds, however these did not appear in the top ten up- or down-regulated DEGs in aged *Atwhy1why3* seeds. Seed storage proteins have been found to act as a protectant for important proteins required for seed germination and seedling formation as they more highly

oxidised after accelerated ageing compared to the WT (Nguyen *et al*., 2015). This data correlates with the theory that WHY1 and WHY3 proteins are involved with seed protection during the ageing process that comes with long-term storage.

Moreover, the most commonly differentially expressed gene was HSP70, which was downregulated in *Atwhy1why3* aged seed transcripts yet upregulated in two lists of WT aged seed transcripts. This is a novel finding, but perhaps not a surprising one; many studies have reported increases in the levels of WHY transcripts in plants exposed to abiotic stress, including heat stress (Zhuang, *et al.,* 2020). Overexpression of WHY1 in tomato led to increased levels of SlHSP21.5A transcripts compared to WT which raises the possibility that HSPs are direct targets for WHY transcription. These plants also had reduced wilting phenotypes, greater membrane stability, higher soluble sugar contents and decreased ROS levels under heat stress compared to WT, whilst the RNAi knock down lines had the opposite phenotype (Zhuang, *et al.*, 2020).

Furthermore, overexpression of a sunflower heat shock transcription factor (*HaHSFA9*) in tobacco (*Nicotiana tabacum*) seeds was found to have accumulated levels of heat shock proteins and were more tolerant to accelerated ageing treatment compared to wild type (Prieto-Dapena *et al.*, 2006). More recent studies have attributed the enhanced longevity of seeds through the interaction of heat shock factors, namely *HaHSFA9*, which when overexpressed with *DROUGHT-RESPONSIVE ELEMENT-BINDING FACTOR2* resulted in significantly more germinated seeds after seed ageing compared to WT and oe*HaHSFA9* lines (Almoguera *et al.*, 2009). RNA-seq data showed an increase in a drought stress tolerance transcripts encoding Dehydration-Responsive Element Binding Protein 2 in aged WT compared to *Atwhy1why3* seeds. Additionally, transcripts encoding the protein MAPKKK18, which is involved in drought stress tolerance, was expressed at significantly lower levels after ageing in imbibed *Atwhy1why3* seeds. Recent studies have found enhanced levels of SlWHY1 and SlWHY2 in tomato plants under drought stress (Akbudak and Filiz, 2019), and RNAi knock down SlWHY2 lines had severe wilting phenotypes under drought in addition to numerous factors that indicate decreases in photosynthetic activity (Meng *et al.,* 2020). Furthermore, cassava plants lacking in any MeWHYs were more sensitive to drought stress (Yan *et al.*, 2020). These results indicate reduced HSP transcripts in *why* mutants, as observed in Arabidopsis seeds, is a feature shared with other plant species (Yan *et* al., 2020), for example, WHY1 binds to promoter elements of HSPs in tomato to activate the heat shock response (Zhuang *et al.*, 2020; Taylor *et al*., 2022). However, there were no consistent findings in

photosynthetic capacity measurements of the *Atwhy* that suggested a heat shock phenotype throughout vegetative development. This matches with the theory that the HSP is linked to drought stress response which is more critical in seeds where the proper imbibition of water is important for seeds to properly germinate. Perhaps WHY proteins shift their function to other developmental changes in later development such as the significant delay in time to flowering seen in *Atwhy3* mutants and most of the *Tawhy* lines. A lack of correlation between flowering timing, plant biomass and delayed chlorophyll degradation have previously been suggested to be a stress-escape behaviour phenotype which is inhibited by endogenous basal ABA (Negin *et al*., 2019) in which *Atwhy1* mutants have reduced sensitivity (Isemer *et al.*, 2012).

Modifications in developmental processes and gas exchange function in response to drought stress are stimulated by the SHORT VEGETATIVE PHASE (SVP) MADS-domain protein which has also been found to suppress the onset of flowering (Castelán-Muñoz *et al*., 2019). There was drought sensitivity in *Atsvp* loss-of-function mutants, whilst lines overexpressing SVP were drought tolerant (Castelán-Muñoz *et al*., 2019). Moreover, a positive effect on stomatal opening has been found to be exerted by a gene, *SOC1*, which also integrates multiple flowering signals in MADS box transcription (Kimura *et al*., 2015). The cumulative effect of these signals may be the cause of the significant delays in bolting phenotype and significant increases in VpdL seen in plants lacking WHY3. Therefore, it is likely that the *Atwhy1*, *Atwhy3* and *Atwhy1why3* phenotypes reported in this study are synonymous with stress response behaviour. It is interesting that despite the effect on flowering timing, none of the *Atwhy* nor the *Tawhy* mutants had a consistent significant difference in seed yield. Perhaps the consistent growth conditions used in this study maintain seed development even when flowering is earlier which would not be the case for plants in the field which have to deal with a range of abiotic and biotic stresses during development.

Furthermore, this study did not identify any changes in senescence of the leaves and therefore the end of development remains undefined in *Tawhy* mutants. Thus, the shift in time to anthesis could be a shift in the entire final stage of development and this may be why seed yields were unaffected. There have been no published studies on the functions of the WHY1 protein in wheat leaf or whole plant development, as such, the flowering timing results provided in this study are completely novel. However, this phenotype may not carry to crops in the field where earlier flowering could mean loss of yields as the temperature is not at optimum conditions to properly develop seeds. Days with average temperatures greater than 34 °C, which are of increasing incidence, accelerate leaf senescence in wheat which significantly reduces grain

yield (Wardlaw and Moncur, 1995). Plants with increased tolerance to high heat are likely to become critical for southern and central Europe soon to achieve high yields to continue feeding the growing global population (Stratonovitch and Semenov, 2015). Therefore, long-term research into the functions of the WHY protein in flowering timing and heat stress could be vital for maintaining food security. The heat-stress experiments in this project were time limited and more oriented towards photosynthetic capacity than long-term heat stress, therefore more work must be carried out to determine the growth changes of *why* mutants in both Arabidopsis and wheat under hours or even days of stress as crops would experience in the field.

8.4 Recommendations for Implementation in Future Research

For future research the function of multiple WHY proteins must be explored in wheat to fully determine the effect on plant phenotypes as the research in Arabidopsis suggests that multiple WHY proteins function together to create an additive effect on the phenotype. The most significant results were seen in the *Atwhy1why3* mutants; while the single mutants often had a shift in growth phenotypes, there was an additive effect when both WHY proteins were removed. The practical application of these findings is that WHY could be a suitable target for plant breeding strategies as a consistent effect on plant fertility and flowering timing was seen in both the Arabidopsis and wheat plants. However, when implementing this is crops, particularly wheat, it is important that all copies of WHY proteins are monitored for these changes.

Another interesting avenue of future research for the functions of WHY in plants would be to further explore the interactions that WHY proteins have during abiotic and biotic stress to further understand their mechanism in plant development. Many recent studies have revealed that plants exposed to abiotic stresses such as high salinity, drought (Akbudak and Filiz, 2019; Yan *et al.*, 2020), heat (Zhuang *et al.*, 2020a), and biotic stress such as Botrytis fungal infection (Akbudak and Filiz, 2019) had increases in WHY transcript levels. Furthermore, research into tomato and cassava suggests that WHY1 binds to the ERE-like element in response to heat and drought [Figure 35] and this could be an interesting avenue of research to further explore with more abiotic stresses.

8.5 CONCLUDING SUMMARY

This research finds that WHY1 and WHY3 work together to integrate multiple environmental factors to elucidate a reduced seed viability and vigour germination response under accelerated ageing which is implicated in heat shock response and drought response which is likely mediated by the phytohormone ABA. Additionally, mutants lacking WHY3 in Arabidopsis and WHY1 in wheat have a delay in flowering timing compared to the wild type. These data show novel functions of WHY proteins which contribute to the wider research of future food security in crops where environmental changes due to climate change are likely to be more extreme. Therefore, WHY proteins are a good target for crop breeding programmes.

9 BIBLIOGRAPHY

Afgan, E., Baker, D., van den Beek, M., Blankenberg, D., Bouvier, D., Čech, M., Chilton, J., Clements, D., Coraor, N., Eberhard, C., Grüning, B., Guerler, A., Hillman-Jackson, J., Von Kuster, G., Rasche, E., Soranzo, N., Turaga, N., Taylor, J., Nekrutenko, A., and Goecks, J. 2016. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. *Nucleic acids research*, **44**(Web Server Issue), pp.W3–W10.

Akbudak, M. and Filiz, E. 2019. Whirly (Why) transcription factors in tomato (*Solanum lycopersicum* L): genome-wide identification and transcriptional profiling under drought and salt stresses. *Molecular Biology Reports*, **46**(4), pp.4139-4150.

Almoguera, C., Prieto-Dapena, P., Díaz-Martín, J., Espinosa, J. M., Carranco, R., and Jordano, J. 2009. The HaDREB2 transcription factor enhances basal thermotolerance and longevity of seeds through functional interaction with HaHSFA9. *BMC plant biology*, **9**, article no:75 [no pagination].

Arabidopsis Biological Resource Center (ABRC). 2013 Handling Arabidopsis plants and seeds. [Online] Available: https://abrc.osu.edu/seed-handling [Accessed: 15/01/2018].

Asdal, Å., and Guarino, L. 2018. The Svalbard Global Seed Vault: 10 Years-1 Million Samples. *Biopreservation and Biobanking*, **16**(5), pp.391–392.

Bawono, P., and Heringa, J. 2014. PRALINE: a versatile multiple sequence alignment toolkit. In: Russell D. ed. *Methods in molecular biology Multiple Sequence Alignment Methods*, **1079**, pp.245–262.

Bechtold, N., Ellis, J., and Pelletier, G. 1993. *In planta Agrobacterium*- mediated gene transfer by infiltration of adult *Arabidopsis thaliana* plants. C. R. Acad. Sci. Paris, Life Sciences 316:1194-1199

Belamkar, V., Weeks, N. T., Bharti, A. K., Farmer, A. D., Graham, M. A., and Cannon, S. B. 2014. Comprehensive characterization and RNA-Seq profiling of the HD-Zip transcription factor family in soybean (*Glycine max*) during dehydration and salt stress. *BMC genomics*, **15**, article no.: 950 [no pagination].

Benjamini, Y. and Hochberg, Y. 2000. On the Adaptive Control of the False Discovery Rate in Multiple Testing With Independent Statistics. *Journal of Educational and Behavioral Statistics*, **25**(1), pp.60-83.

Bewley, J. D., and Black, M. 1994. Seeds: Germination, Structure, and Composition. In: *Seeds: Physiology of Development and Germination*. Second Edition. Springer Science + Business Media, Plenum Press, New York, p.1.

Bolger, A. M., Lohse, M., and Usadel, B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics (Oxford, England)*, **30**(15), pp.2114–2120.

Bonthala, V. S., Mayes, K., Moreton, J., Blythe, M., Wright, V., May, S. T., Massawe, F., Mayes, S., and Twycross, J. 2016. Identification of Gene Modules Associated with Low Temperatures Response in Bambara Groundnut by Network-Based Analysis. *PloS one*, **11**(2), article no.: e0148771 [no pagination].

Boyes, D. C., Zayed, A. M., Ascenzi, R., McCaskill, A. J., Hoffman, N. E., Davis, K. R., and Görlach, J. 2001. Growth stage-based phenotypic analysis of Arabidopsis: a model for high throughput functional genomics in plants. *The Plant cell*, **13**(7), pp.1499–1510.

Breeze, E., Harrison, E., McHattie, S., Hughes, L., Hickman, R., Hill, C., Kiddle, S., Kim, Y., Penfold, C., Jenkins, D., Zhang, C., Morris, K., Jenner, C., Jackson, S., Thomas, B., Tabrett, A., Legaie, R., Moore, J., Wild, D., Ott, S., Rand, D., Beynon, J., Denby, K., Mead, A. and Buchanan-Wollaston, V., 2011. High-Resolution Temporal Profiling of Transcripts during Arabidopsis Leaf Senescence Reveals a Distinct Chronology of Processes and Regulation. *The Plant Cell*, **23**(3), pp.873-894.

Bucholc, M., and Buchowicz, J. 1992. Synthesis of extrachromosomal DNA and telomere-related sequences in germinating wheat embryos. *Seed Science Research*, **2**(3), pp.141-146.

Cao, F. Y., Yoshioka, K. and Desveaux, D. 2011. The roles of ABA in plant-pathogen interactions. *Journal of Plant Research*, **124**(4), pp.489-499.

Cappadocia, L., Maréchal, A., Parent, J. S., Lepage, E., Sygusch, J., and Brisson, N. 2010. Crystal structures of DNA-Whirly complexes and their role in Arabidopsis organelle genome repair. *The Plant cell*, **22**(6), pp.1849–1867.

Cappadocia, L., Parent, J. S., Sygusch, J., and Brisson, N. 2013. A family portrait: structural comparison of the Whirly proteins from Arabidopsis thaliana and Solanum tuberosum. *Acta crystallographica. Section F, Structural biology and crystallization communications*, **69**(11), pp.1207–1211.

Cappadocia, L., Parent, J. S., Zampini, E., Lepage, E., Sygusch, J., and Brisson, N. 2012. A conserved lysine residue of plant Whirly proteins is necessary for higher order protein assembly and protection against DNA damage. *Nucleic acids research*, **40**(1), pp.258–269.

Castelán-Muñoz, N., Herrera, J., Cajero-Sánchez, W., Arrizubieta, M., Trejo, C., García-Ponce, B., Sánchez, M. P., Álvarez-Buylla, E. R., and Garay-Arroyo, A. 2019. MADS-Box Genes Are Key Components of Genetic Regulatory Networks Involved in Abiotic Stress and Plastic Developmental Responses in Plants. *Frontiers in Plant Science*, **10**, article no.: 853 [no pagination].

Chevigny, N., Schatz-Daas, D., Lotfi, F., and Gualberto, J. M. 2020. DNA Repair and the Stability of the Plant Mitochondrial Genome. *International journal of molecular sciences*, **21**(1), p.328.

Clough, S. J. and Bent, A. F., 1998 Floral dip: a simplified method for Agrobacterium mediated transformation of Arabidopsis thaliana. Plant J 16:735-43.

Comadira, G., Rasool, B., Kaprinska, B., García, B. M., Morris, J., Verrall, S. R., Bayer, M., Hedley, P. E., Hancock, R. D., and Foyer, C. H. 2015. WHIRLY1 Functions in the Control of Responses to Nitrogen Deficiency But Not Aphid Infestation in Barley. *Plant physiology*, **168**(3), pp.1140–1151.

Dangl, J. L., and Jones, J. D. 2001. Plant pathogens and integrated defence responses to infection. *Nature*, **411**(6839), pp.826–833.

De Giorgi, J., Piskurewicz, U., Loubery, S., Utz-Pugin, A., Bailly, C., Mène-Saffrané, L., and Lopez-Molina, L. 2015. An Endosperm-Associated Cuticle Is Required for Arabidopsis Seed Viability, Dormancy and Early Control of Germination. *PLoS genetics*, **11**(12), article no.: e1005708 [no pagination].

Debeaujon, I., Léon-Kloosterziel, K. M., and Koornneef, M. 2000. Influence of the testa on seed dormancy, germination, and longevity in Arabidopsis. *Plant physiology*, **122**(2), pp.403–414.

Debeaujon, I., Lepiniec, L., Pourcel, L. and Routaboul, J.-M. 2007. Seed coat development and dormancy. K. Bradford, H. Nonogaki. eds. *Seed Development, Dormancy and Germination*, Blackwell, pp.25-49

Despres, C., Subramaniam, R., Matton, D. P., and Brisson, N. 1995. The Activation of the Potato PR-10a Gene Requires the Phosphorylation of the Nuclear Factor PBF-1. *The Plant Cell*, **7**(5), pp.589–598.

Desveaux D, Maréchal A and Brisson N. 2005. Whirly transcription factors: defense gene regulation and beyond. *Trends Plant Sci*. **10**(2), pp.95-102.

Desveaux, D., Allard, J., Brisson, N. and Sygusch, J., 2002. A new family of plant transcription factors displays a novel ssDNA-binding surface. *Nature Structural Biology*, **9**(7), pp.512-517.

Desveaux, D., Després, C., Joyeux, A., Subramaniam, R. and Brisson, N. 2000. PBF-2 Is a Novel Single-Stranded DNA Binding Factor Implicated in PR-10a Gene Activation in Potato. *The Plant Cell*, **12**(8), pp.1477-1490.

Desveaux, D., Maréchal, A. and Brisson, N. 2005. Whirly transcription factors: defense gene regulation and beyond. *Trends in Plant Science*, **10**(2), pp.95-102.

Desveaux, D., Subramaniam, R., Després, C., Mess, J., Lévesque, C., Fobert, P., Dangl, J. and Brisson, N. 2004. A "Whirly" Transcription Factor Is Required for Salicylic Acid-Dependent Disease Resistance in Arabidopsis. *Developmental Cell*, **6**(2), pp.229-240.

Dhaka, N., Krishnan, K., Kandpal, M., Vashisht, I., Pal, M., Sharma, M. and Sharma, R., 2020. Transcriptional trajectories of anther development provide candidates for engineering male fertility in sorghum. *Scientific Reports*, **10**(1), p.897.

Dietz, K. J., and Pfannschmidt, T. 2011. Novel regulators in photosynthetic redox control of plant metabolism and gene expression. *Plant physiology*, **155**(4), pp.1477–1485.

Donà, M., Balestrazzi, A., Mondoni, A., Rossi, G., Ventura, L., Buttafava, A., Macovei, A., Sabatini, M. E., Valassi, A., and Carbonera, D. 2013. DNA profiling, telomere analysis and antioxidant properties as tools for monitoring ex situ seed longevity. *Annals of botany*, **111**(5), pp.987–998.

Duan, S., Hu, L., Dong, B., Jin, H. L., and Wang, H. B. 2020. Signaling from Plastid Genome Stability Modulates Endoreplication and Cell Cycle during Plant Development. *Cell reports*, **32**(6), article no.: 108019 [no pagination].

Erenstein, O., Chamberlin, J. and Sonder, K. 2021. Estimating the global number and distribution of maize and wheat farms, *Global Food Security*, **30**, article no: 100558 [no pagination].

Escoubas, J. M., Lomas, M., LaRoche, J., and Falkowski, P. G. 1995. Light intensity regulation of cab gene transcription is signaled by the redox state of the plastoquinone pool. *Proceedings of the National Academy of Sciences of the United States of America*, **92**(22), pp.10237–10241.

Fauser F., Schiml S. and Puchta H. 2014 Both CRISPR/Cas-based nucleases and nickases can be used efficiently for genome engineering in *Arabidopsis thaliana. The Plant Journal*, 79: 348–359. doi: 10.1111/tpj.12554

Finch-Savage, W. E., and Bassel, G. W. 2016. Seed vigour and crop establishment: extending performance beyond adaptation. *Journal of Experimental Botany*. **67**(3), pp.567–591.

Foyer, C. H., and Noctor, G. 2009. Redox regulation in photosynthetic organisms: signaling, acclimation, and practical implications. *Antioxidants & redox signaling*, **11**(4), pp.861–905.

Foyer, C. H., Karpinska, B., and Krupinska, K. 2014. The functions of WHIRLY1 and REDOX-RESPONSIVE TRANSCRIPTION FACTOR 1 in cross tolerance responses in plants: a hypothesis. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, **369**(1640), article no.: 20130226 [no pagination].

Foyer, C. H., Neukermans, J., Queval, G., Noctor, G., and Harbinson, J. 2012. Photosynthetic control of electron transport and the regulation of gene expression. *Journal of experimental botany*, **63**(4), pp.1637–1661.

García-Medel, P. L., Baruch-Torres, N., Peralta-Castro, A., Trasviña-Arenas, C. H., Torres-Larios, A., and Brieba, L. G. 2019. Plant organellar DNA polymerases repair double-stranded breaks by microhomology-mediated end-joining. *Nucleic acids research*, **47**(6), pp.3028–3044.

Garciarrubio, A., Legaria, J. P., and Covarrubias, A. A. 1997. Abscisic acid inhibits germination of mature Arabidopsis seeds by limiting the availability of energy and nutrients. *Planta*, **203**(2), pp.182–187.

Giardine, B., Riemer, C., Hardison, R. C., Burhans, R., Elnitski, L., Shah, P., Zhang, Y., Blankenberg, D., Albert, I., Taylor, J., Miller, W., Kent, W. J., and Nekrutenko, A. 2005. Galaxy: a platform for interactive large-scale genome analysis. *Genome research*, **15**(10), pp.1451–1455.

Golin, S., Negroni, Y. L., Bennewitz, B., Klösgen, R. B., Mulisch, M., La Rocca, N., Cantele, F., Vigani, G., Lo Schiavo, F., Krupinska, K., and Zottini, M. 2020. WHIRLY2 plays a key role in mitochondria morphology, dynamics, and functionality in *Arabidopsis thaliana*. *Plant direct*, **4**(5), article no.: e00229 [no pagination].

Grabowski, E., Miao, Y., Mulisch, M., and Krupinska, K. 2008. Single-stranded DNA-binding protein Whirly1 in barley leaves is located in plastids and the nucleus of the same cell. *Plant physiology*, **147**(4), pp.1800–1804.

Graña, E., Díaz-Tielas, C., Sánchez-Moreiras, A. M., Reigosa, M. J., Celeiro, M., Abagyan, R., Teijeira, M., Duke, M. V., Clerk, T., Pan, Z., and Duke, S. O. 2019. Transcriptome and binding data indicate that citral inhibits single strand DNA-binding proteins. *Physiologia Plantarum*, **169**(1), pp.99–109.

Graña, E., Sotelo, T., Díaz-Tielas, C., Araniti, F., Krasuska, U., Bogatek, R., Reigosa, M. J., and Sánchez-Moreiras, A. M. 2013. Citral induces auxin and ethylene-mediated malformations and arrests cell division in Arabidopsis thaliana roots. *Journal of chemical ecology*, **39**(2), pp.271–282.

Greenup, A. G., Sasani, S., Oliver, S. N., Walford, S. A., Millar, A. A., and Trevaskis, B. 2011. Transcriptome analysis of the vernalization response in barley (Hordeum vulgare) seedlings. *PloS one*, **6**(3), article no.: e17900 [no pagination].

Grossiord, C., Buckley, T. N., Cernusak, L. A., Novick, K. A., Poulter, B., Siegwolf, R., Sperry, J. S., and McDowell, N. G. 2020. Plant responses to rising vapor pressure deficit. *The New phytologist*, **226**(6), pp.1550–1566.

Guan, Z., Wang, W., Yu, X., Lin, W., and Miao, Y. 2018. Comparative Proteomic Analysis of Coregulation of CIPK14 and WHIRLY1/3 Mediated Pale Yellowing of Leaves in *Arabidopsis*. *International journal of molecular sciences*, **19**(8), article no: 2231 [no pagination]

Hanson, M. R., and Hines, K. M. 2018. Stromules: Probing Formation and Function. *Plant physiology*, **176**(1), pp.128–137.

He, H., Zhu, S., Jiang, Z. Ji, Y., Wang, F., Zhao, R. and Bie, T. 2016. Comparative mapping of powdery mildew resistance gene Pm21 and functional characterization of resistance-related genes in wheat. *Theoretical and Applied Genetics*, **129**(4), pp.819–829.

Hu, Y., Jiang, Y., Han, X., Wang, H., Pan, J., Yu, D. 2017. Jasmonate regulates leaf senescence and tolerance to cold stress: crosstalk with other phytohormones. *Journal of experimental botany*, **68**(6), pp.1361–1369.

Huala, E; Dickerman, A; Garcia-Hernandez, M; Weems, D; Reiser, L; LaFond, F; Hanley, D; Kiphart, D; Zhuang, J; Huang, W; Mueller, L; Bhattacharyya, D; Bhaya, D; Sobral, B; Beavis, B, Somerville, C; and Rhee, SY. 2001. The Arabidopsis Information Resource (TAIR): A comprehensive database and web-based information retrieval, analysis, and visualization system for a model plant. Nucleic Acids Res.; **29**(1) pp.102-105.

Huang, C., Yu, J., Cai, Q., Chen, Y., Li, Y., Ren, Y., and Miao, Y. 2020. Triple-localized WHIRLY2 Influences Leaf Senescence and Silique Development via Carbon Allocation. *Plant Physiology*, **184**(3), pp.1348–1362.

Huang, D., Lin, W., Deng, B., Ren, Y., and Miao, Y. 2017. Dual-Located WHIRLY1 Interacting with LHCA1 Alters Photochemical Activities of Photosystem I and Is Involved in Light Adaptation in Arabidopsis. *International journal of molecular sciences*, **18**(11), p. 2352.

Hundertmark, M., Buitink, J., Leprince, O., and Hincha, D. 2011. The reduction of seed-specific dehydrins reduces seed longevity in Arabidopsis thaliana. *Seed Science Research*, **21**(3), pp.165-173.

Huot, B., Yao, J., Montgomery, B. L., and He, S. Y. 2014. Growth-defense tradeoffs in plants: a balancing act to optimize fitness. *Molecular plant*, **7**(8), pp.1267–1287.

International Wheat Genome Sequencing Consortium (IWGSC), IWGSC RefSeq principal investigators:, Appels, R., Eversole, K., Feuillet, C., Keller, B., Rogers, J., Stein, N., IWGSC whole-genome assembly principal investigators:, Pozniak, C. J., Stein, N., Choulet, F., Distelfeld, A., Eversole, K., Poland, J., Rogers, J., Ronen, G., Sharpe, A. G., Whole-genome sequencing and assembly:, Pozniak, C., … Uauy, C. (2018). Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science (New York, N.Y.)*, **361**(6403), article no: eaar7191.

Isemer, R., Krause, K., Grabe, N., Kitahata, N., Asami, T., and Krupinska, K. 2012 Plastid located WHIRLY1 enhances the responsiveness of Arabidopsis seedlings toward abscisic acid. *Frontiers in Plant Science*, **3**, article no.:283, [no pagination]

Isemer, R., Mulisch, M., Schäfer, A., Kirchner, S., Koop, H. U., and Krupinska, K. 2012. Recombinant Whirly1 translocates from transplastomic chloroplasts to the nucleus. *FEBS letters*, **586**(1), pp.85–88.

Jagadish, S., Way, D. A., and Sharkey, T. D. 2021. Plant heat stress: Concepts directing future research. *Plant, cell & environment*, **44**(7), pp.1992–2005.

Janack, B., Sosoi, P., Krupinska, K., and Humbeck, K. 2016. Knockdown of WHIRLY1 Affects Drought Stress-Induced Leaf Senescence and Histone Modifications of the Senescence-Associated Gene HvS40. *Plants*, **5**(3), article no: 37 [no pagination]

Job, C., Rajjou, L., Lovigny, Y., Belghazi, M. and Job, D. 2005. Patterns of Protein Oxidation in Arabidopsis Seeds and during Germination, *Plant Physiology*, **138**(2), pp.790–802.

Jolliffe, I. T., and Cadima, J. 2016. Principal component analysis: a review and recent developments. *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences*, **374**(2065), article no.: 20150202 [no pagination].

Kaur, H., Petla, B. P., Kamble, N. U., Singh, A., Rao, V., Salvi, P., Ghosh, S., and Majee, M. 2015. Differentially expressed seed aging responsive heat shock protein OsHSP18.2 implicates in seed vigor, longevity and improves germination and seedling establishment under abiotic stress. *Frontiers in plant science*, **6**, article no.: 713 [no pagination].

Kimura, Y., Aoki, S., Ando, E., Kitatsuji, A., Watanabe, A., Ohnishi, M., Takahashi, K., Inoue, S., Nakamichi, N., Tamada, Y., and Kinoshita, T. 2015. A flowering integrator, SOC1, affects stomatal opening in Arabidopsis thaliana. *Plant & cell physiology*, **56**(4), pp.640–649.

Kitajima, K. and Hogan, K. P. 2003. Increases of chlorophyll a/b ratios during acclimation of tropical woody seedlings to nitrogen limitation and high light. *Plant, cell & environment*, **26**(6), pp.857–865.

Kobayashi, Y., Takusagawa, M., Harada, N., Fukao, Y., Yamaoka, S., Kohchi, T., Hori, K., Ohta, H., Shikanai, T., and Nishimura, Y. 2015. Eukaryotic Components Remodeled Chloroplast Nucleoid Organization during the Green Plant Evolution. *Genome biology and evolution*, **8**(1), pp.1–16.

Kobs, G. 1998. Isolation of RNA from plant, yeast and bacteria. Promega Notes **68**, pp.28-29.

Krause, K., and Krupinska, K. 2009. Nuclear regulators with a second home in organelles. *Trends in plant science*, **14**(4), pp.194–199.

Krause, K., Herrmann, U., Fuß, J., Miao, Y. and Krupinska, K. 2009 Whirly proteins as communicators between plant organelles and the nucleus? *Endocytobiosis Cell Research*, **19**, pp.51-62.

Krause, K., Kilbienski, I., Mulisch, M., Rödiger, A., Schäfer, A., and Krupinska, K. 2005. DNA-binding proteins of the Whirly family in Arabidopsis thaliana are targeted to the organelles. *FEBS letters*, **579**(17), pp.3707–3712.

Kretschmer, M., Croll, D., and Kronstad, J. W. 2017. Chloroplast-associated metabolic functions influence the susceptibility of maize to Ustilago maydis. *Molecular Plant Pathology*, **18**(9), pp.1210–1221.

Krupinska, K., Oetke, S., Desel, C., Mulisch, M., Schäfer, A., Hollmann, J., Kumlehn, J., and Hensel, G. 2014. WHIRLY1 is a major organizer of chloroplast nucleoids. *Frontiers in plant science*, **5**, article no: 432 [no pagination]

Kucharewicz, W., Distelfeld, A., Bilger, W., Müller, M., Munné-Bosch, S., Hensel, G. and Krupinska, K., 2017. Acceleration of leaf senescence is slowed down in transgenic barley plants deficient in the DNA/RNA-binding protein WHIRLY1. *Journal of Experimental Botany*, **68**(5), pp.983-996.

Kushwaha, S. K., Grimberg, Å., Carlsson, A. S., and Hofvander, P. 2019. Charting oat (Avena sativa) embryo and endosperm transcription factor expression reveals differential expression of potential importance for seed development. *Molecular genetics and genomics : MGG*, **294**(5), pp.1183–1197.

Langstroff, A., Heuermann, M. C., Stahl, A., and Junker, A. 2022. Opportunities and limits of controlled-environment plant phenotyping for climate response traits. *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik*, **135**(1), pp.1–16.

Leijten, W., Koes, R., Roobeek, I., and Frugis, G. 2018. Translating Flowering Time From Arabidopsis thaliana to Brassicaceae and Asteraceae Crop Species. *Plants (Basel, Switzerland)*, **7**(4), p.111.

Lepage, É., Zampini, É., and Brisson, N. 2013. Plastid genome instability leads to reactive oxygen species production and plastid-to-nucleus retrograde signaling in Arabidopsis. *Plant physiology*, **163**(2), pp.867–881.

Lichtenthaler, H. K. 1987. Chlorophylls and carotenoids: Pigments of photosynthetic biomembranes. *Methods in Enzymology*. 148: p. 350-382.

Lièvre, M., Granier, C., and Guédon, Y. 2016. Identifying developmental phases in the Arabidopsis thaliana rosette using integrative segmentation models. *The New phytologist*, **210**(4), pp.1466–1478.

Lin, W., Huang, D., Shi, X., Deng, B., Ren, Y., Lin, W., and Miao, Y. 2019. H_2O_2 as a Feedback Signal on Dual-Located WHIRLY1 Associates with Leaf Senescence in *Arabidopsis*. *Cells*, **8**(12), p.1585.

Lin, W., Zhang, H., Huang, D., Schenke, D., Cai, D., Wu, B., and Miao, Y. 2020. Dual-Localized WHIRLY1 Affects Salicylic Acid Biosynthesis via Coordination of ISOCHORISMATE SYNTHASE1, PHENYLALANINE AMMONIA LYASE1, and *S*-ADENOSYL-L-METHIONINE-DEPENDENT METHYLTRANSFERASE1. *Plant physiology*, **184**(4), pp.1884–1899.

Ling, H. Q., Ma, B., Shi, X., Liu, H., Dong, L., Sun, H., Cao, Y., Gao, Q., Zheng, S., Li, Y., Yu, Y., Du, H., Qi, M., Li, Y., Lu, H., Yu, H., Cui, Y., Wang, N., Chen, C., Wu, H., … Liang, C. 2018. Genome sequence of the progenitor of wheat A subgenome Triticum urartu. *Nature*, **557**(7705), pp.424–428.

Liu, J., Li, L., Yuan, F., and Chen, M. 2019. Exogenous salicylic acid improves the germination of *Limonium bicolor* seeds under salt stress. *Plant signaling & behavior*, **14**(10), article no.: e1644595 [no pagination].

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

Lobo, F., de Barros, M., Dalmagro, H., Dalmolin, Â., Pereira, W., de Souza, É., Vourlitis, G. and Ortíz, C., 2013. Fitting net photosynthetic light-response curves with Microsoft Excel - a critical look at the models. *Photosynthetica*, **51**(3), pp.445-456.

Lu H. 2009. Dissection of salicylic acid-mediated defense signaling networks. *Plant signaling & behavior*, **4**(8), pp.713–717.

Mackenzie, S. A., and Kundariya, H. 2019. Organellar protein multi-functionality and phenotypic plasticity in plants. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, **375**(1790), article no: 20190182 [no pagination]

Maréchal, A., Parent, J. S., Véronneau-Lafortune, F., Joyeux, A., Lang, B. F., and Brisson, N. 2009. Whirly proteins maintain plastid genome stability in Arabidopsis. *Proceedings of the National Academy of Sciences of the United States of America*, **106**(34), pp.14693–14698.

Melonek, J., Mulisch, M., Schmitz-Linneweber, C., Grabowski, E., Hensel, G., and Krupinska, K. 2010. Whirly1 in chloroplasts associates with intron containing RNAs and rarely co-localizes with nucleoids. *Planta*, **232**(2), pp.471–481.

Meng, C., Yang, M., Wang, Y., Chen, C., Sui, N., Meng, Q., Zhuang, K., and Lv, W. 2020. SlWHY2 interacts with SlRECA2 to maintain mitochondrial function under drought stress in tomato. *Plant Science: an international journal of experimental plant biology*, **301**, article no.:110674, [no pagination].

Mersereau, M., Pazour, G. and Das, A., 1990. Efficient transformation of *Agrobacterium tumefaciens* by electroporation. *Gene*, **90**(1), pp.149-151.

Miao, Y., Jiang, J., Ren, Y., and Zhao, Z. 2013. The single-stranded DNA-binding protein WHIRLY1 represses WRKY53 expression and delays leaf senescence in a developmental stage-dependent manner in Arabidopsis. *Plant Physiology*, **163**(2), pp.746–756.

Mochida, K., Yamazaki, Y., and Ogihara, Y. 2003. Discrimination of homoeologous gene expression in hexaploid wheat by SNP analysis of contigs grouped from a large number of expressed sequence tags. *Molecular genetics and genomics : MGG*, **270**(5), pp.371–377.

Moreno, A.S., Margarit, E., Morales, L.M., Montecchiarini, M.L., Bello, F., Vázquez, D., Tripodi, K.E., and Podestá, F.E. 2020. Immediate- and long-term proteomic responses of epicarp from two heat conditioned tangor cultivars stored at low temperature differing in their susceptibility to infection. *Postharvest biology and technology*, **161**, article no: 111091 [no pagination].

Murashige, T. and Skoog, F., 1962. A Revised Medium for Rapid Growth and Bio Assays with Tobacco Tissue Cultures. *Physiologia Plantarum*, 15(3), pp.473-497.

Myers, P. 2008. Synteny: Inferring ancestral genomes. *Nature Education* **1**(1), p.47.

Narsai, R., Law, S. R., Carrie, C., Xu, L., and Whelan, J. 2011. In-depth temporal transcriptome profiling reveals a crucial developmental switch with roles for RNA processing and organelle metabolism that are essential for germination in Arabidopsis. *Plant physiology*, **157**(3), pp.1342–1362.

Nath, K., Phee, B. K., Jeong, S., Lee, S. Y., Tateno, Y., Allakhverdiev, S. I., Lee, C. H., and Nam, H. G. 2013. Age-dependent changes in the functions and compositions of photosynthetic complexes in the thylakoid membranes of Arabidopsis thaliana. *Photosynthesis research*, **117**, pp.547–556.

Negin, B., Yaaran, A., Kelly, G., Zait, Y., and Moshelion, M. 2019. Mesophyll Abscisic Acid Restrains Early Growth and Flowering But Does Not Directly Suppress Photosynthesis. *Plant physiology*, **180**(2), pp.910–925.

Nguyen, T. P., Cueff, G., Hegedus, D. D., Rajjou, L., and Bentsink, L. 2015. A role for seed storage proteins in Arabidopsis seed longevity. *Journal of experimental botany*, **66**(20), pp.6399–6413.

Parent, J. S., Lepage, E., and Brisson, N. 2011. Divergent roles for the two PolI-like organelle DNA polymerases of Arabidopsis. *Plant physiology*, **156**(1), pp.254–262.

Pérez Di Giorgio, J. A., Lepage, É., Tremblay-Belzile, S., Truche, S., Loubert-Hudon, A., and Brisson, N. 2019. Transcription is a major driving force for plastid genome instability in Arabidopsis. *PloS one*, **14**(4), article no.: e0214552 [no pagination].

Porter, R. H., Durrell, M., & Romm, H. J. 1947. The Use of 2,3,5-Triphenyl-Tetrazoliumchloride as a Measure of Seed Germinability. *Plant Physiology*, **22**(2), pp.149–159.

Preston, J. C., and Hileman, L. C. 2013. Functional Evolution in the Plant SQUAMOSA-PROMOTER BINDING PROTEIN-LIKE (SPL) Gene Family. *Frontiers in plant science*, **4**(80), [no pagination].

Prieto-Dapena, P., Castano, R., Almoguera, C. and Jordano, J. 2006 Improved resistance to controlled deterioration in transgenic seeds. *Plant Physiology* **142**(3), pp.1102–1112

Prikryl, J., Watkins, K. P., Friso, G., van Wijk, K. J., and Barkan, A. 2008. A member of the Whirly family is a multifunctional RNA- and DNA-binding protein that is essential for chloroplast biogenesis. *Nucleic acids research*, **36**(16), pp.5152–5165.

Provart, N and Zhu, T. 2003. A Browser-based Functional Classification SuperViewer for Arabidopsis Genomics. Currents in Computational Molecular Biology, 271-272.

Qiagen (User-Developed Protocol) 2006. Isolation of total RNA from plants [Online] Available: http://www.qiagen.com/resources/download.aspx?id=1333be8d-6568-4577-a1eb-8c6029d5dcc0&lang=en&ver=1 [Accessed: 20/10/ 2019].

Razak, N. B. 2019. *The characterisation of WHIRLY1 functions in chloroplast development*. Doctor of Philosophy, University of Leeds.

Redza-Dutordoir, M., and Averill-Bates, D. A. 2016. Activation of apoptosis signalling pathways by reactive oxygen species. *Biochimica et biophysica acta*, **1863**(12), pp.2977–2992.

Ren, Y., Li, M., Wang, W., Lan, W., Schenke, D., Cai, D., and Miao, Y. 2022. MicroRNA840 (MIR840) accelerates leaf senescence by targeting the overlapping 3'UTRs of PPR and WHIRLY3 in Arabidopsis thaliana. *The Plant journal : for cell and molecular biology*, **109**(1), pp.126–143.

Ren, Y., Li, Y., Jiang, Y., Wu, B., and Miao, Y. 2017. Phosphorylation of WHIRLY1 by CIPK14 Shifts Its Localization and Dual Functions in Arabidopsis. *Molecular plant*, **10**(5), pp.749–763.

Rhoades, M. W., Reinhart, B. J., Lim, L. P., Burge, C. B., Bartel, B., and Bartel, D. P. 2002. Prediction of plant microRNA targets. *Cell*, **110**(4), pp.513–520.

Romeuf, I., Tessier, D., Dardevet, M., Branlard, G., Charmet, G. and Ravel, C. 2010. wDBTF: an integrated database resource for studying wheat transcription factor families. *BMC Genomics*, **11**, article no: 185 [no pagination].

Schmidt EM, Zhang J, Zhou W, Chen J, Mohlke KL, Chen YE, Willer CJ. 2015. GREGOR: evaluating global enrichment of trait-associated variants in epigenomic features using a systematic, data-driven approach. *Bioinformatics*, 31 (16), 2601-2606.

Schneider, C.A.; Rasband, W.S. and Eliceiri K.W., NIH Image to ImageJ: 25 years of image analysis. Nature Methods, 2012. 9: p. 671.Vishwanath, S.J., Domergue, F., and Rowland, O. 2014. Seed Coat Permeability Test: Tetrazolium Penetration Assay. Bioprotocol, **4** (13). [Online] Available: https://bio-protocol.org/e1173 [Accessed: 23/02/2021].

Sels, J., Mathys, J., De Coninck, B. M., Cammue, B. P., and De Bolle, M. F. 2008. Plant pathogenesis-related (PR) proteins: a focus on PR peptides. *Plant physiology and biochemistry : PPB*, **46**(11), pp.941–950.

Simossis, V. A., and Heringa, J. 2005. PRALINE: a multiple sequence alignment toolbox that integrates homology-extended and secondary structure information. *Nucleic acids research*, **33**(Web Server issue), pp.289–294.

Spetea, C., Rintamäki, E., and Schoefs, B. 2014. Changing the light environment: chloroplast signalling and response mechanisms. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, **369**(1640), article no:20130220 [no pagination].

Stratonovitch, P., and Semenov, M. A. 2015. Heat tolerance around flowering in wheat identified as a key trait for increased yield potential in Europe under climate change. *Journal of Experimental Botany*, **66**(12), pp.3599–3609.

Tada, Y., Spoel, S. H., Pajerowska-Mukhtar, K., Mou, Z., Song, J., Wang, C., Zuo, J., and Dong, X. 2008. Plant immunity requires conformational changes [corrected] of NPR1 via S-nitrosylation and thioredoxins. *Science*, **321**(5891), pp.952–956.

Tamaoki, D., Seo, S., Yamada, S., Kano, A., Miyamoto, A., Shishido, H., Miyoshi, S., Taniguchi, S., Akimitsu, K., and Gomi, K. 2013. Jasmonic acid and salicylic acid activate a common defense system in rice. *Plant signaling & behavior*, **8**(6), article no: e24260 [no pagination].

Tanaka, Y., Sugano, S. S., Shimada, T., and Hara-Nishimura, I. 2013. Enhancement of leaf photosynthetic capacity through increased stomatal density in Arabidopsis. *The New phytologist*, **198**(3), pp.757–764.

Taylor, R.E., West, C.W., Foyer, C. H. 2022. WHIRLY Protein Functions in Plants. *Food and Energy Security*, article no: e379. [Online] Available: https://doi.org/10.1002/fes3.379

The R Core Team, The R Project. 2018: University of Auckland, New Zealand.

Till, B. J., Reynolds, S. H., Greene, E. A., Codomo, C. A., Enns, L. C., Johnson, J. E., Burtner, C., Odden, A. R., Young, K., Taylor, N. E., Henikoff, J. G., Comai, L. and Henikoff,

S. 2003. Large-Scale Discovery of Induced Point Mutations with High-Throughput TILLING. *Genome Research*, 13, p. 524-530.

Topham, A. T., Taylor, R. E., Yan, D., Nambara, E., Johnston, I. G., and Bassel, G. W. 2017. Temperature variability is integrated by a spatially embedded decision-making center to break dormancy in *Arabidopsis* seeds. *Proceedings of the National Academy of Sciences of the United States of America*, **114**(25), pp.6629–6634.

Travella, S., Klimm, T. E., and Keller, B. 2006. RNA interference-based gene silencing as an efficient tool for functional genomics in hexaploid bread wheat. *Plant Physiology*, **142**(1), pp.6–20.

Tunc-Ozdemir, M., Miller, G., Song, L., Kim, J., Sodek, A., Koussevitzky, S., Misra, A. N., Mittler, R., and Shintani, D. 2009. Thiamin confers enhanced tolerance to oxidative stress in Arabidopsis. *Plant Physiology*, **151**(1), pp.421–432.

Uddling, J., Gelang-Alfredsson, J., Piikki, K. and Pleijel, H., 2007. Evaluating the relationship between leaf chlorophyll concentration and SPAD-502 chlorophyll meter readings. *Photosynthesis Research*, 91(1), pp.37-46.

Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A., and Speleman, F. 2002. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome biology*, **3**(7), article no: RESEARCH0034.

Vanhaeren, H., Gonzalez, N., and Inzé, D. 2015. A Journey Through a Leaf: Phenomics Analysis of Leaf Growth in Arabidopsis thaliana. *The Arabidopsis Book*, **13**, article no.: e0181 [no pagination].

Verma, P., Kaur, H., Petla, B. P., Rao, V., Saxena, S. C., and Majee, M. 2013. PROTEIN L-ISOASPARTYL METHYLTRANSFERASE2 is differentially expressed in chickpea and enhances seed vigor and longevity by reducing abnormal isoaspartyl accumulation predominantly in seed nuclear proteins. *Plant physiology*, **161**(3), pp.1141–1157.

Vishwanath, S. J., Domergue, F. and Rowland, O. 2014. Seed Coat Permeability Test: Tetrazolium Penetration Assay. *Bio-protocol* **4**(13), article no.: e1173 [no pagination]. [Online] Available: https://bio-protocol.org/e1173 [Accessed: 23/02/2021].

Vishwanath, S. J., Kosma, D. K., Pulsifer, I. P., Scandola, S., Pascal, S., Joubès, J., Dittrich-Domergue, F., Lessire, R., Rowland, O., and Domergue, F. 2013. Suberin-associated fatty alcohols in Arabidopsis: distributions in roots and contributions to seed coat barrier properties. *Plant physiology*, **163**(3), pp.1118–1132.

Walkowiak, S., Gao, L., Monat, C., Haberer, G., Kassa, M. T., Brinton, J., Ramirez-Gonzalez, R. H., Kolodziej, M. C., Delorean, E., Thambugala, D., Klymiuk, V., Byrns, B., Gundlach, H., Bandi, V., Siri, J. N., Nilsen, K., Aquino, C., Himmelbach, A., Copetti, D., Ban, T., … Pozniak, C. J. 2020. Multiple wheat genomes reveal global variation in modern breeding. *Nature*, **588**(7837), pp.277–283.

Wang, W., Li, K., Yang, Z., Hou, Q., Zhao, W. W., and Sun, Q. 2021. RNase H1C collaborates with ssDNA binding proteins WHY1/3 and recombinase RecA1 to fulfill the DNA damage repair in Arabidopsis chloroplasts. *Nucleic Acids Research*, **49**(12), pp.6771-6787.

Wardlaw, I.F., and Moncur, L. 1995. The Response of Wheat to High Temperature Following Anthesis. I. The Rate and Duration of Kernel Filling. *Australian Journal of Plant Physiology*, **22**, pp.391-397.

Waterworth, W. M., Bray, C. M., and West, C. E. 2015. The importance of safeguarding genome integrity in germination and seed longevity. *Journal of Experimental Botany*, **66**(12), pp.3549–3558.

Waterworth, W. M., Bray, C. M., and West, C. E. 2019. Seeds and the Art of Genome Maintenance. *Frontiers in plant science*, **10**, p.706.

Waterworth, W. M., Footitt, S., Bray, C. M., Finch-Savage, W. E., and West, C. E. 2016. DNA damage checkpoint kinase ATM regulates germination and maintains genome stability in seeds. *Proceedings of the National Academy of Sciences of the United States of America*, **113**(34), pp.9647–9652.

Wen-jun, S. and Forde, B., 1989. Efficient transformation of Agrobacterium spp. by high voltage electroporation. *Nucleic Acids Research*, 17(20), pp.8385-8385.

Weraduwage, S. M., Chen, J., Anozie, F. C., Morales, A., Weise, S. E., and Sharkey, T. D. 2015. The relationship between leaf area growth and biomass accumulation in Arabidopsis thaliana. *Frontiers in plant science*, **6**, p.167.

Wharton, M. J. 1955. The use of tetrazolium test for determining the viability of seeds of the genus Brassica. *Proc. Int. Seed Test Assoc*. **20**(1), pp.81-88.

Winfield, M. O., Lu, C., Wilson, I. D., Coghill, J. A., and Edwards, K. J. 2009. Cold- and light-induced changes in the transcriptome of wheat leading to phase transition from vegetative to reproductive growth. BMC plant biology, **9**, p.55.

Xiong, J. Y., Lai, C. X., Qu, Z., Yang, X. Y., Qin, X. H., and Liu, G. Q. 2009. Recruitment of *Atwhy1* and *Atwhy3* by a distal element upstream of the kinesin gene AtKP1 to mediate transcriptional repression. *Plant Molecular Biology*, **71**(4-5), pp.437–449.

Xiong, J. Y., Lai, C. X., Qu, Z., Yang, X. Y., Qin, X. H., and Liu, G. Q. 2009. Recruitment of AtWHY1 and AtWHY3 by a distal element upstream of the kinesin gene AtKP1 to mediate transcriptional repression. *Plant Molecular Biology*, **71**(4-5), pp.437–449

Xu, W., Li, R., Zhang, N., Ma, F., Jiao, Y., and Wang, Z. 2014. Transcriptome profiling of Vitis amurensis, an extremely cold-tolerant Chinese wild Vitis species, reveals candidate genes and events that potentially connected to cold stress. *Plant molecular biology*, **86**(4-5), pp.527–541.

Yan, Y., Liu, W., Wei, Y., and Shi, H. 2020. MeCIPK23 interacts with Whirly transcription factors to activate abscisic acid biosynthesis and regulate drought resistance in cassava. *Plant Biotechnology Journal*, **18**(7), pp.1504–1506.

Yang, X.-Y., Chen, Z.-W., Xu, T., Qu, Z., Pan, X.-D., Qin, X.-H., Ren, D.-T. and Liu, G.-Q. 2011. Arabidopsis Kinesin KP1 Specifically Interacts with VDAC3, a Mitochondrial Protein, and Regulates Respiration during Seed Germination at Low Temperature. *The Plant Cell*, **23**(3), pp.1093-1106.

Yobi, A., Schlauch, K. A., Tillett, R. L., Yim, W. C., Espinoza, C., Wone, B. W., Cushman, J. C., and Oliver, M. J. 2017. Sporobolus stapfianus: Insights into desiccation tolerance in the resurrection grasses from linking transcriptomics to metabolomics. *BMC plant biology*, **17**(1), p.67.

Yoo, H. H., Kwon, C., Lee, M. M., and Chung, I. K. 2007. Single-stranded DNA binding factor AtWHY1 modulates telomere length homeostasis in Arabidopsis. *The Plant journal: for cell and molecular biology*, **49**(3), pp.442–451.

Zampini, É., Lepage, É., Tremblay-Belzile, S., Truche, S., and Brisson, N. 2015. Organelle DNA rearrangement mapping reveals U-turn-like inversions as a major source of genomic instability in Arabidopsis and humans. *Genome research*, **25**(5), pp.645–654.

Zhao, S.Y., Wang, G.D., Zhao, W.Y., Zhang, S., Kong, F.-Y., Dong, X.C. and Meng, Q.-W. 2018. Overexpression of tomato WHIRLY protein enhances tolerance to drought stress and resistance to *Pseudomonas solanacearum* in transgenic tobacco. *Biologia Plantarum* **62**(1), pp.55–68.

Zhao, X., Chen, T., Feng, B., Zhang, C., Peng, S., Zhang, X., Fu, G. and Tao, L., 2017. Non-photochemical Quenching Plays a Key Role in Light Acclimation of Rice Plants Differing in Leaf Color. *Frontiers in Plant Science*, **7**, article no: 1968 [no pagination].

Zhuang, K., Gao, Y., Liu, Z., Diao, P., Sui, N., Meng, Q., Meng, C., and Kong, F. 2020a. WHIRLY1 Regulates HSP21.5A Expression to Promote Thermotolerance in Tomato. *Plant & Cell Physiology*, **61**(1), pp.169–177.

Zhuang, K., Kong, F., Zhang, S., Meng, C., Yang, M., Liu, Z., Wang, Y., Ma, N., and Meng, Q. 2019. Whirly1 enhances tolerance to chilling stress in tomato via protection of photosystem II and regulation of starch degradation. *The New phytologist*, **221**(4), pp.1998–2012.

Zhuang, K., Wang, J., Jiao, B., Chen, C., Zhang, J., Ma, N., and Meng, Q. 2020b. WHIRLY1 maintains leaf photosynthetic capacity in tomato by regulating the expression of RbcS1 under chilling stress. *Journal of Experimental Botany*, **71**(12), pp.3653–3663.

10 APPENDIX: ARABIDOPSIS MUTANT GENOTYPING

10.1 IDENTIFICATION OF THE *ATWHY1* MUTANTS

The mutant Arabidopsis genotypes, *Atwhy1 Atwhy3* and *Atwhy1why3* were determined using PCR with specific primers designed to identify the presence of WHY1 and WHY3, as described in Materials and Methods: 3.2.3. Following the establishment of the primer efficacy, the nature of the mutations in each line was verified. The wild-type AtWHY1 primers for genotyping are shown in Figure 71a, whilst the mutant At*why1* primers are shown in Figure 71b. These results are summarised in Table 13 which was used to determine which mutants to carry forward for future experiments. It is predicted that the wild-type and *Atwhy3* plants will only have a band at 1100 bp in Figure 71a and no band in Figure 71b. Whilst, the *Atwhy1* and *Atwhy1why3* mutant lines will have no band in Figure 71a and a band in Figure 71b at 750 bp corresponding to the At*why1* mutant gene. Further more specific At*WHY3* primers for the insertion mutation were designed and a similar experiment was carried out on the At*why3* mutant lines in Figure 72.

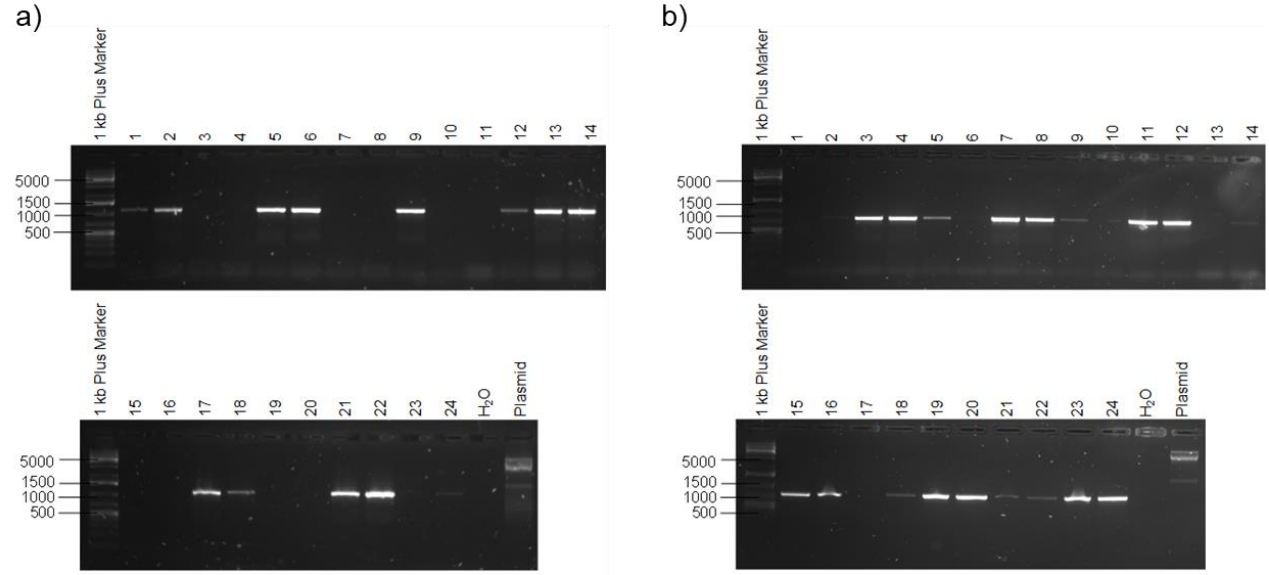


Figure 71: A UV light photograph of 1.4% agarose gel from PCR to determine a) wild-type *WHY1* genotypes of Arabidopsis plants using the *AtWHY1* left and right primers to give a 1100 bp band in wild-type plants; and b) mutant *why1* genotypes of Arabidopsis plants using the At*WHY1* right primer and the LBB1-3 left border primer. A positive control of the *WHY1* primer was used along with a negative H_2O control. Sample numbers of individual plants are indicated above the gel corresponding to the column below. The numbers at the side are the base pairs of the DNA ladder.

A strong band was predicted to be visible at the 1100 bp mark for wild-type At*WHY1* in Figure 71. This is seen in columns 2, 5, 6, 9, 13, 14, 17, 21 and 22. Samples 1, 12, 18 and 24 had visible bands, however these are not as strong as the predicted wild-type band. Furthermore, samples, 3, 4, 7, 8, 10, 11, 15, 16, 19, 20 and 23 had no wild-type band [Figure 71a]. These same samples were tested for mutant *why1* with the LBB1-3 primer which gave a predicted 750 bp band for mutant lines [Figure 71b]. The samples with the predicted At*why1* mutant 750 bp band were 3, 4, 7, 8, 11, 12, 15, 16, 19, 20, 23 and 24 [Figure 71b]. However, as before some lines had a weak band at 750 bp these were samples 2, 5, 9, 14, 18, 21 and 22 and other samples had no band at 750 bp, these were 1, 6, 10, 13 and 17. The information obtained in the gel electrophoresis studies was summarised in Table 13. A visibly strong band indicated that the sample had the specific sequence identified by the primers; the opposite is true for no band visible. The *Atwhy1* lines that were highlighted in green in Table 13 matched to the genotype and were selected for further analysis and grown for seed production for the phenotyping studies.

Table 13: Table of genotypes match to genotype observed of all samples. Correlations to predicted genotype were from the labelled tube that they originated and written in green if they match to reported genotype, and red if they do not match to the reported genotype.

Sample number	Predicted genotype	Band Visible?		Match to predicted genotype?
		WT WHY1	Mutant *why1*	
1	WT	Weak	No	No
2	WT	Strong	Weak	No
3	WT	No	Strong	Yes
4	*Atwhy1*	No	Strong	Yes
5	*Atwhy3*	Strong	Weak	No
6	*Atwhy3*	Strong	No	Yes
7	*Atwhy1why3*	No	Strong	Yes
8	*Atwhy1why3*	No	Strong	Yes
9	WT	Strong	Weak	No
10	WT	No	Weak	No
11	*Atwhy1*	No	Strong	Yes
12	*Atwhy1*	Weak	Strong	No
13	*Atwhy3*	Strong	No	Yes
14	*Atwhy3*	Strong	Weak	No
15	*Atwhy1why3*	No	Strong	Yes
16	*Atwhy1why3*	No	Strong	Yes
17	WT	Strong	No	Yes
18	WT	Weak	Weak	No
19	*Atwhy1*	No	Strong	Yes
20	*Atwhy1*	No	Strong	Yes
21	*Atwhy3*	Strong	Weak	No
22	*Atwhy3*	Strong	Weak	No
23	*Atwhy1why3*	No	Strong	Yes
24	*Atwhy1why3*	Weak	Strong	No

10.2 IDENTIFICATION OF THE *ATWHY3* MUTANTS

To identify plants with mutated AtWhy3, plants 6, 8 and 16 which were found to match their predicted genotype based on the WHY1 PCR genotyping [Table 13] were next analysed using At*WHY3* primers as shown in Figure 72 to determine whether they are the At*why3* mutant lines. A predicted 925 bp band is seen in samples of all three plants [Figure 72]. Additionally, a faint band was visible in the plasmid positive control therefore these plants were sent off for sequencing.

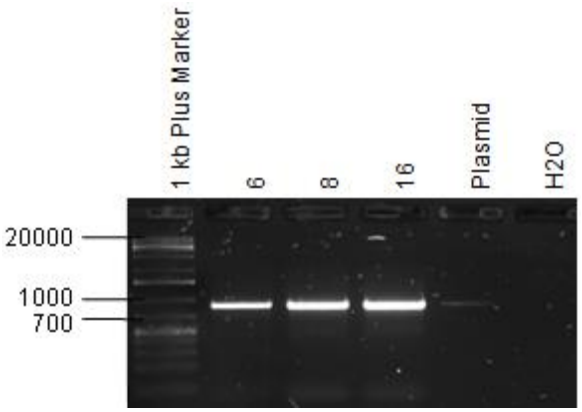


Figure 72: UV light photograph of 1.4% agarose gel from the PCR to determine mutant *why3* genotypes in selected samples. This used *AtWHY3* specific left and right primers to give a predicted 925 bp band in *Atwhy3* mutant lines. A positive control of the *WHY3* primer was used along with a negative H_2O control. Sample numbers of individual plants are indicated above the gel corresponding to the column below.

In order to genotype the *Atwhy3* and *Atwhy1why3* mutant plants the single nucleotide polymorphism of TGG to TGA at position W138 relative to the initial ATG start codon which created a premature stop codon was identified through sequencing of the plants. A multiple sequence analysis highlighted a sample of each of *Atwhy3* and *Atwhy1why3* plants compared to the Col-0 WT [Figure 73]. The presence of the TGA codon identified the plants as mutants in *Atwhy3*. These plants were then grown for seed production and used for the phenotyping studies in comparison to the WT.

```
WHY3                          ------------------------------------------------------------  0
why1-1why1-3-why3_C12.ab1     NNNNNNNNNNNTTTGTAAGANNNTAAATATNATTGGGTAAAAAACGAATTAACAGAAAGT  60
why1-3-why3_A12.ab1           -NNNNNNNNNNNTTGTAANANGNTAAATATGATTGGGTAAAAAACGAATTAACAGAAAGT  59
Col-0-why3_E12.ab1            -NNNNNNNNNNTTTGTAAGANNNTAAATATNATTGGGTAAAAAACGAATTAACAGAAAGT  59


WHY3                          -------------------------ATCGTAATCCCGCGGTGAAGAAGACGAAGAAGCATT  36
why1-1why1-3-why3_C12.ab1     ATTCTCTCCTAAACCCTTTTCGTCATCGTAATCCCGCGGTGAAGAAGACGAAGAAGCATT  120
why1-3-why3_A12.ab1           ATTCTCTCCTAAACCCTTTTCGTCATCGTAATCCCGCGGTGAAGAAGACGAAGAAGCATT  119
Col-0-why3_E12.ab1            ATTCTCTCCTAAACCCTTTTCGTCATCGTAATCCCGCGGTGAAGAAGACGAAGAAGCATT  119
                                                  ************************************


WHY3                          GAGAGTGAAAAGAAAAACAATGTCGCAGCTGTTATCTTCTCCTCCAATGGCGGTTTTCTC  96
why1-1why1-3-why3_C12.ab1     GAGAGTGAAAAGAAAAACAATGTCGCAGCTGTTATCTTCTCCTCCAATGGCGGTTTTCTC  180
why1-3-why3_A12.ab1           GAGAGTGAAAAGAAAAACAATGTCGCAGCTGTTATCTTCTCCTCCAATGGCGGTTTTCTC  179
Col-0-why3_E12.ab1            GAGAGTGAAAAGAAAAACAATGTCGCAGCTGTTATCTTCTCCTCCAATGGCGGTTTTCTC  179
                              ************************************************************


WHY3                          CAAAACCTTCATAAACCATAAGTTTTCAGATGCTCGTTTTCTCTCTTCTCACTCAATTCT  156
why1-1why1-3-why3_C12.ab1     CAAAACCTTCATAAACCATAAGTTTTCAGATGCTCGTTTTCTCTCTTCTCACTCAATTCT  240
why1-3-why3_A12.ab1           CAAAACCTTCATAAACCATAAGTTTTCAGATGCTCGTTTTCTCTCTTCTCACTCAATTCT  239
Col-0-why3_E12.ab1            CAAAACCTTCATAAACCATAAGTTTTCAGATGCTCGTTTTCTCTCTTCTCACTCAATTCT  239
                              ************************************************************


WHY3                          CACCTCCGGTGGATTCGCCGGAAAAATTATTCCTTTAAAACCGACGGCGAGGTTGAAATT  216
why1-1why1-3-why3_C12.ab1     CACCTCCGGTGAATTCGCCGGAAAAATTATTCCTTTAAAACCGACGGCGAGGTTGAAATT  300
why1-3-why3_A12.ab1           CACCTCCGGTGAATTCGCCGGAAAAATTATTCCTTTAAAACCGACGGCGAGGTTGAAATT  299
Col-0-why3_E12.ab1            CACCTCCGGTGGATTCGCCGGAAAAATTATTCCTTTAAAACCGACGGCGAGGTTGAAATT  299
                              *********** ********************************************


WHY3                          AACGGTGAAGTCTCGACAAAGTGATTACTTTGAGAAACAGAGGTTTGGTGACTCGTCTTC  276
why1-1why1-3-why3_C12.ab1     AACGGTGAAGTCTCGACAAAGTGATTACTTTGAGAAACAGAGGTTTGGTGACTCGTCTTC  360
why1-3-why3_A12.ab1           AACGGTGAAGTCTCGACAAAGTGATTACTTTGAGAAACAGAGGTTTGGTGACTCGTCTTC  359
Col-0-why3_E12.ab1            AACGGTGAAGTCTCGACAAAGTGATTACTTTGAGAAACAGAGGTTTGGTGACTCGTCTTC  359
                              ************************************************************


WHY3                          TTCGCAAAACGGTAGGGTTTTCACTCATTGTTTGAGCTTTGAAAAATGGGAAGCTTTGAA  336
why1-1why1-3-why3_C12.ab1     TTCGCAAAACGGTAGGGTTTTCACTCATTGTTTGAGCTTTGAAAAATGGGAAGCTTTGAA  420
why1-3-why3_A12.ab1           TTCGCAAAACGGTAGGGTTTTCACTCATTGTTTGAGCTTTGAAAAATGGGAAGCTTTGAA  419
Col-0-why3_E12.ab1            TTCGCAAAACGGTAGGGTTTTCACTCATTGTTTGAGCTTTGAAAAATGGGAAGCTTTGAA  419
                              ************************************************************


WHY3                          ATTCAGAAAGTTTGAAATTTTGTTTGTTTGTTTGTTTTGGTTTTATAGCTGAAGTATCTT  396
why1-1why1-3-why3_C12.ab1     ATTCAGAAAGTTTGAAATTTNGNNNGTTTGTTTGTTTTGGTTTTATAGCTGAAGTATCTT  480
why1-3-why3_A12.ab1           NTTCANANNNNNNGANATNNNNNNNGNNNGTTTGNTNTGNNTNNNNNNNNNNNNNCTNTN  479
Col-0-why3_E12.ab1            ATTCANAAAGTTTGAANTTTTNNNNGNNTGTTTGNTTTGNTTTNATAGNTNANNTANCNT  479
                               **** *      **  *       *   ***** * **  *
```

Figure 73: Multiple sequence analysis of WHY3 genomic sequence compared to *Atwhy3*, *Atwhy1why3* and WT plant sequences. The ATG start codon was highlighted in blue, the presence of the TGA SNP was highlighted in red and the TGG in the WHY3 genomic sequence was highlighted in yellow. Asterisks identify the presence of the same nucleotide in all sequences (obtained through: https://www.ebi.ac.uk/Tools/msa/clustalo/).