The characterisation of WHIRLY1 functions in chloroplast development

Nurhayati Binti Razak

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

The University of Leeds
Faculty of Biological Sciences

April, 2019

The candidate confirms that the work submitted is his/her own and that appropriate credit has been given where reference has been made to the work of others.

Abstract

Chloroplast biogenesis requires coordinated expression of plastome and nuclear genes. The single-stranded DNA binding protein, WHIRLY1 (WHY1), which is localised in chloroplasts and nuclei has important but poorly characterised roles in this process. WHY1 functions in barley chloroplast development were assessed in the base, middle and tip sections of two RNAi-knockdown lines (W1-1 and W1-7) with less than 5% of the wild type protein. RNA-seq analysis revealed that transcripts encoding photosynthetic proteins were highly expressed in the embryos of the dry seeds of the W1-7 compared to the wild type. The greening of the developing leaves was delayed in the WHY1-deficient seedlings relative to the wild type, with slower pigment accumulation and attainment of photosynthetic capacity in the WHY1-deficient leaves. However, the leaves of all lines reached a similar stage of chloroplast development at 14 days after germination. Transcript and metabolite profiling analysis showed changes in RNA and amino acid metabolism, TCA cycle, photosynthesis and photorespiration, particularly in the basal sections of the WHY1-deficient leaves. The expression of the plastid-encoded ribosomal genes was greatly increased in the WHY1-deficient lines, including transcripts involved in RNA processing such as pentatricopeptide repeat proteins, redox-associated proteins and transcription factors of the MYB, bHLH and WRKY families. The levels of transcripts encoding FAR1, Val-tRNA synthetase and chloroplast 50S and 30S ribosomal subunits were significantly higher in the basal sections of the W1-7 leaves than the wild type. The WHY1-deficient leaves had twice the amount of plastid DNA as the wild type. Nevertheless, plastome-encoded transcripts and proteins were significantly lower than the wild type. Conversely, the levels of nuclear-encoded photosynthetic transcripts and proteins were significantly higher that the wild type. Developing WHY1-deficient leaves showed aberrant splicing of plastid ribosomal RNAs of 23S and 4.5 ribosomal RNAs. The Arabidopsis WHY1 protein interacted with the RH22, which is required for the splicing of chloroplast rRNAs. The LEA5 protein was also shown to interact with RH22 in the chloroplasts. WHY1 therefore has multiple roles in chloroplasts. In particular, plastid-encoded ribosomal transcripts are not effectively translated into ribosomal subunits in the absence of WHY1 during early leaf development. WHY1 is required for the transcription and translation of plastome genes that are required for the transition from plastids to chloroplasts in the developing barley leaf.

I dedicate this thesis to Mariah Mohd, someone who I called 'mother' and my other family members for their generous support and love because they are the reason I have completed this journey.

Acknowledgements

In the name of Allah, the Most Gracious, the Most Merciful. Praise to Allah, praise to Him for giving me strength and patience to complete this study.

My deepest appreciations go to my supervisor, Prof Christine Foyer, for her guidance, encouragement and knowledge have given me forte to finish my PhD. I am thankful for her patience, attention and immeasurable amount of time to help me complete this study. I also thank Dr Chris West and Prof Alison Baker who helped me a lot especially in my first year of PhD.

I would like to express my gratitude to Barbara Karpinska who became my closest friend. Her invaluable helps and companionship are highly appreciated in hard and good times. I also particularly thank James Cooper for a number of great discussions. Thanks for that!

A special thanks to Dorothee Wonthzee from Max Planck Institute for Plant Breeding Research, Cologne, Germany, for RNA-seq analysis, Michael Wilson for helping me in bioinformatics analysis and Dr Rob Hancock, Dr Pete Hadley, Katie Schulz and Diane for their helps at James Hutton Institute, Dundee, Scotland.

Thanks a lot to all my lab mates both past and present, Sarah, Brwa, Gloria, Henrick, Yuosef, Ambra, Rachel, and Rakesh for being great colleagues to work with. Not to forget my closest friends for good laugh, love and talk; Kaisan, Azuwa, Syikin, Fatin and to all FBS Malaysian mates especially to Jiha and Jiji.

Special appreciation to the Majlis Amanah Rakyat (MARA), or the Council of Trust for the People, an agency under the Ministry of Rural and Regional Development Malaysia for the 4 years PhD funding at the University of Leeds. Thank you!

Table of Contents

List of Abbreviations

ATP	adenosine triphosphate
ABA	abscisic acid
PR	pathogenesis-relate
Fm	Maximal fluorescence
Fv	Variable fluorescence
BiFc	Bimolecular fluorescence complementation
TBT	Tris-buffered tween
DNA	Deoxyribonucleic acid
RH22	DEA (D/H)-box RNA helicase 22
ROS	reactive oxygen species
TOC	Translocon at the outer envelope membrane
TOM	Translocase at the outer membrane
TIC	Translocon at the inner membrane
TIM	Translocase at the inner membrane
LEA	late embryogenesis abundant
RNA	Ribonucleic acid
PCR	polymerase chain reaction
qRT-PCR	quantitative reverse transcription PCR
RH22	DEA (D/H)-box RNA helicase 22
WHY1	WHIRLY1
W1-7	WHIRLY1 RNAi line 7
W1-1	WHIRLY1 RNAi line 1
WT	wild type
GFP	green fluorescent protein
YFP	yellow fluorescent protein

List of Tables

List of Figures

Chapter 1

1.1 Introduction

Plant cells are organised into compartments that have different functions, such as chloroplasts and mitochondria. The role of the chloroplast is to generate reducing power (e.g. for ferredoxins and nicotinamide adenine dinucleotide phosphate hydrogen (NADPH) and adenosine triphosphate (ATP), which are used to drive metabolism and maintain cell functions as well as to generate biomass for plant growth and development. Mitochondria use the carbon fixed in photosynthesis to generate energy through respiration. Chloroplasts and mitochondria are thus the two 'energy converting' organelles of the cells that house the processes that are essential to plant life such as photosynthesis and respiration. Photosynthesis supplies substrates for mitochondrial respiration, and mitochondrial metabolism is important in maintaining photosynthetic carbon assimilation. In addition, mitochondrial respiration can protect photosynthesis against light-induced damage, a process called photoinhibition by removing excess reducing equivalents produced by the chloroplasts (Blanco et al., 2014). These organelles are dependent on each other for the exchange of metabolites and energy (Blanco et al., 2014). Therefore, there is considerable metabolic communication between the mitochondria and chloroplasts.

Chloroplasts and mitochondria are also able to sense specific environmental stresses that can affect their functional activities, passing this information to the nucleus in order to modulate nuclear gene expression (Barajas-López et al., 2013). The signalling between the chloroplasts and mitochondria includes the metabolites and reactive oxygen species (ROS), as well as proteins, which are targeted at both compartments, depending on the stress condition (Suzuki et al., 2012). Exposure to environmental stresses can cause the processes that facilitate energy conversion in chloroplasts and mitochondria to malfunction,

leading to the generation of ROS and other signals that are transmitted to the nucleus to regulate gene expression.

1.2 The chloroplast

Chloroplasts belong to the plastid family of organelles (Figure 1.1). They generate chemical energy and assimilate carbon through the process known as photosynthesis. They synthesise carbon skeletons that are used to produce carbohydrates, amino acids, lipids and a wide range of secondary compounds including phytohormones, and they store starch and oils (Sakamoto et al., 2008). In addition, plastids sense environmental changes and are highly responsive to light fluctuations and other cues, such as gravity, pathogen infection and stomatal opening and closure (Sakamoto et al., 2008).

According to the endosymbiotic theory, chloroplasts are derived from an ancient cyanobacterium-like ancestor that was taken up by mitochondriate eukaryotic cells approximately 1 billion years ago (Zoschke and Bock, 2018). All plastids evolved from this single endosymbiotic event (Kovács-Bogdán et al., 2010). During endosymbiosis, most plastid genes were lost and transferred to the nucleus, undergoing a reduction in genome size (Bock and Timmis, 2008). As a result, today's plastids contain about 3000 proteins; however, only 50–200 of these genes are encoded in the plastid genome, as most of them are nuclear-encoded. The proteins are synthesised in the cytosol and post-translationally imported into the organelle (Leister, 2003, Martin et al., 2002).

Plastid genes are important for plant viability because they encode multiple components required for photosynthesis apparatus, such as: the large subunit of ribulose bisphosphate carboxylase (RuBisCO) and subunits of the thylakoid protein complexes involved in the light reactions of photosystems I and II (PSI, PSII); the cytochrome b6f complex (Cyt b_6f) and ATP synthase and chloroplast

gene expression systems, such as the bacterial-type ribonucleic acid (RNA) polymerase core subunits, rRNAs, tRNAs; and some ribosomal proteins (Allen et al., 2011, Green, 2011).

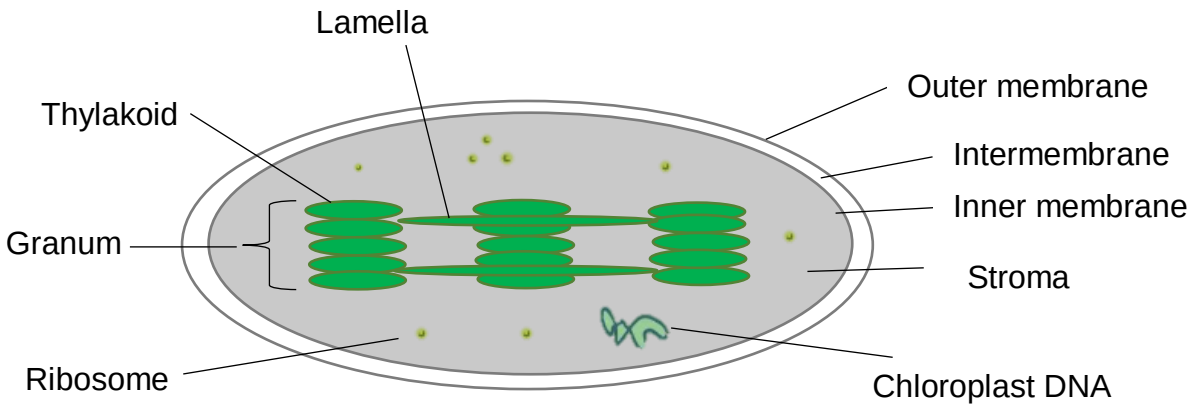


Figure 1.1: Diagram illustrating basic chloroplast structure.

The chloroplast structure consists of an outer membrane, intermembrane, inner membrane, stroma, thylakoid, ribosome, lamella and granum.

1.2.1 Chloroplast functions

Photosynthesis is often divided into two stages that are light-dependent (light reactions) and light-independent (dark reactions). The light-dependent reactions occur on the thylakoid membrane between PSII, Cyt b6f and PSI. Water is split into protons and electrons, and O_2 is produced as a by-product in this photosynthetic electron-transfer reaction (Ruban, 2014). The protons and electrons are moving across the thylakoid membrane to produce ATP and NADPH (Ruban, 2014). The ATP and NADPH is later consumed in the light-independent reaction that takes place in the chloroplast stroma, where CO2 is fixed by RuBisCO to generate sugars (Ruban, 2014). This carbohydrate is then exported to the cytosol or stored as starch. In addition, the core protein of PSII particularly DI protein is sensitive to light-induced damage (Theis and Schroda, 2016). The light-induced damage of PSII leads to the reduction in the photosynthetic capacity known as photo-inhibition (Tyystjarvi, 2013). The PSII

repair processes and recovery from photo-inhibition requires disassembly of damaged D1 and fast cycles of D1 turn-over (Theis and Schroda, 2016).

In addition to photosynthesis, chloroplasts are also involved in other metabolic processes such as amino acid biosynthesis, fatty acid biosynthesis, sulfur metabolism and also hormones synthesis such as abscisic acid, jasmonic acid and salicylic acid. Chloroplasts are also responsible for the synthesis of purine and pyrimidine bases, terpenoids, nitrogen and sulphur assimilation (Neuhaus and Emes, 2000). The chloroplast envelope contains metabolite transporter that help with the metabolic activities in chloroplasts and integrate with cellular compartments (Rolland et al., 2012). The interorganellar cooperation is necessary for lipid synthesis, photorespiration and other processes. Changes in biotic and abiotic stresses alter the redox status and excitation balance in PSI and PSII and affect light-harvesting complex II (LHCII) phosphorylation and antenna size in chloroplast (Bellafiore et al., 2005). Under severe stress such as high light, plants are unable to eliminate undesirable reactive oxygen species (ROS) due to the oxidation of proteins and lipids in PSII (Pospisil, 2016).

1.3 Chloroplast biogenesis

Chloroplast biogenesis is a light-triggered process that leads to the formation of fully-differentiated and photosynthetically-competent plastids. The chloroplast develops from small, undifferentiated, non-photosynthetic proplastids that are present in meristematic cells (Sakamoto et al., 2008). This process involves a rapid accumulation of chlorophyll and photosynthetic proteins (Harpster et al., 1984). During the conversion of proplastids into chloroplasts, there is a concomitant increased in the transcription and translation of photosynthesis-related proteins. Photosynthetic pigment–protein complexes are embedded in the thylakoid membranes, where they serve as highly developed sites for photosynthetic energy transduction, comprising PSI and PSII (Pribil et al., 2014).

Chloroplast biogenesis is tightly correlated with leaf development. In monocotyledonous plants, the leaves grow on a basal meristem. Hence, there is a developmental gradient of chloroplast biogenesis along the leaf blade (Pogson et al., 2015, Vothknecht and Westhoff, 2001). The base of the monocotyledonous leaf houses cells that contain proplastids, while the cells in the tips of the leaves house fully-developed chloroplasts (see Figure 1.2) (Vothknecht and Westhoff, 2001). The development of chloroplasts is triggered by photomorphogenesis, a process that requires light. In addition, the cotyledons open due to the inhibition of hypocotyl growth (Waters and Langdale, 2009).

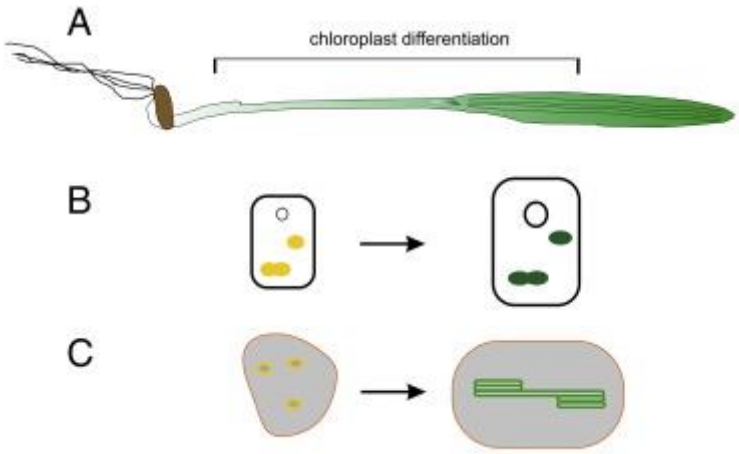


Figure 1.2: General overview of chloroplast biogenesis.

A) Chloroplast differentiation takes place at the base of the monocotyledonous leaves, where proplastids, located in meristematic cells, develop into photosynthetic green tissue. B) Proplastids (yellow dots) develop into photosynthetic chloroplasts (green dots). The nucleus is shown as an empty black circle. C) Later, prothylakoid vesicles (yellow) further develop into thylakoid vesicles (green). After Sun and Zerges (2015).

The establishment of functional chloroplasts is a complicated process, requiring the coordination in gene expression in both the nuclear and chloroplast genomes, followed by the assembly of proteins in response to developmental and environmental signals. Most chloroplast proteins are encoded by nuclear genes, translated in cytosol, and post-translationally imported into the stroma (Leister, 2003). Some intermediate precursors are translocated into, or across, the thylakoid membrane (Di Cola et al., 2005).

The establishment of functional chloroplasts requires the expression of plastid-encoded photosynthetic genes. The initiation of chloroplast gene expression depends on the expression of nuclear-encoded factors, such as the RNA polymerase sigma factors (SIGs) and polymerase-associated proteins (PAPs) (Kindgren and Strand, 2015). In addition, the establishment of fully-functional chloroplasts requires the exchange of information between the plastid and nuclear genomes, the import of nuclear-encoded proteins and the establishment of thylakoid networks embedded with photosynthetic electron transport complexes (Vothknecht and Westhoff, 2001).

Chloroplast gene expression and protein assembly requires components encoded by chloroplast genome and the nuclear genome (see Figure 1.3). Gene expression in the chloroplast and nucleus must be coordinated to achieve a balanced stoichiometric assembly (Nelson and Yocum, 2006). This requires extensive communication between the nucleus and chloroplast (Chan et al., 2010). The plastid-nucleus communication includes the bi-directional signalling pathways. The anterograde pathway involves signals arising in the nucleus travelling to the plastid. Conversely, the retrograde signalling pathways involves signals generated in the chloroplasts/plastids travelling to the nucleus (Pogson et al., 2015). These communication pathways are important for the plant developmental processes under ambient and environmental stress conditions (Kacprzak et al., 2019).

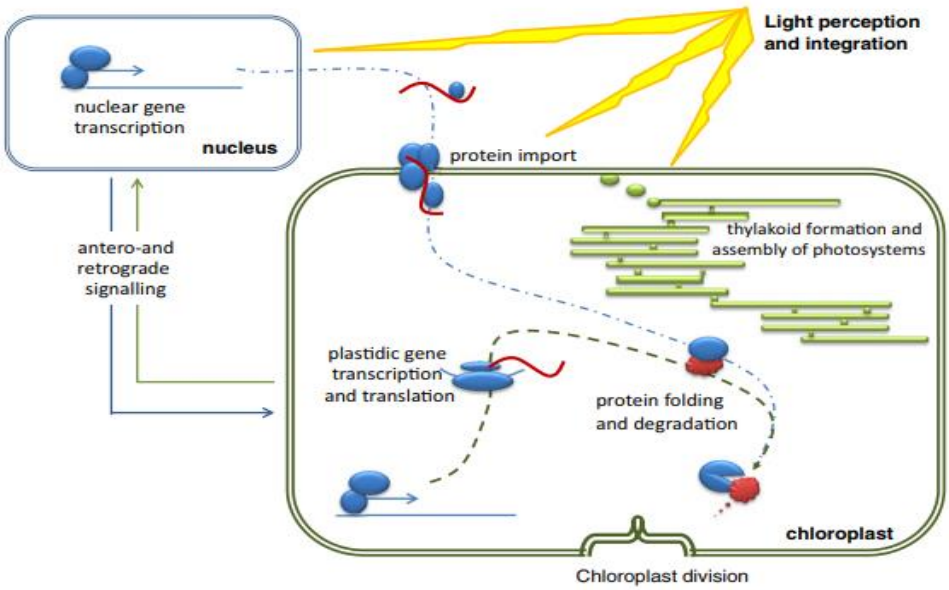


Figure 1.3: Communication and transport between chloroplast and nucleus that is important in chloroplast biogenesis and development.

The establishment of a functional chloroplast requires protein transcription, translation, import and turnover, as well as communication between the nucleus and chloroplasts. Orchestrated communication helps stoichiometric assembly between the nuclear- and plastid-encoded proteins with chlorophylls and carotenoids, which are important in reducing oxidative damage caused by photoreactive pigments. Moreover, the formation of thylakoids and the assembly of photosystems requires metabolite import and synthesis. This process required communication between the chloroplast and nucleus. After Pogson et al. (2015).

The establishment of photosynthesis in *Arabidopsis* cells in culture is controlled by a two-phase process that allows coordination of the activities of the nuclear and plastid genomes (Dubreuil et al., 2018). The first step occurs when light initiates changes in gene expression and the cellular metabolite profile (Dubreuil et al., 2018). The second phase, which is initiated by the activation of the chloroplast functions as a result of changes in nuclear gene expression, is required for the transition into fully-functional chloroplasts (Dubreuil et al., 2018).

1.4 Chloroplast gene expression

Due to the bacterial ancestry of chloroplasts, the plastids retain a prokaryotic-gene expression apparatus, including polycistronic transcripts that undergo post-transcriptional maturation steps, making chloroplast gene expression a complicated process (Yagi and Shiina, 2014). As a result of the polycistronic nature of primary transcripts, the control of chloroplast gene expression relies mainly on post-transcriptional regulation. The polycistronic RNAs are transcribed by plastid polymerases and undergo extensive post-transcriptional RNA processing, including 5' and 3' trimming, intercistronic cleavage, intron splicing and RNA editing, to produce mature, functional RNAs (Barkan, 2011, Börner et al., 2015, Stern et al., 2010, Lyska et al., 2013). Translation in plastids occurs on bacterial-type (70S ribosomes), using a set of tRNAs that are encoded by the plastid genome (Tiller and Bock, 2014, Sun and Zerges, 2015). The plastid ribosome consists of large (50S) and small (30S) multi-component ribosomal subunits (Yamaguchi and Subramanian, 2000). The summary of chloroplast transcription and maturation of chloroplast RNAs can be read here (see Figure 1.4).

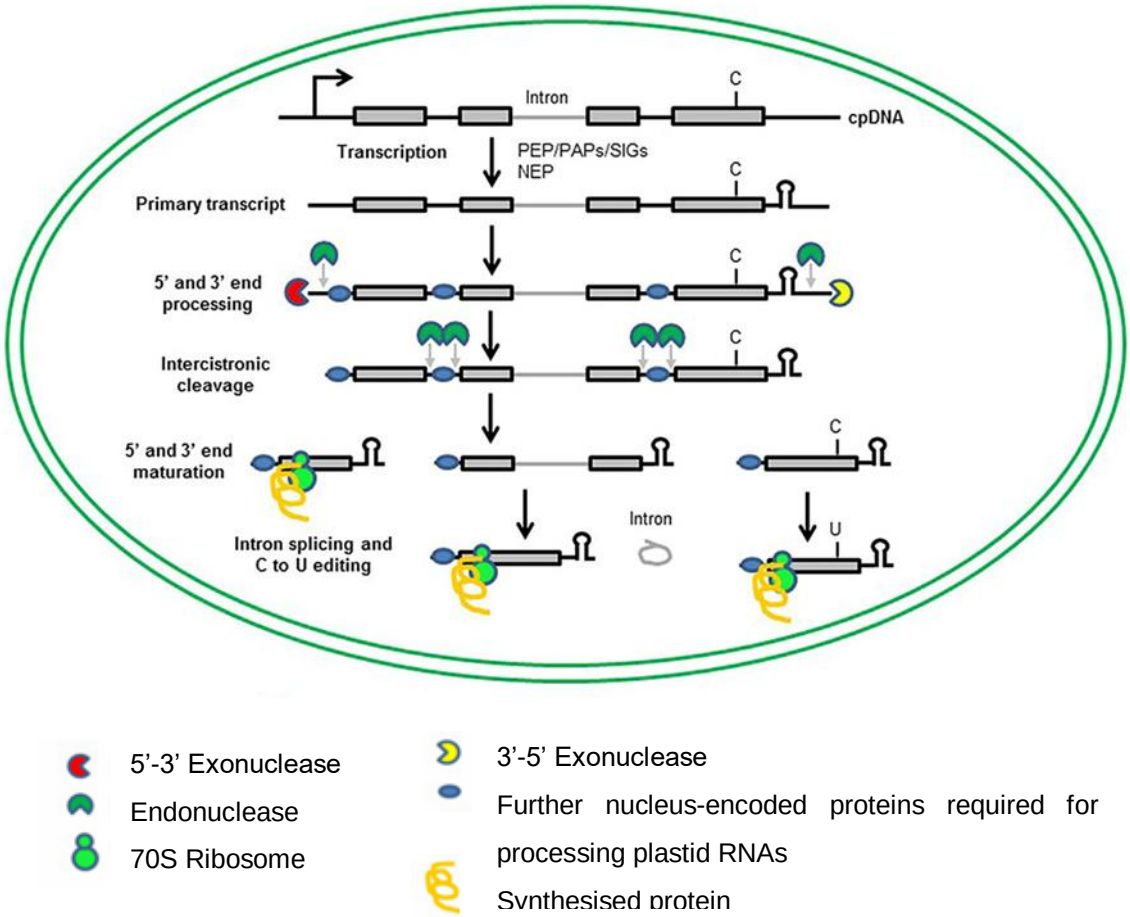


Figure 1.4: **Overview of chloroplast transcription and maturation of chloroplast RNAs.**

In prokaryotes, chloroplast genes are organised in operons and transcribed as polycistronic RNAs that undergo several post-transcriptional maturation steps. Chloroplast transcription is carried out by plastid-encoded polymerases (PEPs), together with one or two nucleus-encoded polymerases (NEPs) because genes encoded in chloroplast genomes are not sufficient to regulate chloroplast gene expression. Several maturation steps are encoded by the primary transcript that includes 5' and 3' trimming, intercistronic cleavage, 5' and 3' end maturation, intron splicing and RNA editing to produce mature RNAs. To produce functional RNAs, nucleus-encoded proteins are required for the processing of plastid RNAs (blue/segmented circles). Translation occurs on bacterial-type (70S ribosomes), using a set of tRNAs encoded by the plastid genome. After Leister et al. (2017).

The rates of transcription and translation regulate the production and accumulation of photosynthetic proteins such as RuBisCO and the light harvesting and electron transfer components (Leister, 2003). Transcription is the first step in chloroplast gene expression. In higher plants, chloroplast genes are transcribed by two different RNA polymerases (see Figure 1.5). One is the nuclear-encoded plastid RNA polymerase (NEP), a single-subunit of the type T3-T7 bacteriophage, encoded by the *RPOT* gene that is responsible for transcribing the housekeeping gene during early phase of plant development (Hedtke et al., 1997). The second type of RNA polymerase is a bacterial-type multi-subunit enzyme called the plastid-encoded RNA polymerase (PEP), which is inherited from a cyanobacterial ancestor and transcribes the photosynthesis-related genes (Börner et al., 2015; Lerbs-Mache, 2011). In chloroplast gene expression, there is a shift in the primary RNA polymerase from NEP to PEP during chloroplast development (Díaz et al., 2018, Hernández-Verdeja and Strand, 2018). However, this process is not well understood. In green leaves, PEP is the major polymerases in the transcription machinery, and over 80% of plastid genes are transcribed by PEP (Zhelyazkova et al., 2012). The PEP subunits are encoded by the *rpo* plastid genes (*rpoA, rpoB, rpoC1* and *rpoC2*) (Börner et al., 2015). PEP activity requires nuclear-encoded sigma factors for promoter specificity (Hanaoka et al., 2003). A large number of nuclear-encoded proteins have been found to be associated with the PEP subunits (Steiner et al., 2011). There is also a large number of PEP-associated proteins required in plastid transcription, suggesting that chloroplast gene expression mechanisms are complex (Kindgren and Strand, 2015). Chloroplast gene expression can be divided into three categories: i) photosynthesis-related genes transcribed by PEP; ii) housekeeping genes (clpP and the rrn operon) transcribed by both PEP and NEP; and iii) genes (accD and the rpoB operon) exclusively transcribed by NEP (Allison et al., 1996, Hajdukiewicz et al., 1997).

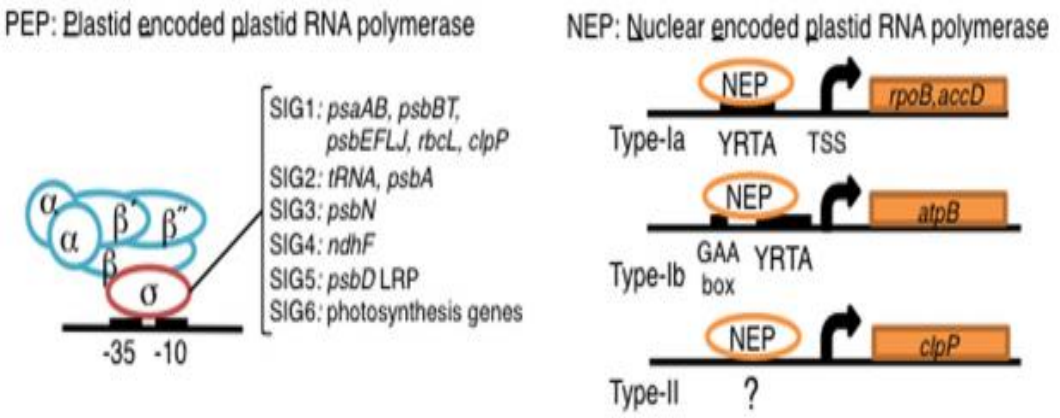


Figure 1.5: The transcriptional machinery in the higher plants. Chloroplast genes are transcribed by two different types of RNA polymerases

PEP is a bacterial-type multi-subunit RNA polymerase, composed of four core enzymatic subunits α, β, β', β'' (blue) and a sigma subunit (red) that recognise bacterial σ70-type promoters with -10 and -35 consensus elements. There are six subgroups of plastid sigma factors (SIG1–SIG6) that recognise bacterial-type promoters in the plastid. NEP is involved in the transcription of housekeeping genes, such as rpo genes, for PEP core subunits and ribosomal protein-coding genes. The upstream regions of genes transcribed by NEP are known as Types Ia, Ib and II). Mainly, NEP promoters such as rpoB, rpoA, and accD share a core sequence, the YRTA motif (type-Ia), with transcription start sites (TSSs). The YRTA motif is also typical for type-Ib NEP promoters together with GAA-boxes. However, type-II NEP promoter mapped upstream of clpP gene lacks the YRTA motif. After Yagi and Shiina (2014).

1.5 Communication between the nucleus and organelles

The majority of organellar proteins are nuclear-encoded. Hence, both the mitochondria and chloroplasts are dependent on the nucleus for the provision of structural and other proteins, requiring an extensive and complex organisation and co-ordination of processes involved in gene expression, translation and protein trafficking (Blanco et al., 2014). As mentioned previously, organelle-to-nucleus (retrograde signalling) and nucleus-to-organelle (anterograde signalling) communication is required (see Figure 1.6). Retrograde signalling is a key to this coordination (Woodson and Chory, 2008). However, relatively little is known about the orchestration of the signals and pathways, involved in this communication. Several plastid signals have been identified, of which changes in the redox state of the chloroplasts and mitochondria leading to ROS accumulation are thought to be particularly important (Fernández and Strand, 2008). Mitochondrial retrograde regulation can be triggered by the disruption of respiratory electron transport leading to ROS accumulation (Ho et al., 2008).

A number of new candidate signalling molecules have been identified in the chloroplast-to-nucleus retrograde pathway, including a number of metabolites. *Arabidopsis* mutants that over-express potential plastid signals, such as heme (Woodson et al., 2011), methylerythritol cyclodiphosphate (Xiao et al., 2012) and 3'- phosphoadenosine 5'phosphate (Estavillo et al., 2011), had smaller rosette sizes and altered rosette morphologies. Many of these components have been identified in screens using carotenoid synthesis inhibitors like norflurazon, or plastid translation inhibitors like lincomycin that inhibits chloroplast development, producing a bleached phenotype. The *genomes uncoupled (gun)* mutants, which were isolated using such screens show higher nuclear gene expression, particularly genes encoding the light-harvesting chlorophyll *a*/*b*-binding proteins (Kacprzak et al., 2019). The biogenic pathway of retrograde signalling originates in the plastid early in chloroplast biogenesis (Kacprzak et al., 2019). In this pathway, loss of chloroplast functions leads to a reduced expression of nuclear

genes associated with photosynthesis and chloroplast development (Koussevitzky et al., 2007, Woodson et al., 2013).

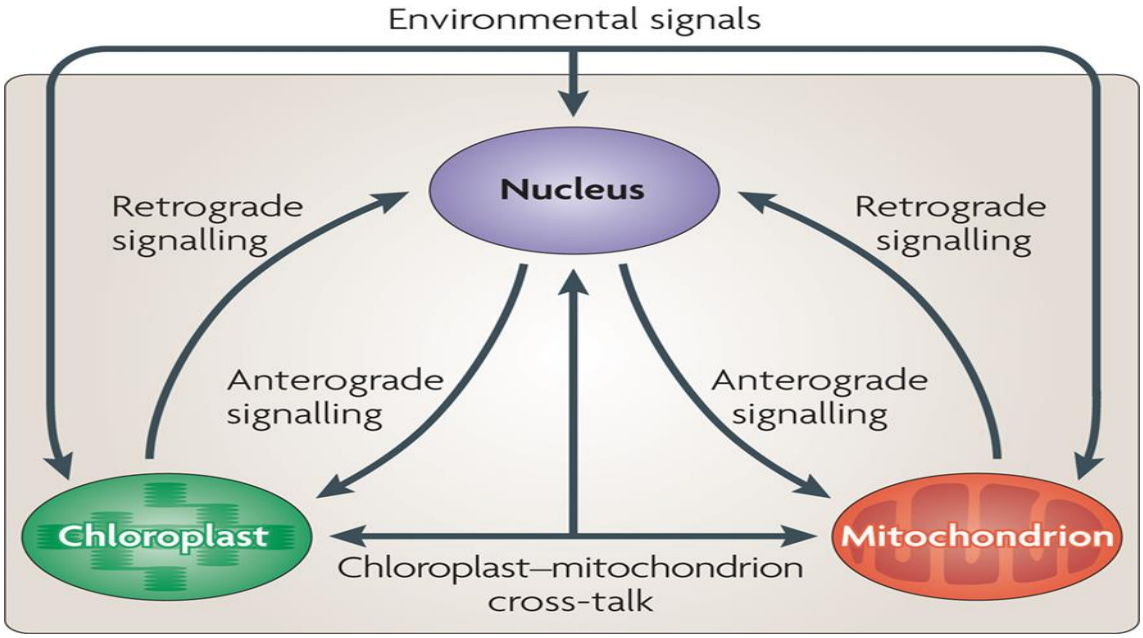


Figure 1.6: Summary of genome coordination with the nucleus and intracellular organelles.

The chloroplasts and nucleus communicate with each other in order to maintain growth and adapt to environmental stresses. Signalling from the nucleus to the chloroplast/mitochondria is called anterograde signalling, and signalling from the chloroplast/mitochondria to the nucleus is known as retrograde signalling. Chloroplast-mitochondrion cross-talk is signalling between these two organelles. After Woodson and Chory (2008).

1.6 The transport of proteins into organelles

Chloroplasts and mitochondria are responsible for the ATP synthesis associated with photosynthesis and respiration, respectively. Since the vast majority of proteins found in the mitochondria and chloroplast are encoded by nuclear genes and synthesised as precursor forms on cytosolic ribosomes, some require specific targeting signals in the amino acid sequence to transport proteins into specific organelles. These processes share some similarities in the mitochondria and chloroplasts: i) the preproteins are bound to chaperones in order to prevent premature folding; ii) the targeting of proteins to both organelles requires specific signals, such as presequences or transit peptides; and iii) both organelles have special translocon complexes in their outer and inner membranes.

1.6.1 The transport of proteins into mitochondria

Baker's yeast (*Saccharomyces cerevisiae*) has been used extensively as a model organism for studying the principles of protein import into mitochondria (Dudek et al., 2013). However, the mechanisms discovered in *S. cerevisiae* were later found in most other higher eukaryotes. The majority of mitochondrial proteins are synthesised as precursor proteins in the cytosol. The import of these precursor proteins into the mitochondria occurs by post-translational mechanisms. However, there are exceptions in some proteins, such as Sod2 and fumarase, because they involve a co-translational import mechanism instead of a post-translational mechanism (Luk et al., 2005, Yogev et al., 2007).

The mitochondrial import signals are N-terminal extensions of the mature proteins, known as presequences (Dudek et al., 2013). These presequences are amphipathic α-helical segments with a net positive charge, and contain about 15 to 55 amino acids. They are recognised by the translocase of the outer membrane (TOM) and translocase of the inner membrane (TIM) 23 complexes on the mitochondrial membrane. Thus, N-terminal presequences function as targeting

signals that interact with the mitochondrial import receptors, and direct the preproteins across the outer and inner membranes. The preproteins directed to the inner membrane contain a hydrophobic sorting signal that allows them to be inserted into the membrane (Schmidt et al., 2010). The translocation of preproteins into mitochondria requires ATP hydrolysis and an electrochemical H^+ gradient across the inner membrane. Cytosolic chaperones are also important because they allow the precursor proteins to be targeted at the outer mitochondrial surface (Wiedemann et al., 2004)

After the import, the N-terminal polypeptides are proteolytically removed by a mitochondrial-processing peptidase and other proteases (Mossmann et al., 2012, Taylor et al., 2001). In addition, some precursor proteins are synthesised without cleavable extensions. These internal targeting signals remain a part of the mature protein (Wiedemann et al., 2004). In all proteins directed to the outer membrane, intermembrane space and inner membrane, these precursor proteins contain the most internal targeting signals.

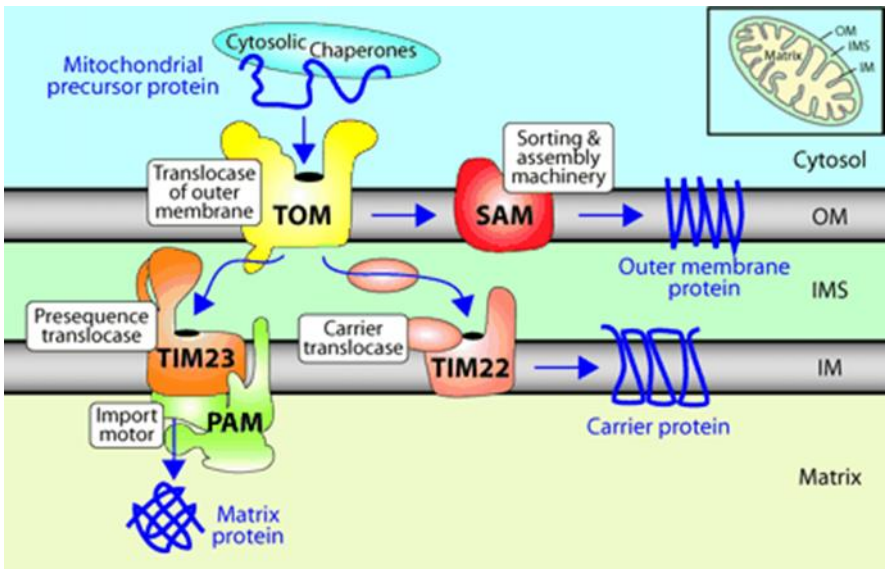


Figure 1.7: Protein import into mitochondria.

Mitochondrial proteins are mainly synthesised in the cytosol. With the help of cytosolic chaperones, mitochondrial precursor proteins are moved to the TOM complex. Biogenesis β-barrel outer membrane proteins require the sorting and assembly machinery (SAM) complex. Preproteins targeted at the matrix depend on the presequence translocase (TIM23 complex) and its associated import presequence translocase-associated motor (PAM) complex) for their transport across the inner mitochondrial membrane. Carrier proteins are inserted into the inner membrane with the help of the carrier translocase (TIM22 complex). After Wiedemann et al. (2004).

The TOM is important in the importation of all nucleus-encoded mitochondrial proteins (see Figure 1.7). This translocase carries preproteins with appropriate signal sequences into the intermembrane space. It also mediates the insertion of transmembrane proteins into the outer membrane. After passing through the TOM complexes, the precursor proteins follow different pathways to reach specific mitochondrial sub-compartments.

(i) Transport pathways of outer membrane proteins

Proteins of the outer mitochondrial membrane are synthesised as non-cleavable precursors that contain internal targeting signals (Schmidt et al., 2010). There are two types of membrane-integrated proteins in the outer membrane – α-helical proteins that are attached to the outer membrane by transmembrane α-helical segments, and pore-forming β-barrel proteins with transmembrane β-strands (Schmidt et al., 2010). The β-barrel membrane proteins were discovered in the outer membranes of bacteria, mitochondria and chloroplasts. In contrast, the mitochondrial inner membrane consists only of α-helical membrane proteins. The mitochondrial import pathway for β-barrel proteins involves the TOM complex, intermembrane space chaperones and the SAM complex of the outer membrane (Gentle et al., 2004, Kozjak et al., 2003, Wiedemann et al., 2004).

(ii) Presequence pathway to the matrix and inner membrane

Preproteins with cleavable N-terminal presequences are allocated from TOM complexes to the presequence (TIM23) complex (Chacinska et al., 2009, Dolezal et al., 2006, Neupert and Herrmann, 2007). Subsequently, preproteins are either targeted at the inner membrane or the mitochondrial matrix. Preproteins imported into the matrix are translocated through a channel in the inner membrane with the help of the matrix-localised heat-shock protein 70 (mtHsp70), which is a core protein of the PAM complex. Most metabolic enzymes localised in the matrix are synthesised with such cleavable presequences (Schmidt et al., 2010).

(iii) Carrier pathway to the inner membrane

Hydrophobic preproteins of the inner membrane, such as metabolite carriers (ADP and ATP) are imported via the TOM complex (Schmidt et al., 2010), the chaperone complexes of the intermembrane space and the carrier translocase (TIM22 complex), which then transfer them into the inner membrane (Koehler, 2004, Rehling et al., 2003).

1.6.2 The transport of proteins into chloroplasts

The majority of chloroplast proteins are synthesised as precursor proteins (preproteins) and imported into the organelles after translation on cytosolic ribosomes. They contain N-terminal transit peptides, which are important for transport to specific plastid sub-compartments (Bruce, 2000, Lee et al., 2008). The presequences or transit sequences are proteolytically removed after import (Soll, 2002). The transit sequence is important for organelle recognition and translocation initiation (Soll, 2002).

The transit peptide contains about 13 to 146 amino acids (Zhang and Glaser, 2002). Preproteins contain a cleavable transit peptide that is regulated by guanosine triphosphate (GTP)-binding (Soll, 2002) and by the receptor translocon on the outer envelope membrane that is known as the translocon outer membrane (TOC) complex (see Figure 1.8) (Kovács-Bogdán et al., 2010). Within the TOC complex, TOC 34 acts as an initial receptor for the preproteins. It is regulated by GTP-binding and phosphorylation (Kessler et al., 1994, Schleiff et al., 2002, Sveshnikova et al., 2000). These preproteins cross the outer envelope through an aqueous pore, and are imported into the translocon in the inner membrane of the chloroplast (TIC) complex (Soll and Schleiff, 2004).

During the translocation process, the TOC and TIC translocons cooperate to facilitate the passage of polypeptides across both membranes (see Figure 1.8) (Kovács-Bogdán et al., 2010). A stromal-processing peptidase then cleaves the transit sequence to yield mature proteins (Soll, 2002). Although the protein-targeting pathways into the mitochondria and chloroplasts share some similarities, there are some major differences between these processes. For example, mitochondria use the electrochemical hydrogen gradient and ATP generated by the respiratory electron transport chain for the translocation of preproteins into the mitochondrial matrix. However, the transport of plastid preproteins into the stroma is driven exclusively by the hydrolysis of GTP and ATP in the cytosol. Moreover, the transport of chloroplast peptides from the

stroma to other sub-compartments, such as the thylakoid and the lumen, requires a second targeting sequence, which is subsequently cleaved.

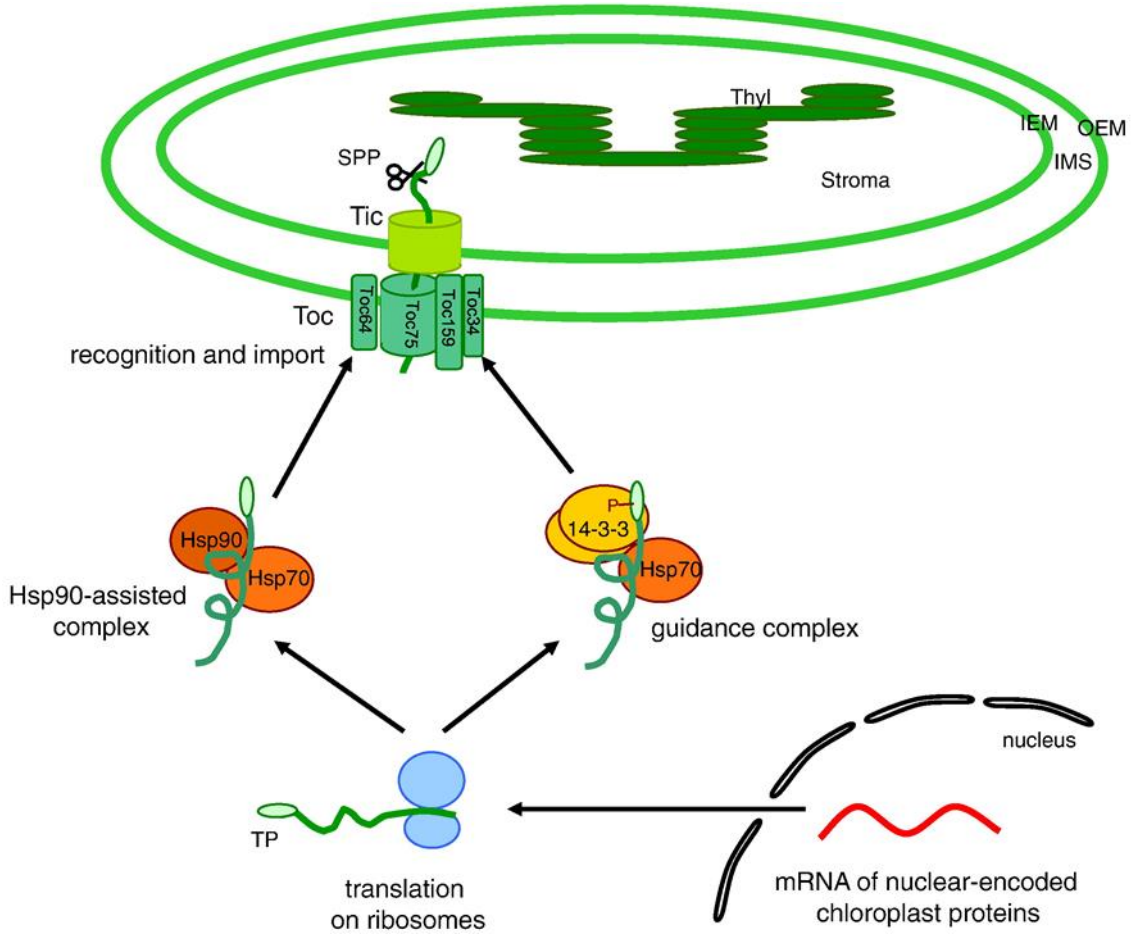


Figure 1.8: Protein import into chloroplasts

Nuclear-encoded chloroplastic proteins are synthesised on cytoplasmic ribosomes and then imported into the chloroplast with the help of molecular chaperones. Transit peptides with a specific binding site for the 14-3-3 dimers bind selectively to phosphorylated transit peptide with the help of Hsp70 chaperones. Another cytosolic complex (Hsp 90 and Hsp70) also had chaperons associated with preprotein. TOC34 and TOC64 function as gates for the guidance and Hsp90-assisted complexes, respectively. TOC159 not only functions as a receptor in the outer envelope membrane but also as a GTP-driven motor that helps preproteins into the TOC-channel. The import channel of the TOC complex is formed by a β-barrel protein, TOC75. After being imported across the outer and inner envelope membranes of the chloroplasts, through the TOC and TIC complexes, the transit peptide is cleaved by zinc-binding metallopeptidase and the stromal processing peptidase after its import into the stroma, yielding the mature proteins. IMS – intermembrane space, Thyl – thylakoids. After Kovács-Bogdán et al. (2010).

1.6.3 Dual-targeted proteins

Some proteins can be dual-targeted at two or more compartments. For example, glutathione reductase in *Pisum sativum* (pea) has been identified as a dual-targeting protein to chloroplasts and mitochondria (Creissen et al., 1995). This study demonstrated that proteins can be imported into two organelles and that protein import is not limited to one specific location (Creissen et al., 1995). Other dual-targeted proteins in *Arabidopsis* and rice include methionine aminopeptidase, monodehydroascorbate reductase glutamyl-transfer RNA synthetase and tyrosyl-transfer RNA synthetase (Morgante et al., 2009). To date, more than 100 dual-targeted proteins have been identified in a variety of plants (Carrie et al., 2009a, Carrie and Small, 2013). In *Arabidopsis*, 72 proteins are dual-targeted (Carrie et al., 2009b, Carrie and Small, 2013). In addition, 500 proteins containing ambiguous signals are predicted to be dual-targeted (Mitschke et al., 2009). The reason for dual-targeting remains unknown, but it is suggested to be necessary for the coordination of organellar functions in retrograde and anterograde signalling pathways (Carrie et al., 2009a).

1.7 Reactive oxygen species

Plants metabolism is regulated to maintain an appropriate balance between energy production and consumption (Sharma et al., 2012). The steady-state of the cellular energy balance depends on a signalling network that links key processes such as photosynthesis, dark respiration and photorespiration with multiple points of reciprocal control (Foyer and Noctor, 2009, Suzuki et al., 2012). Many of the common forms of ROS such as superoxide anions ($O_2^{\bullet -}$), hydroxyl radicals ($^{\bullet}OH$) and hydrogen peroxide (H_2O_2) are produced by the partial reduction of atmospheric triplet oxygen ($^3O_2$). In addition, singlet oxygen ($^1O_2$) is produced by the direct transfer of energy from chlorophylls in PSII to ground state molecular oxygen (Foyer and Noctor, 2009). ROS are generated at many sites during steady-state cellular metabolism, particularly through photosynthesis and respiratory electron transport processes. ROS production is unavoidable in aerobic metabolism (Moller, 2001, Shapiguzov et al., 2012). The main sources of ROS in photosynthetic cells are the chloroplasts, peroxisomes and mitochondria (Pospisil, 2009, Foyer and Noctor, 2009).

Despite very high rates of ROS production, these important metabolites do not normally accumulate in cells because of a very effective antioxidant network that rapidly metabolises ROS as soon as they are produced. However, increased ROS accumulation can be triggered by exposure to environmental stresses, such as drought, salinity, cold, metal toxicity and UV-B radiation (Gill and Tuteja, 2010). The accumulation of ROS triggers key cellular signalling pathways (Foyer and Noctor, 2009, Suzuki et al., 2012). The accumulation of ROS has the potential to cause damage to DNA, lipids and proteins generate further oxidative signals that can trigger cell suicide programs and programmed cell death. As mentioned previously, ROS are widely considered to participate in retrograde signal cascades (Chan et al., 2010).

1.8 WHIRLY1 protein

The WHIRLY (WHY) proteins have putative DNA-binding domains consists of 6 amino acids, Lys-Gly-Lys-Ala-Ala-Leu (KGKAAL), which allow them to bind single-stranded DNA (ssDNA-binding) (Desveaux et al., 2000). In general, ssDNA-binding proteins are involved in multiple cellular processes, such as DNA replication, repair and recombination (Cappadocia et al., 2010). The WHY ssDNA-binding protein name is derived from the whirligig-like appearance of the tetramers and specific to the plant kingdom (Desveaux et al., 2002, Desveaux et al., 2005). The tetrameric forms can further assemble into 24-oligomers upon binding to the thylakoid membrane (Cappadocia et al., 2010, Cappadocia et al., 2012)

WHY proteins were first described as transcriptional activators that bind to an elicitor response element in the promoter region of pathogenesis-related genes in the nucleus of potato (Despres et al., 1995, Desveaux et al., 2000). The DNA binding of WHY1 is induced by pathogen elicitors and salicylic acid (SA) (Desveaux et al., 2004). Elicitor-induced gene expression of the pathogenesis-related nuclear gene (*PR-10a*) in potato is mediated by the transcriptional activator PBF-2, which has a DNA-binding component of 24 kDa (Despres et al., 1995). This protein was therefore first called a nuclear factor (p24) (Desveaux et al., 2000). In addition to the role of the WHY1 protein as a transcriptional regulator, it maintains telomere homeostasis through the regulation of telomerase activity, as shown in a study of telomere length in *Arabidopsis* lines with altered levels of AtWHY1 expression (Yoo et al., 2007a).

In all plants studied to date, there are at least two WHY proteins (WHY1 and WHY2) that are targeted at the mitochondria or plastids (Isemer et al., 2012b, Krause and Krupinska, 2009) depending on the plant species (Marechal et al., 2009). For example, there are three *WHY* genes encoding WHY proteins in *Arabidopsis* (Desveaux et al., 2005). The AtWHY1 and AtWHY3 proteins have

77% and 82% sequence similarity, respectively, and are targeted at plastids, while AtWHY2 is targeted at mitochondria (Desveaux et al., 2005, Krause and Krupinska, 2009). In barley, there are two WHY proteins. WHY1 is targeted at the nucleus and chloroplasts and WHY2 is located in the mitochondria (Melonek et al., 2010). The functions of the WHY proteins might vary depending on the stage of plant development and their localisation within the cell.

WHY1 was first thought to be exclusively located in the nucleus (Desveaux et al., 2000, Desveaux et al., 2004). However, it was later demonstrated that WHY1 is also located in the chloroplasts (Krause et al., 2005). WHY1 is thus dual-targeted at the plastids and nucleus of the same cell (Grabowski et al., 2008). The WHY1 protein has the same molecular weight in the chloroplasts and nucleus, and is synthesised on the 80S ribosomes, imported to the chloroplasts and processed by a cleavage of N-terminal plastid transit peptide (Grabowski et al., 2008). Transplastomic studies using a recombinant form of WHY1 suggested that WHY1 could move from chloroplasts to the nucleus (Isemer et al., 2012b, Foyer et al., 2014)

The chloroplast form of WHY1 is present in the nucleoids (Krause et al., 2005), as well as being linked to the thylakoid membranes (Grabowski et al., 2008). The WHY1 protein is located at the boundary between thylakoids and nucleoids in the chloroplasts (Foyer et al., 2014). It is possible that the WHY1 protein participates in chloroplast to nucleus signalling (see Figure 1.9).

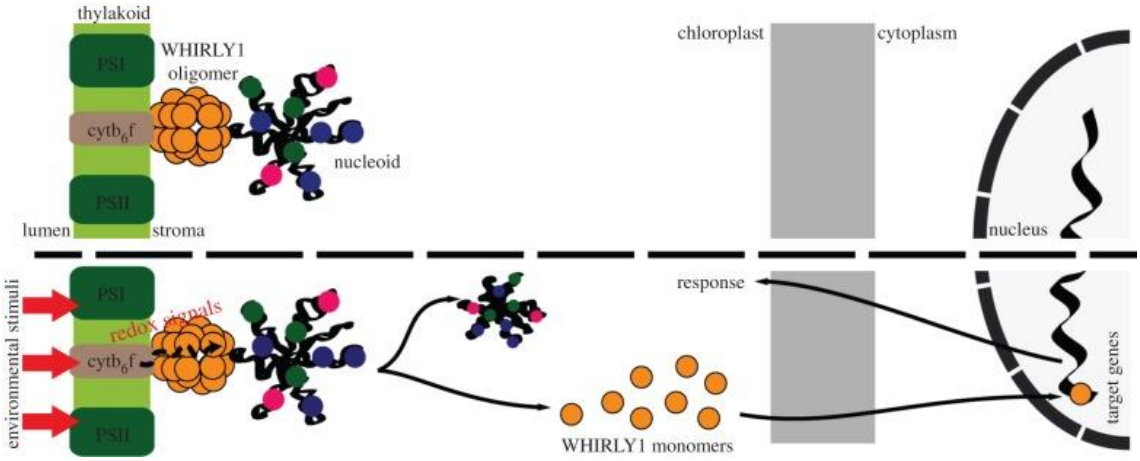


Figure 1.9: Schematic model of the WHY1-dependent perception and transduction of redox signals from chloroplast to nucleus.

WHY1 forms 24-oligomers that form a bridge between the thylakoid and nucleoid under normal conditions. The redox state of the photosynthetic apparatus responds to extreme environmental conditions. After Foyer et al. (2014).

Arabidopsis mutants defective in WHY1 have a similar phenotype to the wild type (Yoo et al., 2007a). The seeds of *why1 Arabidopsis* mutants were shown to be less sensitive to salicylic acid (SA) and abscisic acid (ABA) than the wild type in germination assays (Isemer et al., 2012a). The *Arabidopsis* seeds were insensitive towards ABA when the WHY1 protein was expressed in the nuclei. In contrast, the seeds were sensitive towards ABA when the WHY1 protein was targeted to both plastids and nuclei (Isemer et al., 2012a).

Plastid-targeted WHY proteins were important for genome stability in *Arabidopsis* (Marechal et al., 2009). A small percentage of *Atwhy1 Atwhy3* double knockout mutants had altered chloroplast development (Marechal et al., 2009). There is a high level of variation in the mutant phenotypes observed in the double mutant line, with a small percentage exhibiting chlorosis, or yellow/white sectors on the leaves (Marechal et al., 2009). The WHY1 and WHY3 proteins are therefore considered to contribute to plastid genome stability by preventing illegitimate recombination (Marechal et al., 2009). Most of the *Arabidopsis Atwhy1 Atwhy3* mutants however, had similar phenotypes to the wild type (Cappadocia et al., 2010, Marechal et al., 2009).

Crossing the *Arabidopsis Atwhy1 Atwhy3* double mutants, with a mutant impaired in DNA polymerase IB (pollB) (*atwhy1 atwhy3 pollb-1*) resulted in a severe yellow-variegated phenotype (Lepage et al., 2013). The *atwhy1 atwhy3 pollb-1* mutants had a higher level of illegitimate recombination between repeated sequences and high plastid genome instability compared to the wild type (Lepage et al., 2013). These findings suggested that the DNA polymerase IB and WHY proteins function synergistically to maintain plastid genome stability (Krupinska et al., 2014). Moreover, the *atwhy1 atwhy3 pollb-1* mutants exhibited a low level of photosynthetic electron transport efficiency than the wild type, and have high ROS levels (Lepage et al., 2013). The high level of ROS accumulation observed in these mutants links these proteins to chloroplast and nucleus signalling, and show its ability to tolerate oxidative stress, suggesting a role for WHY1 in DNA polymerase (Lepage et al., 2013).

The functions of WHY1 in chloroplasts have been analysed by mutants and transgenic plants lacking, or low in, WHY1 protein abundance. Maize knockout mutants (*zmwhy1-1*) have shown the most extreme phenotype (see Figure 1.10). Experiments with knockdown mutations of *ZmWhy1*, produced by transposon insertion, have shown impaired plastid gene expression, resulting in ivory or pale plant leaves that lack plastid ribosomes, thus suggesting a role of WHY1 in chloroplast RNA metabolism (Prikryl et al., 2008). Analysis of these mutants has shown that WHY1 is associated with both DNA and RNA in chloroplasts, and also co-immunoprecipitates with CRS1, a protein involved in the splicing of a specific set of chloroplast introns, suggesting that WHY1 may play an accessory function in intron splicing (Prikryl et al., 2008). Studies have also indicated that knockout and knockdown maize lines have similar amounts of cpDNA, suggesting that WHY1 is not needed for cpDNA replication (Prikryl et al., 2008).

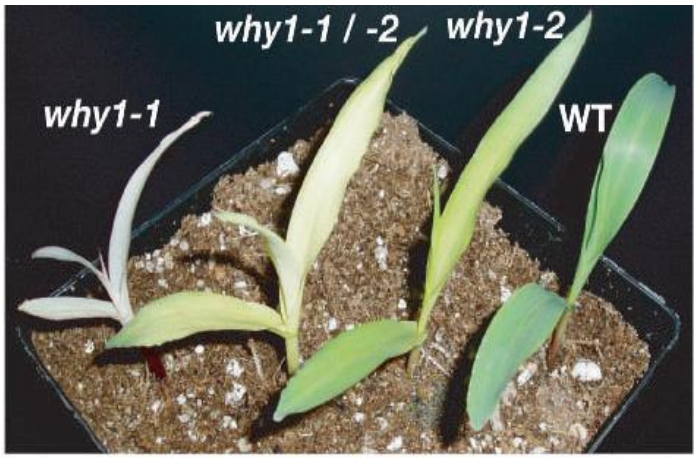


Figure 1.10: Comparison of the phenotype in wild and ZmWHY1 mutant seedlings lacking the WHY1 protein at 9 days old.

The seedlings shown are homozygous for either the ZmWHY1-1 and ZmWHY1-2 allele, or are the heteroallelic progeny of a complementation cross. After Prikryl et al. (2008).

1.9 The WHY1 protein in barley

In earlier studies, WHY1 was not found to associate with plastid nucleoids (Melonek et al., 2010). Later, it has been shown that WHY1 associates with chloroplast nucleoids in barley (Krupinska et al., 2014). The loss of the WHY1 protein in this transgenic barley line increased the chloroplast copy number with an increased expression of an organellar DNA polymerase (Krupinska et al., 2014). Analysis using nucleic acid staining has shown that a nucleoid population in chloroplasts of the W1-7 was more heterogeneous than in the wild type (see Figure 1.11) (Krupinska et al., 2014).

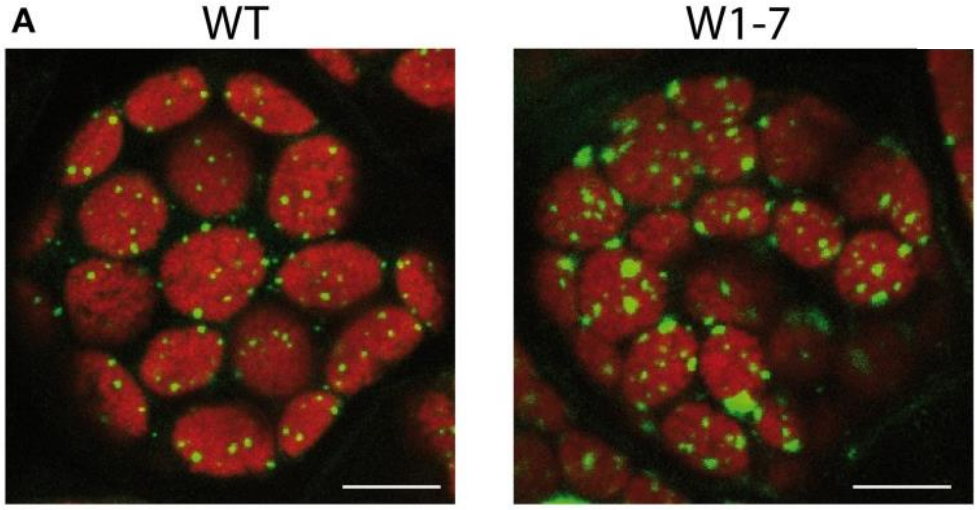


Figure 1.11: Morphology and distribution of chloroplast nucleoids in leaves of WT and WHY1-deficient barley plants (W1-7).

The DNA was stained using YO-PRO®-1 (green). Scale bar = 5 µm. After Krupinska et al. (2014).

The barley WHY1 protein was shown to bind to one of the two ERE motifs on the *HvS40* gene, suggesting that it might act as a promoter of this senescence-associated gene (Krupinska et al., 2002, Krupinska et al., 2014). Experiments with the RNAi transgenic barley lines (W1-1, W1-7 and W1-9) that have very low levels of the WHY1 protein had much greater levels of transcripts encoding components of the thylakoid NADH complex, the chloroplast RNA polymerase (RPOC2, RPOB and RPOC1), the cytochrome b/f complex (PETA, PETD and

YCF5) and the chloroplast ribosomes (RPL20, RPL23.2, RPL33 and RPS2) was higher in the WHY1-deficient lines than the wild type (Comadira et al., 2015). The transgenic barley lines W1-1, W1-7 and W1-9 exhibited phenotypes similar to the wild type (see Figure 1.12), with comparable photosynthesis rates, although they contained significantly more chlorophyll and less sucrose than the wild type, while also exhibiting similar phenotypes to the wild type and having no effect on aphid infestation (Comadira et al., 2015).

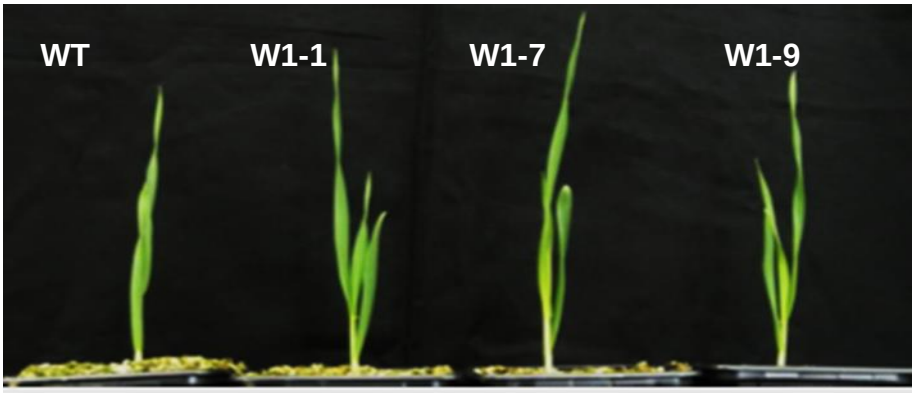


Figure 1.12: A comparison of phenotypes in the wild type and WHY1-deficient barley lines (W1-1, W1-7 and W1-9).

Representative appearance of lines 22 days after germination. (Comadira et al., 2015).

The WHY protein functions are summarised in Table 1.1.

Table 1.1: WHY protein functions

Protein	Function	Reference
StWHY1 ***Solanum tuberosum***	Pathogen response ssDNA binding Transcriptional activator	Despres et al. (1995) Desveaux et al. (2000) Desveaux et al. (2002)
AtWHY1 ***Arabidopsis thaliana***	SA-mediated pathogen response Regulator of telomere homeostasis Plastid genome stability Negative regulator of *AtKP1* Organelle genome repair ABA-sensitive in seed germination WRKY53 repressor, delays senescence Plastid DNA stability with polIB U-turn-like rearrangement suppression	Desveaux et al. (2004) Yoo et al. (2007) Marechal et al. (2009) Xiong et al. (2009) Cappadocia et al. (2010) Isemer et al. (2012) Miao et al. (2013) Lepage et al. (2013) Zampini et al. (2015)
ZmWHY1 ***Zea mays***	Plastid biogenesis, Ribosomal RNA metabolism and RNA splicing	Prikryl et al. (2008)
HvWHY1 ***Hordeum vulgare***	Plastid nucleoids compaction HvS40 promoter binding Redox sensor Leaf development	Krupinska et al. (2014) Krupinska et al. (2014) Foyer et al. (2014) Unpublished work

1.10 Late embryogenesis abundant proteins

Interest in the functions of late embryogenesis abundant 5 (LEA5) began when a PhD student in the lab, Daniel Shaw, undertook a tandem affinity purification (TAP) tagging study to identify proteins that interacted with LEA5 in the lab of Geert De Jaeger in the Department of Plant Systems Biology, VIB, Gent. These studies consistently showed that LEA5 interacted with chloroplast proteins. Moreover, chloroplast DEA (D/H)-box RNA helicase 22 (RH22) was identified as a LEA5-binding partner in all experiments.

LEA proteins were first identified as proteins that were abundant in the later stages of seed development (Grzelczak et al., 1982). Later, LEA proteins were found to be expressed in many of the plant vegetative and reproductive tissues (Hundertmark and Hincha, 2008) and were found to be present in other organisms, such as eukaryotes and prokaryotes (Garay-Arroyo et al., 2000). Abiotic stresses, such as drought and cold, can induce LEA proteins (Thomashow, 1999). This suggests that the LEA proteins in transgenic plants provides resistance to extreme environmental stresses, such as drought, extreme cold and freezing. They are thought to act as molecular chaperones in order to protect the plant against the aggregation of protein under water stress (Goyal et al., 2005). In *Arabidopsis*, LEA proteins make up to 51 members of a family that can be further divided into nine different groups, with different constitutive and inducible expression patterns (Hundertmark and Hincha, 2008). These proteins are hydrophilic, mostly intrinsically disordered, proteins, which play major roles in stress tolerance. However, the functions of many of the 51 genes encoding LEA proteins in *Arabidopsis* remain uncharacterised.

LEA5 (LEA38: At4g02380) is a member of the LEA-3 group in *Arabidopsis*. AtLEA5 protein is likely to be localised either in the plastids or the mitochondria (Hundertmark and Hincha, 2008). The expression of *AtLEA5* shows a diurnal pattern of regulation that is abundant in the dark, but is suppressed in light (Mowla

et al., 2006). Canonical LEAs are normally expressed in seeds, but AtLEA5 is unique compared to the other LEA family members because it is expressed in the roots and reproductive organs (Mowla et al., 2006). Previously, LEA5 was known as a senescence-associated gene that was expressed transiently early in leaf senescence, as the leaves began to yellow, being induced by darkness (Weaver et al., 1998). There have been many reports on LEA5, including data from open-access microarrays (Zimmermann et al., 2004), showing that other stresses, such as ozone-induced leaf senescence in *Arabidopsis* (Miller et al., 1999), cold (Seki et al., 2001), nitrogen deficiency (Wang et al., 2000) and pathogen attacks, including *Colletotrichum higginsianum* (Liu et al., 2007), increase *LEA* transcripts. Similarly, treatment with hormones, such as ethylene (Weaver et al., 1998) and jasmonate (Jung et al., 2007) and also sugar (Xiao et al., 2000) can induce *LEA5* expression.

Arabidopsis genes associated with oxidative stress tolerance were identified using the complementation of an oxidant-sensitive yeast mutant (Δyap1) strain of *Saccharomyces cerevisiae* (Mowla et al., 2006). The expression of *AtLEA5* allows the yeast mutant strain (Δyap1) to grow in the presence of oxidants, such as H_2O_2, diamide, menadione and tert-butyl hydroperoxide light (Mowla et al., 2006). This study suggests that LEA5 may play a role in the response to oxidative stress. AtLEA5 proteins were also found to be localised to mitochondria, using a yellow fluorescent protein (YFP) fusion (Salleh et al., 2012) and separately to the mitochondrial matrix, using a green fluorescent protein (Candat et al., 2014). This suggests that LEA5 may have a role in protecting mitochondrial functions that relate to respiration and oxidative stress tolerance or signalling in mitochondria in plants exposed to stress.

1.11 DEAD-box RNA helicase 22

The plastid DEAD-box RNA helicase 22 (RH22) is a putative DEAD RNA helicase known as heavy seed 3 (HS3) (Kanai et al., 2013). RH22 is localised to plastids (Kanai et al., 2013) and is known to be a plastid-specific helicase (Chi et al., 2012). Seedlings of *Arabidopsis* defective in RH22 exhibit a pale green phenotype in young seedlings, but later on, the mature leaves are similar to the wild type (Kanai et al., 2013). The level of *RH22* expression is high in young seeds and seedlings, but not in the stems, rosette leaves or flowers (Kanai et al., 2013). The plastid gene expression of *rh22* mutant young seeds and seedlings differs from the wild type. The expression levels of the gene encoding the β-subunit of carboxyltransferase, a component of acetyl-CoA carboxylase in plastids, is low in RH22 seeds. This β subunit of carboxyltransferase is required in fatty-acid biosynthesis for the conversion of acetyl-CoA to malonyl-CoA by acetyl-CoA carboxylase (Cahoon et al., 2007, Konishi et al., 1996, Ke et al., 2000, Sasaki and Nagano, 2004). RNA helicase is important in regulating a variety of plant growth and development stages through the regulation of RNA metabolism (Kanai et al., 2013). RH22 is also required for chloroplast ribosome biogenesis, where a knockdown of *RH22* has resulted in a delayed-greening phenotype in *Arabidopsis* (Chi et al., 2012). RH22 has also been found to be involved in the biogenesis of 50S ribosomal subunits in *Arabidopsis* because the precursors of 23S and 4.5 rRNA accumulated in *rh22* mutants (Chi et al., 2012).

A large number of RNA helicases have been identified to be localised in plastids, using proteome analysis (Zybailov et al., 2008). RNA helicases are important in the rearrangement of ribonucleoproteins and gene expression (Cruz et al., 1999). The *Arabidopsis* genome encodes more than 100 putative RNA helicases, some of which have been identified and characterised (Mingam et al., 2004; Umate et al., 2010). Plastids and mitochondria have their own genomes that are separated from the nuclear genome, and there are organelle-localised helicases encoded in the nuclear genome (Kanai et al., 2013). Other examples of RNA helicases are AtSUV3 and ISE1. Both are mitochondrial helicases, and have different functions.

AtSUV3 has ATP hydrolytic activity, while the loss of ISE1 leads to dysfunction in the mitochondria and plasmodesmata (Gagliardi et al., 1999, Stonebloom et al., 2009). MH1 and PMH2 are linked to large RNA-containing complexes in the mitochondria (Matthes et al., 2007), and PMH2 is important for intron splicing in mitochondrial genes (Kohler et al., 2010). RH39 is a plastid-specific helicase required in the post-maturation processing of 23S rRNA in chloroplasts (Nishimura et al., 2010).

1.12 Hypothesis and project objectives

WHY1 is a dually targeted protein that is localised in the chloroplasts and nucleus (Figure 1.13). This protein is therefore a candidate for the study of organelle communication with the nucleus, particularly during plant development and exposure to environmental stress. Chloroplast biogenesis requires the coordinated expression of plastome and nuclear genes. The switch from the nuclear-encoded RNA polymerase (NEP) to the plastid-encoded RNA polymerase (PEP) early in chloroplast development is essential for the establishment of photosynthesis. However, the mechanisms that facilitate this switch remain poorly understood. The hypothesis that forms the basis of this study is that WHY1 plays an important role in chloroplast development and that it interacts with other proteins in the chloroplast such as LEA5 to regulate plastome gene expression and translation. Since the precise functions of WHY1 in chloroplast development are poorly understood, the following experiments were performed to characterise the function of WHY1 in barley leaf development. Earlier studies had revealed that LEA5 is not expressed in leaves in the light except under conditions of biotic or abiotic stress. This protein is localised in the mitochondria and is important in plant responses to oxidative and other stresses. However, its precise functions and mechanisms of action are unknown. A previous study using tandem affinity purification (TAP) tagging had revealed that LEA5 can interact with the chloroplast protein (RH22). A part of this study was therefore dedicated to the characterisation of LEA5 binding to RH22 i*n vivo* using a split-YFP system. A preliminary hypothesis that was tested in the following studies was that both WHY1 and LEA5 bind to RH22 and that this binding is important in the regulation of the functions of these proteins in the chloroplasts.

The specific objectives of this thesis were as follows:

1) To investigate the differences in the transcript profiles of the embryos of the dry seeds of WHY1-deficient seedlings compared to the wild type. These studies aim to understand the function of WHY1 during seed germination. This analysis will provide a better understanding of early events that are important during the germination process.

2) To characterise the development of barley leaves that are deficient in WHY1 compared to the wild type in terms of biochemistry and physiology (chlorophyll content and protein content) as well transcriptome and metabolome profiles at early stages of seedling development, i.e. at 7 and 14 days after germination. These studies aim to provide a better understanding on the roles of WHY1 in barley leaf development.

3) To characterise LEA5 functions in *Arabidopsis*. Firstly, the subcellular localisation of the AtLEA5 protein was performed using transgenic lines expressing a LEA5-YFP fusion protein driven by the constitutive 35S promoter. The second approach was to determine whether LEA5 binds to RH22 in a transient expression system. *Arabidopsis* leaf protoplasts were transfected with a range of constructs designed to interrogate the interactions between LEA5 with RH22.

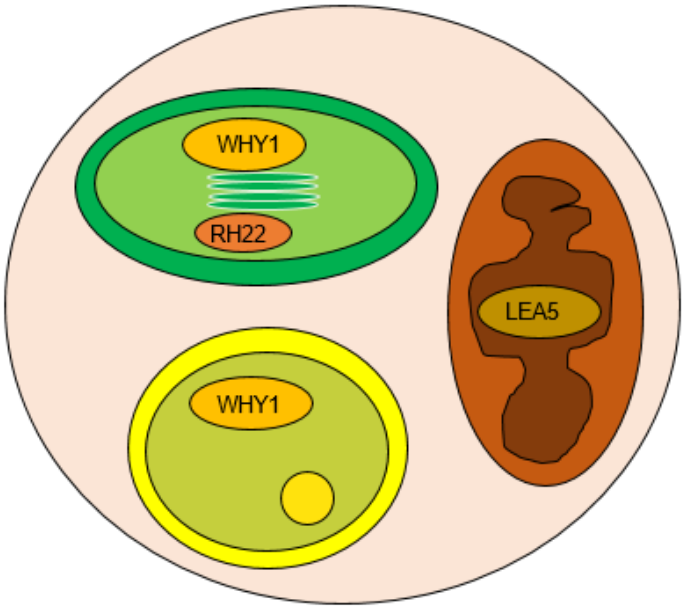


Figure 1.13: WHY1, RH22 and LEA5 localisation in the cell.

The WHY1 is localised to chloroplast and nucleus in the same cell (Grabowski et al., 2008), while the LEA5 protein is located in the mitochondria (Salleh et al., 2012) and the RH22 is located in the chloroplast (Chi et al., 2012).

Chapter 2 . Materials and Methods

2.1 Plant material and growth conditions

2.1.1 Barley

Wild type control seeds (Hordeum vulgare L.cv. Golden Promise) and two independent WHIRLY1 RNAi knockdown transgenic lines (W1-1 and W1-7) produced and characterised in a previous study by Dr Karin Krupinska (University of Kiel) were used in this study.

2.1.2 Arabidopsis

Wild type *Arabidopsis*, accession Columbia 0 (Col-0), widely available in this laboratory for protoplast transformation studies, was used unless stated. *Arabidopsis* expressing a *LEA5*-YFP fusion construct was provided by Dr Hilary Rogers (Cardiff University). Seeds of the T-DNA insertion line, DEA (D/H)-box RNA helicase 22 (RH-22), were obtained from Masatake Kanai (National Institute for Basic Biology, Japan).

2.1.3 Growth conditions

Arabidopsis seeds expressing a *LEA5*-YFP fusion construct were surface-sterilised by exposing the seeds to commercial bleach (100 ml) and 100% ethanol (3 ml) for 2 h. The seeds were then sown on Petri dishes containing 0.5% Murashige and Skoog (MS) basal salts, 1% agar, pH 5.7 and sealed with micro pore tape. They were then cold-stratified for 2 days at 4°C, after which they were placed in a controlled environment chamber for 5 days. The growth conditions were as follows: 20ºC/16ºC temperature regime corresponding to a 16 h light/8 h dark photoperiod with 250 µmol $m^{-2}s^{-1}$ irradiance and 60% relative humidity. All plants were grown under these conditions unless otherwise stated.

Arabidopsis (Col-0) plants used in the protoplast study were placed into pots (5 cm x 5 cm) and after 10 days, were transplanted using forceps into a new medium potting tray (William Sinclair Horticulture Ltd, UK). The plants were grown for 3–5 weeks in a controlled environment chamber.

The barley plants (1 per pot) were sown in compost pots (SHL professional potting compost) in a controlled environment chamber with a 16 h light/8 h dark photoperiod, with an irradiance of 250 µmol $m^{-2}s^{-1}$, 20ºC/16ºC day/night temperature regime and 60% relative humidity.

2.1.3.1 Seed Production

The barley plants were grown to maturity in compost in a 22°C greenhouse at the James Hutton Institute (Scotland) with supplementary lighting provided by high-pressure sodium vapour lamps (Powertone SON-T AGRO 400W; Philips Electronics, UK) to maintain a 16 h light /8 h dark photoperiod.

2.2 Barley leaf sample preparation

The wild type and transgenic barley (W1-1 and W1-7) seedlings were harvested at 7 days and 14 days after sowing. The first leaves of 7-day old and 14-day old seedlings were excised, weighed and divided into the following sections: base, middle and tip, as illustrated in Figure 2.1. The leaves were weighed into 100 mg fresh weight samples and ground in liquid nitrogen to be stored at −80°C until analysis. A minimum of three biological replicates were used in all experiments.

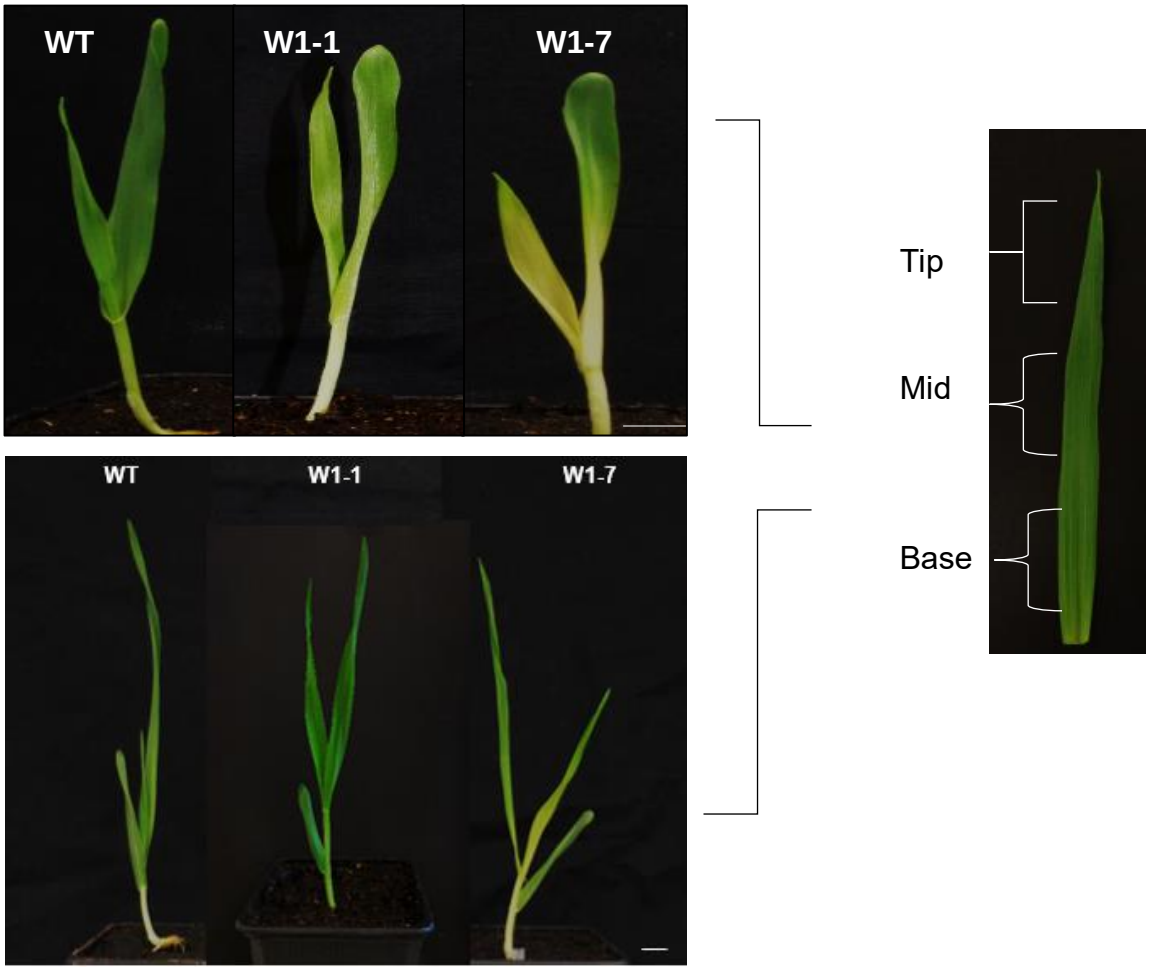


Figure 2.1 Experimental design used in the analysis of the WHY1-deficient lines and wild type barley seedlings.

Barley plants were grown for 7 days and 14 days. The first leaves were excised and divided into three sections: base, mid and tip; the leaves were stored in −80°C until analysis.

2.3 Leaf Pigments

Base, middle and tip sections (100 mg fresh weight) of the first leaves of 7- and 14-day old barley seedlings were ground in liquid nitrogen. Then, ice-cold 95% ethanol (1 ml) was added to the sample and the mixture was ground again. The extracts were later centrifuged (Centrifuge 5804R, Eppendorf, UK) at maximum speed of 14,000 × g for 10 min at 4ºC. The supernatant fractions were collected and used for pigment analysis. The absorbance values were determined at 470 nm, 649 nm and 664 nm on a Fluostar Omega plate reader (BMG Labtech GmbH, Ortenberg, Germany) using a 95% ethanol solution as blank. Pigment content was determined according to the method of (Lichtenthaler, 1987) using equations:

Chlorophyll *a* = 13.36 A_{664} − 5.19 A_{649}

Chlorophyll *b* = 27.43 A_{649} − 8.12 A_{664}

Total chlorophyll (*a*+*b*) = 5.24 A_{664} + 22.24 A_{649}

Carotene = (1000 A_{470} − 2.13 C*a* − 97.64 C*b*)/209

2.4 Chlorophyll Fluorescence Measurements

Chlorophyll fluorescence measurements were performed on the base, middle and tip sections of the first leaves of the barley seedlings, starting from the first day of leaves emerging at 3-day old until 10 days after sowing. Plants were adapted in the dark for an hour for the measurement of Fv'/Fm'. The chlorophyll fluorescence parameters were obtained using a Fluorimager imaging system with automated camera control and image processing scripts provided by the manufacturer (Technologica Ltd, Colchester, UK). Dolphin camera (Allied Vision Technologies, UK) was used to capture and process the chlorophyll fluorescence images. Fluorescence measurements were completed with the help of Dr Tracy Lawson (University of Essex).

2.4.1 Measurement of chlorophyll *a* fluorescence quenching parameters

Chlorophyll *a* fluorescence parameters were measured on the base, middle and tip sections of the first leaves of 7- to 14-day old barley seedlings. The barley plants were kept in the dark for 1 hour. The ratio of dark adapted chlorophyll *a* fluorescence variable (Fv) to the maximal value of chlorophyll *a* fluorescence (Fm) was measured using Fluorometer (FP 100-SN-FP-680, Drasov, Czech Republic, http://www.psi.cz).

2.5 Nucleic acid extraction

2.5.1 DNA extraction

DNA was extracted from the base, middle and tip sections of frozen first leaves of the barley seedlings (100 mg fresh weight per sample) using the DNeasy Plant Mini Kit (Qiagen, Manchester, UK). The samples were ground in liquid nitrogen in a pestle and mortar. Buffer AP1 (400 µl) and RNase A (4 µl) were added to the samples and the mixtures were incubated for 10 min at 65°C with 2 or 3 inversions. Buffer P3 (130 µl) was then added and the mixtures were incubated for a further 5 min on ice. Samples were pipetted onto QIAshredder spin columns and collected by centrifugation at 20,000 × g for 2 min. Samples were transferred into new tubes without disturbing the pellets. Buffer AW1 (1.5 volume) was added to each sample and mixed well. The samples (650 µl) were transferred to DNeasy Mini spin columns with 2 ml collection tubes and centrifuged for 1 min at ≥6000 × g. The flow-through was discarded and the spin columns were placed into new 2 ml collection tubes. Wash buffer (500 µl AW1) was added to each tube and centrifuged for 1 min at ≥6000 × g. The final wash buffer (500 µl AW2) was added to each tube and centrifuged for 2 min at 20,000 × g; this step was repeated twice. The spin columns were transferred to new 1.5 ml microcentrifuge

tubes and the DNA was eluted in AE buffer (100 µl) and incubated for 5 min at room temperature before centrifugation for 1 min at ≥ 6000 × g. The DNA was stored at −20°C for further analysis after quantification using Nanodrop (ND-1000 Spectrophotometer, Labtech International, UK).

2.5.2 RNA extraction

RNA was extracted from the base, middle and tip sections of frozen first leaves of the barley seedlings (100 mg fresh weight per sample) using plant total RNA isolation kit (Sigma-Aldrich, Haverhill, United Kingdom). Samples were ground in liquid nitrogen with a pestle and mortar. The samples were then lysed in Lysis Solution (500 µl) and 2-mercaptoethanol (10 µl), vortexed vigorously following incubation at 56°C for 3–5 min, and centrifuged at 20,000 × g for 3 min to pellet cellular debris. The lysate supernatants were transferred into a filtration column with a 2-ml collection tube and centrifuged at 20,000 × g for 1 min to remove residual debris and collect the clarified flow-through lysate. A binding solution (500 µl) was added into the clarified lysate and mixed thoroughly by brief vortex. The mixture (700 µl) was added to a binding column in a 2-ml collection tube and centrifuged at 20,000 × g for 1 min to bind RNA. The flow-through was discarded. In each digestion, DNase I (10 µl) and DNase digestion buffer (70 µl) were combined and the mixture was added directly onto the centre of the filter inside the binding column. The sample was then incubated at room temperature for 15 minutes. The bound RNA was washed with wash solution 1 (500 µl) and centrifuged at 20,000 × g for 1 min. Final wash solution 2 (500 µl) was added into the column and centrifuged at 20,000 × g for 30 s; this step was repeated twice. RNA was eluted in elution solution (50 µl) for 1 min and centrifuged at 20,000 × g for 1 min. The RNA was stored at −80°C for further analysis after quantification using a Nanodrop (ND-1000 Spectrophotometer, Labtech International, UK), via ratio of absorbance at 260 nm and 280 nm that is used to assess the purity of RNA. The ratio of absorbance at 260 nm and 230 nm is used as a secondary measurement for nucleic acid purity. The ratios of 260/230 of RNA were in the range of 2.0-2.2 and the ratio of 2.0 was considered pure for RNA. RNA quality

was also assessed by running the non-denaturing and glyoxal-denaturing agarose gel electrophoresis in which the ratio of the 28S and 18S rRNA bands were estimated of approximately 2:1. Moreover, the 2100 Bioanalyzer (Agilent Technologies) was used to estimate the RNA Integrity Number (RIN).

2.5.2.1 cDNA synthesis

RNA (1 µg) was used to synthesise cDNA using QuantiTect® Reverse Transcriptase Kit (Qiagen, Manchester, UK) following the manufacturer's protocol. gDNA Wipeout Buffer (2 µl) was added to the RNA sample and made up to a final volume of 14 µl with RNase-free water. Samples were incubated for 2 min at 42°C and placed immediately on ice. Quantiscript Reverse Transcriptase (1 µl), 1X final concentration of Quantiscript RT Buffer (4 µl) and RT primer mix (1 µl) was added to each RNA sample. Reverse transcription was performed on a thermal cycler (Biorad, Hemel Hempstead, UK) as follows: 42°C for 30 min followed by 95°C for 3 min. For a negative reverse transcriptase control, a reverse transcription step was done in the absence of reverse transcriptase. This control was to check the amount of DNA contamination present in the RNA preparation.

2.6 Quantitative Real-Time Reverse Transcription PCR (qRT-PCR)

cDNA synthesised using QuantiFast SYBR® Green PCR Kit (Qiagen, Manchester, UK) was quantified using Quantitative real-time PCR (qRT-PCR), using a C1000TM Thermal Cycler (BIO-RAD) real-time PCR system according to the manufacturer's instructions. A PCR reaction mixture (20 µl) containing 2x QuantiFast® SYBR® Green PCR master mix (10 µl), 0.5 µM final concentration forward primer, 0.5 µM final concentration reverse primer, 10 ng of DNA and RNase-free water was prepared. Three biological replicates of each sample were used in all experiments. Low-profile 96-well plates (STARLAB, Milton Keynes, UK) were used in three technical replicates for each sample. The same master mix without cDNA was used as the negative control. The two-step cycling protocol was programmed as follows: initial denaturation at 95°C for 5 min followed by 40 amplification cycles of 95°C for 10 s, 60°C for 30 s and 72°C for 30 s and a final extension step at 72°C for 5 minutes. Data analysis was performed using Delta-delta CT method (Livak and Schmittgen, 2001). Relative expression was

normalised using actin11 and 16S rRNA. The second standard (GADPH) was used as a second internal control but was varied throughout the analysis.

Primers were designed using a primer designing tool, SDSC Biology WorkBench software (http://workbench.sdsc.edu/), which includes the prediction of the formation of the self-dimers, heterodimers and hair-pin structures (http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi). The cDNA sequence was obtained from the Plant Genome and Systems Biology database. The amplicon length was between 100–160 bp with melting temperature (T_m) of 60–62°C and difference in T_m of primers pair was within 1°C. Primer length was approximately 22 bp and the GC content was set to ~50% as optimal. In most cases, the specific primers were spanning the region between CDS and 3' UTRs. Primer-BLAST was also used as a tool to confirm the primers specificity (https://www.ncbi.nlm.nih.gov/tools/primer-blast/).

Table 2.1 List of primer sequences used for qRT-PCR.

Forward (Fwd) and reverse (Rev) primers are shown. *Primers were taken from Krupinska et al., (2014).

Primer	Sequence 5'- 3'
HvWHIRLY1*	Fwd 5´-GATGGGAATGGTCGCTTTTT -3´
	Rev 5´-CCATGATGTGCGGTATGATG -3´
Hv18S rRNA*	Fwd 5´-CAGGTCCAGACATAGCAAGGATTGACAG-3´
	Rev 5´-TAAGAAGCTAGCTGCGGAGGGATGG-3´
HvRbcS*	Fwd 5´-CTACCACCGTCGCACCCTTCC-3´
	Rev 5´-TGATCCTTCCGCCATTGCTGAC-3´
HvpsbA*	Fwd 5´-CAGAAAAGCTTCCTTGACCA-3´
	Rev 5´-CAATGGTGGTCCTTATGAGC-3´
HvpetD*	Fwd 5´-GGGCGTTCTCTTAATGGTTT-3´
	Rev 5´-AATGGGTAGTGTTGCTCCAA-3´
Hv16S	Fwd 5'-TTAAGTATCCCGCCTGGGGAGT-3'
	Rev 5'-TCTCTTTCAAAGAGGATTCGCGG-3'
rpoTp	Fwd 5'-TCCTGTTGATGGGAACTGTTGGT-3'
	Rev 5'-GAGACAGCAGCATGAGGTGATGA-3'
rps16	Fwd 5'- TTCTACCGGGTGATGGCTGC-3'
	Rev 5'- CTGGACCTTGCGACGAAACAT-3'

rpoC2	Fwd 5’- CCCGCGGTTTTGAAATAAGGA-3’
	Rev 5’- TATGGCCGGTAGGAATTTGCC-3’
Hvactin11	Fwd 5’- CGACAATGGAACCGGAATG-3’
	Rev 5’CCCTTGGCGCATCATCTC-3’
HvBeta-Tub	Fwd 5’- CAAGGAGGTGGACGAGCAGATG-3’
	Rev 5’- GACTTGACGTTGTTGGGGATCCA-3’
HvLHCB1	Fwd 5’-CCCGAGACCTTTGCCAAGAA-3’
	Rev 5’- CCCTCGCTGAAGATCTGGGA-3’
HvLHCB1.1	Fwd 5’- GCAGAGCATCCTGGCCATCT-3’
	Rev 5’-TCCTTGACCTTCAGCTCGGC-3’
HvLHCA4	Fwd 5’- GGCAGGACATCAAGAACCCG-3’
	Rev 5’- TGGCGAGCTCCTTCTCCTTG-3’
HvrbcL	Fwd 5’-TTGGGTTCAAAGCCCTACGTGC-3’
	Rev 5’-ACATCCCAATAAAGGACGGCCA-3’
AK25216	Fwd 5’- TTCTACCGGGTGATGGCTGC-3’
	Rev 5’- CTGGACCTTGCGACGAAACAT-3’
MLOC_76327	Fwd 5’- CCCGCGGTTTTGAAATAAGGA-3’
	Rev 5’- TATGGCCGGTAGGAATTTGCC-3’
MLOC_58312	Fwd 5’- GAGCCCTTGAAAAGCTTCGGA-3’
	Rev 5’- CAAGCCTGGACTTGCGATGAT-3’
MLOC_59016	Fwd 5’- CAACACCCGTTTCGTCGAGTC-3’
	Rev 5’- CTGCAGCCCTCGCTTCATCTA-3’
MLOC_64606	Fwd 5’- TTGGTGTGCCTTTGGTTCTTCA-3’
	Rev 5’- GCCCAGTCCTCACGGTATTGA-3’
AK362199	Fwd 5’- GGACTGCCTTGGGTTCGACTT-3’
	Rev 5’- CTTTGGGTTGAGCCTGTGGTG-3’
MLOC_33258	Fwd 5’- TGAGAAGGCATGGTGGGACAT-3’
	Rev 5’- TGCTCTCACTGCGTTGCGTAG-3’
MLOC_77244	Fwd 5’- AAGACGGATGACAATAGCTTGGA-3’
	Rev 5’- TACCATCTCCTCCCCCTGGAA-3’

2.6.1 Determination of relative ptDNA levels

qRT-PCR analysis on relative ptDNA levels was performed on DNA (see 2.5.1) using QuantiFast SYBR® Green PCR Kit (Qiagen, Manchester, UK) on a C1000TM Thermal Cycler (BIO-RAD) real-time PCR system according to the manufacturer's instructions (see 2.6). Data was normalised to the 18S rDNA gene to determine relative ptDNA levels.

2.7 Protein extract preparation

2.7.1 Leaf protein extraction

The leaf sections were ground in liquid nitrogen using chilled mortars and pestles (100 mg fresh weight per sample). Leaf samples were then extracted in 1X protein extraction buffer (Agrisera, Vannas, Sweden). Total soluble protein content was quantified in the supernatants after centrifugation for 3 min at 10,000 × g using Pierce™ bicinchonic acid (BCA) protein assay kit (Thermo Scientific, Rockford, USA). A freshly prepared reducing agent (5 mM dithiotreitol) was added to the protein prepared for loading. Leaf extract (10 µg of protein) was mixed into 1/3 the protein volume of 4X Laemmli sample buffer (Bio-Rad,CA, USA). The samples were incubated at 70°C for 5 min.

2.7.2 Western Blots

Each sample was loaded onto a 4–20% Mini-PROTEAN® TGX™ protein gel (Bio-Rad, Herefordshire, UK), together with 5 µl of PageRuler™ Prestained Protein Ladder (Thermoscientific, Paisley, UK). The proteins were separated according to size using sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) at 100 V for 60 min. Proteins were transferred to a nitrocellulose membrane (Trans-Blot® Turbo™ Midi Nitrocellulose) using the Trans-Blot Turbo System (Bio-Rad, UK) according to the manufacturer's guidelines. The membrane was incubated in 5% skimmed milk in Tris-Buffered Tween (TBT: 20 mM Tris-HCl, 150 mM NaCl, 0.1% (v/v) Tween®) for 60 min with shaking at room temperature. Then, the membrane was incubated with primary antibody in 5% skimmed milk in TBT for 2 h with shaking at room temperature and the blot was washed 3 times in 5 min in TBST. The nitrocellulose membranes were then incubated with secondary antibodies conjugated to horse-radish peroxidase (HRP) in 5% skimmed dried milk in TBST for 1 h with shaking at room temperature. Following incubation, the nitrocellulose membranes were again washed in TBST 6 times, 5 min each time. The proteins were visualised by washing the nitrocellulose membranes in chemiluminescent substrate (SuperSignal™ West Pico PLUS, Thermo Scientific, Leicestershire, UK) and recorded using an INGENIUS gel imager (Syngene, Cambridge, UK). All proteins apart from WHY1 were detected with rabbit polyclonal primary antibody (Agrisera) and secondary HRP-linked anti-rabbit (1:10000, Agrisera AS09 602). For immunological detection of WHY1, the antibodies were directed toward the synthetic peptide of recombinant HvWhy1 protein (PRQYDWARKQVF) in rabbits and antibodies were affinity-purified (Generon, UK). The specificity of immunodetection was validated using pre-immune sera.

2.8 Northern blot analysis

Total RNA (5 µg) from the leaf section was denatured to an equal volume of NorthernMax® glyoxal dye (Ambion, MA USA), heated at 50°C for 30 min and chilled on ice. Samples were separated on a 1.2% agarose gel with 1X MOPS Buffer (0.2 M MOPS, 0.05 M sodium acetate, 0.01 M Na_2EDTA pH 7.5) and run at 80 V until 2/3 of the gel. Picture of the gel was taken using INGENIUS gel imager (Syngene, Cambridge, UK). Gel was washed with deionised water and incubated in 75 mM NaOH for 20 min, followed by Tris/NaCl (0.5 M Tris-HCL pH 7.0, 1.5 M NaCL) for 15 min and lastly, washed with 6X SSC (3 M NaCL, 0.3 M sodium citrate). The RNA was blotted onto a positively charged nylon membrane overnight and fixed by crosslinking. The membrane was then prehybridised with hybridisation buffer (ULTRAhyb® Ultrasensitive, AMBION, MA USA) for 30 min at 42°C. The primers were designed spanning the region of 23S rRNA and 4.5S rRNA. The details of primer sequences (*23S-4.5S rrn*) are as follows: forward sequence (TTCAGAACGTCGTGAGACAGTTCGGTC) and reverse sequence (CAAATCGTTCGTTCGTTAGGATGCCTC). The hybrisation probes were amplified via PCR, using a master mix comprising the following (per reaction): cDNA (15 µl), forward primer and reverse primer (2 µl), and BioMix™ Red Mix (17 µl; Bioline, London, United Kingdom). PCR amplifications were run as described in Table 2.2 using a thermal cycler (BioRad, Hemel Hempostead, United Kingdom).

Table 2.2: PCR thermal-cycling conditions

Step	Temperature, °C	Time	Number of cycles
Initial denaturation	94	5 s	1
Denaturation	94	15 s	
Annealing	60	15 s	38
Extension	72	40 s	
Final extension	72	5 min	1

PCR products were run on agarose gel at 60 V for 1 h and gel was excised as DNA fragments for the next step. Radiolabelled probes were prepared using DNA fragments (25 ng) excised from the agarose gel (PCR products) dissolved in distilled water (5-20 µl) by heating for 5 min in boiling water bath, then chilled on ice. DNA was mixed using Random Primers DNA Labelling System (Life Technologies, Paisley, UK) as per manufacturer's instruction as follows: random primers buffer mixture (15 µl), dATP solution (2 µl), dGTP solution (2 µl), dTTP solution (2 µl), ^{32}P-dCTP labelled probes (5 µl) (Pelkin Elmer, Bucks, UK), Klenow fragment (1 µl) and distilled water to total volume (49 µl). After 1h incubation at 25°C, stop buffer (5 µl) was added to the probes. Hybridisation buffer (ULTRAhyb® Ultrasensitive Hybridization buffer, Thermo Scientific, Paisley, UK) was heated to 68°C. The blot was prehybridised for 30 min at 42°C to keep the membrane thoroughly wet. Double-stranded DNA probes were denatured before hybridisation. Probe (10^6 cpm/ml) was added to the prehybridised blot. Incubation with the membrane was carried out overnight at 48°C. The membrane was washed two times in 2X SSC with 0.1% SDS, followed by high-stringency washing in 0.1X SSC and 0.1% SDS at 48°C. Detection of signal was performed using x-ray film (FUJIFILM, Tokyo, Japan)

2.9 RNA extraction and sequencing

Embryo extraction and RNA-seq analysis of excised embryos was carried out by Dorothee Wozny at the Max Planck Institute for Plant Breeding Research, Cologne, Germany. RNA was isolated from mature and dry seeds of the wild type and WHY1-deficient barley lines (W1-7) using RNA Qiagen Kit as per manufacturer's protocol. Three biological replicates per genotype were used in RNA extraction. The total of 20 embryos were excised from mature barley seeds per genotype and was ground with sterile mortars and pestles. The RNA was extracted using the Qiagen RNAEasy® Mini Kit (50) following manufacturer´s instructions (Qiagen, Hilden, Germany). RNA was stored at −80°C after DNase treatment (Ambion, Carlsbad, USA). Quality of the RNA was evaluated before library preparation using Bioanalyzer (Agilent). Illumina TruSeq libraries were prepared using the manufacturer's protocol (version 2, Illumina). Single-end sequencing was performed on the HiSeq 2000 (Illumina®) platform of the Max Planck Genome Centre, Cologne. For each library, a minimum of 15 million reads were generated by multiplexing eight libraries. Initial quality control of the raw reads was performed using the FastQC software. Reads were trimmed using the Trimmomatic platform, embedded within the Trinity pipeline (Grabherr et al., 2011) using the following the default criteria: phred 33, leading and trailing 3, sliding window 4:15 and minimum read length 36. Sequence alignments were performed with Bowtie2 (version 2.1.0; http://bowtie-bio.sourceforge.net/bowtie2/index.shtml) using a merged dataset of high confidence (HC) and low confidence (LC) predicted barley genes (The International Barley Genome Sequencing Consortium, 2012) as reference. SAMtools; version 1.2; http://www.htslib.org/ (Li et al., 2009) was used for BAM format conversion; sorting and indexing and read duplicate removal was conducted with the Picard command-line tool MarkDuplicate (version 1.110; https://broadinstitute.github.io/picard/command-line-overview.html). To correct misalignments, Genome Analysis ToolKit (GATK, version 3.1) re-aligner was used with the recommended settings. Trimmed Mean of M values (TMM) method was used to normalise library size. TMM factor was computed for each lane, with

one lane being the reference sample and the others test samples. Normalised read counts were obtained by dividing raw read counts by the re-scaled normalisation factors. Variants were obtained via the GATK UnifiedGenotyper platform (minimum phred score of 30). The variants were refined using GATK Variant Filtration tool (Fisher Strand values FS >30.0; Qual By Depth values QD <2.0) to reduce false positive SNPs. Resulting SNP calls were kept for further analysis if they passed the filtration step and their read coverage exceeded four reads. Transcripts containing filtered homozygous SNPs were mapped to their respective positions along the barley POPSEQ map (Mascher et al., 2013) using R (https://www.r-project.org). For expression analysis, the reads were aligned to high confidence (HC) and low confidence (LC) gene sets (The International Barley Genome Sequencing Consortium, 2012) as described above and only in this case, the read duplicates were not removed from the BAM file. Raw counts were extracted from the BAM file using Salmon (Patro et al., 2017). Differentially expressed genes were those showing fold change of >2 and a false discovery rate (FDR) corrected *p*-value of 0.05 or less using the R bioconductor package limma-vroom (Ritchie et al., 2015).

2.9.1 2.9.1 GO analysis

The raw data processing procedures of RNAseq were done by Michael Wilson, University of Leeds, UK. An enhanced set of Genome Ontology terms for the IGSB v1 (2012) transcripts used for RNA seq mapping was created, by using GO annotations from IGSB v1, combined with GO annotations for IGSB v2 (2015) where mappings existed in uniport, enhanced by identifying missing transcripts using blastn from the BLAST+ package. GO slim terms were created using these full GO annotations using owltools (https://github.com/owlcollab/owltools) and goslim_plant (http://www.geneontology.org/ontology/subsets/goslim_plant.obo) from the GO Consortium. GO term enrichment was performed using TopGo (Alexa and Rahnenfuhrer, 2018) in R/Bioconductor (Huber et al., 2015). RNA seq reads are available in ArrayExpress Archive of Functional Genomics

Data(https://www.ebi.ac.uk/arrayexpress/) under accession number E-MTAB-7015.

2.10 Microarray processing and analysis

Microarray processing was carried out by Jenny Morris at the James Hutton Institute, Dundee, Scotland. Data extraction, quality control analysis and initial statistical analysis were performed by Pete Hedley at the James Hutton Institute, Dundee.

Base, mid and tip sections (see section 2.2) of the first leaves of 7-day old and 14-day old barley seedlings were used for microarray analysis. The Qiagen® RNeasy Plant Mini Kit was used for total RNA extraction according to the manufacturer's protocol (see section 2.5.2). Microarray processing was performed on leaf RNA extracts from four biological replicates per genotype (WT, W1-1 and W1-7) using a custom-designed barley Agilent microarray (MEXP-2357; www.ebi.ac.uk/arrayexpress) which represents transcripts of the entire barley genome reference (as described in Comadira et al. (2015). The microarray contains approximately 61,000 60-mer probes derived from predicted barley transcripts and full-length cDNAs. The 'One-Color Microarray-Based Gene Expression Analysis' protocol (v. 6.5; Agilent Technologies) was used to run microarray processing. In brief, cDNA was transcribed into cRNA which was amplified, labelled with Cy3 dye, purified and hybridised to the array slides overnight (Figure 2.2). The next day, hybridised slides were washed twice and dried as recommended. The hybridised slides were scanned using a Agilent G2505B scanner at a resolution of 5 µm at 532 nm.

Feature Extraction (FE) software (v. 10.7.3.1; Agilent Technologies) was used for quality control and data extraction using the default settings as recommended. Extracted data for each microarray was subsequently imported into GeneSpring GX (v. 7.3; Agilent Technologies) software for statistical analysis. Data

normalisation was performed using default Agilent FE one-colour settings in GeneSpring and a filter function was used to remove inconsistent probe data, flagged as absent in two or more replicates per sample. Principal Component Analysis (PCA) was performed using default settings to identify relationships between replicate samples. Probes were identified as significantly changing between the WT and WHY1-deficient lines across leaf position using 2-way Analysis of Variance (ANOVA), with 'genotype' (WT, W1-1 and 1-7) and 'position' (base, mid, tip) as factors, and Bonferroni multiple-testing correction at a *p*-value ≤0.05. Pairwise comparisons between WT and W1-7, for each leaf position, were performed using volcano plots, combining cut-offs of *t*-test *p*-value ≤0.05 and fold-change ≥2x. Heatmaps were generated from selected genelists using the Genetree function in Genespring, with clustering based upon Pearson's correlation and default parameters.

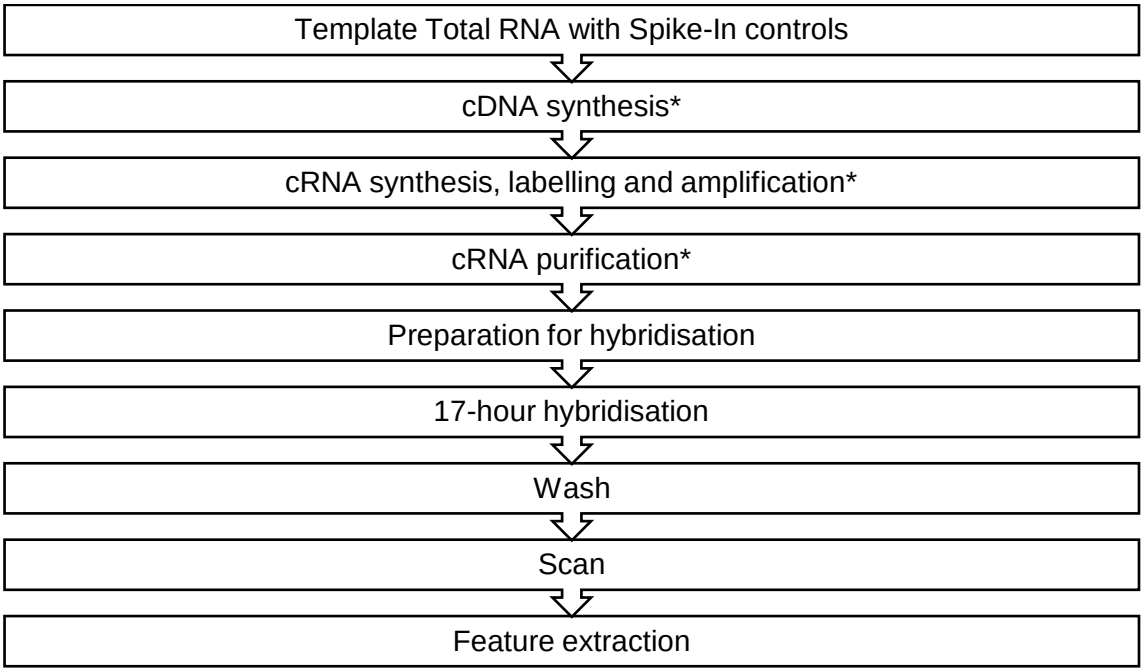


Figure 2.2: A simple flowchart of sample preparation and array processing for microarray processing.

2.11 Metabolite analysis

Metabolomics processing was carried out by Katie Schulz at the James Hutton Institute, Dundee. Data extraction and the initial statistical analysis were perfomed by Susan R. Verrall at the James Hutton Institute, Dundee.

The base, middle and tip sections of frozen first leaves of the barley seedlings grown under normal conditions were weighed (>100 mg fresh weight per sample) and freeze-dried for 24 h. The samples were then lyophilised using a Gamma 1-16 LSC freeze drier (Martin Christ Gefriertrocknungsanlagen GmbH, Germany) at a pressure of 0.7 mbar, with a shelf temperature of 25°C and a condenser temperature of −50°C. To extract and derivatise polar and non-polar metabolites from freeze-dried samples, sequential extraction with methanol, chloroform and water, and the presence of ribitol and nonadecanoic acid methyl ester as internal standards were used. Retention standards were added to aliquots of the derivatised polar and non-polar samples and later were measured by GC-MS. The protocol used for the extraction and analysis are described in detail below.

2.11.1 Extraction of polar and non-polar fractions

Freeze-dried barley leaves were weighed out in a culture tube (150 × 16 mm). Extraction volumes were adjusted according to sample weight. Methanol (1.5 ml) was added to each sample, then tubes were capped and transferred to a vortex-type shaker and shaken at 1500 revolutions min^{-1} at 30°C for 30 min in an incubator. Polar ribitol 2 g L-2 in water (50 µl) and non-polar nonadecanoic acid methyl ester 0.2 g L-2 in methanol (50 µl) internal standards were added to samples with 0.375 ml water. The samples were subjected to shaking for a further 30 min at 1500 revolutions min^{-1} at 30°C. Chloroform (3 ml) was added and the mixtures were shaken at 2500 revolutions min^{-1} at 30°C for 30 min. Next, water (0.75 ml) was added to each mixture and the samples were vigorously shaken by hand. The polar and non-polar phases were separated by centrifugation at 1200 g for 10 min. The polar (upper layer) and non-polar (lower layer) were transferred to amber vials using Pasteur pipettes and stored in −20°C until next day.

2.11.2 Derivatisation of polar fraction

The polar extracts were removed from freezer and left to warm to room temperature. The polar fractions (250 µl) were pipetted into culture tubes (100 × 16 mm) which were transferred to a centrifugal extractions evaporator until dry. Methoxylamine hydrochloride (80 µl) (20 mg methoxylamine hydrochloride/ml anhydrous pyridine) were added to the dried polar fraction to oximate the carbonyl functional groups at 50°C for 4 h in an incubator. During incubation, the retention standard mixture (50 µl) (undecane, tridecane, hexadecane, eicosane, tetracosane, triacontane, tetratriacontane and octatriacontane) which were dissolved in isohexane were added to amber autosampler vials (300 µl fixed glass inserts with PTFE-coated snap caps) and the isohexane was allowed to evaporate from the vials at room temperature. After oximation, N-methyl, N-trimethylsilyl trifluroacetamide (MSTFA) (80 µl) was added to samples, which were then incubated for 30 min at 37°C. Finally, the derivatised polar fractions (40 µl) and dry pyridine (40 µl) were added to the amber autosampler vials

containing the dried retention standards. The polar fraction was then ready for GC-MS analysis.

2.11.3 Derivatisation of non-polar fraction

The non-polar fraction was dried in a centrifugal evaporator for 30 min. Then, chloroform (1 ml) and1% methanolic sulphuric acid (2 ml) was added. The mixture was incubated at 50°C for 16 h to release free fatty acids. The tubes were then cooled down to room temperature. Aqueous sodium chloride (5 ml of 5% (w/v)) and chloroform (3 ml) were added to each tube, which were then shaken vigorously to allow the polar and non-polar layers to settle. The top aqueous layer was discarded and aqueous potassium hydrogen carbonate (3 ml of 2% (w/v)) was added to the lower chloroform:methanol layer. Samples were vigorously shaken and then, the layers were allowed to settle. After settling, the upper layer was discarded again and the chloroform:methanol layer (lower layer) was pipetted through columns containing anhydrous sodium sulphate (3 cm). Columns were Pasteur pipettes plugged with cotton wool and prewashed with chloroform (4 ml) to remove all residual water. The columns were washed with extra chloroform (2 ml) that was collected with the fractions. The fractions were dried down in the centrifugal evaporator for 60 min. Next, chloroform (50 µl), anhydrous pyridine (10 µl) and MSTFA (40 µl) were added and incubated at 37°C for 30 min in an incubator. Retention standard mixture (50 µl) was added into amber autosampler vials and the isohexane was allowed to evaporate from the vials at room temperature. Finally, the derivatised non-polar fraction (40 µl) and anhydrous pyridine (40 µl) were added to autosampler vials that had been prepared with retention time standards as described in 2.11.1 The non-polar fraction was then ready for analysis by GC-MS.

2.11.4 Sample analysis

Samples were analysed using DSQ II Single Quadrupole GC-MS system (Thermo). The samples (1 µl) were injected with a split ratio of 40:1 into a programmable temperature vaporising injector according to the following conditions: injection temperature of 132°C for 1 min, transfer rate 14.5°C/s, transfer temperature 320°C for 1 min, clean rate 14.5°C/s and clean temperature 400°C for 2 min. Analytes were chromatographed on a DB5-MSTM column (15 m × 0.25 mm × 0.25 µm; J&W, Folsom, USA) using helium at 1.5 ml/min in constant flow mode as the mobile phase. The temperature gradient was 100°C for 2.1 min, 25°C /min to 320°C and isothermal for 3.5 min and the interface temperature was 250°C. Mass data were attained at 70 eV electron impact ionisation conditions over a 35–900 a.m.u mass range at 6 scans per sec with a source temperature 200°C and a solvent delay of 1.3 min. Acquisition rates were set to give approximately ten data points across a chromatographic peak. XcaliburTM v1.4 and XcaliburTM v2.0.7 software packages were used to analyse the data. A processing method at James Hutton Institute was used to determine identities of each metabolite to the peaks. The software uses the retention times and masses of known standards. TheGenesis algorithm (part of the XcaliburTM package)was used to measure peak integration. The expected retention time for each peak was adjusted using the retention times of the retention standards. The integrated area of the annotated peaks was normalised against the integrated area of the respective fractions for internal standards, ribitol (polar) and nonadecanoic acid (non-polar). The peak area ratios were normalised based on a dry weight basis. Statistical analysis for metabolite data was performed by two-way analysis of variance (ANOVA) with a *p*-value of <0.05.

2.12 Subcellular localisation of LEA5.

Subcellular localisation of LEA5 protein was predicted using MitoProt II (version 1.101) Target P and IPSORT where the sequence of LEA5 proteins were used to predict LEA5 intracellular localisation.

2.13 Protein-protein interaction analysis

In this study, protein-protein interaction analysis was performed using Bimolecular fluorescence complementation (BiFC) assays. Full length LEA5 (AtLEA5-nYFP or AtLEA5-cYFP), AtWHY1 (WHY1-nYFP and WHY1-cYFP) and full length DEA (D/H)-box RNA helicase cDNAs were cloned into the pDH51-GW-YFPn vector and pDH51-GW-YFPc vectors, respectively with appropriate controls. These were expressed in *Arabidopsis* protoplast. YFP fluorescence was analysed 24 h after transfection using a confocal laser scanning microscope.

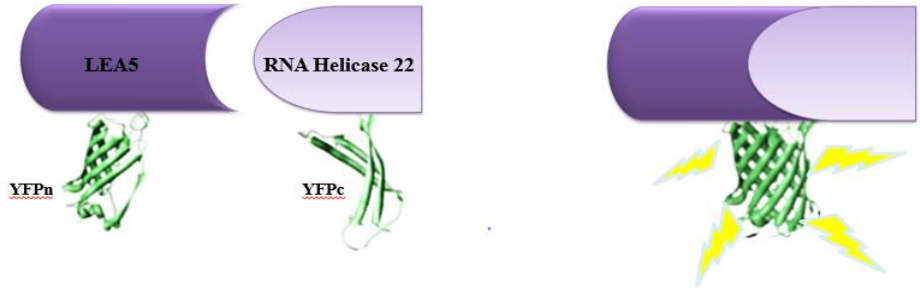


Figure 2.3 : Schematic diagram of split YFP/BiFC analysis showing interaction of AtLEA5-YFP and DEA (D/H)-box RNA helicase 22 constructs.

YFP is split into two non-fluorescent halves (N-terminal half (YFPn) and C-terminal half (YFPc)) which fuse to the protein of interest. The YFP molecule is reconstituted upon interaction between the two different protein halves when the molecules are excited with the correct wavelength.

2.13.1 Primer design of DEA (D/H)-box RNA helicase

The primers used to amplify DEA (D/H)-box RNA helicase are shown in Table 2.3.

Table 2.3: PCR primers for DEA (D/H)-box RNA helicase 22.

Kozak sequence (underlined) and DEA (D/H)-box RNA helicase sequences (bold). Four guanine (G) residues at the 5 end followed by attB1 site (italics).

Primer		Sequence
Fwd DEA (D/H)-box RNA helicase 22	*att*B-	5'-GGGG*ACAAGTTTGTACAAA*AAAGCAGGCTCCACC**ATGATTCTCTCACGCTCTGTCTCC**-3'
Rev DEA (D/H)-box RNA helicase 22	*att*B-	5'-GGGG*ACCACTTTGTACAAGAAAGCTGGGTC***ATA TCTCACAGCTTGAGGCTCCTC-3'**

2.13.2 LEA5 interactions with DEA (D/H)-box RNA helicase 22

LEA5-pDONR201 plasmid used in the following studies was kindly provided by Daniel Shaw (Faculty of Biological Sciences, University of Leeds). With the help of Dr Christopher West (Faculty of Biological Sciences, University of Leeds), the full length DEA (D/H)-box RNA helicase cDNA was cloned into split YFP vectors (pDH51-GW-YFPn and pDH51-GW-YFPc) that contained an ampicillin resistance gene used for selection.

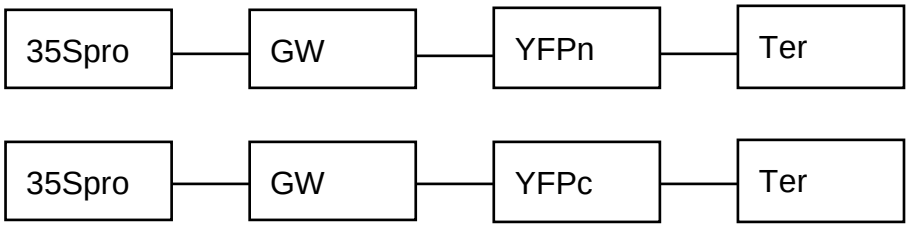


Figure 2.4: Features of pDH51-GW-YFPn and pDH51-GW-YFPc plasmid.

pDH51-based vectors; GW, gateway cassette with attR1 and attR2 recombination sites; cauliflower mosaic virus 35S promoter and terminator; Amp, ampicillin resistance; Cm, chloramphenical resistance; YFPN, N-terminus of Venus (aa 1-154); YFPC, C-terminus of Venus (aa 155-238). AM779183; AM779184.

2.13.3 Construction of the LEA5-YFPc and LEA5-YFPn vectors

The pDONR201+*LEA5* plasmid was used in LR reaction with N9842 to produce LEA5-YFPc plasmid and used in LR reaction with N9843 to produce LEA5-YFPn plasmid. LR recombination reaction was performed according to the handbook manual (Thermofisher Gateway® Technology manual, Paisley, UK). A reaction volume containing 50-150 ng of the entry clone, 1 µl (150 ng/µl) of the destination vector and TE buffer (8 µl) pH 8.0 were mixed briefly. LR clonase™ II (2 µl) was added to each sample and mixed well by vortexing briefly twice. Then, the reaction was incubated at 25°C for 1 h. After that, proteinase K solution (1 µl) was added to terminate the reaction and incubated for 10 min at 37°C. The reaction was stored at −20°C prior to protoplast transfection.

2.13.3.1 Competent cells transformation

Each LR reaction (1 µl) was transformed into DH5α competent cells (100 µl). The cells were incubated on ice for 10 min and swirled gently every 2 min. The cells were incubated on ice for 30 min and heat-shocked by incubation at 42°C for 30 s. Next, the cells were incubated on ice for 2 min. After that, preheated SOC medium (0.9 ml) was added and incubated at 37°C for 1 h with shaking at 200 rpm. Each transformation was plated (20 µl and 100 µl) onto selective plates with appropriate antibiotic selection and incubated at 37°C overnight. The next day, six colonies from each plate were selected. Colonies were inoculated in LB broth (2 ml) containing appropriate antibiotics of 50 µg/ml kanamycin and 100 µg/ml ampicillin, incubated at 37°C with shaking at 200 rpm overnight.

2.13.3.2 Plasmid extraction

Overnight cultures (2 ml) incubated with vigorous shaking at 200 rpm at 37°C were used for plasmid purification using a Qiagen Plasmid Maxi Kit for large-scale plasmid extraction (up to 250 pg) (Qiagen, Manchester, UK).

Table 2.4: List of plasmids used in protein-protein interaction studies using split-YPF.

Name	Selection	Origin
pGREEN-35S-LEA5-YFP	Kan	Salleh et al., 2012.
35S-WHIRLY1-YFPn 35S-WHIRLY1-YFPc	Amp	Barbara Karpinska, Faculty of Biological Sciences, University of Leeds.
pGREEN::AOX::RFP	Kan	Estavillo et al., 2011.
pGREEN::SSU::RFP	Kan	Estavillo et al., 2011.

2.13.4 Protoplast isolation

Protoplasts were isolated according to protocol with minor modifications (Wu et al., 2009). Leaves (width: 2 cm, length: 5 cm) were collected from 3 to 5-week old plants grown under long photoperiod light conditions (see section 2.1.3). The upper epidermal leaves of *Arabidopsis* were placed by attaching a strip of Time tape (Time Med, Burr Ridge, IL), while the lower epidermal layer was affixed to a strip of Magic tape (3M, St. Paul, MN). The Magic tape was carefully peeled from the leaf to remove the lower epidermal layer. The peeled leaves (7 to 10 optimal-light-growth leaves, about 1-2 g, up to 5 g) still attached to the Time tape was transferred to a Petri dish with enzyme solution (20 ml). The enzyme solution composition was as follows: (1% (w/v) cellulase 'Onozuka' R10 (Yakult, Tokyo, Japan), 0.25% (w/v) macerozyme 'Onozuka' R10 (Yakult), 0.4 M mannitol, 10 mM $CaCl_2$, 20 mM KCl, 0.1% BSA and 20 mM MES, pH 5.7). The leaves were gently shaken (40 rpm) in light for 60 min until the protoplasts were released into the solution. The protoplasts were then centrifuged at 100 × g for 3 min in an Eppendorff A-4-44 rotor (Hamburg, Germany), washed with 50 ml pre-chilled modified W5 solution (154 mM NaCl, 125 mM $CaCl_2$, 5 mM KCl, 5 mM glucose and 2 mM MES, pH 5.7) and incubated on ice for 30 min. During the incubation period, a sample of mesophyll protoplasts was viewed under a light microscope to capture the image and visually count for intact and round protoplasts. After 30 min of the incubation period, the protoplasts were then centrifuged (100 × g for 3 min) and resuspended in modified MMg solution (0.4 M mannitol, 15 mM $MgCl_2$ and 4 mM MES, pH 5.7) to a final concentration (2 to 5 × 10^5 cells/ml).

Table 2.5: List of plasmids used in protoplast transfection.

Amino peptidase P2 (APP2-GFP) as chloroplast marker and amino peptidase P2 (APP1-GFP) as cytosolic marker were used as positive control for localisation. The same halves of split-YFP (nYFP), in this case, LEA5-nYFP and DEA (D/H)-box RNA helicase 22-nYFP was fused and used as the negative control for split-YFP system.

Positive control	i) ii)	APP2-GFP (At3g05350) APP1-GFP (At4g36760)
Split-YFP	i) ii) iii) iv)	LEA5-nYFP + DEA (D/H)-box RNA helicase 22-cYFP LEA5-cYFP + DEA (D/H)-box RNA helicase 22-nYFP WHIRLY1-nYFP + DEA (D/H)-box RNA helicase 22-cYFP WHIRLY1-cYFP + DEA (D/H)-box RNA helicase 22-Nyfp

2.13.5 Protoplast Transfection Assays

Transfection of protoplast was performed using the method described by (Yoo et al., 2007b), with minor modifications. Plasmid DNA (10 µg) was mixed gently with protoplasts (200 µl), and an equal volume of freshly prepared PEG was added (MW 4000; Fluka; 40 % (w/v)) containing 0.1 M $CaCl_2$ and 0.2 M Mannitol. Samples were mixed and incubated at room temperature for 5 min and W5 solution (3 ml) was added slowly; next, the solution was mixed and protoplasts pelleted (100 × g for 1 min). The supernatant fractions were removed and transformed protoplasts were resuspended gently in W5 solution (1 ml), then transferred to 6-well plates. The protoplasts were left at room temperature for 16 h in the dark to allow transfection.

2.14 Imaging

2.14.1 Light microscopy

Protoplast solution (5 µl) was placed on a slide and covered using a cover slip for examination by light microscopy (Leica DM 2500). Images were captured using Nikon (D5000) and Nikon Camera Control Pro V.1 Software.

2.14.2 Confocal Laser Scanning Microscopy

2.14.2.1 Confocal microscopy of intact LEA5-YFP leaves

YFP expression was recorded in at 5-day old *Arabidopsis* seedlings in the light period by Zeiss LSM700 inverted confocal microscopy (Faculty of Biological Sciences, University of Leeds). The YFP signal was detected using an excitation wavelength range of 488 nm and an emission wavelength of 530 nm. The chlorophyll auto-fluorescence was excited simultaneously with the 555 nm laser, emission was detected at 650–710 nm.

2.14.2.2 Visualisation of interaction

Protoplasts were observed using a LSM700 inverted confocal microscope using 20X/0.8 PlanApochromat or 40X/1.2 WC-Apochromat in multi-track channel mode. Excitation wavelength was 488 nm and emission filter of YFP was 530 nm, with simultaneous excitation a band-pass of 650–710 nm for the detection of chloroplast auto-fluorescence. Image processing was performed using Zeiss ZEN 2011 (Black Edition) v7.1 and ImageJ v1.46r Translational research

Chapter 3 . Characterisation of WHIRLY1-deficient barley seeds

3.1 Introduction

The WHY1 protein, which is localised in the chloroplasts and nucleus of the same cell, has been shown to play a number of diverse and important roles in the regulation of chloroplast nucleosome structure and nuclear (Grabowski et al., 2008; Krupinska et al., 2014). WHY1 binds to both ssDNA and RNA and functions in intron splicing in maize chloroplasts (Prikryl et al., 2008). WHY1 is also associated with intron-containing RNA in barley chloroplasts (Melonek et al., 2010). WHY proteins were first described as a nuclear transcriptional activator that binds to an elicitor response element in the promoter region of pathogenesis-related genes in potato (Desveaux et al., 2000). In addition to pathogen-responsive genes, WHY proteins also bind to promoter regions of several genes in the nucleus associated with senescence including *WRKY53* (Miao et al., 2013) and *HvS40* in barley (Krupinska et al., 2014). WHY proteins also regulate the length of telomeres (Yoo et al., 2007a) and bind to the distal element upstream of the kinesin gene (Xiong et al., 2009) which is necessary for repression of promoter activity fully and partially in the cotyledon and roots in *Arabidopsis*. Such findings suggest that WHY proteins have different functions at different developmental stages, depending on their intracellular localisation (Ren et al., 2017). The regulated partitioning of WHY1 between chloroplasts and nuclei at different developmental stages has been shown to be phosphorylated by the MAP kinase, calcineurin B-Like-Interacting Protein Kinase14 (CIPK14) and transported into the nucleus (Ren et al., 2017). The WHY1 protein is localised in chloroplasts in young leaves but it accumulates more in the nuclei of senescence leaves (Ren et al., 2017). Overexpression of *CIPK14* showed the stay-green phenotype, which was recovered by overexpression of plastid-form WHY1,

linking CIPK14 functions to senescence (Ren et al., 2017). In the chloroplasts and mitochondria, the WHY proteins function as anti-recombinant factors that are required for accurate DNA repair and the maintenance of organellar genome stability (Cappadocia et al., 2010, Lepage et al., 2013).

A triple mutant (*Atwhy1why3polIb-1)* that lacks WHY1, WHY3 and type I chloroplast DNA polymerase 1B (Pol1B) exhibits a severe yellow-variegated phenotype (Lepage et al., 2013). This mutant had a lower photosynthetic electron transport efficiency and accumulates reactive oxygen species with altered expression of redox-regulated genes compared to the wild type (Lepage et al., 2013). In the previous studies, the RNAi knockdown lines with less than 5% of wild type WHY1 protein (W1-1, W1-7 and W1-9) were shown to influence the expression of specific subsets of genes encoding chloroplast proteins (Comadiraet al., 2015). Several transcripts that encode chloroplast-localised proteins, such as ribosomal proteins, subunits of the RNA polymerase, and thylakoid nicotinamide adenine dinucleotide (reduced) and cytochrome b6/f complexes were much higher in the W1-7 barley leaves than the wild type, resulting in lower sensitivity of photosynthesis to low nitrogen (Comadira et al., 2015). In chloroplasts, WHY1 is not only associated with nucleoids but also with the thylakoid membrane (Foyer et al., 2014, Huang et al., 2017). However, the mechanisms by which WHY1 influences chloroplast function and participates in chloroplast signalling remain poorly understood.

Previously, it was shown that WHY1 overexpressing *Arabidopsis* seeds had an altered response to ABA during seed germination. In these experiments, the *Arabidopsis why1* mutants had been transformed with constructs that allowed expression of WHY1 targeted to either the plastids or the nuclei (Isemer et al., 2012a). The plastid-localised WHY1 enhanced the responsiveness of germination of the *Arabidopsis* seeds to ABA (Isemer et al., 2012a). In contrast, when WHY1 was expressed only in the nuclei there was no change in the sensitivity of the seeds to ABA (Isemer et al., 2012a). The following studies were performed to determine whether the absence of WHY1 had an effect on the

transcriptome profile of the embryos of the dry seeds. This information would provide a better understanding of how transcripts encoding components of hormone signalling pathways were affected by the absence of WHY1. The embryos of the dry seeds of the W1-7 and the wild type were harvested and subjected to RNA-seq analysis. In this chapter, the characterisation of WHY1-deficient barley seeds is described. These data are discussed both in terms of the effects of WHY1 deficiency in the mother plant, which determines the composition and abundance of transcripts in the embryo, and the effects of these changes on germination. These data lay the foundations for the subsequent studies on the functions of WHY1 during chloroplast biogenesis. The RNA-seq analysis was performed in collaboration with Professor Maarten Koorneef at the Max Planck Institute in Cologne, Germany. Data were analysed in collaboration with Dr Michael Wilson at the University of Leeds. In the following analysis, only transcripts that were differentially changed above the 1.5 fold cut-off threshold ($p<0.05$) were selected.

3.2 Results

3.2.1 Characterisation of WHY1- deficient barley seeds

3.2.1.1 Seedling phenotype of WHY1-deficient barley

The first true leaves of the W1-7 showed a delayed in greening relative to the wild type, 4 days after germination (Figure 3.1).In contrast, the first true leaves of the W1-1 seedlings showed no phenotypic difference to the wild type, 4 days after germination (Figure 3.1).

Figure 3.1: Comparison of 4-day old seedlings of transgenic W1-1 and W1-7 line phenotypes to the wild type.

Seeds were kept at 4°C for 3 days before the seedlings were sown in pots in soil in controlled environment chambers with a 16h light/ 8h dark photoperiod, irradiance of 200 µmol $m^{-2}s^{-1}$, 20°C/16°C day/night temperature regime and 60% relative humidity. (Scale bar =1 cm).

3.2.1.2 Seed germination

Seeds germination were significantly lower in the W1-7 than the wild type at 24, 48 and 72 hours after imbibition (Figure 3.2). Furthermore, seeds germination were also lower in the W1-1 relative to the wild type at 24, 48 and 72 hours after imbibition (Figure 3.2). In general, the WHY1-deficient barley lines had lower germination rates compared to the wild type at 24, 48 and 72 hours after imbibition (Figure 3.2).

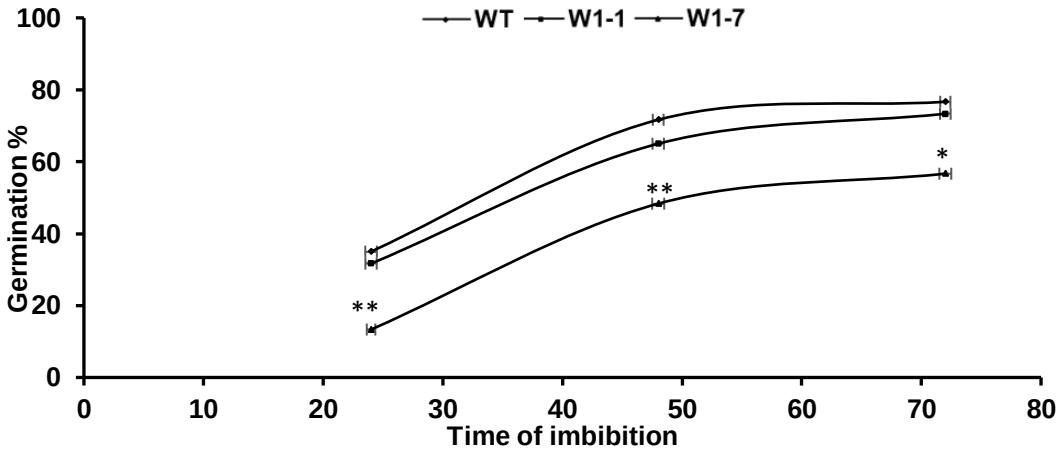


Figure 3.2: Germination characteristics of transgenic W1-1 and W1-7 lines to the wild type (WT).

Percentage of seed germination in the wild type and WHY1-deficient barley seedlings (W1-1 and W1-7) during imbibition. Error bars illustrate standard deviation (n=60). Asterisks indicate significant differences between WHY1-deficient and wild type plants as estimated by the Student's *t*-test (*p<0.05) and (**p<0.01).

3.2.1.3 Seed characteristics

Seed size was compared in the WHY1-deficient barley seeds and the wild type. There were no significant differences in the WHY1-deficient barley seeds compared to the wild type in seed length (Table 3.1). However, the seed width was significantly lower in the W1-1 seeds compared to the wild type. In contrast, seed width of the W1-7 was similar to the wild type (Table 3.1).

Table 3.1: Characteristics of seeds produced by WT, W1-1 and W1-7 barley.

Seed length and width were determined as mean values of 15 seeds from each line. Values are reported as mean ± SE. Values are reported as mean ± SE. Values that were significantly different determined using the Student's *t*-test are indicated by asterisks (****p<0.0001).

Seeds	Length (cm) (%)	Width (cm) (%)
WT	0.88 ±0.02 (100)	0.36±0.01 (100)
W1-1	0.83±0.02 (94.60)	0.31±0.01****(85.60)
W1-7	0.83±0.01 (94.70)	0.35±0.01(97.42)

3.2.1.4 Yield parameters in WHY1-deficient (line W1-7) and wild type barley plants

Seeds of transgenic barley (*Hordeum vulgare* L. cv. Golden Promise) line W1-7, with RNAi knockdown of the *WHY1* expression and wild type controls were obtained as described by Comadira et al. (2015). For the following analysis of yield parameters, only plants of the W1-7 line and the wild type were grown in compost to maturity in a standard heated greenhouse at the James Hutton Institute (Scotland) under a 16-h photoperiod regime at 22°C, where supplementary lighting was provided by high-pressure sodium vapour lamps.

Seeds of the T4 generation of the WHY1-deficient (W1-7) plants were divided into 4 groups at random [W1-7(1), W1-7(2), W1-7(3) and W1-7(4)]. Up to 20 plants were grown to maturity in each group together with the wild type. Measurements of number of tillers and total seed weight were performed once the seeds had matured. Some variations were observed in the data obtained from each set of W1-7 plants (Table 3.2). For example, the W1-7(1), W1-7(2) and W1-7(3) plants had the same number of tillers as the wild type. However, the W1-7(4) plants had significantly fewer tillers (Table 3.2). Total seed yield was also higher in the W1-7(2) and W1-7(3) lines than the W1-7(1), W1-7(4) and wild type plants (Table 3.2). However, the mean values of the combined W1-7(1), W1-7(2), W1-7(3) and W1-7(4) data provide an accurate assessment of the W1-7 phenotype.

Table 3.2: : A comparison of yield parameters in the wild type barley and WHY1-deficient plants.

Seeds from 4 different plants of the T4 generation transgenic line WHY1-7 (Comadira, 2015) were grown to maturity. In each case, 30 plants were grown from seeds from 4 plants. The number of fertile tillers were counted and total seed yield quantified in plants grown to maturity in controlled environment glasshouses. Data is presented as mean values ± SE (n=30). Values that were significantly different between wild type and W1-7 plants as determined using the Student's *t*-test are indicated by asterisks (*p<0.05) and (**p<0.01).

	Wild type	**W1-7(1)**	**W1-7(2)**	**W1-7(3)**	**W1-7(4)**
Number of fertile tillers	17.25±3.09	15.74±0.76	13.68±0.64	13.50±0.58	12.21±0.67*
Seed yield per fertile tiller (g)	0.80±0.13	1.07±0.06	1.27±0.08*	1.26±1.05**	1.47±0.21
Total seed yield (g)	12.66±0.26	16.14±0.64*	16.25±0.63*	16.70±0.76	15.43±0.67

Average data from all four WHY1-7 plants was compared to wild type barley plants grown under the same conditions (Table 3.3). Although the number of fertile tillers tended to be lower in the WHY1-7 plants, the values obtained were not significantly different to wild type. In contrast, total seed yield was significantly higher in the WHY1-7 plants than the wild type (Table 3.3). Seed yield per tiller was the same in both lines, as previously reported (Comadira et al., 2015). The data in Table 3.3 show a trend for higher seed yields per tiller in the W1-7 plants. The difference in the yield of the T4 generation in WHY1-deficient plants (Table 3.3) is marked compared to the T3 generation plants (Comadira et al., 2015). The difference in yield of the W1-7 plants in the T3 and T4 generations could be due to the different soil quality, humidity, light and intensity. These results may be explained by generation to generation variations. Interestingly, seed yield of the 4th generation plant was higher than the wild type. It may be deduced therefore that the WHY1-deficient plants have the potential to produce greater seed yields than the wild type. However, this analysis must be repeated in future generations to determine generic trends and need to be repeated in the W1-1 line as only W1-7 line was available during this study.

Table 3.3: A comparison of yield parameters in W1-7 and wild type barley.

Plants were grown to maturity from seeds in controlled environment glasshouses. The number of fertile tillers were counted and total seed yield quantified. Data is presented as mean values ± SE (n=30). Values that were significantly different between wild type and W1-7 plants as determined using the Student's *t*-test are indicated by asterisks (*p<0.05).

	Wild type	**W1-7**
Number of fertile tillers	17.25 ± 3.09	13.82 ± 0.35
Total seed yield (g)	12.66 ± 0.26	16.16 ± 0.34*
Seed yield per fertile tiller (g)	0.80 ± 0.13	1.26 ± 0.05

3.2.2 Functional categorisation of differentially-regulated transcripts

Transcripts that were differentially altered in the embryos of the W1-7 seeds relative to the wild type using RNA-seq analysis were first analysed using gene ontology (GO) enrichment analysis. This is an efficient tool providing a first overview of the overrepresented functional gene groups. Genes related to enzyme activity, mostly hydrolytic activities, were significantly changed by the loss of WHY1 (Figure 3.3). In addition, transcripts encoding proteins associated with DNA/RNA binding and nucleotide binding were differentially changed in the W1-7 embryos relative to the wild type (Figure 3.3). Analysis of functional groups associated with biological processes shows an altered abundance of transcripts involved in metabolism, as well as responses to abiotic and biotic stresses (Figure 3.4). A large number of differentially-expressed transcripts were targeted to either the nucleus or plastids (Figure 3.5).

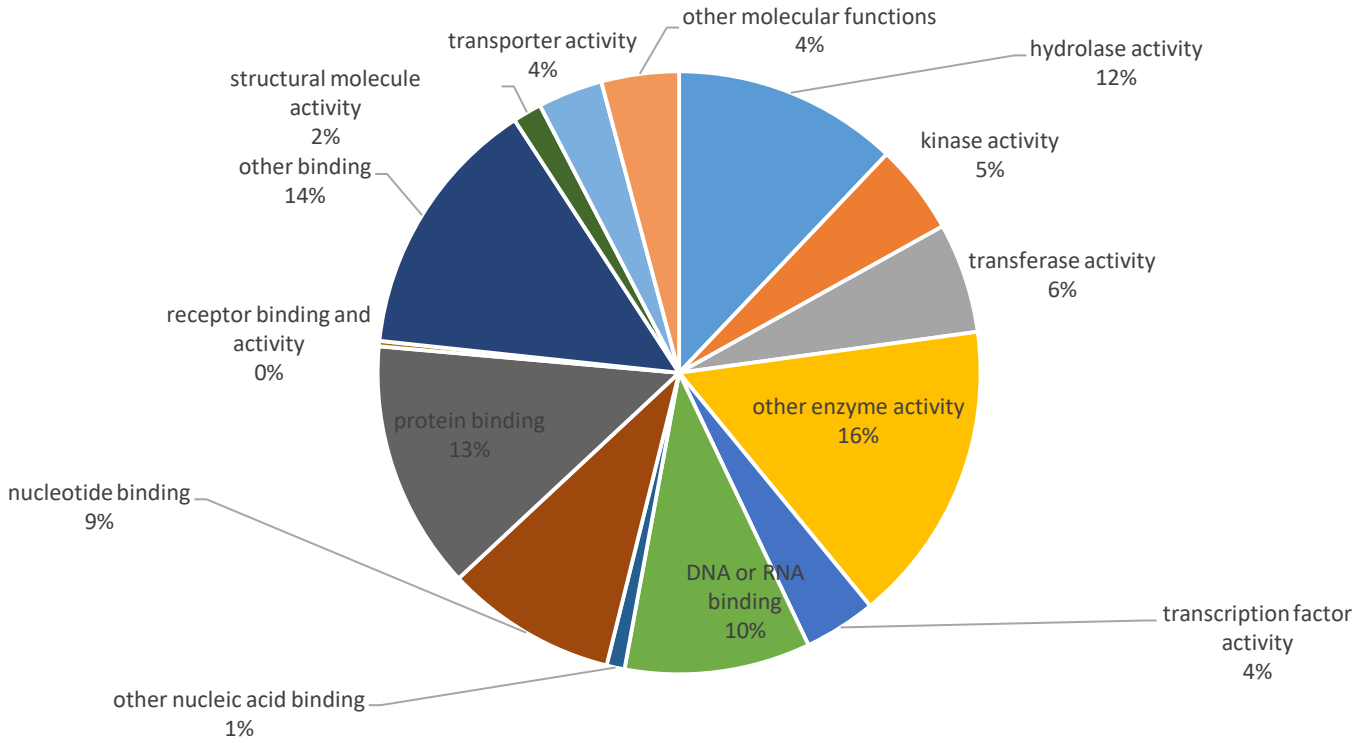


Figure 3.3: Gene ontology enrichment analysis for transcripts that were increased in abundance.

GO enrichment analysis was performed based on molecular function in the embryos of W1-7 barley seeds relative to the wild type.

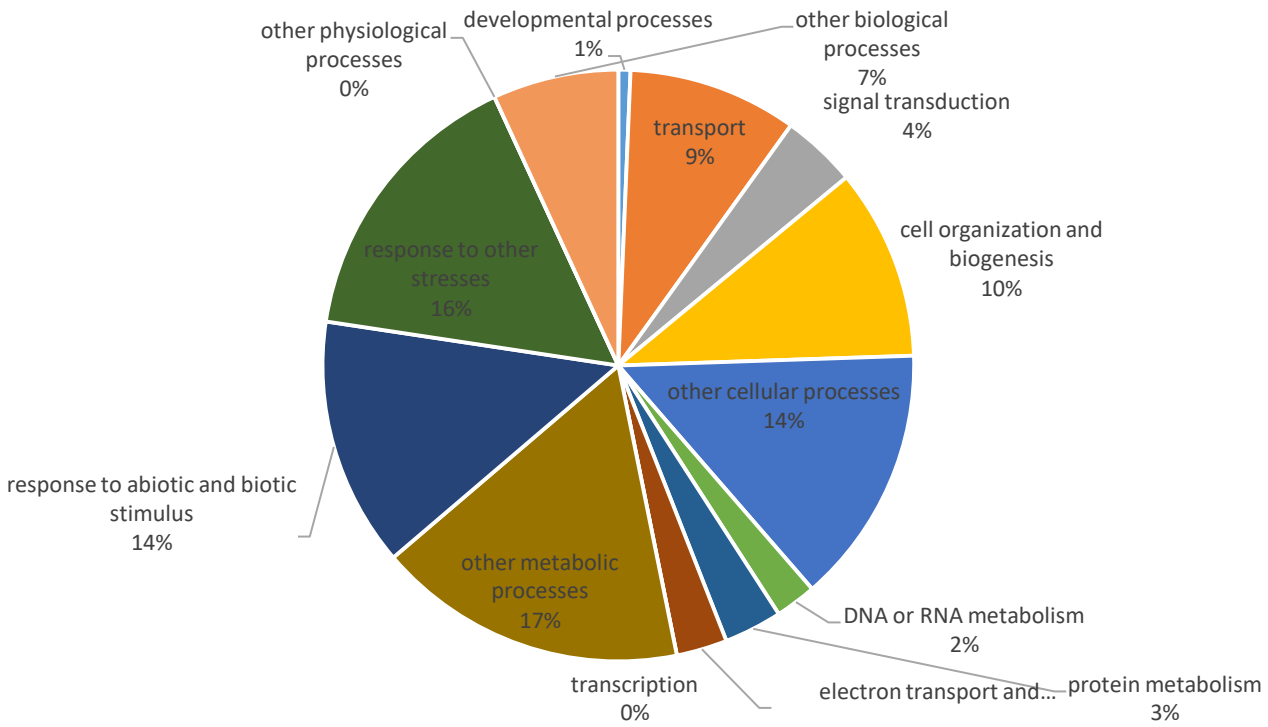


Figure 3.4: Gene ontology enrichment analysis for transcripts that were increased in abundance.

GO enrichment analysis was performed based on biological processes in the embryos of W1-7 barley seeds relative to the wild type.

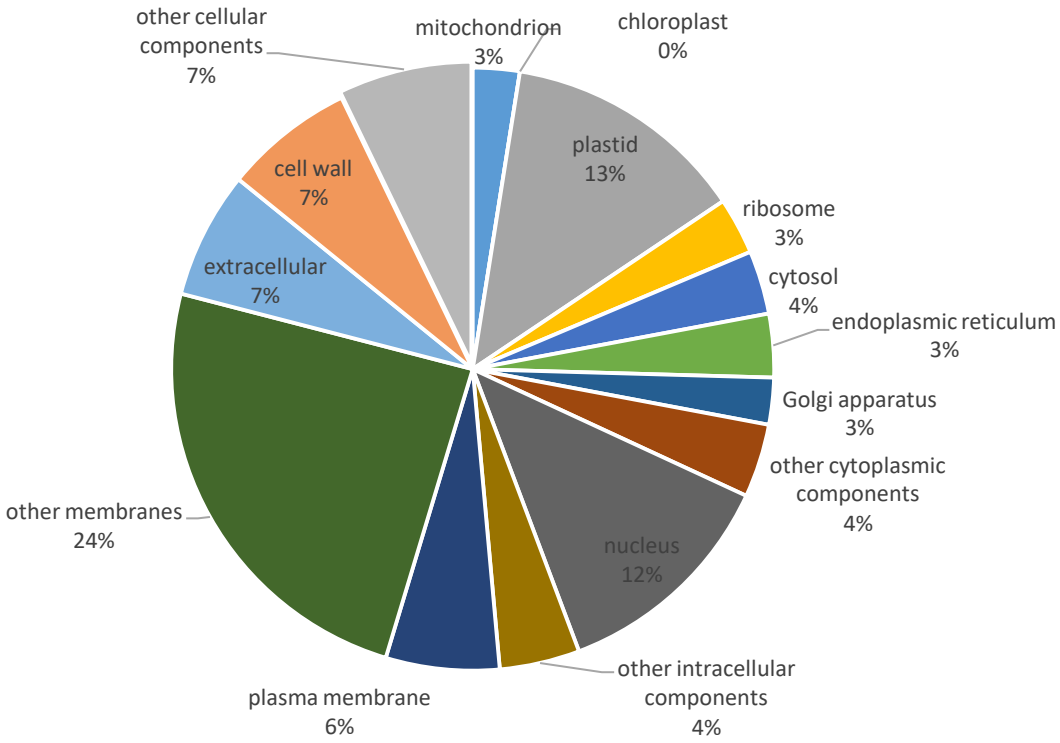


Figure 3.5: Gene ontology enrichment analysis for transcripts that were increased in abundance.

GO enrichment analysis was performed based on cellular localisation in the embryos of W1-7 barley seeds relative to the wild type.

3.2.3 Transcripts that were increased in abundance in the embryos of the W1-7 barley seeds relative to the wild type

The ten most abundant transcripts in embryos of the dry WHY1-deficient seeds are shown in Figure 3.6. Many of these transcripts are associated with plant adaptation to stressful environmental conditions. For example, transcripts encoding a disease resistance protein (AK361511) were abundant in the WHY1-deficient embryos compared to the wild type (Figure 3.6). This finding is important because stress tolerance may be important both in seed storage and in germination (Sreenivasulu et al., 2008).

Moreover, transcripts associated with the plant growth hormone gibberellic acid (MLOC_56462; gibberellin-20) were increased in the dry seeds of the WHY1-deficient embryos compared to the wild type (Figure 3.6). Gibberellins are important phytohormones with key roles in seed maturation, germination and also post-germination growth (Sreenivasulu et al., 2008). This finding is consistent with the hypothesis that gibberellins are increased in the seed embryos (Bewley, 1997).

The levels of MLOC_67727 transcripts (encoding a leucine-rich repeat protein) were significantly increased in the dry seeds of WHY1-deficient embryos compared to the wild type (Figure 3.6). In *Arabidopsis*, this protein is required for growth and increased seed production (Shahollari et al., 2007). Such leucine-rich repeat proteins were shown to be involved in nitrogen reallocation in near-isogenic barley lines under low nitrogen (Jukanti et al., 2008).

Two transcripts that encode cysteine proteases (MLOC_76470 and AK248416) were also more abundant in the WHY1-deficient embryos compared to the wild type (Figure 3.6). Cysteine proteases are responsible for the mobilisation of seed

storage proteins upon germination (Malgorzata Grudkowska, 2004, Sreenivasulu et al., 2008).

The abundance of the transcripts encoding a calcium-binding EF-hand family protein (MLOC_40019) were more abundant in the WHY1-deficient embryos relative to the wild type (Figure 3.6). The calcium-binding EF-hand family are central regulators of cellular calcium signalling pathways. Calcium is an important mediator of hormonal and environmental stress signals that underpin plant responses to biotic and abiotic threats. Changes in cytosolic calcium are also important in the regulation of developmental processes.

The abundance of transcripts encoding a MYB family transcription factor, AK356219 was significantly higher in WHY1-deficient embryos relative to the wild type (Figure 3.6). This transcription factor has been reported to play a role in hormonal responses during seed development and germination. For example, gibberellin-regulated MYB transcription factors in barley are important in the expression of α-amylase in the aleurone in response to gibberellin signals (Gubler et al., 1995).

DNA-related group proteins such as AAA-type ATPase family proteins (AK358288) were found to be highly expressed in WHY1-deficient embryos (Figure 3.6). DNA-related group proteins are involved in the protein degradation process. Nodulin (AK25271) transcripts were highly expressed in WHY1-deficient embryos (Figure 3.6). It has been reported that transgenic rice overexpressing an early nodulin gene had increased nitrogen-use efficiency (Bi et al., 2009).

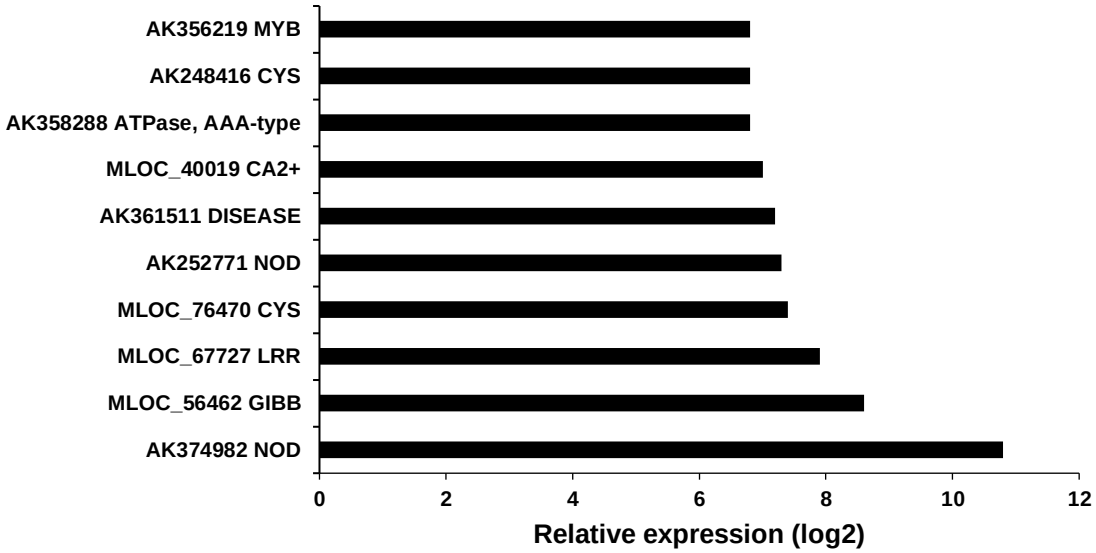


Figure 3.6: Ten most abundant transcripts in the embryos of W1-7 barley seeds relative to the wild type.

The relative expression (log2) fold change-range of genes annotated with the accession numbers and description, for all top 10 differentially regulated up-regulated transcripts. Differentially expressed genes were those showing fold changes of >1.5 and an FDR-corrected *p*-value of 0.05 or less.

3.2.4 Transcripts that were highly decreased in abundance in the embryo of the W1-7 barley seeds relative to the wild type

The ten transcripts showing lowest levels in embryos of the W1-7 seeds relative to the wild type are shown in Figure 3.7. Transcripts encoding expression of several retrotransposons (MLOC_3219, MLOC_41306, MLOC_79536 and MLOC_8485) were significantly decreased in the embryo of the W1-7 barley seeds relative to the wild type (Figure 3.7). Retrotransposons are regulatory components in plant and animal genomes. Retrotransposon function can be silenced by epigenetic processes and amplified by reverse transcription and reintegration into the genome (Cavrak et al., 2014). Retrotransposons are associated with repressive chromatin modifications in plants that are controlled by RNA-directed DNA methylation. However, retrotransposons can use several strategies to avoid this epigenetic silencing (Cavrak et al., 2014).

Transcripts encoding putative histidine kinases that are involved in cell signalling (MLOC_32926 and MLOC_42368) were also significantly decreased in the WHY1-deficient embryos relative to the wild type (Figure 3.7). Plant histidine kinases are involved in signal transduction pathways associated with phytohormones such as ethylene and cytokinin as well as environmental stress responses. Histidine kinases are protein kinases that are responsible for intracellular signal transduction (Urao et al., 2001). In *Arabidopsis*, five two-component histidine kinase-like proteins (ETR1, ETR2, EIN4, ERS1, and ERS2) function in ethylene perception. For example, CRE1 is a hybrid histidine kinase that functions as a cytokinin receptor. Together with CKI1 and CKI2, CRE1 is involved in cytokinin signal transduction (Urao et al., 2001).

Transcripts encoding a FAR-RED IMPAIRED RESPONSE1 (FAR1) -like protein (MLOC_30557) were significantly decreased in the WHY1-deficient embryos relative to the wild type (Figure 3.7). FAR1 is a positive regulator of the phytochrome A pathway of ABA signalling in *Arabidopsis* that is important in

responses to abiotic stresses and is involved in stomatal opening (Tang et al., 2013).

A helicase transcript encoding a protein potentially involved in RNA processing (MLOC_77244) was significantly decreased in the WHY1-deficient embryos relative to the wild type (Figure 3.7). Helicases are important in plant stress responses, acting through effects on RNA metabolism (Kanai et al., 2013).

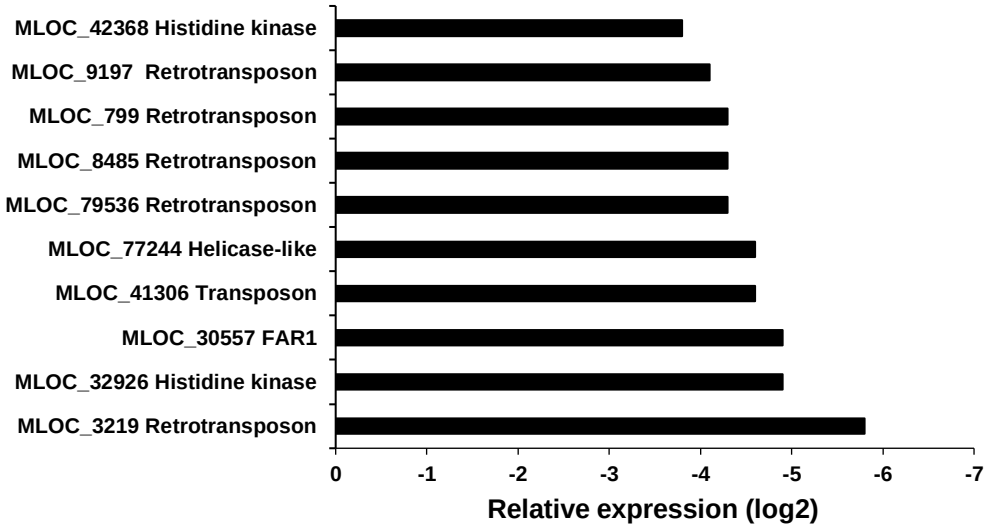


Figure 3.7: Ten most decreased transcripts in the embryo of the W1-7 seeds relative to the wild type.

The relative expression (log2) fold change-range of genes annotated with the accession numbers and description, for all top 10 differentially regulated down-regulated transcripts. Differentially expressed genes were those showing fold changes of >1.5 and an FDR-corrected *p*-value of 0.05 or less.

3.2.5 Transcripts associated with plastid biogenesis that are differentially-regulated in W1-7 seeds relative to the wild type

A number of transcripts encoding nuclear-encoded photosynthetic proteins were significantly increased in the WHY1-deficient embryos relative to the wild type (Table 3.4). Transcripts encoding transcriptionally-active chromosome proteins (pTAC18 and pTAC7) and a sigma factor (SIG6), were more abundant in the WHY1-deficient embryos relative to the wild type. These transcripts are involved in the establishment of plastid transcription and translation systems. Transcripts encoding an RNA polymerase associated factor were significantly increased in the WHY1-deficient embryos relative to the wild type (Table 3.4). Transcripts encoding components of the photosynthetic complexes were significantly higher in the W1-7 embryos relative to the wild type, particularly the photosystem II (PsbR, PsbQ, PsbP, PsbR, Psb28, PsbW, PsbO, psbS, violaxanthin de-epoxidase and LHCB7) and photosystem I (PsaH, PsaG, PsaF, PsaD and PsaG/PsaK) reaction centres (Table 3.4). Transcripts encoding components of the cyclic electron transport pathway (TMP14, TSP9, PIF1 and PGR5) were also higher in the W1-7 barley embryos relative to the wild type (Table 3.4). The levels of transcripts encoding the small subunit (RBCS) of ribulose-1, 5-bisphosphate carboxylase-oxygenase (AK248995.1, AK249082.1, AK369652 and AK249588.1) were also higher in the WHY1-deficient embryos relative to the wild type. Interestingly, only one plastid-encoded photosynthetic protein (MLOC_61558.1) was more abundant in the W1-7 embryos relative to the wild type (Table 3.4).

Table 3.4: Transcripts involved in plastid biogenesis and photosynthesis that were significantly changed in the W1-7 embryos relative to the wild type.

The transcript abundance in the W1-7 embryos relative to the wild type were shown as in (log2FC). The differentially expressed genes were identified as those showing fold changes of >1.5 and FDR-corrected *p*-value of 0.05 or less.

Description	Accession	Relative expression (log2FC)	FDR (*p*-value)
Plastid transcriptionally active 18 (PTAC18)	AK360045	1.9	0.0002442
RNA polymerase sigma-70 (SIG6)	AK251756	3.7	4.03E-09
Plastid transcriptionally active 7 (PTAC7)	AK365977	2.1	1.25E-14
50S ribosomal protein L35	AK358025	2.4	2.46E-10
50S ribosomal protein L28	AK376786	2.2	1.79E-16
50S ribosomal protein L13	MLOC_64398.1	2.4	5.46E-22
50S ribosomal protein L12	AK370491	2.1	3.76E-12
50S ribosomal protein L4	AK354785	2	2.59E-08
50S ribosomal protein L17	MLOC_57719	1.8	1.95E-09
50S ribosomal protein L11	AK375360	1.7	8.63E-09
50S ribosomal protein L1	MLOC_59016.1	1.5	3.59E-24
50S ribosomal protein L40	MLOC_67764.1	2.3	2.34E-10
50S ribosomal protein L34	MLOC_14202.1	2.2	8.5E-09
50S ribosomal protein L9	AK370836	2.2	1.44E-17
30S ribosomal protein S20	MLOC_58312.1	1.6	6.1E-16
30S ribosomal protein S16	AK252167.1	3.1	4.08E-10
30S Ribosomal protein PSRP-3/Ycf65	AK353736	2.7	1.09E-13
Photosystem II (PSII)			
PsbR	AK354522	5.4	0.0000693
PsbQ	MLOC_72691.1	2.5	1.48E-10
PsbQ	AK359673	2.5	1.48E-10
PsbQ	MLOC_17228.1	2.1	2.53E-25
PsbP	MLOC_54528.1	2.1	6.58E-12

PsbP	MLOC_38413.2	1.8	1.79E-07
PsbP	MLOC_75514.2	1.6	0.0000197
PsbR	AK250934.1	5	0.001044319
Psb28	MLOC_15730.2	4.2	6.37E-25
PsbP	AK365640	3.8	0.00009
PsbW	AK369292	2.9	1.6E-08
PsbO	MLOC_78630.1	2.7	1.71E-38
psbS	AK359183	2.8	0.0000132
Violaxanthin de-epoxidase	MLOC_61961.2	3.2	1.81E-09
LHCB7	MLOC_60073.2	2.5	0.000226768
Photosystem I (PSI)			
PsaH	MLOC_53469.2	1.9	0.00103612
PsaG	MLOC_6738.1	1.7	1.64E-21
PsaF	MLOC_66074.2	5.4	0.0000689
PsaD	AK376369	4.8	0.002063455
PsaG/PsaK	AK362139	2.9	4.4E-10
CET (Cyclic electron transport)			
TMP14	MLOC_54334.1	1.7	0.010488295
TSP9	MLOC_67228.1	3.9	0.0000238
PIF1	MLOC_23394.2	3.7	1.6E-09
PGR5	MLOC_7826.2	3	0.00046604
Stroma			
Ribulose bisphosphate carboxylase small	AK248995.1	3.1	0.000000416
Ribulose bisphosphate carboxylase large	MLOC_61558.1	2.4	0.00001
Ribulose bisphosphate carboxylase small	AK249082.1	2.6	0.0000149

Ribulose bisphosphate carboxylase small	AK369652	2.5	0.000000731
Ribulose bisphosphate carboxylase small	AK249588.1	2.3	0.0000016

3.2.6 Transcripts involved in RNA and DNA binding that are differentially changed in the WHY1-deficient embryos relative to the wild type

The abundance of transcripts encoding several transcription factors was significantly changed in the WHY1-deficient embryos relative to the wild type. For example, MYB transcription factors and a basic helix-loop-helix (bHLH) were more abundant in the WHY1-deficient embryos compared to the wild type (Appendix A.1). Similarly, transcripts encoding several WRKY transcription factors were more abundant in the WHY1-deficient embryos relative to the wild type. Transcripts encoding zinc finger proteins that are involved in DNA binding were also more abundant in the absence of WHY1 (Appendix A.1). However, transcripts encoding WRKY40 (MLOC_10687) were significantly lower in the WHY1-deficient embryos relative to the wild type (see Appendix A.1). Only two transcripts encoding for zinc finger proteins were significantly lower in the WHY1-deficient embryos (MLOC_61611 and MLOC_57307). As discussed above, FAR1 (MLOC_30557) transcripts were lower in the WHY1-deficient embryos relative to the wild type. Another transcript encoding FAR1 (MLOC_33258.3) was also lower in the WHY1-deficient embryos relative to the wild type (Appendix A.1).

Transcripts encoding proteins involved in DNA binding such as basic leucine zipper (bZip) proteins were also higher in the WHY1-deficient embryos compared to the wild type. In addition, transcripts involved in RNA splicing and processing such as AK356654 and MLOC_16173 were significantly higher in the WHY1-deficient embryos relative to the wild type (Appendix A.1). Other transcripts encoding proteins involved in telomere binding, DNA repair and chromatin organisation were higher in the WHY1-deficient embryos relative to the wild type. Only one transcript encoding a protein involved in DNA mismatch repair (MLOC_50820) was decreased in the absence of WHY1 (see Appendix A.1).

3.2.7 Changes in redox-related regulated transcripts in the WHY1-deficient embryos

Transcripts encoding enzymes involved in hydrogen peroxide metabolism such as ascorbate peroxidases (APX) and other peroxidases (Caverzan, 2012) were increased in WHY1-deficient embryos (Table 3.5). Similarly, the levels of transcripts encoding thioredoxins (TRX) and glutaredoxins (GRX) were higher in the WHY1 embryos (Table 3.5). A large number of transcripts encoding proteins involved in disease resistance such as nucleotide-binding site leucine-rich repeat (NBS-LRR) proteins were increased in abundance in WHY1-deficient embryos. In general, MYB transcription factor levels were increased in WHY1-deficient embryos. Moreover, transcripts encoding histones were also increased in WHY1-deficient embryos (Table 3.5).

Table 3.5: Differentially expressed transcripts involved in redox processing and defence in the WHY1-deficient embryos relative to the wild type.

The transcript abundance in the W1-7 embryos relative to the wild type are shown as in (log2FC). The differentially expressed genes were identified as those showing fold changes of >1.5 and FDR-corrected *p*-value of 0.05 or less.

Description	Accession	Relative expression (log2FC)	FDR (*p*-value)
Peroxidase 1	AK370705	4.0	0.0000129
Peroxidase 52	MLOC_64966.1	5.1	3.62E-34
Peroxidase 66	AK353768	3.0	1.95E-49
Amine oxidase	AK366005	2.1	0.0000121
Peroxidase	AK375268	4.2	4.93E-17
Peroxidase 12	MLOC_54893.1	2.7	2.64E-09
Ascorbate peroxidase	MLOC_56459.2	2.5	2.92E-46
Thioredoxin	AK359722	3.8	3.66E-13
TPX2	MLOC_55674.1	2.2	2.41E-08
Thioredoxin-like fold	MLOC_47648.1	2.2	5.41E-09
Thioredoxin-like fold	MLOC_15839.1	2.0	2.38E-34
Thioredoxin-like fold	AK376466	1.5	0.008132951
Glutaredoxin	MLOC_21098.1	1.5	0.0000335
Glutaredoxin	AK360350	2.9	0.0000574
Myb transcription factor	AK367954	4.8	0.002063455
Myb domain protein	MLOC_7426.1	4.0	6.73E-61
Myb domain protein	MLOC_52439.6	2.3	0.000374566
MYB-related transcription factor	MLOC_7981.1	2.1	0.0000357
Myb family transcription factor	MLOC_8187.2	1.7	3.05E-08
Myb family transcription factor	AK356219	6.6	6.14E-11
Myb transcription factor	MLOC_9835.2	-1.9	6.6E-56

Single myb histone 6	MLOC_34636.1	-1.9	0.0000181
Myb family transcription factor	MLOC_53628.1	-2.5	0.000991191
Histone H4	AK377086	2	6.66E-13
Histone H2A	AK373742	1.9	0.00000408
Histone H2A	AK252540.1	1.9	0.00000032
Histone H2A	AK373523	1.8	1.37E-09
Histone H2A	MLOC_68568.1	1.7	1.44E-31
Histone H2B	AK251792.1	1.7	3.54E-08
Histone H2A	MLOC_64906.1	1.6	0.000408853
Histone H4	AK252102.1	1.5	0.000014
Histone H2A	AK252601.1	4.4	1.84E-34
Histone H2A	AK250752.1	4.5	1.66E-18
Histone H1	MLOC_63569.1	3.7	2.87E-61
Histone H2A	AK250581.1	3.4	1.19E-47
Histone H2A	AK376310	3	0.0000363
Histone H2A	MLOC_43244.1	2.8	0.02224912
Histone H2A	AK369538	2.7	4.59E-10
Histone H2B	AK250385.1	2.6	1.13E-14
Histone H3	AK358538	2.5	8.68E-31
Histone H2A	AK353773	2.5	5.52E-14
Histone H2A	AK251633.1	2.3	0.0000126
Histone H3	AK375327	2.2	1.54E-38
Histone H2A	AK374191	2.2	0.000000224
Histone H3	AK353900	2.2	9.69E-14
Histone H1	MLOC_74808.3	2.2	3.71E-110
Single myb histone 6	MLOC_34636.1	-2	0.0000181

3.2.8 Changes in transcripts encoding pathogen-related proteins and cold and wound responses in the WHY1-deficient embryos relative to the wild type

A large number of transcripts related to disease and stress responses were increased in abundance in the WHY1-deficient embryos relative to the wild type (Table 3.6). Of these, transcripts encoding proteins involved in pathogen, nematode and disease resistance, such as NBS-LRR proteins were significantly higher in the WHY1-deficient embryos (Table 3.6). NBS-LRR proteins have key functions in plant defence responses to pathogen attack (McHale et al., 2006). Some transcripts encoding wound-induced proteins were significantly higher in WHY1-deficient embryos relative to the wild type (Table 3.6). Additionally, two transcripts encoding cold response proteins were significantly higher in the WHY1-deficient embryos (Table 3.6).

Table 3.6: Change in abundance of transcripts involved in pathogen response and cold and wound response in WHY1-deficient embryos relative to the wild type.

The transcript abundance in the W1-7 embryos relative to the wild type are shown as in (log2FC). The differentially expressed genes were identified as those showing fold changes of >1.5 and FDR-corrected *p*-value of 0.05 or less.

Description	Accession	Relative expression (log2FC)	FDR (p-value)
Pathogen-related protein	MLOC_79446.2	2.9	1.03E-11
Pathogenesis-related	MLOC_23000.1	2.7	3.57E-07
NPR1 interactor	AK362984	3	7.79E-07
Cold shock protein	MLOC_73670.1	3	0.00078022
Wound induced protein	MLOC_65985.1	3	1.69E-61
Wound induced protein	MLOC_4511.1	2	1.2E-10
Universal stress protein	MLOC_77034.3	1.5	1.56E-52
Heavy metal transport	MLOC_60983.1	2.8	0.000000125
Heat stress transcription factor	MLOC_11286.1	1.7	1.44E-23
Heavy metal transport	MLOC_22984.1	2.6	2.18E-14
Heavy metal transport	AK357922	2.2	1.43E-53
Wound induced protein	MLOC_57090.1	5.3	2.14E-74
Wound induced protein	MLOC_43430.2	4.7	5.41E-21
Wound induced protein	AK376360	2.8	6.73E-69
Wound induced protein	MLOC_64888.1	2.8	7.64E-60
Wound induced protein	MLOC_76721.1	2.6	6.42E-50
Wound induced protein	AK355456	2.2	2.57E-68
Cold shock protein	MLOC_17065.1	2.5	8.65E-09
Cold shock protein	MLOC_75604.1	2.5	0.000220579
Pathogenesis-related	AK369614	2	0.00000292
Pathogenesis-related	AK356356	1.7	1.04E-11
Pathogenesis-related	MLOC_72965.1	1.7	1.05E-13
Pathogenesis-related	AK371567	2.6	2.03E-18
Disease resistance	AK361511	7.2	9.84E-87
Disease resistance	MLOC_77921.1	1.5	0.0000624

Disease resistance	AK370420	-1.6	0.000341971
Disease resistance	AK249642.1	-1.9	0.000000303
Disease resistance	MLOC_5504.3	-2	2.88E-11
Disease resistance	MLOC_78491.4	-2	9.8E-15
Defend	MLOC_19160.1	5.1	1.58E-30
Nematode-resistance protein	AK365897	2	6.07E-115
Disease resistance	MLOC_34953.1	-2.6	9.08E-08
Disease resistance	MLOC_75347.1	-3.3	0.000000298
Disease resistance	MLOC_79257.1	4.7	1.12E-09
Disease resistance	MLOC_18373.1	4	0.0000138
Disease resistance	MLOC_66163.1	2.7	9.08E-18
NBS-LRR	MLOC_54830.2	2.4	0.000149577
Disease resistance	MLOC_4541.1	2.2	1.52E-08
Disease resistance	MLOC_70111.1	1.9	1.4E-16
Disease resistance	MLOC_60393.3	1.9	2.45E-12
Disease resistance	MLOC_43035.1	1.6	0.00000634
Disease resistance	MLOC_44924.2	1.6	0.004171635
Disease resistance	MLOC_52055.2	1.5	0.000614468
Disease resistance	MLOC_34954.1	-1.8	5.5E-20
Disease resistance	MLOC_24654.2	-3.4	2.17E-10

3.2.9 Transcripts encoding phytohormones -related pathways in the WHY1-deficient embryos relative to the wild type

A number of transcripts encoding proteins involved in pathways associated with hormones such as gibberellins (GA), abscisic acid (ABA), ethylene and auxin were significantly higher in the WHY1-deficient embryos relative to the wild type (Table 3.7). For example, transcripts encoding 1-aminocyclopropane-1-carboxylate oxidase (ACC oxidase) were significantly higher in the WHY1-deficient embryos. ACC oxidase is a key precursor in the synthesis of ethylene (Van de Poel and Van Der Straeten, 2014). Transcripts encoding 9-cis-epoxycarotenoid dioxygenase 1 (NCED) which cleaves 9-cis xanthophylls to xanthoxin in the ABA synthesis pathway (Tan et al., 2003) were significantly higher in the WHY1-deficient embryos relative to the wild type.

Table 3.7: Differentially expressed transcripts involved in phytohormones pathways in the WHY1-deficient embryos relative to the wild type.

The transcript abundance in the W1-7 embryos relative to the wild type are shown as in (log2FC). The differentially expressed genes were identified as those showing fold changes of >1.5 and FDR-corrected *p*-value of 0.05 or less.

Description	Accession	Relative expression (log2FC)	FDR (p-value)
Gibberellin-20 oxidase-2	MLOC_56462.1	8.7	1.22E-44
Gibberellin-20	AK373555	5	3.92E-56
Gibberellin-regulated protein 2	MLOC_68372.1	4.5	0.008084918
Gibberellin receptor GID1L2	AK368846	4.4	1.7E-16
Gibberellin-regulated protein 3	MLOC_62101.1	3.4	0.0000139
Gibberellin receptor GID1L2	AK357963	3	8.95E-10
Gibberellin 20 oxidase	MLOC_16059.1	2.9	4.48E-16
Gibberellin receptor GID1	AK356233	2.7	0.0000231
Gibberellin receptor GID1L2	MLOC_72547.2	2.4	0.000643323
Gibberellin 2-oxidase	AK373885	1.8	0.011717323
Gibberellin-regulated protein	AK358265	1.7	5.89E-10
Gibberellin receptor GID1L2	AK363140	1.7	4.02E-17
Gibberellin receptor GID1L2	MLOC_53106.1	1.7	9.89E-12
Gibberellin receptor GID1L2	MLOC_59654.1	1.6	0.002474983
Gibberellin receptor GID1	MLOC_70507.1	-1.6	0.00000497
Gibberellin 2-oxidase	MLOC_72016.2	6.3	2.16E-27
Gibberellin-regulated protein 2	AK373320	5	0.001044319
ABA-responsive protein-related	MLOC_11331.2	2.3	2.77E-25
Abscisic acid receptor PYR1	MLOC_80832.1	1.5	0.0000425
Abscisic acid receptor PYR1	MLOC_72289.1	2.4	1.01E-09

9-cis-epoxycarotenoid dioxygenase 1 (NCED)	AK365103	3.3	6.4E-39
1-aminocyclopropane-1-carboxylate oxidase (acc)	MLOC_61547.1	3.1	1.04E-13
Ethylene-responsive transcription factor	MLOC_59305.1	3.1	9.23E-19
Ethylene responsive transcription factor 2b	MLOC_51143.1	3	3.4E-61
Ethylene-responsive transcription factor	MLOC_6403.2	2.8	2.12E-23
Ethylene responsive transcription factor 2a	AK367525	2.7	0.00000492
Ethylene-responsive transcription factor	MLOC_7255.1	2.4	0.0000033
Ethylene responsive transcription factor 2b	MLOC_64636.1	2.1	1.21E-29
Ethylene responsive transcription factor 2a	AK374826	2	8.84E-25
Ethylene-responsive transcription factor 1	MLOC_3095.1	2	0.0000159
Ethylene-responsive transcription factor 4	AK354662	1.7	2.92E-52
Ethylene responsive transcription factor 2b	AK367417	4.7	2.37E-60
Ethylene responsive transcription factor 2a	AK356313	4.3	0.000000319
Ethylene responsive transcription factor 2a	AK248243.1	3.8	9.15E-08
Auxin response factor	MLOC_63261.1	1.5	0.000000605
Auxin responsive	MLOC_58508.9	-1.5	0.008628122
Auxin induced-like protein	MLOC_18171.1	5.6	0.000018
Auxin response factor	AK374546	5.3	0.000132747
Auxin responsive protein	MLOC_75265.1	3.6	4.03E-09

Auxin responsive SAUR protein	AK363815	3.1	0.006645499
GH3 auxin-responsive	MLOC_64672.1	2.6	2.68E-14
Auxin response factor	AK366601	2.2	0.008185636
Auxin response factor	MLOC_11014.7	2.2	0.008185636
Auxin-responsive protein	MLOC_56819.1	2.2	1.79E-17
Auxin responsive SAUR protein	MLOC_33053.1	2.1	0.000000459
GH3 family protein	MLOC_63528.1	1.9	1.56E-17
Auxin responsive protein	MLOC_36987.1	1.8	0.00000589

3.3 Discussion

Germination is an important step in plant development. The embryos housed in the mature seeds require the right environment and conditions to grow. Nutrients stored in the seeds support seedling growth after germination. Germination was lower in both the W1-1 and W1-7 seeds that the wild type (Figure 3.2). However, after germination, the seedlings of all lines show a similar growth rate despite the lower seed germination properties (Figure 3.1). The WHY1 protein is required for seed germination but not the subsequent growth of the barley seedlings.

Yield parameters were measured in the T4 generation of the W1-7 line and in the wild type barley plants, as described previously (Comadira, 2015). The W1-7 plants had significantly fewer tillers than the wild type (see Table 3.3). However, the W1-7 plants produced more seeds and had significantly higher yields than the wild type (see Table 3.3). A previous report (Comadira et al., 2015) showed that seed yield per tiller was similar in both lines but data shown in Table 3.3 reveals a higher seed yield per tiller in the W1-7 plants. The yield differences and changes in the shoot phenotypes of the T4 generation WHY1-deficient plants (Table 3.3) compared to the T3 generation plants (Comadira et al., 2015) is interesting. Seed yield in the 4^{th} generation plants was higher than the wild type. The results obtained here may be explained by generation to generation variations. Thus, the WHY1-deficient plants are likely to produce greater seed yields than the wild type. However, this analysis must be confirmed in future generations to confirm this trend.

The data reported in this chapter show significant differences in the embryo transcriptomes of the WHY1-deficient seeds relative to the wild type. In total, 1292 transcripts were increased in abundance in the WHY1-deficient embryos relative to the wild type. In contrast, only 180 transcripts were significantly less abundant in the WHY1-deficient embryos relative to the wild type. These results suggest that WHY1 has important functions in barley that alter the programming

of the embryo and its preparation for germination. However, further experiments will be required to validate the expression levels of some of the transcripts that were changed in the WHY1-deficient embryos relative to the wild type using qRT-PCR.

Many nuclear-encoded transcripts involved in photosynthesis were increased in abundance in the WHY1-deficient seeds relative to the wild type. These transcripts encode key components of the photosynthetic complexes and the establishment of plastid transcription and translation systems (see section 3.2.5). Other transcripts that were increased in the WHY1-deficient embryos encode proteins associated with photosynthetic electron transport complexes, particularly photosystem (PS) II and components involved in cyclic electron transport (Table 3.4). The level of transcripts encoding the small subunit (RBCS) of ribulose-1, 5-bisphosphate carboxylase-oxygenase (Table 3.4) was increased in the WHY1-deficient embryos relative to the wild type. The photosynthetic associated transcripts are present in the embryos of the dry seeds were laid down by the mother plant during seed production and maturation. Only one plastid-encoded transcript encoding a protein involved in photosynthesis (RBCL) was increased in abundance in the WHY1-deficient embryos relative to the wild type (Table 3.4). These findings suggest that the WHY1-deficient embryos are primed to establish photosynthesis more rapidly than the wild type. In addition, transcripts involved in photo-protective and antioxidant metabolism were more abundant in the WHY1-deficient embryos, suggesting that defence processes associated with photosynthesis are also primed by loss of WHY1 functions (see section 3.2.7). In addition, transcripts encoding proteins associated with redox metabolism were more abundant in WHY1-deficient embryos than the wild type, including ascorbate peroxidase (APX) and other peroxidases, thioredoxins (TRX) and glutaredoxins (GRX) (Table 3.5). These changes are interesting because they suggest that ROS levels are higher in the embyos deficient in the WHY1 protein. The reasons why ROS would be more abundant in WHY1-deficient embryos are unknown. However, it is also possible that the flowers and reproductive organs of the mother plants experienced a higher level of oxidation than the wild type during seed production resulting in a greater levels of antioxidant protection in the

WHY1-deficient embryos that would might provide a useful advantage during seed germination.

The levels of transcription factor mRNAs, including MYB and WRKY were increased in the WHY1-deficient embryos relative to the wild type. A large number of zinc finger proteins that are involved in DNA binding were increased in the WHY-deficient embryos relative to the wild type (Appendix A.1). Transcripts encoding proteins involved in telomere binding, DNA repair, chromatin organisation and histones were also higher in the WHY1-deficient embryos. The WHY1 protein therefore has roles that affect the expression of these key regulators of plant growth, development and adaptation to environmental triggers. Moreover, transcripts encoding proteins involved in GA, ABA, ethylene and auxin metabolism were more abundant in the WHY1-deficient embryos relative to the wild type. For example, several transcripts associated with ABA pathways were more abundant in the WHY1-deficient embryos. It has previously been shown that ABA-mediated regulation of seed germination is altered in *Arabidopsis why1* mutants (Isemer et al., 2012a). Localisation of WHY1 in plastids was shown to increase sensitivity to ABA. These findings suggest that hormone signalling pathways associated with seed germination and seedling growth are also primed in response to the loss of WHY1 function. A number of transcripts containing *cis* elements that are recognised by GA and ABA and that encode hydrolases that act on proteins, lipids and sugars were more abundant in the WHY1 mutants (Table 3.7). Both GA and ABA play important roles in during seed maturation and germination. ABA is synthesised during seed maturation and dormancy and stored in the maternal tissue and embryo, and its levels decrease during imbibition (Jacobsen et al., 2002, Millar et al., 2006, Sreenivasulu et al., 2008). GA can be synthesised and stored in the embryo. The levels of GA are higher during imbibition and remains high during and after germination (Bewley, 1997). Hydrolases are required for the mobilisation of seed storage reserves during early seed germination. The level of transcripts encoding a cytochrome P450 (MLOC_69871 was more abundant in the WHY1-deficient embryos relative to the wild type. This cytochrome P450 is involved in the GA synthesis pathway (Davidson et al., 2006). In addition, CYP707A encodes an ABA 8'-hydroxylase

that modulates tissue ABA content to control seed dormancy (Millar et al., 2006). Taken together, these results suggest that seed dormancy and germination are influenced by the loss of WHY1. Moreover, the observed changes in transcripts associated with redox processes and stress responses suggest that redox signalling and associated defences are primed by the loss of WHY1, increasing the fitness of WHY1-deficient seedlings at the earliest stages of development compared to the wild type. This analysis suggests that WHY1 regulates a range of fundamental processes in seeds such as embryogenesis, dormancy and germination.

Chapter 4 . The role of WHIRLY in the establishment of photosynthesis in barley leaves

4.1 Introduction

The WHY proteins are localised in the chloroplasts, mitochondria and nuclei (Krause et al., 2005, Grabowski et al., 2008). To date, in most plant species there are at least two WHY proteins, WHY1 and WHY2 that are known to be targeted to the mitochondria or plastids, respectively (Isemer et al., 2012b, Krause and Krupinska, 2009). In *Arabidopsis*, AtWHY1 and AtWHY3 are targeted to plastids and AtWHY2 is targeted to mitochondria (Krause et al., 2005, Krause and Krupinska, 2009). In contrast to *Arabidopsis*, there are two WHY proteins in barley; WHY1 is targeted in the nucleus and plastids (Melonek et al., 2010) and WHY2 is targeted to the mitochondria (Krause et al., 2005).

WHY1 has been suggested to be important for anterograde and retrograde signalling during plant development and in environmental stress responses (Foyer et al., 2014). WHY1 binds to the chloroplast nucleoids and is essential for DNA maintenance and compactness (Krupinska et al., 2014). WHY1 can control the levels of transcripts of chloroplast-encoded genes (Comadira et al., 2015). However, there are inconsistencies in the literature concerning the phenotypes of WHY1 (*why1*) mutants in different plant species. *Arabidopsis why1* mutants have similar phenotype to the wild type (Yoo et al., 2007a). However, a small percentage of the double-knockout *why1why3* mutants have a variegated green/white/yellow leaf phenotype (Maréchal et al., 2009, Cappadocia et al., 2010). In addition, crossing the *why1why3* mutants with a mutant impaired in organelle DNA polymerase IB (pollB) (i.e. *atwhy1 atwhy3 pollb-1* mutants), one of two type I chloroplast DNA polymerases, resulted in an extreme yellow-variegated phenotype (Lepage et al., 2013). This triple mutant had a higher level

of illegitimate recombination between repeated sequences and greater plastid genome instability than the wild type (Lepage et al., 2013). The *atwhy1 atwhy3 pollb-1* mutants also showed down-regulated photosynthetic electron transport efficiencies with a higher level of ROS accumulations, suggesting an overlap of WHY1 protein and PolIB functions (Lepage, 2013). Moreover, maize WHY1 knockout-mutants (*zmwhy1-1*) produced by transposon insertion tagging had albino phenotype due to defects in chloroplasts ribosome synthesis, which led to incorrect plastid biogenesis (Prikryl et al., 2008). This study reported that WHY1 binds to both DNA and RNA in chloroplasts and it is required for the correct intron splicing of chloroplast ribosomal proteins (Prikryl et al., 2008). However, the knock-out and knock-down maize lines had equivalent amounts of chloroplast DNA and RNAs, suggesting that WHY1 is not required for chloroplast DNA replication or plastid transcription (Prikryl et al., 2008).

The functions of WHY1 have been explored using transgenic barley lines with RNAi-mediated knock-down of the *WHY1* gene. These lines have about 5% of the wild type WHY1 protein (Melonek et al., 2010, Krupinska et al., 2014). The RNA-interference knockdown barley lines (W1-1, W1-7 and W1-9) had reduced levels of WHY1 protein but no marked shoot phenotypes when grown under optimal conditions (Melonek et al., 2010, Krupinska et al., 2014). However, the mature leaves of the WHY1-deficient plants accumulated more chlorophyll but had less sucrose than the wild type (Comadira et al., 2015). Loss of WHY1 protein function influenced the expression of transcripts encoded by chloroplast genes such as ribosomal proteins, subunits of the RNA polymerase and the thylakoid NADH and cytochrome b6/f complexes (Comadira et al., 2015). However, these lines have not yet been fully characterised, particularly with regards to leaf development. The following studies were therefore undertaken to characterise the development of barley leaves that are deficient in WHY1 compared to the wild type.

4.2 Results

4.2.1 Shoot phenotypes of WHY1-deficient barley

The leaves of 7 day old W1-7 seedlings showed a delayed in greening in a strictly developmental manner from base, middle and leaf tip compared to the wild type (Figure 4.1). The delayed greening phenotype was observed in each new emerging leaf (Figure 4.1). However, the developed first leaves of W1-1 and W1-7 appeared visually similar in colour as the wild type (Figure 4.2).

A)

B)

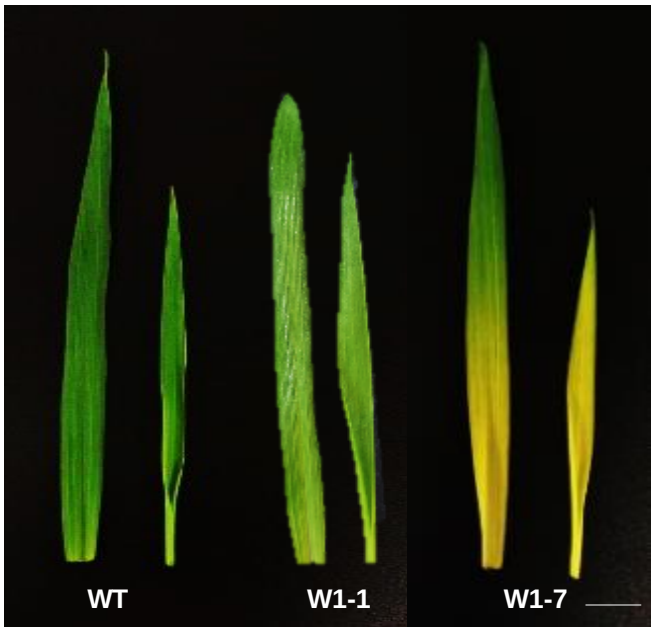


Figure 4.1: A comparison of shoot phenotypes of (A & B) 7-day old seedlings of transgenic W1-1 and W1-7 lines to the wild type (WT).

Seeds were kept at 4°C for 3 days before the seedlings were sown in pots in soil in controlled environment chambers with a 16h light/ 8h dark photoperiod, irradiance of 200 µmol $m^{-2}s^{-1}$, 20°C/16°C day/night temperature regime and 60% relative humidity. (Scale bar =1 cm).

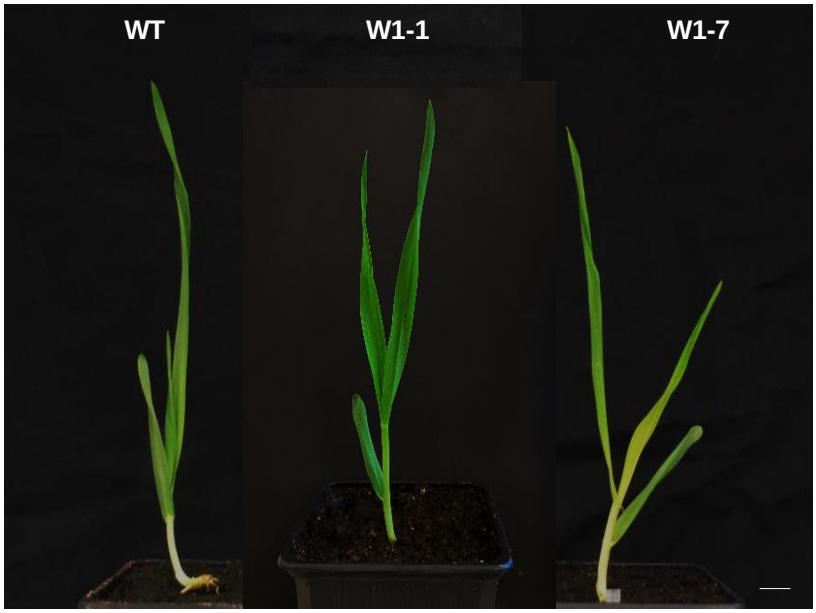


Figure 4.2: A comparison of shoot phenotypes of 14 day-old seedlings of transgenic W1-1 and W1-7 lines to the wild type (WT).

Seeds were kept at 4°C for 3 days before the seedlings were sown in pots in soil in controlled environment chambers with a 16h light/8h dark photoperiod, irradiance of 200 µmol $m^{-2}s^{-1}$, 20°C/16°C day/night temperature regime and 60% relative humidity. (Scale bar =1 cm).

4.2.1.1 Transcript level of WHY1 in RNAi barley lines

Quantitative real-time PCR (qRT-PCR) was performed to measure the levels of *WHY1* transcripts in the base, middle and tip regions of 7 day old barley seedlings. *WHY1* transcripts were very low in all the leaf sections of the W1-1 and W1-7 leaves at 7 days compared to the wild type (Figure 4.3 A). Similarly, *WHY1* transcripts were much lower in all the leaf sections of the W1-1 and W1-7 plants at 14 days compared to the wild type (Figure 4.3 B). While there were small differences in the levels of *WHY1* transcripts in the baser and middle sections of the W1-1 and the W1-7 leaves, the abundance of these transcripts was very low in both lines compared to the wild type. It is therefore difficult to attribute phenotypic differences to the small differences in *WHY1* transcripts between the W1-1 and W1-7 lines (Figure 4.3). Since transcript levels do not always have a direct relationship to protein levels, it is more realistic to relate changes in the phenotype to the abundance of the WHY1 protein in the different lines. The levels of WHY1 protein were lower in all the regions of the W1-7 leaves than the W1-1 and wild type leaves, as is shown later in Figure 4.14.

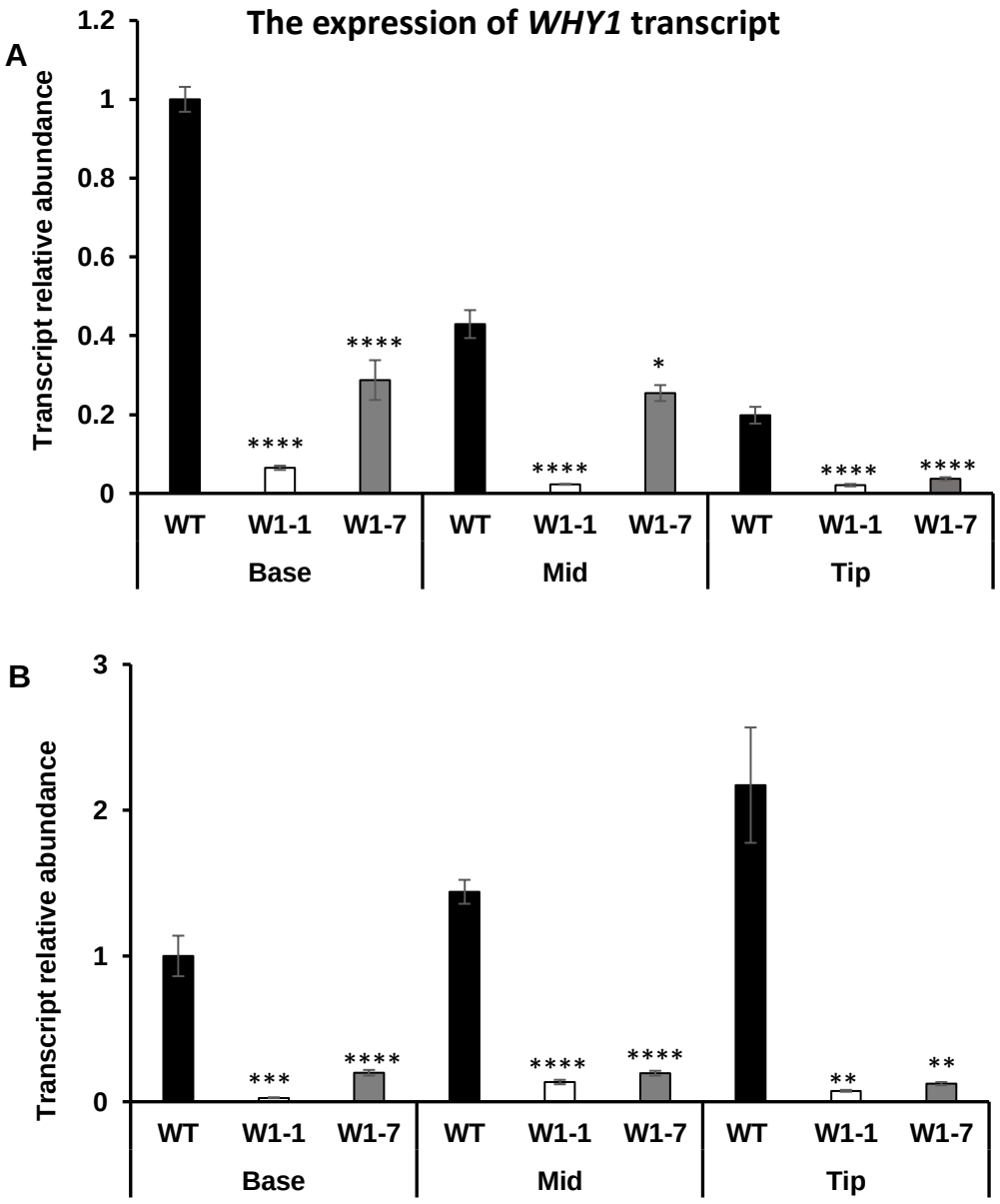


Figure 4.3 : Relative abundance of transcripts encoding WHY1 in the base, middle (Mid) and tip sections of the first leaves of wild type (WT) , W1-1 and W1-7 seedlings of (A) 7 and 14 (B) days after germination.

Data was normalised to ACTIN11. Data for the WT base was set to 1. The data for middle and tip are shown relative to the WT base. Values are represented as mean ± SE (n=6). Asterisks indicate significant differences between WHY1-deficient and wild type plants estimated by the Student's *t*-test (*****p<0.05; **p<0.01; ***p<0.001 and ********p<0.0001).

4.2.2 Leaf pigment content

The base sections of the wild type barley leaves had significantly less chlorophyll than the middle and tip sections at 7 days old, consistent with the presence of greening plastids in the developing monocot leaves (Figure 4.4). The chlorophyll content of the base, middle and tip sections of the first leaves of the 7 day old was lower in the W1-1 and W1-7 seedlings than the wild type (Figure 4.4 A). The W1-7 leaves had the lowest levels of chlorophyll relative to the wild type seedlings. The W1-1 leaves had higher chlorophyll levels than the first leaves of the W1-7 seedlings but values were still significantly lower than the wild type seedlings (Figure 4.4 A). In general, the first leaves of the WHY1-deficient seedlings exhibited a gradual increase in the total chlorophyll content from the base, middle to the tip (Figure 4.4 A).

The base sections of the wild type leaves at 14 days old had low levels of chlorophyll than the middle and tip sections of the leaves (Figure 4.4 B). The chlorophyll content of the base, middle and tip sections of the of the first leaves of the W1-1 and W1-7 seedlings were increased in the 14 day old seedlings compared to the values obtained with the 7 day old seedlings (Figure 4.4). However, the leaves of the W1-1 and W1-7 seedlings still had significantly lower chlorophyll levels than the wild type at 14 days old (Figure 4.4 B). The W1-7 leaves had lower levels of chlorophyll than the W1-1 and to the wild type at 14 days after germination (Figure 4.4 B). The total chlorophyll content increased from the base to the tip in the first leaves of WHY1-deficient barley seedlings (Figure 4.4 B). The chlorophyll contents tended to be lower in the 14 day old W1-1 and W1-7 leaves than the wild type but these differences were much less marked than at 7 day old (Figure 4.4).

The base sections of 7 day old wild type barley leaves had significantly less carotenoid pigments than the middle and tip sections (Figure 4.5 A). The first leaves of the 7 day old W1-1 and W1-7 seedlings had lower carotenoid levels than the wild type leaves at the same age (Figure 4.5 A).

The basal sections of the W1-1 and W1-7 leaves had lower carotenoid pigments than the comparable sections of the wild type leaves at 14 day old (Figure 4.5 B). Moreover, the middle and tip sections of the first leaves of both the W1-1 and W1-7 leaves had significantly lower carotenoid contents than the comparable sections of the wild type leaves (Figure 4.5 B).

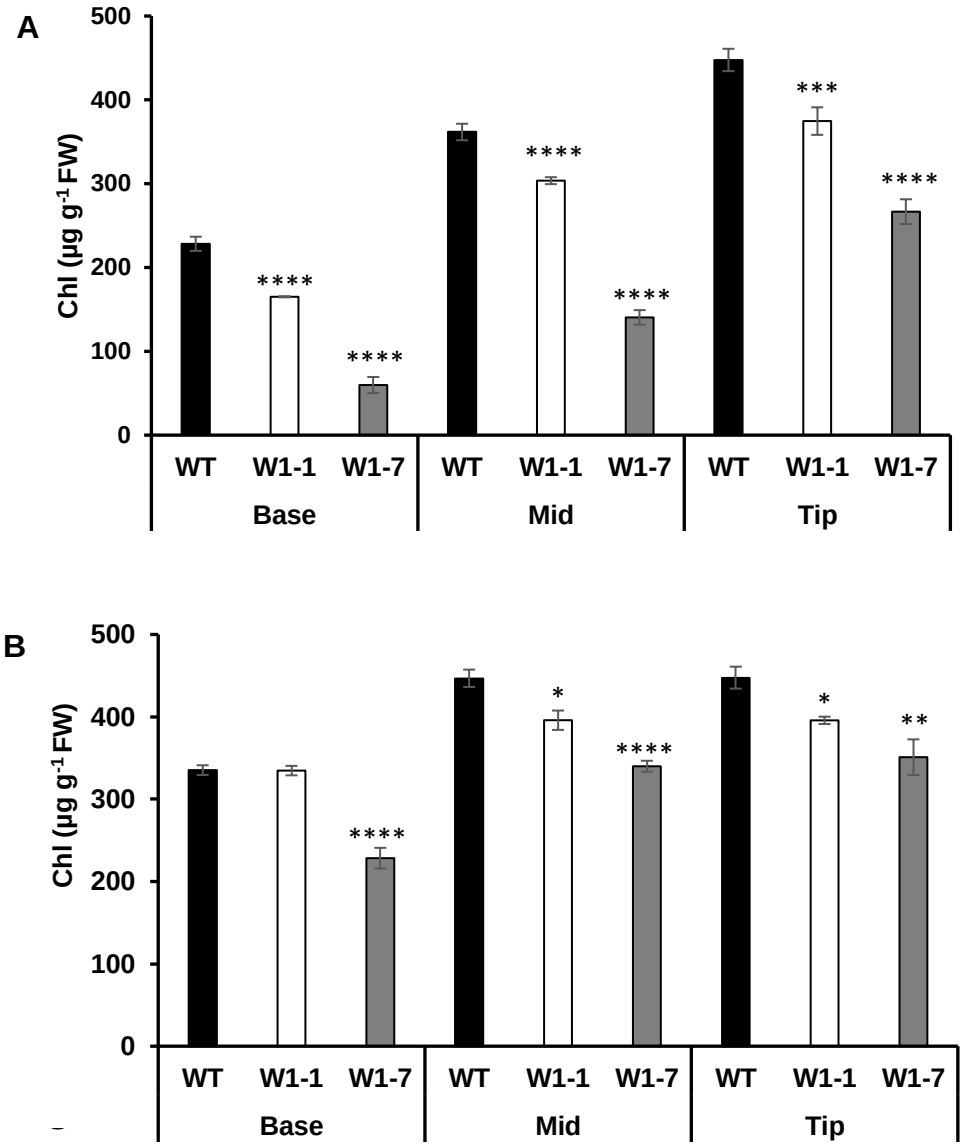


Figure 4.4 : A comparison of the chlorophyll content of first leaves of (A) 7- and (B) 14- day old of the wild type (WT), W1-1 and W1-7 barley seedlings.

Data was expressed as mean values ± SE (n=9). Chl, chlorophyll; FW, fresh weight. Significant differences between the wild type, W1-1 and W1-7 were determined by the Student's *t*-test (*****p<0.05; **p<0.01; ***p<0.001 and ********p<0.0001).

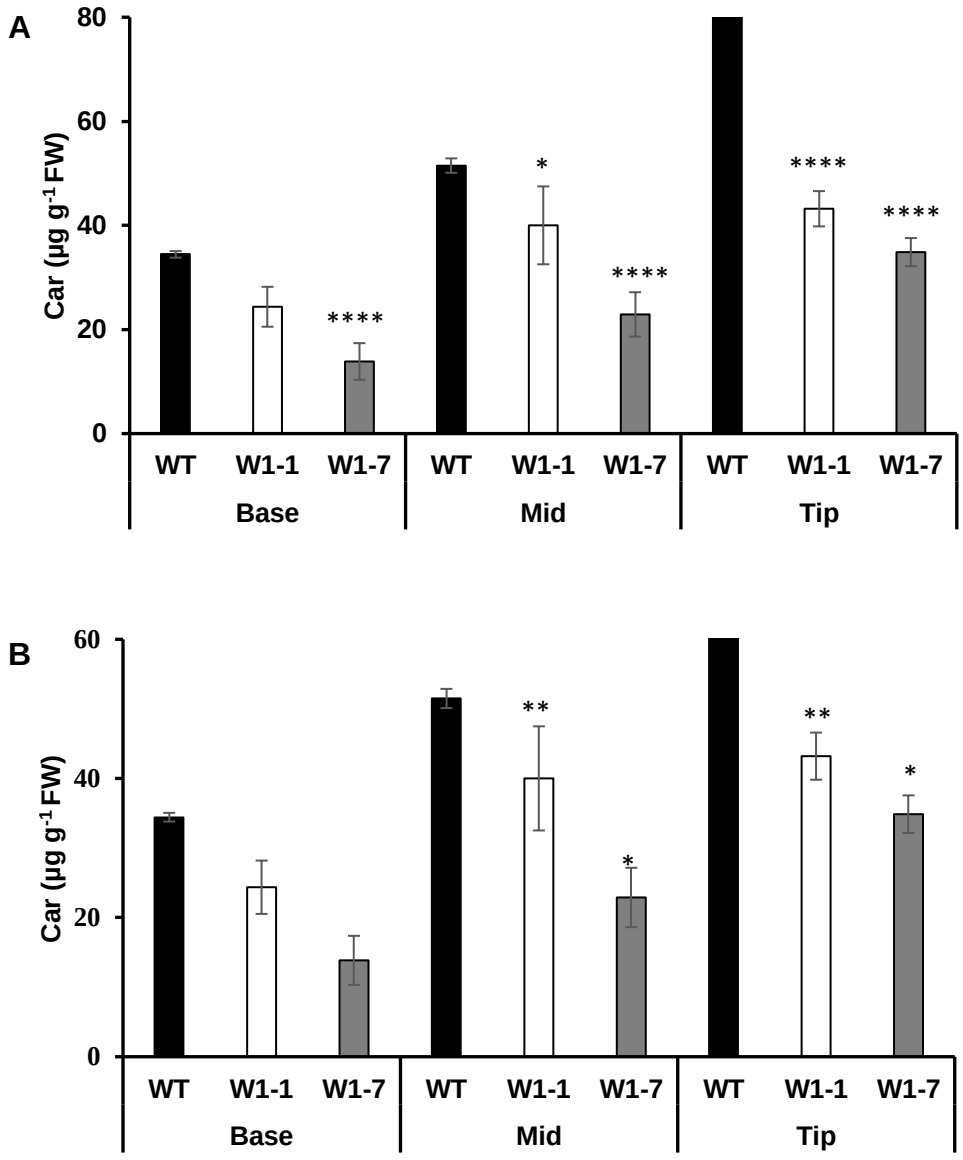


Figure 4.5 : A comparison of carotenoid content in the first leaves of (A) 7- and (B) 14- day old of the wild type (WT), W1-1 and W1-7 barley seedlings.

Data was expressed as mean values ± SE (n=9). Chl, chlorophyll; FW, fresh weight. Significant differences between the wild type, W1-1 and W1-7 was determined by the Student's *t*-test (*****p<0.05; **p<0.01; ***p<0.001 and ********p<0.0001).

4.2.2.1 Chlorophyll *a* fluorescence imaging

The ratio of dark-adapted variable chlorophyll *a* fluorescence (Fv) to the maximal value of chlorophyll *a* fluorescence (Fm) was lower in the W1-7 leaves compared to the wild type (Figure 4.6). Thus, the onset of efficient photosynthesis was delayed in the W1-7 plants compared to the wild type. This data is consistent with the delayed greening observed in the W1-7 leaves (Figure 4.1). Interestingly, the Fv/Fm ratios were similar in the first leaves of the W1-7 and wild type leaves, 9 days after germination, showing that the quantum yield of photosynthesis was similar in both lines at this stage (Figure 4.6). However, the developing second and third leaves had lower Fv/Fm ratios than comparable leaves of the wild type plants (Figure 4.6). The efficiency of photosynthesis was only about 60% of maximum in the base sections and about 80% in the tip sections of the developing W1-7 leaves compared to the wild type (Figure 4.6). This indicates that the efficiency of photosynthesis was slower to reach maximum values in the W1-7-deficient leaves.

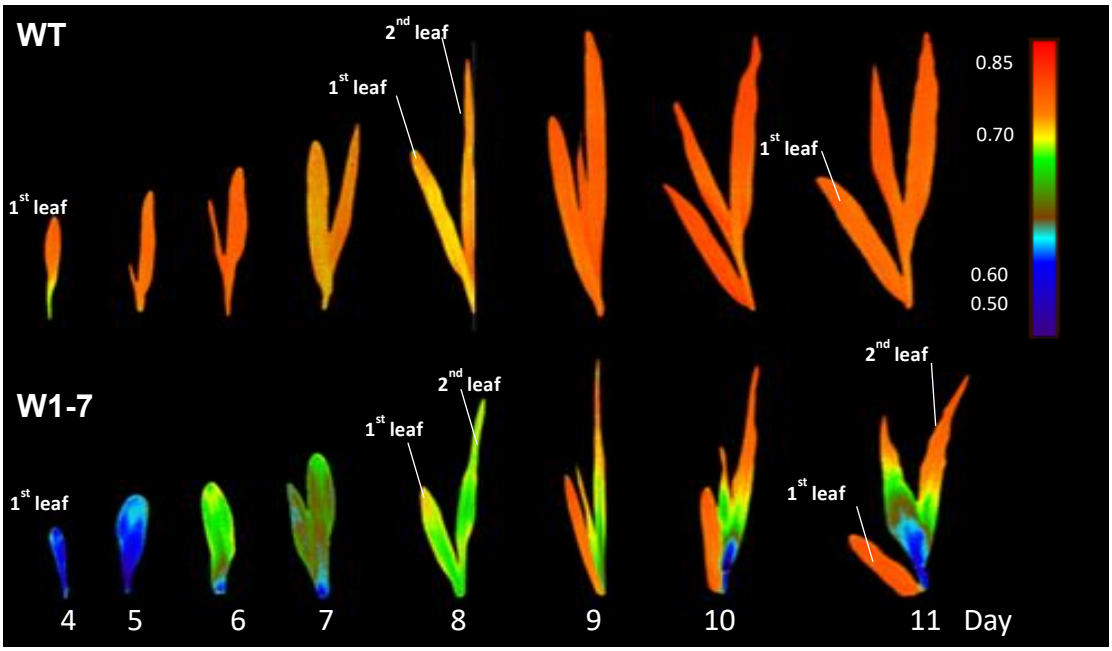


Figure 4.6 : In vivo imaging of the Fv/Fm ratios of wild type and W1-7 seedlings.

Chlorophyll *a* fluorescence quenching measurements were made from the first day of leaf emergence (at 4 days old) up to 10 days after germination. Colour scale represents the values of Fv/Fm ratios.

4.2.2.2 Single time point measurements

Fv/Fm ratios were measured in the base, middle and tip regions of wild type and W1-7 leaves of seedlings from day 5 up to day 14 after germination using a Fluoropen. The Fv/Fm ratios were significantly lower in the base, middle and tip regions of the W1-7 leaves than the equivalent leaf sections from the wild type seedlings for the first 10 days after germination (Figure 4.7). The Fv/Fm ratios were similar in the base, middle and tip sections of both genotypes after day 10 (Figure 4.7). This data is consistent with results obtained by the imaging chlorophyll *a* fluorescence imaging (Figure 4.6).

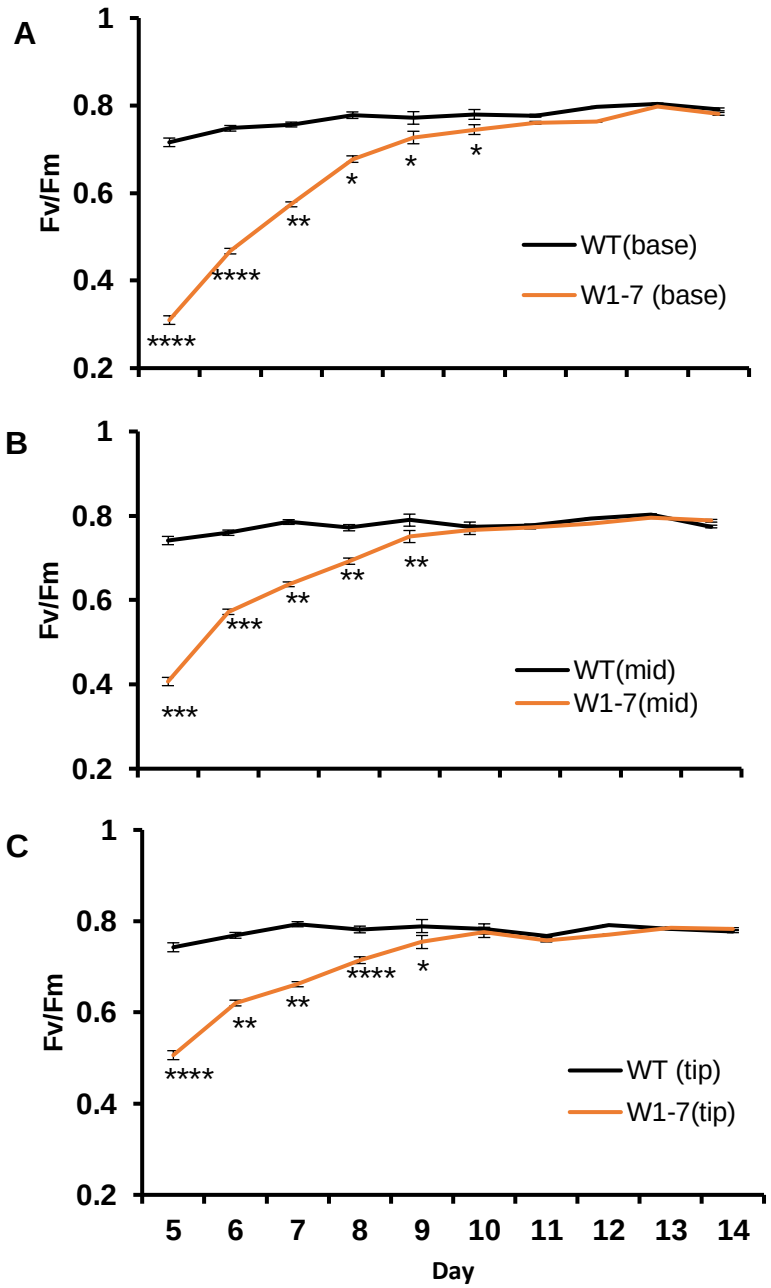


Figure 4.7: A comparison of the dark-adapted Fv/Fm ratios in the (A) base, (B) middle (Mid) and (C) tip regions of the wild type (WT) and W1-7 seedlings.

Data was expressed as mean values ± SE (n=10). Significant differences between the wild type, W1-1 and W1-7 was determined using the Student's *t*-test (*****p<0.05; **p<0.01; ***p<0.001 and ********p<0.0001).

4.2.3 Transcript abundance (plastid-encoded genes)

Genes encoding photosynthetic proteins are found in the nuclear and plastid genomes. The coordinated expression of genes encoded by the plastome and the nuclear genome is required for the efficient assembly of the photosynthetic machinery and the operation of photosynthesis during chloroplast development.

4.2.3.1 Plastid-encoded

The abundance of both plastid-encoded and nuclear-encoded transcripts was assessed by qRT-PCR. A small number of important photosynthetic proteins, such as the D1 of photosystem (PS) II are encoded by the plastid genome. This protein is encoded by the *PSBA* gene. Similarly, the large subunit (RBCL) of ribulose-1, 5-bisphoshate carboxylase (RuBiSCO) is encoded by the plastid genome (Figure 4.8). The levels of *RBCL* and *PSBA* transcripts were significantly lower in the base, middle and tip sections of the first leaves of 7 day old of W1-1 and W1-7 plants compared to the wild type leaves (Figure 4.8). However, the transcript levels of *RBCL* and *PSBA* were significantly higher in the base, middle and tip sections of the first leaves of W1-1 and W1-7 14-day old plants compared to the wild type (Figure 4.9).

The levels of transcripts encoded by plastome genes such as *MLOC_76327* were lower in all sections (base, middle and tip) of the first leaves of 7-day old W1-1 and W1-7 seedlings compared to the wild type (Data in Appendix B.1). The transcript levels of *PETD* were significantly lower in the base and middle sections of the first leaves of W1-1 and W1-7 of 14-day old plants compared to the wild type (Appendix B.1). However, the abundance of *PETD* transcripts was higher in the tip sections of the W-1 leaves compared to the wild type. However, *PETD* transcript levels in the tip section of W1-7 leaves were similar to the wild type (Data in Appendix B.1).

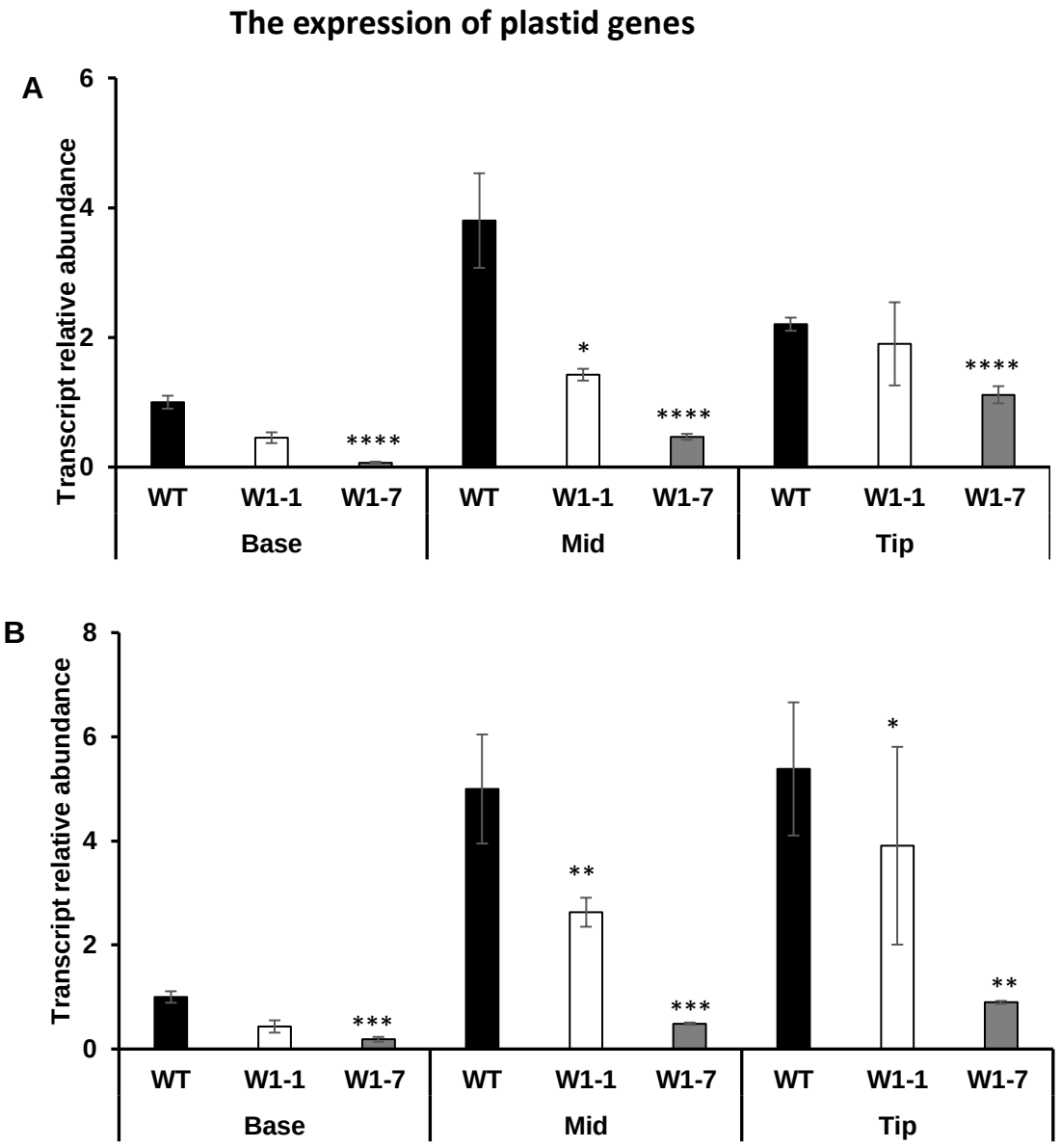


Figure 4.8: Levels of transcripts encoded by plastid genes; (A) the large subunit of ribulose-1, 5-bisphosphate carboxylase (RBCL) and (B) the photosystem II, D1 protein (PSBA) in the base, middle (Mid) and tip sections of the first leaves of wild type (WT), W1-1 and W1-7 seedlings 7 days after germination.

Data was normalised to the 16S. Data was set to 1 and W1-1 and W1-7 were compared to the wild type. Values are represented as mean ± SE (n=6). Asterisks indicate significant differences between WHY1- deficient and wild type plants as estimated by the Student's *t*-test (*****p<0.05; **p<0.01; ***p<0.001 and ********p<0.0001).

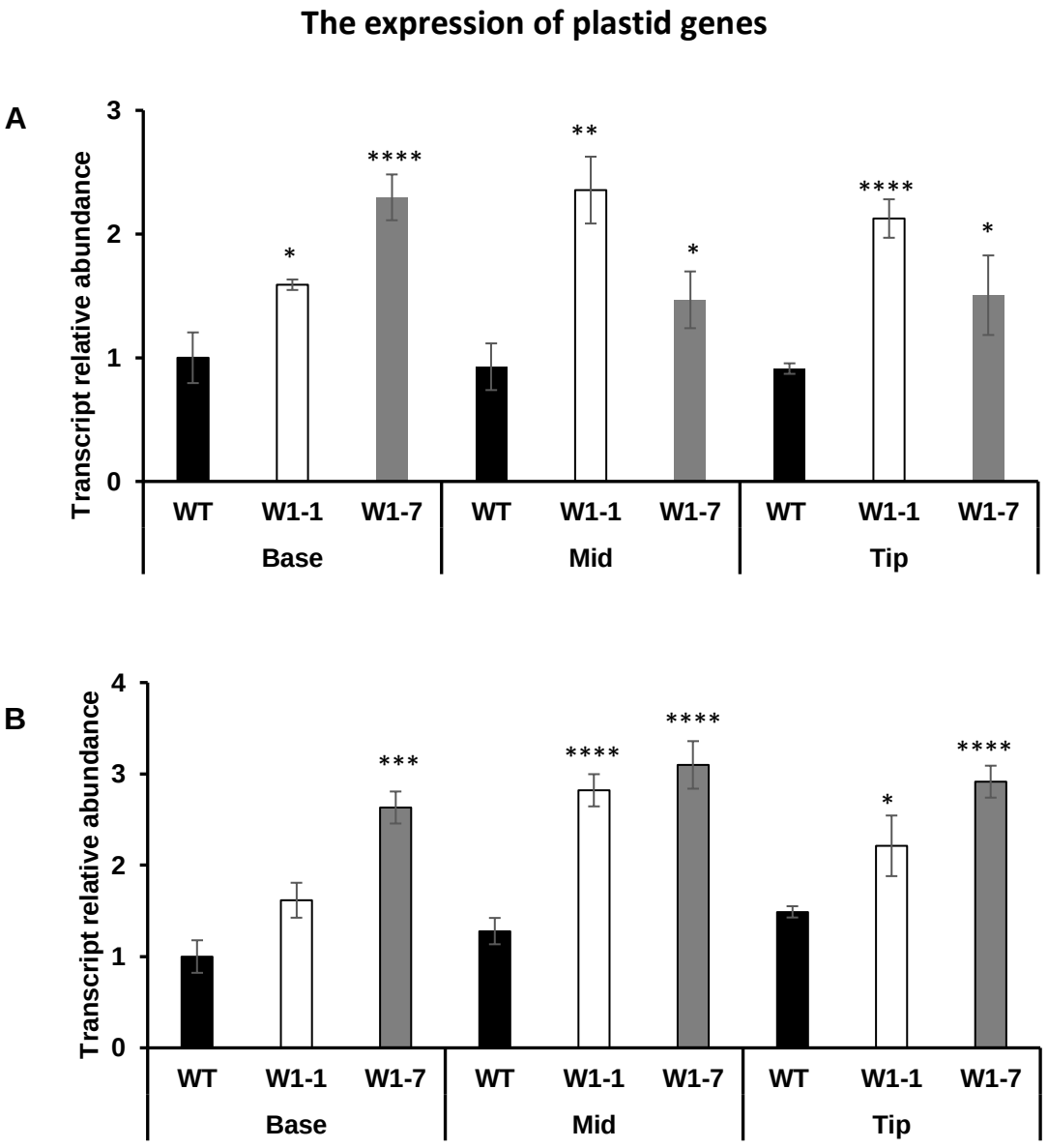


Figure 4.9: Levels of transcripts encoded by plastid genes. (A) The large subunit of ribulose-1, 5-bisphosphate carboxylase (RBCL) and B) the photosystem II, D1 protein (PSBA) in the base, middle (Mid) and tip sections of the first leaves of wild type (WT), W1-1 and W1-7 seedlings 14 days after germination.

Data was normalised to the 16S. Data was set to 1 and W1-1 and W1-7 were compared to the wild type. Values are represented as mean ± SE (n=6). Asterisks indicate significant differences between WHY1- deficient and wild type plants as estimated by the Student's *t*-test (*****p<0.05; **p<0.01; ***p<0.001 and ********p<0.0001).

4.2.3.2 Plastid-encoded RNA polymerases (PEP)

The levels of *RPOC* and *RPS16* transcripts transcribed by the plastid-encoded polymerases (NEP) were significantly lower in the base, middle and tip sections of the first leaves of 7 day old W1-1 and W1-7 seedlings compared to the wild type (Figure 4.10). However, the abundance of *RPOC* and *RPS16* transcripts were higher in the tip sections of the W1-1 leaves than the wild type at this stage (Figure 4.10).

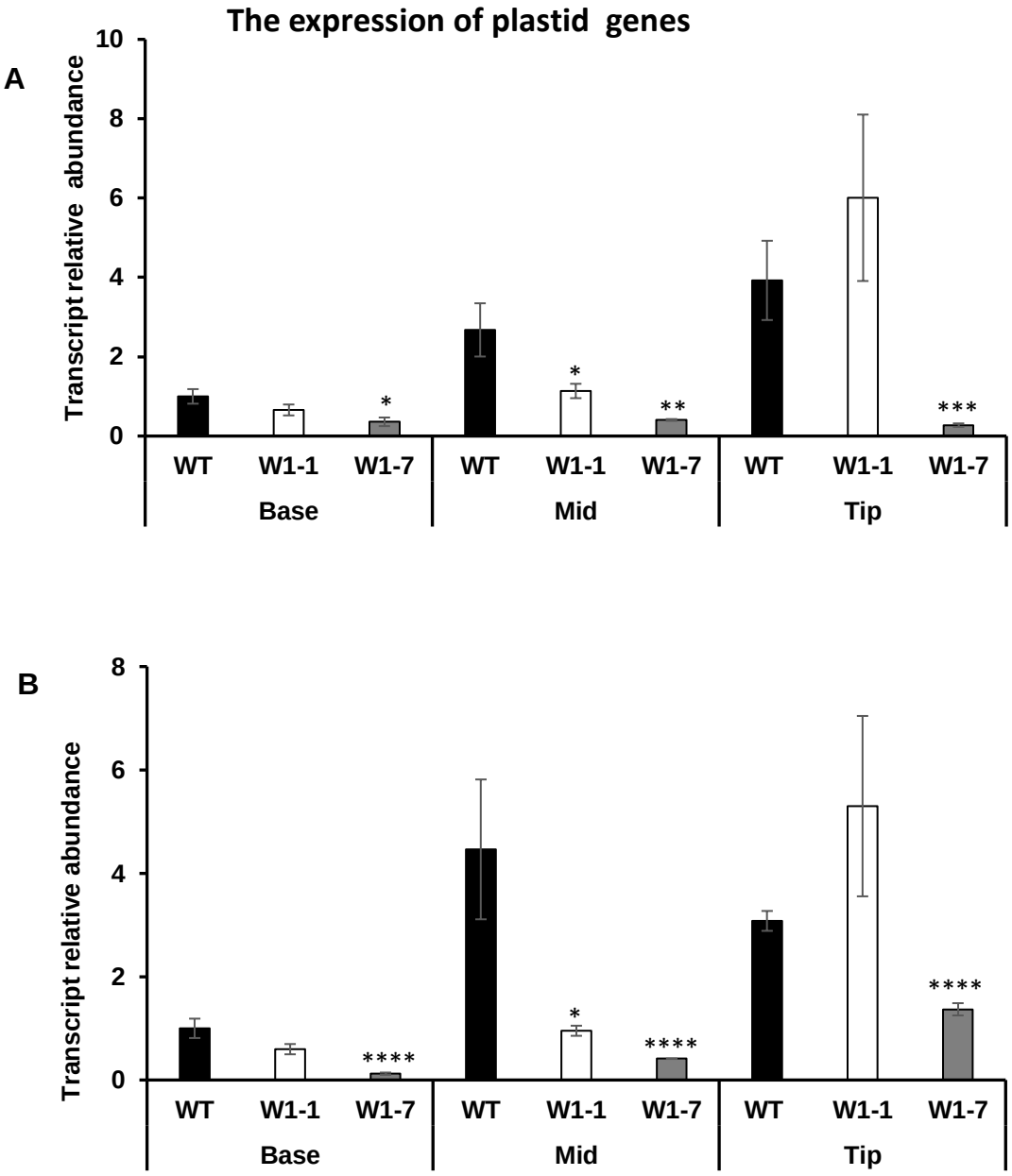


Figure 4.10: Levels of ribosomal photosynthetic transcripts that are encoded by NEP-transcribed plastid genes. (A) *RPOC* and (B) *RPS16* in the base, middle (Mid) and tip sections of the first leaves of wild type (WT), W1-1 and W1-7 seedlings 7 days after germination.

Data was normalised to the 16S. Data was set to 1 and W1-1 and W1-7 were compared to the wild type. Values are represented as mean ± SE (n=6). Asterisks indicate significant differences between WHY1-deficient and wild type plants as estimated by the Student's *t*-test (*****p<0.05; **p<0.01; ***p<0.001 and ********p<0.0001).

4.2.3.3 Nuclear-encoded transcripts

In contrast to the levels of transcripts encoded by plastome genes, the levels of nuclear-encoded transcripts, which encode proteins targeted to the chloroplasts were similar or higher in the W1-1 and W1-7 leaves compared to the wild type at 7 days and 14 days after germination. For example, the levels of transcripts encoding light-harvesting chlorophyll a/b binding proteins (LHCA, LHCB and LHCB1.1) were similar in the base, middle and tip sections of W1-1 leaves and the wild type (Figure 4.11). However, *LHCA, LHCB* and *LHCB1.1* transcripts were higher in the base, middle and tip sections of the first leaves of 7 day old W1-7 seedlings than equivalent sections in the wild type (Figure 4.11).

The levels of a large number of nuclear-encoded photosynthetic transcripts (*MLOC_58312*, *MLOC_64606*, *MLOC_33258*, *MLOC_77244*, and *MLOC_59016* and *AK362199*) were similar in the base, middle and tip sections of the first leaves of 7 day old W1-1 seedlings relative to the wild type (Appendix B.2). However, the levels of these transcripts were significantly higher in the base, middle and tip sections of the W1-7 leaves compared to the wild type at 7 days after germination (Appendix B.2).

The abundance of *RBCS* transcripts was similar in all sections of the first leaves of 7 day old wild type and W1-1 seedlings (Figure 4.12 A). However, *RBCS* transcripts were significantly higher in the middle and tip sections of the first leaves of the W1-7 plants compared to the wild type at 7 days of germination. In contrast, *RBCS* transcripts were significantly lower in the basal sections of the W1-7 leaves relative to the wild type (Figure 4.12 A).

The levels of transcripts encoding the nuclear-encoded plastid targeted RNA polymerase (NEP: *RPOTP*) tended to be higher in all sections of the first leaves of 7 day old W1-1 seedlings and the wild type (Figure 4.12 B). However, the levels of *RPOTP* transcripts were significantly higher in all sections of the W1-7

seedlings than the wild type after 7 days of germination (Figure 4.12 B). The changes in the abundance of nuclear-encoded transcripts were much less marked in the W1-1 and W1-7 leaves compared to the wild type at 14 days after germination (Figure 4.13 and more details in Appendix B.2).

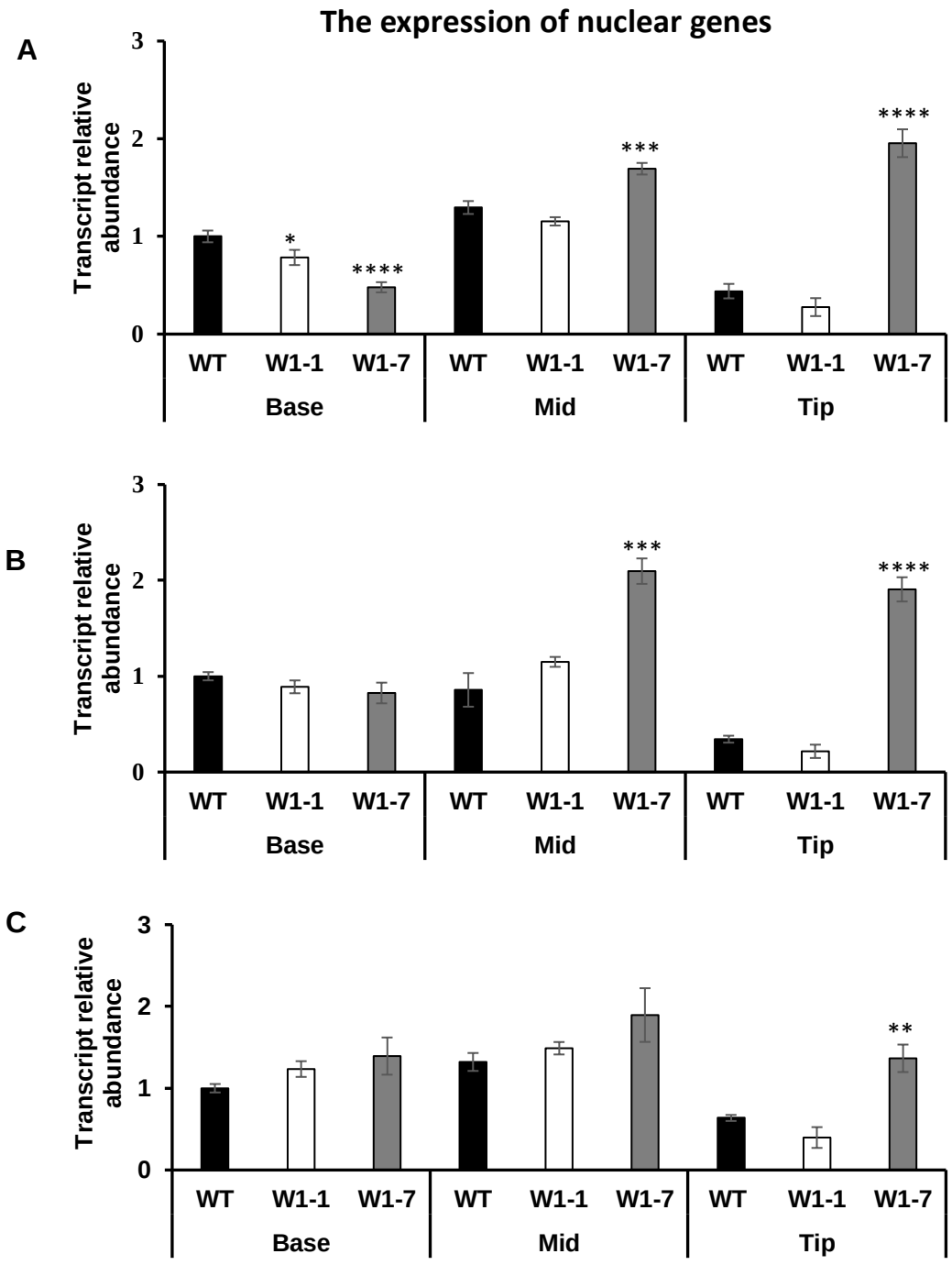


Figure 4.11: Levels of chloroplast-targeted transcripts encoded by nuclear genes. The light harvesting chlorophyll a/b binding complex. (A) LHCA (B) LHCB and (C) LHCB1.1 in the base, middle (Mid) and tip sections of the first leaves of wild type (WT), W1-1 and W1-7 seedlings 7 days after germination. Data was set to 1, and W1-1 and W1-7 were compared to the wild type. Values are represented as mean ± SE (n=6). Asterisks indicate significant differences between WHY1-deficient and wild type plants as estimated by the Student's *t*-test (*****p<0.05; **p<0.01; ***p<0.001 and ********p<0.0001).

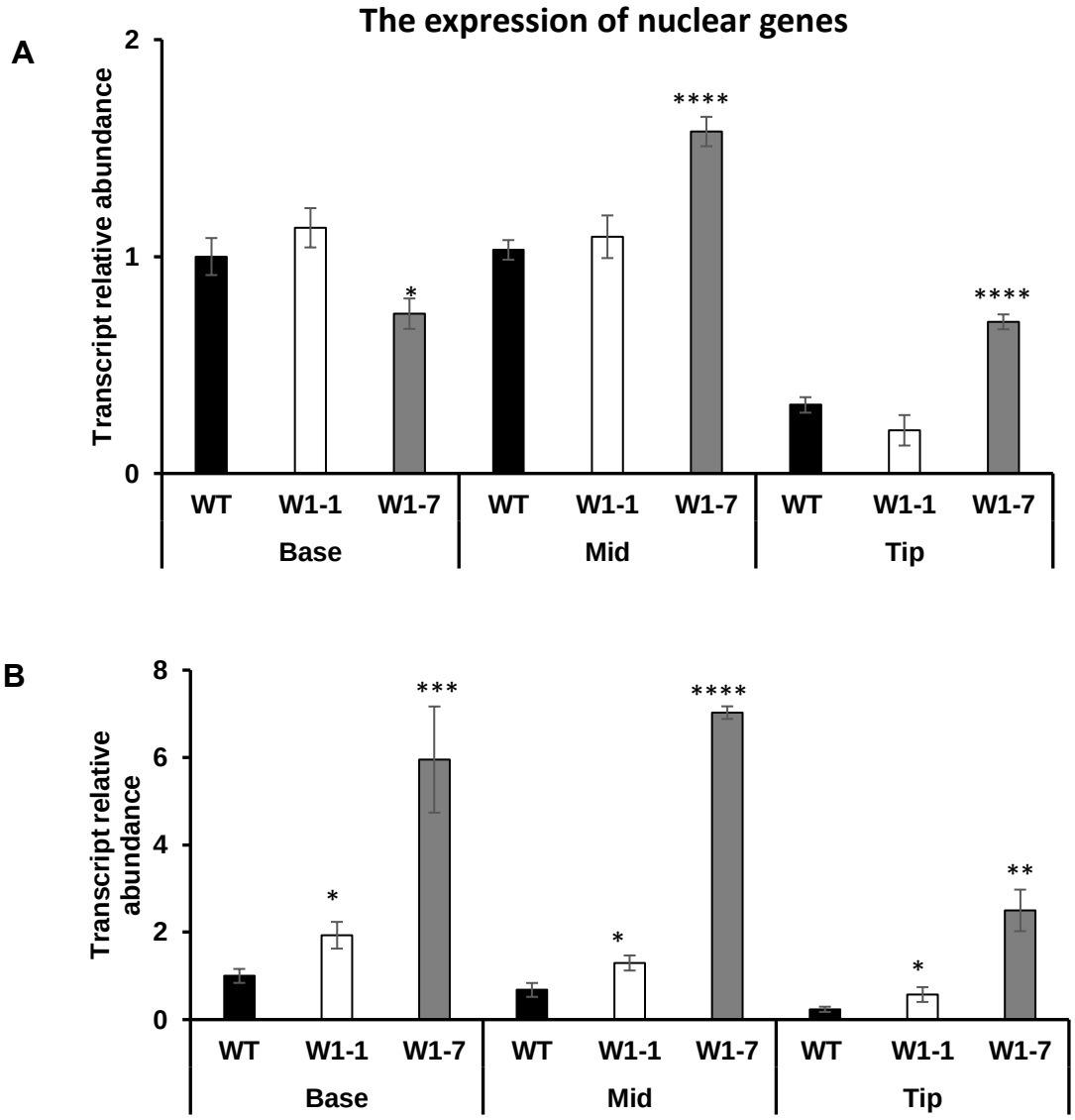


Figure 4.12: Levels of chloroplast-targeted transcripts encoded by nuclear genes. (A) The small subunit of ribulose-1, 5-bisphosphate carboxylase (RBCS) and (B) the nuclear-encoded, plastid targeted RNA polymerases (RpoTp) in the base, middle (Mid) and tip sections of the first leaves of wild type (WT), W1-1 and W1-7 seedlings 7 days after germination.

Data was set to 1 and W1-1 and W1-7 were compared to the wild type. Values are represented as mean ± SE (n=6). Asterisks indicate significant differences between WHY1-deficient and wild type plants as estimated by the Student's *t*-test (*****p<0.05; **p<0.01; ***p<0.001 and ********p<0.0001).

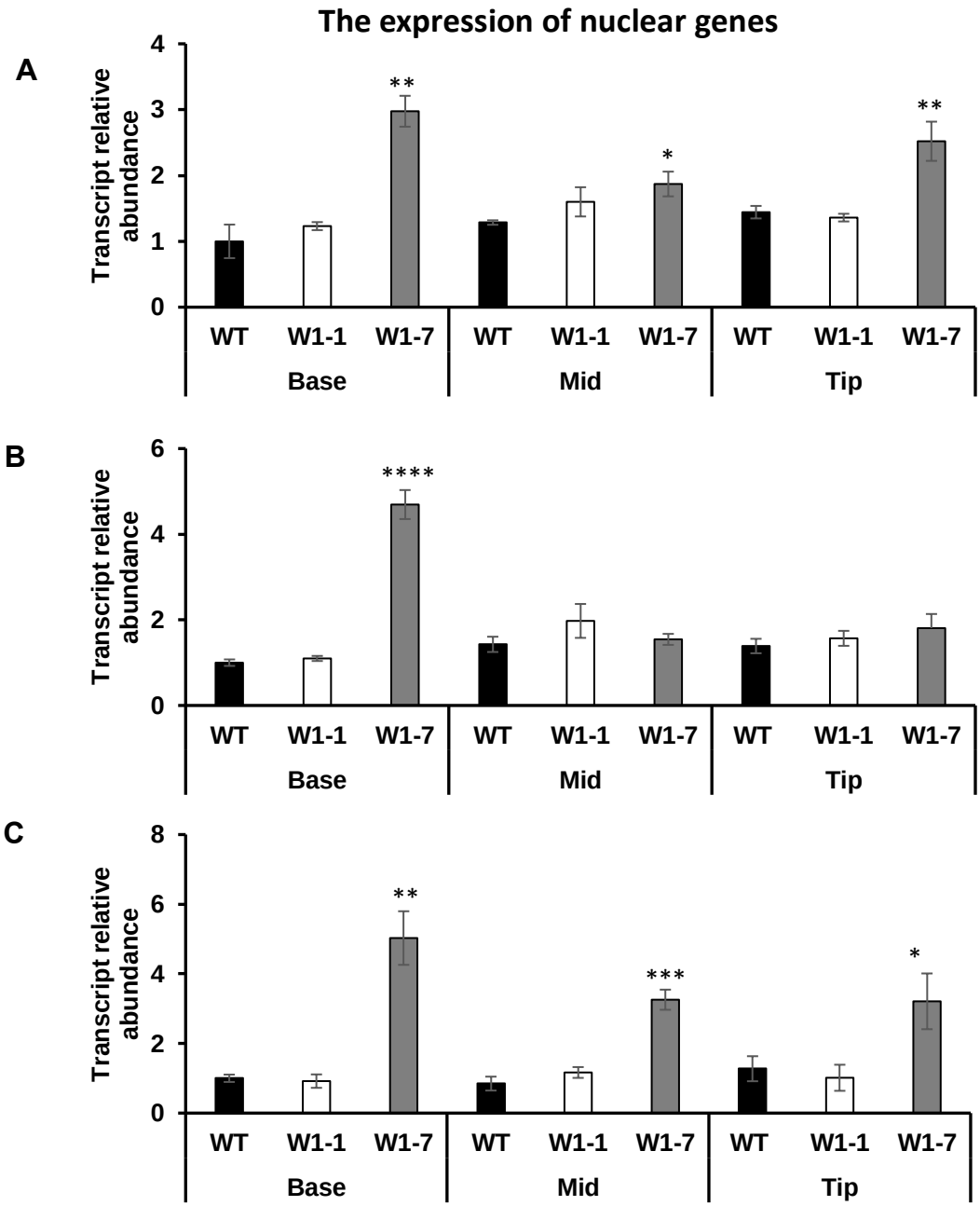


Figure 4.13: Levels of chloroplast-targeted transcripts encoded by nuclear genes. (A) The light harvesting chlorophyll a/b binding complex (LHCA), (B) the small subunit of ribulose-1, 5-bisphosphate carboxylase (RBCS) and (C) the nuclear-encoded photosynthetic transcripts (MLOC_59019) in the base, middle (Mid) and tip sections of the first leaves of wild type (WT), W1-1 and W1-7 seedlings 14 days after germination.

Data was set to 1, and W1-1 and W1-7 were compared to the wild type. Values are represented as mean ± SE (n=6). Asterisks indicate significant differences between WHY1-deficient and wild type plants as estimated by the Student's *t*-test (*****p<0.05; **p<0.01; ***p<0.001 and ********p<0.0001).

4.2.4 Protein accumulation

The W1-1 and W1-7 seedlings had significantly more light harvesting proteins (LHCBII) in all leaf sections than the leaves of the wild type both at 7 and 14 day old (Figure 4.14). The levels of RBCL protein were decreased in all sections of the W1-1 and W1-7 leaves relative to the wild type at 7 days after germination (Figure 4.14 A). The decrease in the levels of the RBCL protein were similar to the observed changes in the abundance of the RBCS protein. While the levels of RBCL proteins were similar in the leaves of the W1-1 and W1-7 seedlings compared to the wild type at 14 days after germination (Figure 4.14.B), the abundance of RBCS proteins remained at low level in the W1-1 and W1-7 leaves compared to the wild type (Figure 4.14.B). The levels of D1 proteins and RPS1 proteins were lower in all leaf sections of the W1-1 and W1-7 seedlings than the wild type leaves 7 days after germination (Figure 4.14 A). In contrast, the abundance of D1 protein was lower in the base and middle section of the W1-1 and W1-7 leaves than the wild type at the 14 day old stage (Figure 4.14.B). The levels of RPS1 protein were similar in all the leaf sections of all genotypes at 14 days after germination (Figure 4.14.B). The levels of WHY1 protein were lower in all regions of the leaves of 7 and 14 day old of W1-1 and W1-7 seedlings than the leaves of the wild type (Figure 4.14). However, the WHY1 protein levels were not in line with the transcript levels as mentioned in section 4.2.1.1.

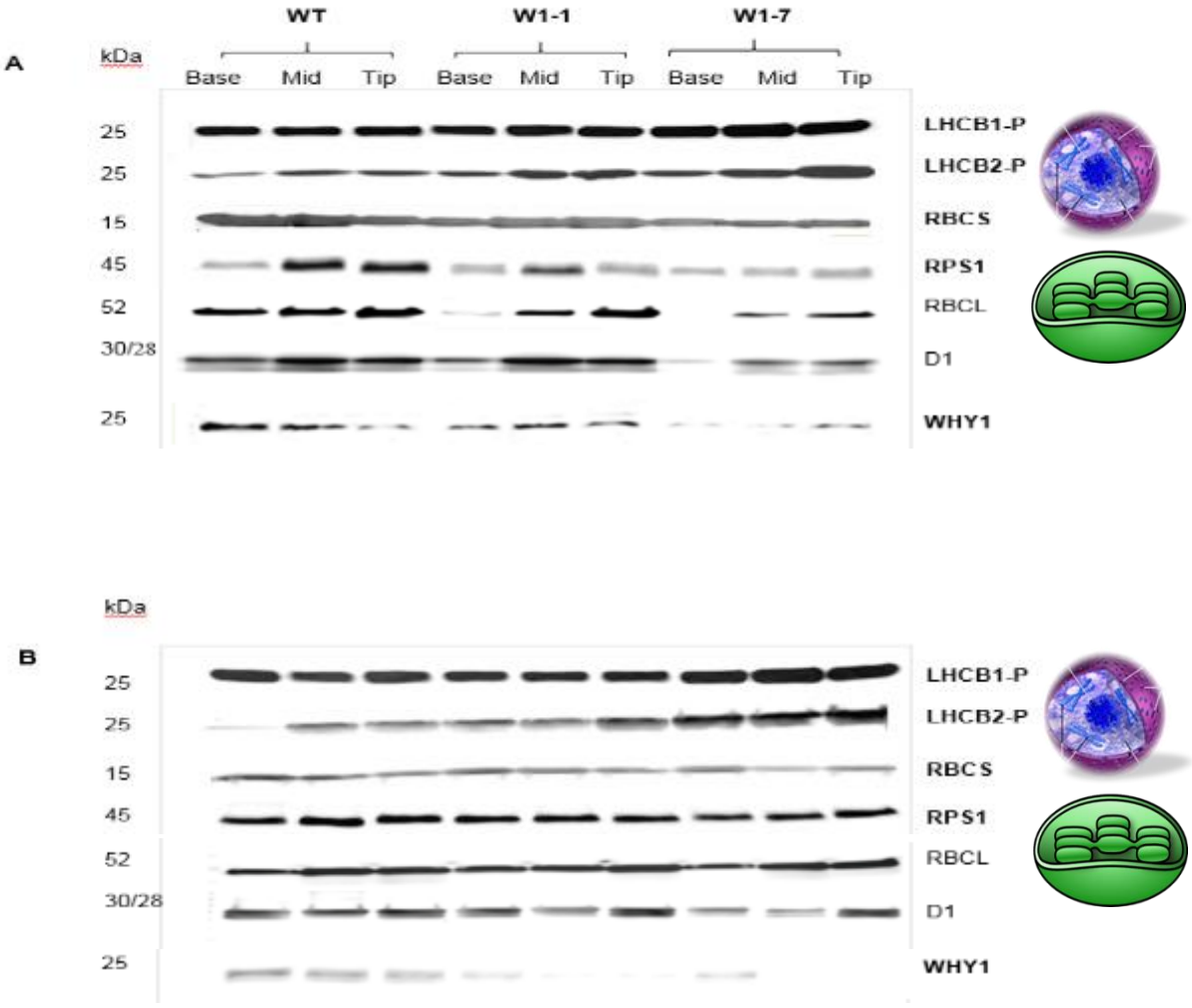


Figure 4.14: Western blot analysis of total proteins in the base, middle (Mid) and tip sections of the first leaves of wild type (WT), W1-1 and W1-7 of seedlings at (A) 7 and (B) 14 days after germination.

Proteins detected are two of the three different individual subtypes of chlorophyll a/b-binding proteins (LHCB1 and LHCB2). LHCB1 is the most abundant form and is encoded by several nuclear genes. The small subunit of ribulose-1, 5-bisphosphate carboxylase (RBCS), chloroplast ribosomal protein S1 (RPS1), WHY1 (WHY1), the large subunit of RBC (RBCL) and the photosystem II, D1 protein.

4.2.5 Plastid DNA content

qRT-PCR with specific primers for plastid genes (*PETD* and *PSBA*) and *RBCS* as reference for the nuclear genome were used to calculate the ratios of plastid DNA levels (ptDNA) to the nuclear levels (nDNA). The levels of *PETD* and *PSBA* genes (ptDNA contents) were significantly higher in all sections of the W1-7 leaves of 7 day old seedlings than the wild type (Figure 4.15 A). The abundance of *PETD* and *PSBA* transcripts were also higher in all sections of the 14 day old W1-7 leaves than the wild type (Figure 4.15 B). The mature leaves of the W1-7 at 3-week old seedlings had higher levels of ptDNA, as determined by the levels of *PETD* and *PSBA* transcripts than the wild type (Data shown in Appendix B.3). In contrast, the levels of *PETD* and *PSBA* transcripts were lower in the roots of 7-day old W1-7 seedlings than the wild type (Data shown in Appendix B.3.1).

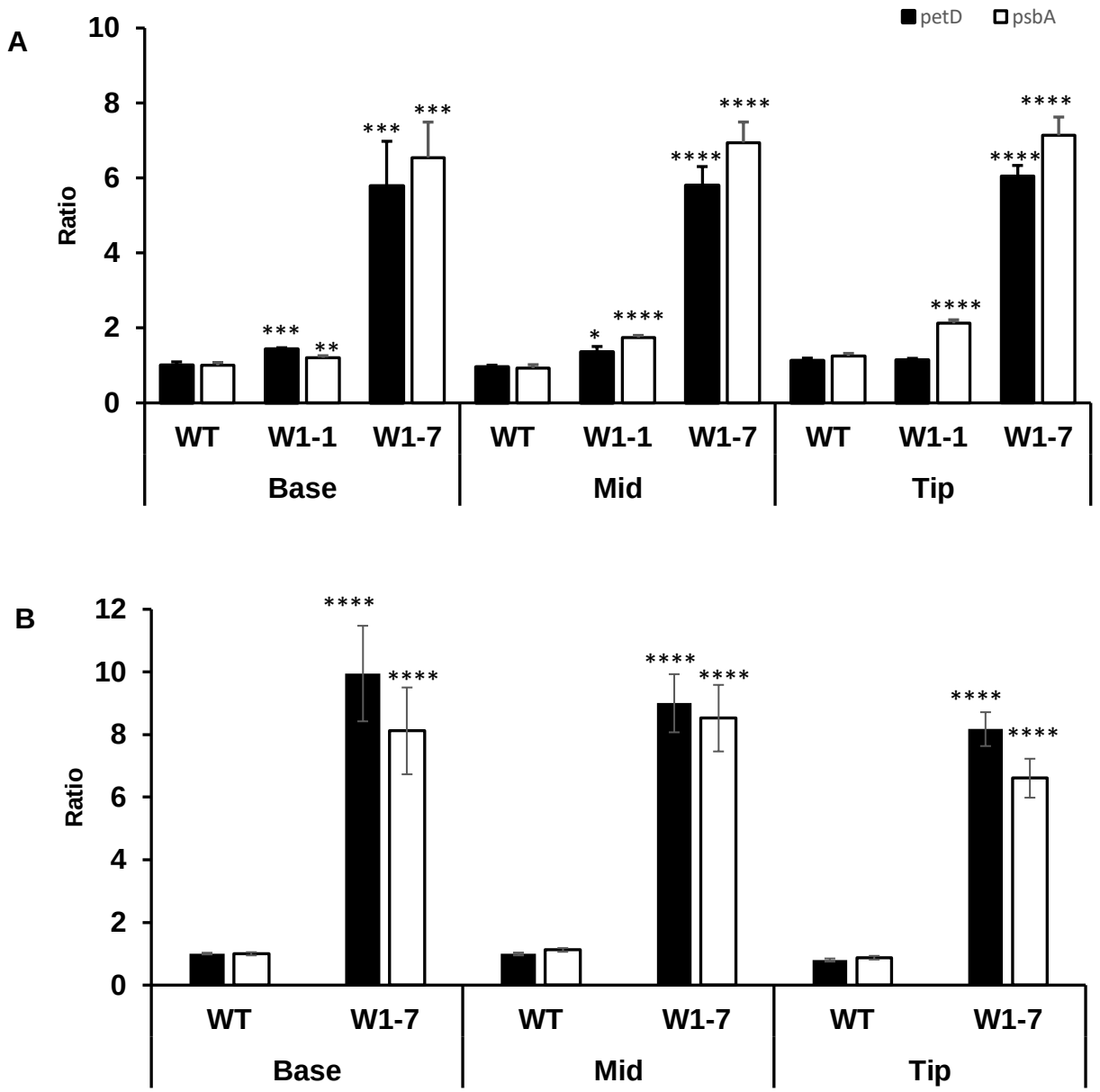


Figure 4.15: The ratios of plastid (pt) DNA levels to nuclear (n) levels (ptDNA/nDNA ratios) in the first leaves of wild type (WT), W1-1 and W1-7 seedlings at (A) 7 days after germination and (B) WT and W1-7, 14 days after germination.

Ratios were measured using specific primers to the plastome targets *petD* and *psbA*, with *rbcS* as a reference for the nuclear genome. Data was normalised to the *18S* rDNA gene and values for the WT were set to 1. Values are represented as mean ± SE (n=6). Asterisks indicate significant differences between WHY1-deficient and wild type plants as estimated by the Student's *t*-test (*****p<0.05; **p<0.01; ***p<0.001 and ********p<0.0001).

4.2.6 Chloroplast rRNA processing

The knock-down of *Arabidopsis RH22* leads to a delayed-greening phenotype with defects in plastid development (Appendix B). The RH22 phenotype is caused by aberrant splicing of rRNAs (23S rRNA and 4.5S rRNA) for the large subunit (50S) of the plastid ribosomes (Chi et al., 2012). Therefore, the accumulation and processing of chloroplast rRNAs was further examined by RNA gel blotting using probes spanning the 23S-4.5S region specific to individual rRNA bands in the base, middle and tip sections of the first leaves of wild type, W1-1 and W1-7 seedlings 7 days after germination to answer the question whether WHY1 is required for chloroplast rRNA processing (Figure 4.16 A). Very low signals were observed for the base, middle and tip regions of the W1-7 leaves compared to the wild type and W1-1 leaves on glyoxal agarose gels with total RNA (Figure 4.16 B). The probes detected equal amounts of 23S rRNA (1.1 kb) in the base, middle and tip regions of the wild type leaves with no detectable unprocessed precursor forms (Figure 4.16 C). However, there were differences in the accumulation of mature 23 S rRNA (1.1 kb) in the W1-1 and W1-7 seedlings relative to the wild type (Figure 4.16 C). Unprocessed precursor forms (3.1 kb) were detected in the W1-1 and W17 seedlings (Figure 4.16 C).

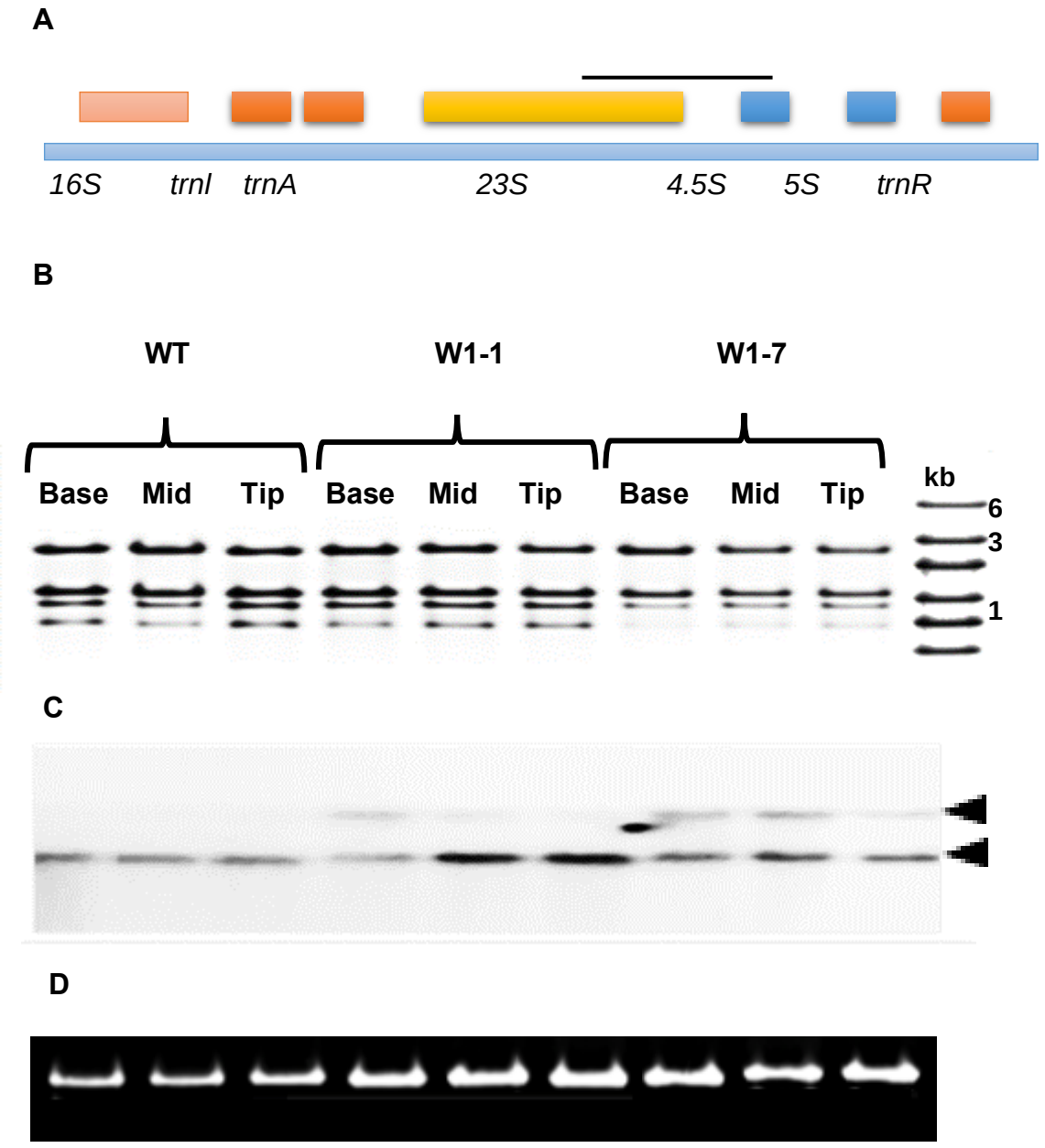


Figure 4.16: Altered splicing of plastid ribosomal RNA spanning the 23S to the 4.5S region in the base, middle (Mid) and tip sections of the first leaves of wild type (WT), W1-1 and W1-7 seedlings 7 days after germination.

(A) Diagram of the chloroplast rRNA operon and the locations of probes used for RNA gel-blot analysis spanning the 23S-4.5S regions (99142-99536 base pairs), (B) total RNA separated on a 1.2% denaturing glyoxal agarose gel stained with ethidium bromide, M-RiboRuller ™ High Range RNA ladder, (C) northern blot hybridisation with a ^{32}P-fragment spanning 23S and 4.5S region and (D) 25S rRNA stained with ethidium bromide as a loading control.

4.3 Discussion

The studies described in this chapter were performed to characterise the molecular phenotypes of WHY1-deficient (W1-1 and W1-7) leaves in relation to the wild type in seedlings grown for 7 and 14 days after germination. The leaves were separated into base, middle and tip sections because barley leaves, like those of other monocotyledonous plants, show a gradient in chloroplast development from the base to the tip. Taken together, the data presented here allow a number of conclusions to be drawn.

4.3.1 WHY1 is required for chloroplast development in barley leaves

Previous studies have reported that the leaves of the transgenic barley lines used in the present study have significantly more chlorophyll than the wild type at 21 days after germination (Comadira et al., 2015). The data presented here shows that the emerging leaves of the W1-1 and W1-7 seedlings show delayed greening at earlier stages of development (7 and 14 days after germination) compared to the wild type (Figure 4.1 and Figure 4.2). Furthermore, the loss of WHY1 delayed leaf pigment accumulation in the WHY1-deficient seedlings in a development-dependent manner (Figure 4.4 and Figure 4.5). These data suggest that the development of mature and fully functional chloroplast was delayed in WHY1-deficient seedlings relative to the wild type (Figure 4.6). The maximal photochemical efficiency of PSII (determined by Fv/Fm ratios) was lower in the developing W1-7 leaves than the wild type (Figure 4.6). The Fv/Fm ratios in the W1-7 leaves gradually increased from the basal sections to the tip, but maximum values were achieved much later in the W1-7 leaves than the wild type (Figure 4.6 & Figure 4.7). However, there was a limitation in measuring Fv/Fm ratios during this study, therefore, this experiment should be repeated in the W1-1 leaves in the future. In addition, there is a gradient of development in monocotyledon leaves that starts in the leaf basal meristem region (Baumgartner et al., 1989, Hess et al., 1993). The findings reported here suggest that WHY1 is required for the timing of greening during barley leaf development, particularly at

the earliest stages of leaf development. Despite the delay in leaf greening, the WHY1-deficient lines produced high seed yield in the T4 generation (Chapter 3, Table 3.3). This observation suggests that the enhanced chlorophyll accumulation and stay green phenotype of the mature leaves induced by loss of WHY more than makes up for the delay in leaf development in terms of seed yield.

4.3.2 WHY1 is essential for chloroplast to nucleus signalling

Plastid transcription requires two different RNA polymerases that recognise distinct types of promoters: a nuclear-encoded plastid RNA polymerase (NEP) and the plastid-encoded RNA polymerase (PEP). A shift from NEP-mediated transcription to PEP-mediated transcription occurs early in chloroplast development. However, little is known about the mechanisms involved in this process. The studies reported here suggest that WHY1 is essential for both NEP and PEP activity during chloroplast development. Like all plastids, chloroplasts develop from proplastids that are present in the immature cells of plant meristems. The differentiation of proplastids into chloroplasts requires the establishment of the PEP complex. The PEP complex is a bacterial-type multi-subunit enzyme involved in the transcription of photosynthesis-related genes. It is composed of a catalytic core comprised of plastid-encoded proteins (*rpoA, rpoB, rpoC1* and *rpoC2*) and additional polymerase-associated proteins (PAP), including other nuclear-encoded polymerase-associated proteins and sigma factors (SIGs), which are required by PEP for promoter recognition (Dietz et al., 2011).

The loss of WHY1 causes differential effects on the levels of photosynthetic transcripts encoded by plastid and nuclear genes. The coordination of nuclear and plastid-encoded gene expression is disrupted in the absence of WHY1. In the present studies, the levels of NEP transcripts (Figure 4.11 & Figure 4.12) were higher in the developing W1-1 and W1-7 leaves than in the wild type, suggesting that NEP functions are limited in the absence of WHY1. In agreement

with this view, plastid-encoded RNA polymerases such as RPOC and RPS16, which are transcribed exclusively by NEP, were very low in WHY1-deficient lines relative to the wild type. This finding indicates that the loss of WHY1 disturbs NEP-dependent transcript accumulation (Figure 4.10). The leaves of the WHY1-deficient lines showed decreased levels of PEP-associated transcripts such as *PSBA* and *RBCL* (Figure 4.8) in 7 day old seedlings. However, the levels of PEP-associated transcripts were higher in the WHY1-deficient lines relative to the wild at 14 days old. These findings are in agreement with studies on other mutants that have high NEP transcript levels but low levels of PEP transcripts (Hess et al., 1994, Díaz et al., 2018).

Evidence suggests that WHY proteins may be polymerase-associated proteins in *Arabidopsis*, allowing the possibility of a functional interaction between these proteins (Díaz et al., 2018). It has recently been suggested that a positive signal generated by PEP activity stimulates the expression of chloroplast genes in the nucleus and promotes photosynthesis associated nuclear genes expression (Díaz et al., 2018). Hence, the PEP complex connects the functional state of the chloroplast to the nucleus, synchronizing the expression of photosynthetic genes in the nuclear and chloroplast genomes during seedling development (Díaz et al., 2018). The data presented here show that WHY1 is required for PEP activation and the establishment of functional chloroplasts and photosynthesis at the earliest leaf developmental stages. The establishment of successful chloroplast differentiation in WHY1-deficient seedlings is delayed because of the faltered or delayed switching from NEP to PEP.

The loss of WHY1 differentially regulates the levels of proteins encoded by plastid and nuclear genes. The RuBiSCO is made up of 8 large subunits (RBCL) that are encoded by plastome genes and 8 small subunits (RBCS) that are encoded by nuclear genes. The abundance of both proteins was co-ordinated in the WHY1-deficient lines. There was a decrease in the overall abundance of both RuBiSCO subunits in WHY1-deficient lines. It is probable that the RBCL protein was turned over rapidly in the absence of sufficient RBCS protein (Figure 4.14 A). The

WHY1-deficient lines have more light harvesting proteins than the wild type at both 7 and 14 days after germination (Figure 4.14). In contrast, photosynthetic proteins encoded by plastome genes such as *RBCL*, *RPS1* and *D1* (*PSBA*) were lower in WHY1-deficient lines than in the wild type at 7 days after germination (Figure 4.14).

Previous studies have shown that an inhibition of either the transcription or translation of photosynthetic proteins elicits retrograde signalling from the chloroplasts to the nucleus to suppress the transcription and translation of LHCBII (Díaz et al., 2018). This finding provides further evidence that lack of WHY1 disrupts the coordination of nuclear and plastid-encoded gene expression and protein synthesis. The decreases in plastid-encoded transcripts occurred despite the fact that the level of ptDNA was increased two to three-fold in the WHY1-deficient lines compared to the wild type seedlings at 7, 14 and 21 days old (Figure 4.15). The increases in ptDNA copy numbers, which normally lead to the increase of the expression of plastid genes, have been reported previously (Grevich and Daniell, 2005). The decrease in the levels of ptDNA copy number was because of the mutation of either of the two closely related organelle-localised DNA polymerases in *Arabidopsis* (Morley and Nielsen, 2016). Furthermore, it would be interesting to know the reason on the differences in the level of ptDNA in the WHY1-deficient lines compared to the wild type that could be due to the number of chloroplasts or the number of DNA per chloroplast. Therefore, this merits further consideration. It has been reported that ptDNA levels increase more than two-fold during chloroplast development in the barley leaves (Baumgartner et al., 1989). The data presented here confirm previous work showing that WHY1-deficient lines have a greater cpDNA copy number than the wild type (Krupinska et al., 2014). These findings suggest that WHY1 is involved in the repression of cpDNA replication during chloroplast development.

4.3.3 WHY1 is required for splicing in WHY1-deficient lines

Chloroplast genes are transcribed as operons to produce mono- and poly-cistronic RNAs. This process is regulated by the RNA metabolism complex that involves various types of RNA polymerases and extensive post-transcriptional RNA processing (Barkan, 2011). The *ZmWhy1* knockdown maize mutants have a pale green phenotype leading to the hypothesis that the absence of a functional WHY1 protein may lead to ribosome deficiency. However, in contrast to the very high levels of ptDNA content in barley lines lacking WHY1, ptDNA levels in *ZmWhy1* mutants did not change (Prikryl et al., 2008). The aberrant 23S and 4.5S rRNA processing observed in *ZmWhy1* suggests that the chloroplast ribosomes are not functional in these mutants. The data presented here shows that the WHY1-deficient barley seedlings had defects in chloroplast ribosomal RNA (rRNA) accumulation. The 23S rRNA and 4.5S rRNA precursors for the large subunit (50S) of the plastid ribosomes accumulated in WHY1-deficient lines (Figure 4.16). Taken together, these data suggest that WHY1 is essential for chloroplast biogenesis in barley because it is required for the correct splicing of chloroplast ribosomal proteins.

Arabidopsis mutants lacking the RH22 show altered splicing of chloroplast ribosomes (Chi et al., 2012, Kanai et al., 2013). The *RH22* knockdown mutants accumulate 23S and 4.5S rRNAs, but not 16S rRNAs (Chi et al., 2012). These rRNAs are the precursors of 50S ribosomal subunits that regulate the RPL24 protein that binds to the 23S subunit. The data presented in Chapter 7 shows that the WHY1 protein interacts with DEA (D/H)-box RNA 22 in *Arabidopsis* mesophyll protoplasts. The barley WHY1 protein may interact with RH22 in the same way it does in *Arabidopsis* and hence may have a function in the assembly of chloroplast ribosomes.

Chapter 5 . Transcript profile of the WHY1-deficient lines during chloroplast development

5.1 Introduction

Like all plastids, chloroplasts develop from proplastids that are present in the immature cells of plant meristems. Chloroplast biogenesis from proplastids requires coordination of the expression of both by nuclear and chloroplast genes, which is regulated by developmental and environmental signals (Kessler and Schnell, 2009). There are around 3000 proteins in the chloroplast, and most of these are encoded by the nucleus (Leister, 2003). Functional processes in chloroplast biogenesis include the import of nuclear-encoded proteins via the TOC and TIC complexes, protein assembly, thylakoid formation, pigment synthesis, plastid division and retrograde signalling (Pogson et al., 2015, Waters and Langdale, 2009). It is important for these processes to be synchronised, as any changes in chloroplast biogenesis can influence leaf development (Pogson et al., 2015). Transcriptome profiling has been used widely to study chloroplast development in a variety of plant species such as green curd cauliflower mutant (Zhou et al., 2011), albino wheat mutant (Shi et al., 2017) and peach (Chen et al., 2018). These studies have shown that transcripts associated with chloroplasts increased in the cauliflower mutant with green curds to the white cauliflower (Zhou et al., 2011) and in albino wheat mutant (Shi et al., 2017).

The analysis of the early stages of leaf development in the WHY1-deficient seedlings reported in Chapter 4 demonstrates a slower rate of pigment accumulation compared to the wild type. Thus, it is important to explore the transcript profiles of the WHY1-deficient barley leaves in order to gain a better understanding of the mechanisms by which the WHY1 protein regulates leaf development. In Chapter 2, it was shown that the loss of WHY1 had a marked

effect on the transcriptome profile of the embryos from dry seeds. RNA sequence comparisons of the transcript profiles revealed that the most abundant transcripts in the WHY1-deficient embryos relative to the wild type were related to plant hormone metabolism and adaptation to stress.

Transcriptome analysis has previously been performed on the leaves of three RNAi-knockdown barley lines (W1-1, W1-7 and W1-9), which have very low levels of *HvWHY1* expression under optimal or nitrogen-deficient conditions (Comadira et al., 2015). The transcript profile analysis of the WHY1-deficient leaves showed that genes encoding photosynthetic proteins were markedly changed relative to the wild type. The leaves of WHY1 had a much greater abundance of transcripts encoding the photosynthetic proteins such as thylakoid NADH complex, the chloroplast RNA polymerase, the cytochrome b/f complexes and chloroplast ribosomes than the wild type (Comadira et al., 2015). However, these analyses were performed on green leaves of plants harvested after the 14-day seedling growth stage. In the present study, therefore, the transcript profile of leaves was characterised in the base, middle and tip sections of the first leaves of 7 and 14-day old WHY1-deficient and wild-type seedlings. The aim of this analysis was to identify:

i) The significantly altered transcripts that were changed with respect to genotype, irrespective of where the leaf section, in 7 and 14 day old leaves;
ii) The significantly altered transcripts that were changed with respect to leaf section (base, middle, tip), irrespective of genotype, in 7 and 14 day old seedlings;
iii) The significantly altered transcripts dependent on both the leaf section and genotype in 7 day old leaves;
iv) The developmental gradient of transcript changes along the leaves of the wild type at 7 days old; and
v) The transcripts changed in the W1-7 relative to the wild type at 7 days old.

5.2 Results

5.2.1 Genotype-dependent transcript changes in the WHY1-deficient barley leaves

5.2.1.1 Ribosomal associated proteins

Many transcripts encoding ribosomal proteins that are present in the plastid genome (plastome) were increased in the base, middle and tip sections of the W1-7 leaves compared to the W1-1 and wild type, 7 days after germination (Figure 5.1 A). In contrast, only four transcripts encoding ribosomal proteins that are nuclear-encoded but targeted to the chloroplasts were significantly changed in abundance at 14 days (Figure 5.2 A). Transcripts encoding ribosomal proteins encoded by the plastome were increased in the base, middle and tip sections of the W1-7 leaves compared to the W1-1 and wild type, 14 days after germination (Figure 5.2 A).

A large number of transcripts encoding ribosomal proteins encoded in the nucleus were greatly increased in the base sections of the W1-7 leaves compared to the W1-1 and wild type, 7 days after germination (Figure 5.1 B). The levels of nuclear-encoded ribosomal transcripts were lower in the middle and tip sections of the leaves of W1-1 and the wild type compared to the W1-7, 7 days after germination (Figure 5.1 B). In contrast, the levels of only one nuclear-encoded ribosomal protein transcript were significantly changed (MLOC_80636.2) with values that were higher in the base, middle and tip sections in the W1-7 leaves compared to the W1-1 and wild type at 14 days (Figure 5.2 B).

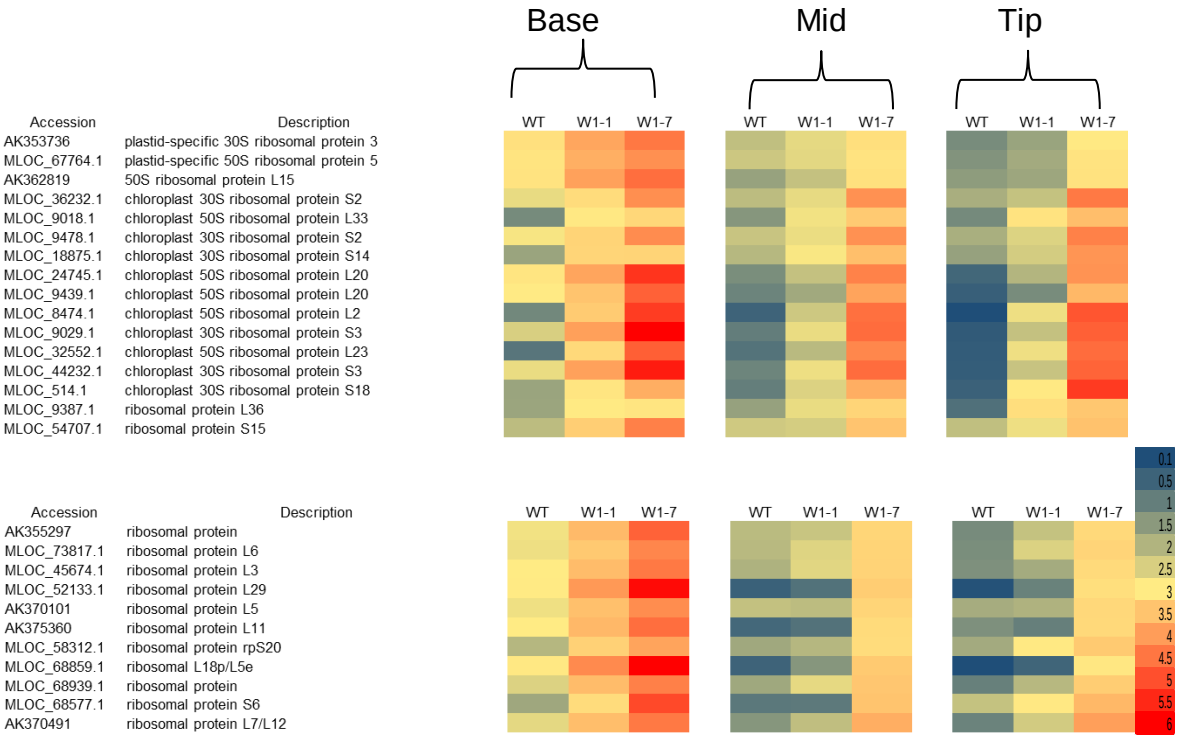


Figure 5.1: Heat map of transcript abundance of key transcripts associated with ribosomal proteins that encoded by the (A) chloroplasts and (B) nuclei in the W1-1, W1-7 and the wild type at 7 days old.

The differentially expressed transcripts were analysed using two-way analysis of variance (ANOVA), with the genotypes (WT, W1-1, W1-7) as factors, and a Bonferroni multiple-testing correction at a *p*-value of ≤0.05, FC>2 (Genespring12, Aligent Technologies). Each column showed transcript abundance in the WT, W1-1 and W1-7 in all the leaf sections. Accession numbers were indicated together with a brief description of the gene lists. The relative abundance of transcripts associated with ribosomal proteins is illustrated according to the scale bar shown.

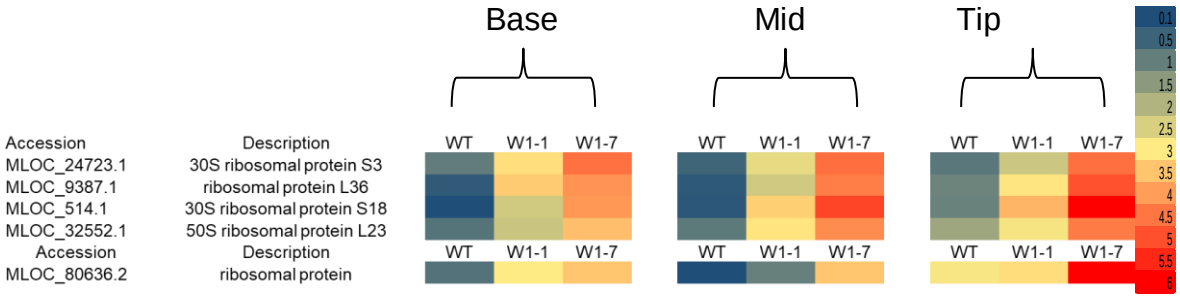


Figure 5.2: Heat map of transcript abundance of key transcripts associated with ribosomal proteins encoded by the (A) chloroplasts and (B) nuclei at 14 days old.

The differentially expressed transcripts were analysed using two-way analysis of variance (ANOVA), with the genotypes (WT, W1-1, W1-7) as factors, and a Bonferroni multiple-testing correction at a *p*-value of ≤0.05, FC>2 (Genespring12, Aligent Technologies). Each column showed transcript abundance in the WT, W1-1 and W1-7 in all the leaf sections. Accession numbers were indicated together with a brief description of the gene lists. The relative abundance of transcripts associated with ribosomal proteins is illustrated according to the scale bar shown.

5.2.1.2 Transcripts associated with photosynthesis

No genotype-dependent transcripts associated with photosynthesis were changed in the WHY1-deficient barley seedlings at 7 days old. Moreover, only two transcripts were higher in the base and middle sections of the W1-1 and W1-7 leaves compared to the wild type 14 days after germination and these encoded a chlorophyll A-B binding protein (Figure 5.3). The levels of transcripts encoding the chlorophyll A-B binding protein were similar in leaves tips in all the genotypes at 14 days old (Figure 5.3).

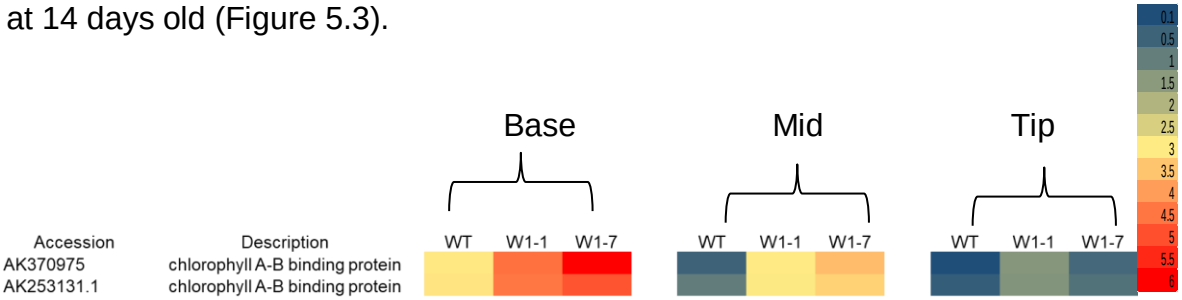


Figure 5.3: Heat map of transcript abundance of key transcripts associated with photosynthesis at 14 days old.

The differentially expressed transcripts were analysed using two-way analysis of variance (ANOVA), with the genotypes (WT, W1-1, W1-7) as factors, and a Bonferroni multiple-testing correction at a *p*-value of ≤0.05, FC>2 (Genespring12, Aligent Technologies). Each column showed transcript abundance in the WT, W1-1 and W1-7 in all the leaf sections. Accession numbers were indicated together with a brief description of the gene lists. The relative abundance of transcripts associated with photosynthesis is illustrated according to the scale bar shown.

5.2.1.3 Transcripts associated with RNA metabolism

The abundance of several transcripts encoding components associated with the RNA-binding proteins that are important in RNA processing in chloroplasts, such as the pentatricopeptide repeat (PPR) proteins, was higher in the base sections of the W1-7 leaves compared to the W1-1 and wild type, 7 days after germination (Figure 5.4 A). Many PPR transcripts were also higher in the middle and tip sections of the W1-7 leaves compared to the W1-1 and wild type at 7 days old (Figure 5.4 A).

The levels of many transcripts encoding the DEAD-box ATP-dependent RNA helicase and chloroplast group IIA intron-splicing facilitator CRS1, which are involved in processing and splicing, were higher in the base, middle and tip sections of the W1-7 leaves than the W1-1 and wild type, 7 days after germination (Figure 5.4 B). At 14 days, all five transcripts encoding PPR proteins were higher in the base, middle and tip sections of the W1-7 leaves compared to the W1-1 and wild type (Figure 5.5 A).

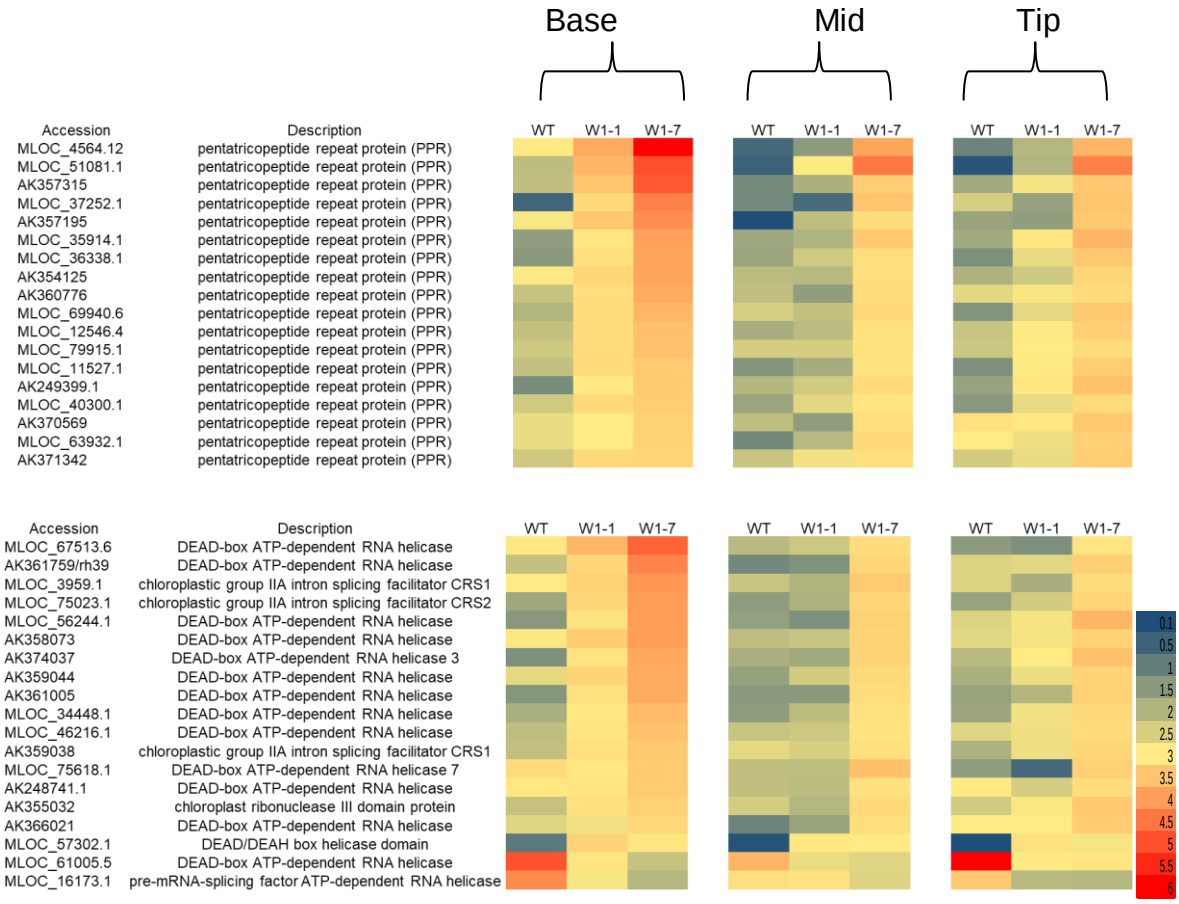


Figure 5.4: Heat map of transcript abundance of key transcripts associated with RNA metabolism such as the (A) PPR and (B) RNA helicases at 7 days old.

The differentially expressed transcripts were analysed using two-way analysis of variance (ANOVA), with the genotypes (WT, W1-1, W1-7) as factors, and a Bonferroni multiple-testing correction at a *p*-value of ≤0.05, FC>2 (Genespring12, Aligent Technologies). Each column showed transcript abundance in the WT, W1-1 and W1-7 in all the leaf sections. Accession numbers were indicated together with a brief description of the gene lists. The relative abundance of transcripts associated with RNA metabolism is illustrated according to the scale bar shown.

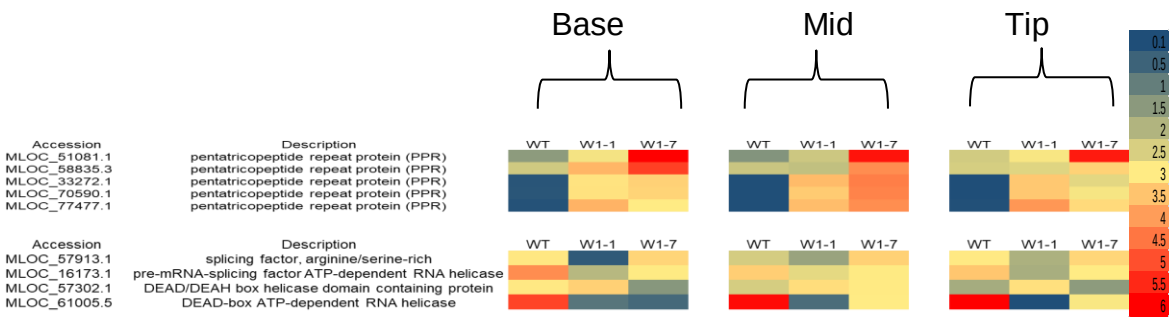


Figure 5.5: Heat map of transcript abundance of key transcripts associated with RNA metabolism, such as (A) pentatricopeptide repeat and (B) RNA helicases at 14 days old.

The differentially expressed transcripts were analysed using two-way analysis of variance (ANOVA), with the genotypes (WT, W1-1, W1-7) as factors, and a Bonferroni multiple-testing correction at a *p*-value of ≤0.05, FC>2 (Genespring12, Aligent Technologies). Each column showed transcript abundance in the WT, W1-1 and W1-7 in all the leaf sections. Accession numbers were indicated together with a brief description of the gene lists. The relative abundance of transcripts associated with RNA metabolism is illustrated according to the scale bar shown.

5.2.1.4 Protein kinase associated transcripts

The levels of transcripts encoding protein kinases (AK359767, AK356512, AK368106, MLOC_1961.1, MLOC_71301.2 and AK375444) were generally increased in the all the sections of the W1-7 leaves compared to the W-1 and wild type, 7 days after germination (Figure 5.6). Two transcripts encoding protein kinases (MLOC_60285.4 and MLOC_56354.3) were increased only in the base sections of the W1-7 leaves compared to the W-1 and wild type, 7 days after germination (Figure 5.6). The abundances of transcripts encoding a serine/theorine protein kinase (MLOC_4459.3), a cysteine-rich protein kinase (AK251994.1) and two other protein kinases (MLOC_4459.3 and AK362262) were similar in all the leaf sections of all of the genotypes 7 days after germination (Figure 5.6). At 14 days, the levels of all the identified transcripts encoding protein kinases were higher in the base, middle and tip sections of the W1-7 leaves compared to the W1-1 and wild type (Figure 5.7). However, one transcript (MLOC_69485.1) encoding an S-locus-like receptor protein kinase was lower in all the sections of W1-7 leaves compared to W1-1 and the wild type (Figure 5.7).

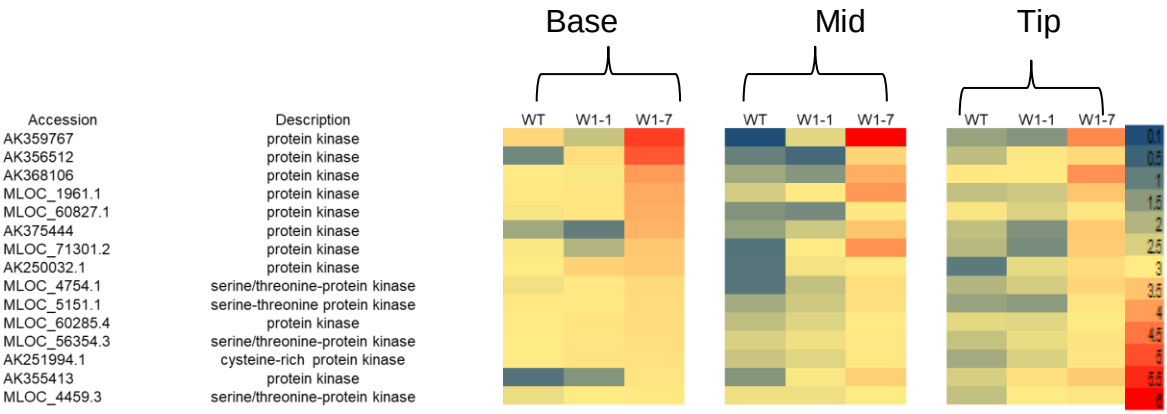


Figure 5.6: Heat map of transcript abundance of key transcripts associated with protein kinases at 7 days old.

The differentially expressed transcripts were analysed using two-way analysis of variance (ANOVA), with the genotypes (WT, W1-1, W1-7) as factors, and a Bonferroni multiple-testing correction at a *p*-value of ≤0.05, FC>2 (Genespring12, Aligent Technologies). Each column showed transcript abundance in the WT, W1-1 and W1-7 in all the leaf sections. Accession numbers were indicated together with a brief description of the gene lists. The relative abundance of transcripts associated with protein kinases is illustrated according to the scale bar shown.

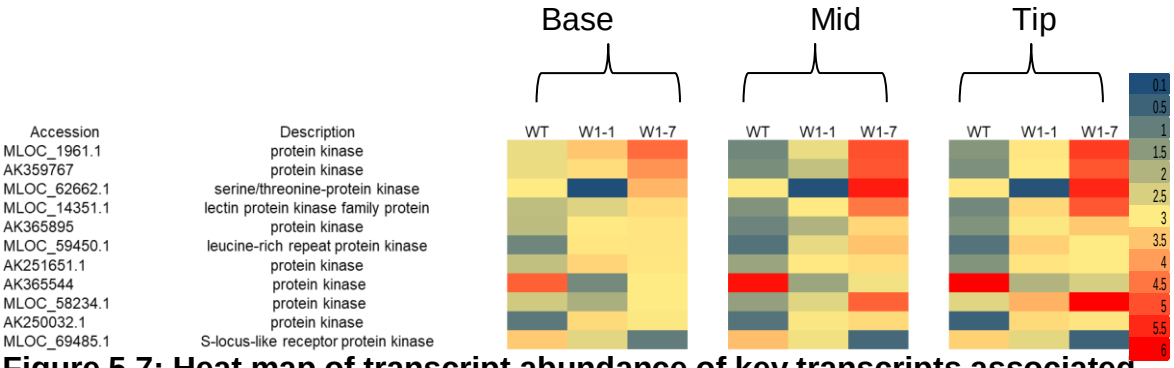


Figure 5.7: Heat map of transcript abundance of key transcripts associated with protein kinases at 14 days old.

The differentially expressed transcripts were analysed using two-way analysis of variance (ANOVA), with the genotypes (WT, W1-1, W1-7) as factors, and a Bonferroni multiple-testing correction at a *p*-value of ≤0.05, FC>2 (Genespring12, Aligent Technologies). Each column showed transcript abundance in the WT, W1-1 and W1-7 in all the leaf sections. Accession numbers were indicated together with a brief description of the gene lists. The relative abundance of transcripts associated with protein kinases is illustrated according to the scale bar shown.

5.2.1.5 Transcripts associated with redox processes and hormone metabolism

The abundance of most transcripts encoding thioredoxins and peroxidases (MLOC_68180.4, MLOC_46333.3, MLOC_49528.1 and AK365489) was higher in the base, middle and tip sections of the W1-7 leaves than the wild type and W1-1 leaves. A marked exception is transcript MLOC_21848.2 that had a lower abundance in all the sections of the W1-7 leaves, 7 days after germination (Figure 5.8). Several transcripts encoding peroxidases (MLOC_55062.1, AK375268, AK360063 and MLOC_65226.3) were higher in the base, middle and tip sections of the W1-7 leaves than the wild type and W1-1 14 days after germination (Figure 5.9). In addition, transcripts associated with phytohormones, such as ethylene, auxin, abscisic acid and gibberellins were changed in the WHY1-deficient lines relative to the wild type both at 7 and 14 day old (Appendix C.1).

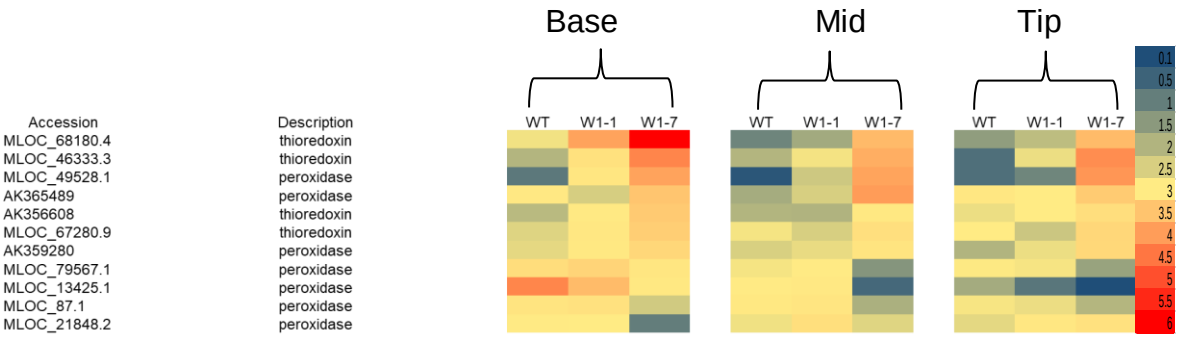


Figure 5.8: Heat map of transcript abundance of key transcripts associated with redox processes at 7 days old.

The differentially expressed transcripts were analysed using two-way analysis of variance (ANOVA), with the genotypes (WT, W1-1, W1-7) as factors, and a Bonferroni multiple-testing correction at a *p*-value of ≤0.05, FC>2 (Genespring12, Aligent Technologies). Each column showed transcript abundance in the WT, W1-1 and W1-7 in all the leaf sections. Accession numbers were indicated together with a brief description of the gene lists. The relative abundance of transcripts associated with redox processes is illustrated according to the scale bar shown.

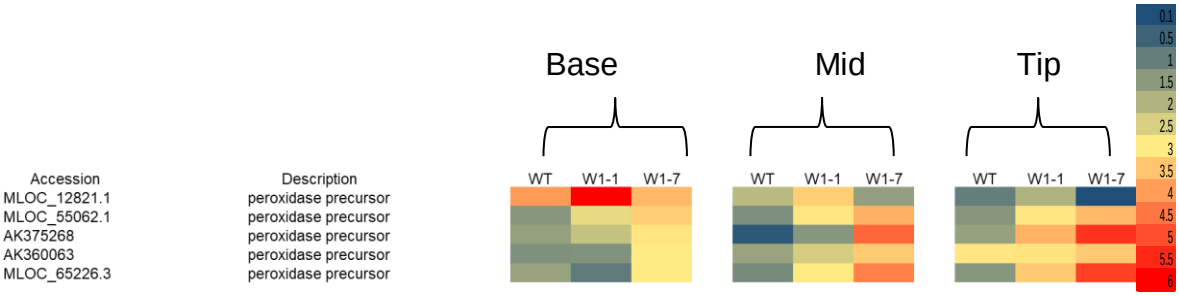


Figure 5.9: Heat map of transcript abundance of key transcripts associated with redox processes at 14 days old.

The differentially expressed transcripts were analysed using two-way analysis of variance (ANOVA), with the genotypes (WT, W1-1, W1-7) as factors, and a Bonferroni multiple-testing correction at a *p*-value of ≤0.05, FC>2 (Genespring12, Aligent Technologies). Each column showed transcript abundance in the WT, W1-1 and W1-7 in all the leaf sections. Accession numbers were indicated together with a brief description of the gene lists. The relative abundance of transcripts associated with redox processes is illustrated according to the scale bar shown.

5.2.1.6 Transcription factors

A transcript encoding the transcription factor AK365452 was higher in all of the leaf regions of the wild type than the other lines both at 7 and 14 days after germination (Figure 5.10,Figure 5.11). A transcript encoding the WRKY family of transcription factor, AK370043, was lower in the base sections of the W1-7 leaves compared to the wild type and W1-1 at 7 days (Figure 5.10). In contrast, a transcript encoding the WRKY family of transcription factor MLOC_62725.1 was higher in the base sections of the W1-7 leaves compared to the wild type and W1-1 at 7 days (Figure 5.10).

The abundance of transcripts encoding MYB transcription factors was similar in all the leaf sections in all genotypes at 7 days after germination (Figure 5.10). In contrast, the levels of transcripts encoding the ethylene-responsive transcription factor (MLOC_71804.1), and the RNA polymerase sigma factor (MLOC_59299.1), were higher in all sections of the W1-7 leaves compared to the wild type and W1-1, 7 days after germination (Figure 5.10). While the abundance of transcripts encoding the MYB family transcription factor (MLOC_37929.1) and WRKY47 (AK371133) was similar in the base sections of the leaves of the genotypes, both transcripts were significantly higher in the middle sections of the W1-7 leaves compared to W1-1 and the wild type 14 days after germination (Figure 5.11). In addition, the tip sections of the WHY1-deficient barley leaves had higher levels of these transcripts compared to the wild type at 14 days (Figure 5.11).

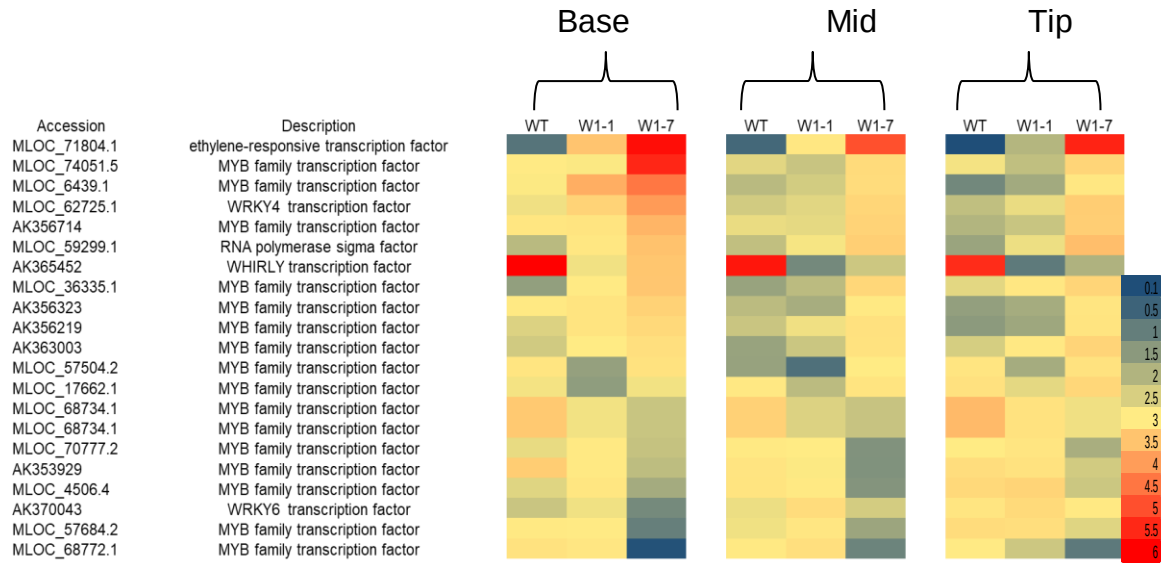


Figure 5.10: Heat map of transcript abundance of key transcripts associated with transcription factors at 7 days old.

The differentially expressed transcripts were analysed using two-way analysis of variance (ANOVA), with the genotypes (WT, W1-1, W1-7) as factors, and a Bonferroni multiple-testing correction at a *p*-value of ≤0.05, FC>2 (Genespring12, Aligent Technologies). Each column showed transcript abundance in the WT, W1-1 and W1-7 in all the leaf sections. Accession numbers were indicated together with a brief description of the gene lists. The relative abundance of transcripts associated with transcription factors is illustrated according to the scale bar shown.

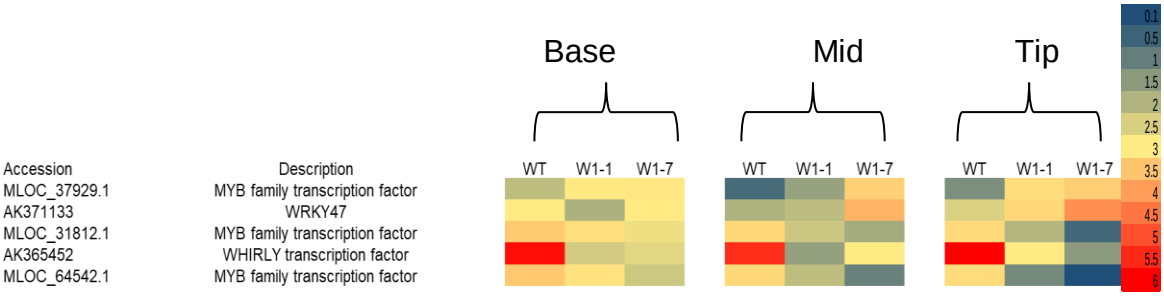


Figure 5.11: Heat map of transcript abundance of key transcripts associated with transcription factors at 14 days old.

The differentially expressed transcripts were analysed using two-way analysis of variance (ANOVA), with the genotypes (WT, W1-1, W1-7) as factors, and a Bonferroni multiple-testing correction at a *p*-value of ≤0.05, FC>2 (Genespring12, Aligent Technologies). Each column showed transcript abundance in the WT, W1-1 and W1-7 in all the leaf sections. Accession numbers were indicated together with a brief description of the gene lists. The relative abundance of transcripts associated with transcription factors is illustrated according to the scale bar shown.

5.2.2 Transcripts changes in the WHY1-deficient barley dependent on leaf region

5.2.2.1 Ribosomal related transcripts

The transcripts encoding ribosomal proteins encoded by the plastome were increased in the base, middle and tip sections of the W1-7 leaves compared to W-1 and the wild type at 7 days after germination (Figure 5.12 A). The levels of several nuclear-encoded ribosomal protein transcripts were higher in the base sections of the W1-7 leaves compared to the W1-1 and wild type at 7 days after germination (Figure 5.12 B).

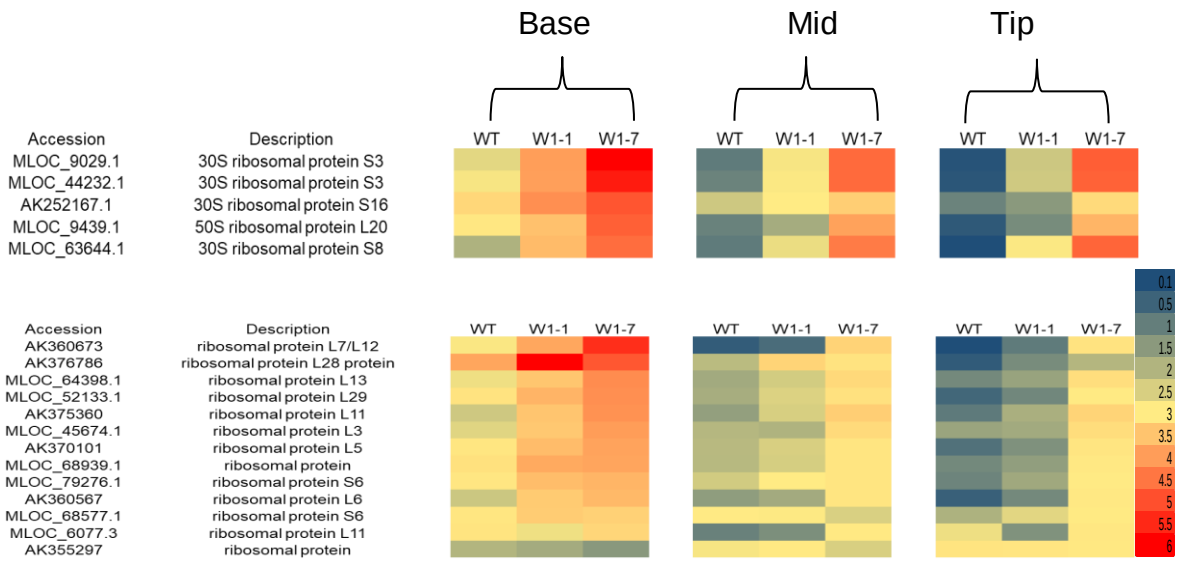


Figure 5.12: Heat map of transcript abundance of key transcripts associated with ribosomal protein encoded by the (A) chloroplasts and (B) nuclei at 7 days old.

The differentially expressed transcripts were analysed using two-way analysis of variance (ANOVA), with the genotypes (WT, W1-1, W1-7) as factors, and a Bonferroni multiple-testing correction at a *p*-value of ≤0.05, FC>2 (Genespring12, Aligent Technologies). Each column showed transcript abundance in the WT, W1-1 and W1-7 in all the leaf sections. Accession numbers were indicated together with a brief description of the gene lists. The relative abundance of transcripts associated with ribosomal proteins is illustrated according to the scale bar shown.

5.2.2.2 Transcripts associated with photosynthesis

The abundance of several transcripts encoding chlorophyll A-B binding proteins was increased in the base sections of the W1-1 and W1-7 leaves compared to the wild type 14 days after germination (Figure 5.13). The levels of these transcripts were also higher in the middle of the WHY1-deficient lines compared to the wild type 14 days after germination. Interestingly, the levels of transcripts encoding the chlorophyll A-B binding proteins were similar in the tips of all leaves of all genotypes 14 days after germination (Figure 5.13).

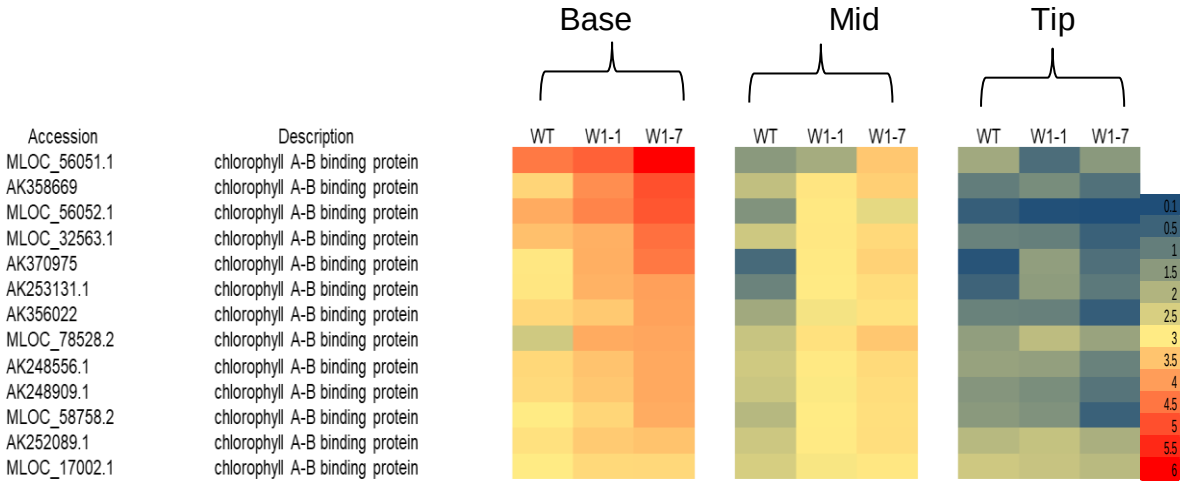


Figure 5.13: Heat map of transcript abundance of key transcripts encoding light-harvesting chlorophyll A-B binding proteins at 14 days old.

The differentially expressed transcripts were analysed using two-way analysis of variance (ANOVA), with the genotypes (WT, W1-1, W1-7) as factors, and a Bonferroni multiple-testing correction at a *p*-value of ≤0.05, FC>2 (Genespring12, Aligent Technologies). Each column showed transcript abundance in the WT, W1-1 and W1-7 in all the leaf sections. Accession numbers were indicated together with a brief description of the gene lists. The relative abundance of transcripts associated with light-harvesting chlorophyll A-B is illustrated according to the scale bar shown.

5.2.2.3 Transcripts associated with RNA metabolism

The levels of transcripts encoding PPR proteins were higher in the base, middle and tip sections of the W1-7 leaves compared to W1-1 and the wild type 7 days after germination (Figure 5.14). The level of only one transcript encoding a PPR proteins (MLOC_75882.1) was similar in the base, middle and tip sections of the leaves of all genotypes (Figure 5.14). However, no differences in the levels of transcripts encoding PPR were found at 14 days.

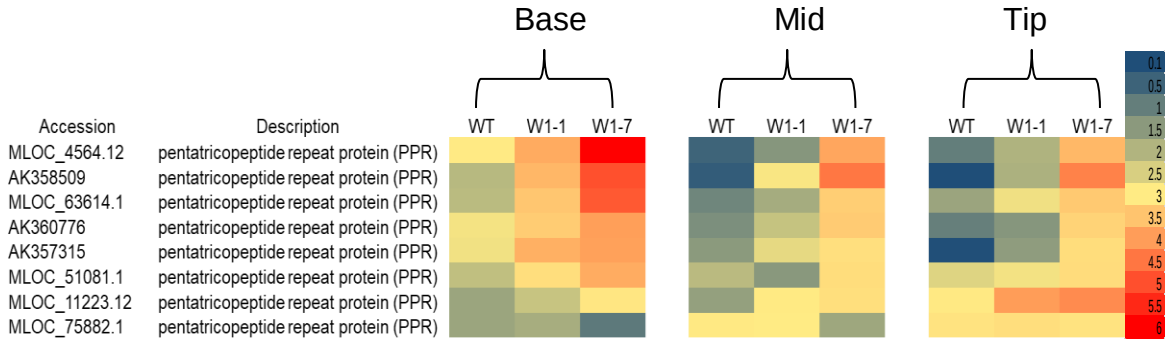


Figure 5.14: Heat map of transcript abundance of key transcripts associated with RNA metabolism, such as PPR, at 7 days old.

The differentially expressed transcripts were analysed using two-way analysis of variance (ANOVA), with the genotypes (WT, W1-1, W1-7) as factors, and a Bonferroni multiple-testing correction at a *p*-value of ≤0.05, FC>2 (Genespring12, Aligent Technologies). Each column showed transcript abundance in the WT, W1-1 and W1-7 in all the leaf sections. Accession numbers were indicated together with a brief description of the gene lists. The relative abundance of transcripts associated with RNA metabolism is illustrated according to the scale bar shown.

5.2.2.4 Protein kinases associated transcripts

The levels of several transcripts related to protein kinases were higher in the base sections of all genotypes 7 days after germination (Figure 5.15). These transcripts may be important during leaf early development. In contrast, no differences in the levels of transcripts encoding protein kinases were found in the middle and tip sections of the leaves of all genotypes 7 days after germination (Figure 5.15). Two transcripts (AK355124 and AK248260.1) had similar levels in the base, middle and tip sections of the leaves of all genotypes 7 days after germination (Figure 5.15).

The abundance of several transcripts encoding protein kinases were higher in the base sections of the leaves of all genotypes 14 days after germination (Figure 5.16). In contrast, there were two transcripts encoding protein kinases (AK355413) and CAMK kinase (AK372880) that were significantly less abundant in the base sections of the leaves of all genotypes14 days after germination (Figure 5.16). The levels of transcripts encoding protein kinases were similar in the middle and tip sections of the leaves of all genotypes 14 days after germination, with the exception of two transcripts (AK355413 and AK372880) (Figure 5.16).

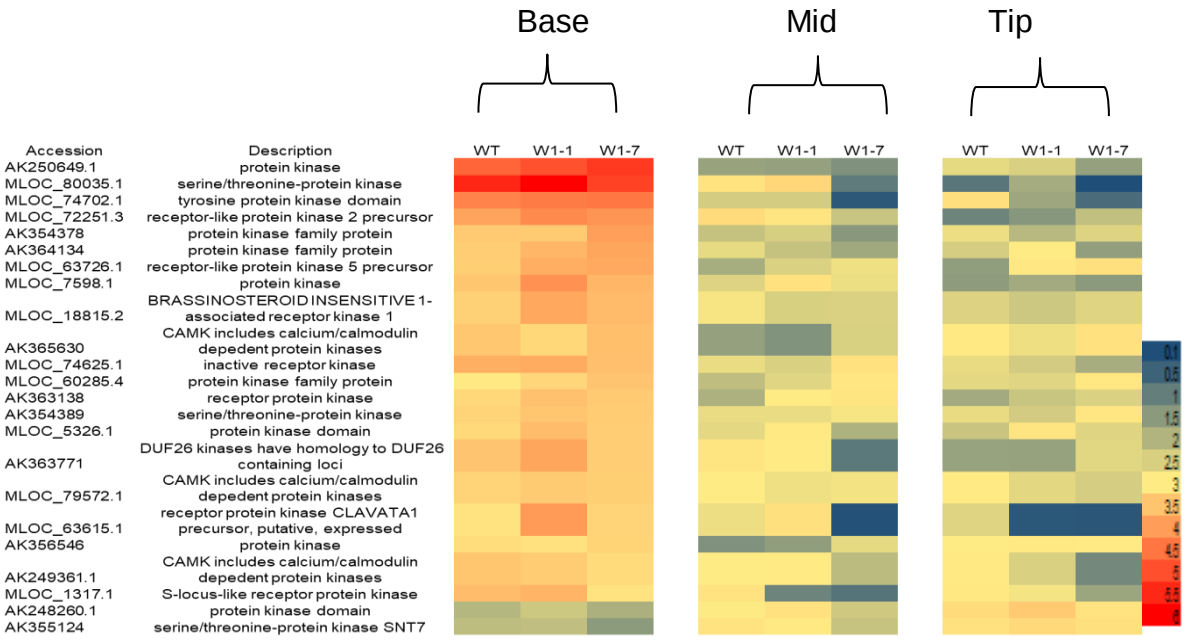


Figure 5.15: Heat map of transcript abundance of key transcripts associated with protein kinases at 7 days old.

The differentially expressed transcripts were analysed using two-way analysis of variance (ANOVA), with the genotypes (WT, W1-1, W1-7) as factors, and a Bonferroni multiple-testing correction at a *p*-value of ≤0.05, FC>2 (Genespring12, Aligent Technologies). Each column showed transcript abundance in the WT, W1-1 and W1-7 in all the leaf sections. Accession numbers were indicated together with a brief description of the gene lists. The relative abundance of transcripts associated with protein kinases is illustrated according to the scale bar shown.

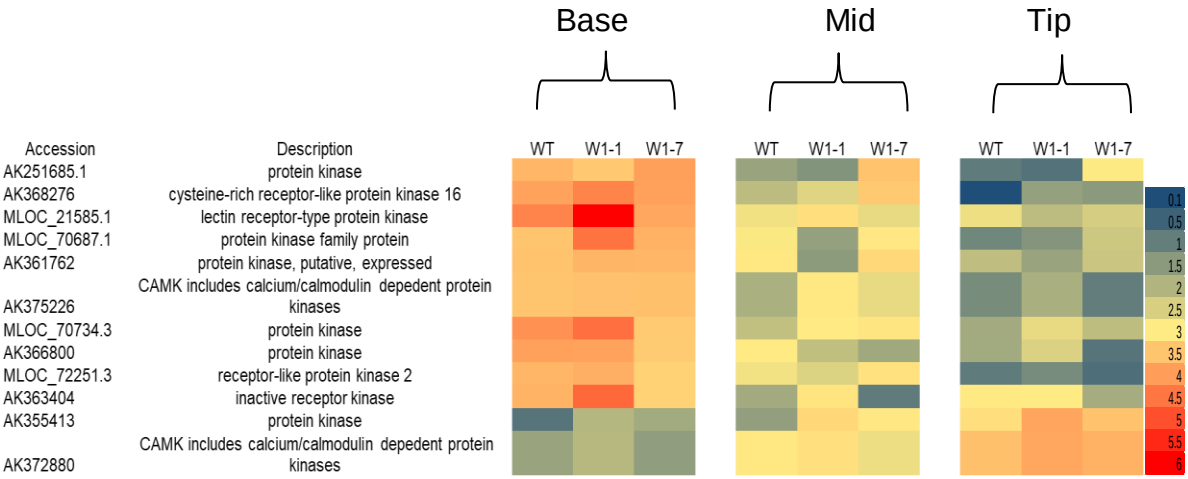


Figure 5.16: Heat map of transcript abundance of key transcripts associated with protein kinases at 14 days old.

The differentially expressed transcripts were analysed using two-way analysis of variance (ANOVA), with the genotypes (WT, W1-1, W1-7) as factors, and a Bonferroni multiple-testing correction at a *p*-value of ≤0.05, FC>2 (Genespring12, Aligent Technologies). Each column showed transcript abundance in the WT, W1-1 and W1-7 in all the leaf sections. Accession numbers were indicated together with a brief description of the gene lists. The relative abundance of transcripts associated with protein kinases is illustrated according to the scale bar shown.

5.2.2.5 Transcripts associated with redox processes and hormone metabolism

The abundances of several transcripts encoding proteins involved in redox processes, such as peroxidases, thioredoxins, glutaredoxins and ascorbate-related enzymes were significantly higher in the base sections of the leaves of all genotypes 7 days after germination (Figure 5.17). However, the levels of these transcripts were significantly lower in the middle and tip sections of the leaves of all genotypes 7 days after germination (Figure 5.17). The levels of transcripts encoding proteins associated with redox processes were also significantly changed 14 days after germination (Figure 5.18).

The levels of the transcripts encoding hormones, such as gibberellin, auxin and ethylene, were higher in the base sections of the leaves of all genotypes 7 days after germination (Appendix C.2). In contrast, the levels of these transcripts was significantly lower in the middle and tip sections of the leaves of all genotypes at this stage. The levels of the transcripts encoding hormones, such as gibberellin and auxin, were also significantly changed 14 days after germination (Appendix C.2.1).

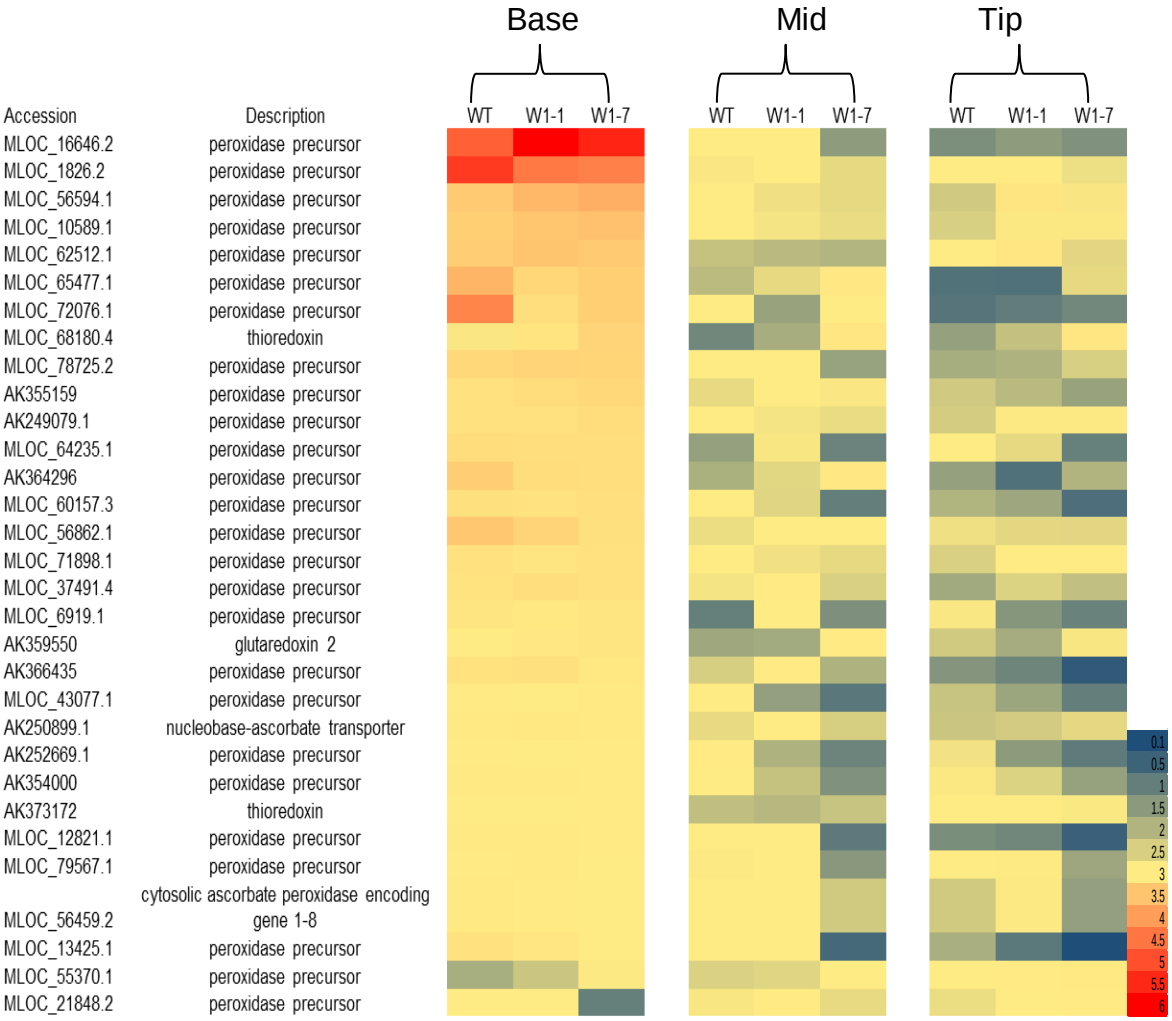


Figure 5.17: Heat map of transcript abundance of key transcripts associated with redox processes at 7 days old.

The differentially expressed transcripts were analysed using two-way analysis of variance (ANOVA), with the genotypes (WT, W1-1, W1-7) as factors, and a Bonferroni multiple-testing correction at a *p*-value of ≤0.05, FC>2 (Genespring12, Aligent Technologies). Each column showed transcript abundance in the WT, W1-1 and W1-7 in all the leaf sections. Accession numbers were indicated together with a brief description of the gene lists. The relative abundance of transcripts associated with redox processes is illustrated according to the scale bar shown.

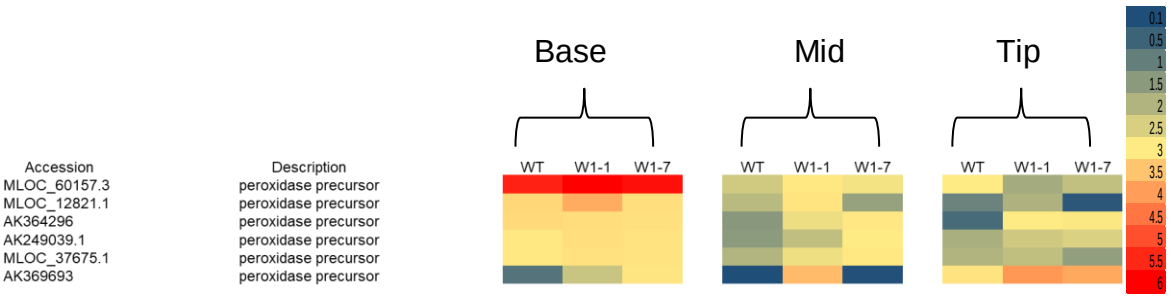


Figure 5.18: Heat map of transcript abundance of key transcripts associated with redox processes at 14 days old.

The differentially expressed transcripts were analysed using two-way analysis of variance (ANOVA), with the genotypes (WT, W1-1, W1-7) as factors, and a Bonferroni multiple-testing correction at a *p*-value of ≤0.05, FC>2 (Genespring12, Aligent Technologies). Each column showed transcript abundance in the WT, W1-1 and W1-7 in all the leaf sections. Accession numbers were indicated together with a brief description of the gene lists. The relative abundance of transcripts associated with redox processes is illustrated according to the scale bar shown.

5.2.2.6 Transcription factors

The abundance of several transcripts encoding basic helix-loop-helix (BHLH) transcription factors was higher in the base sections of the leaves of all genotypes 7 days after germination (Figure 5.19). Moreover, transcripts encoding a MYB family transcription factor were significantly increased in the base sections of the leaves of all genotypes 7 days after germination (Figure 5.19). However, the levels of these transcripts were lower in the middle and tip sections of the leaves of all genotypes than the basal sections.

The abundance of the WRKY39 transcription factor was higher in the base sections of the leaves of all the genotypes 7 days after germination. However, the abundance of this transcript fell to very low levels in the middle and tip sections of the leaves of all genotypes. In contrast to other WRKY transcription factors, whose levels were higher in the base sections of the leaves of all the genotypes, the abundance of the WRKY6 transcription factor transcripts was similar in the base, middle and tip sections of the leaves of all genotypes 7 days after germination. (Figure 5.19).

The levels of the transcripts encoding transcription factors such as bZIP, MYB, BHLH and WRKY were higher in the base sections of the leaves of all genotypes 14 days after germination (Figure 5.20). The levels of transcripts encoding these transcription factors were similar in the middle and tip sections of the leaves of all genotypes 14 days after germination (Figure 5.20).

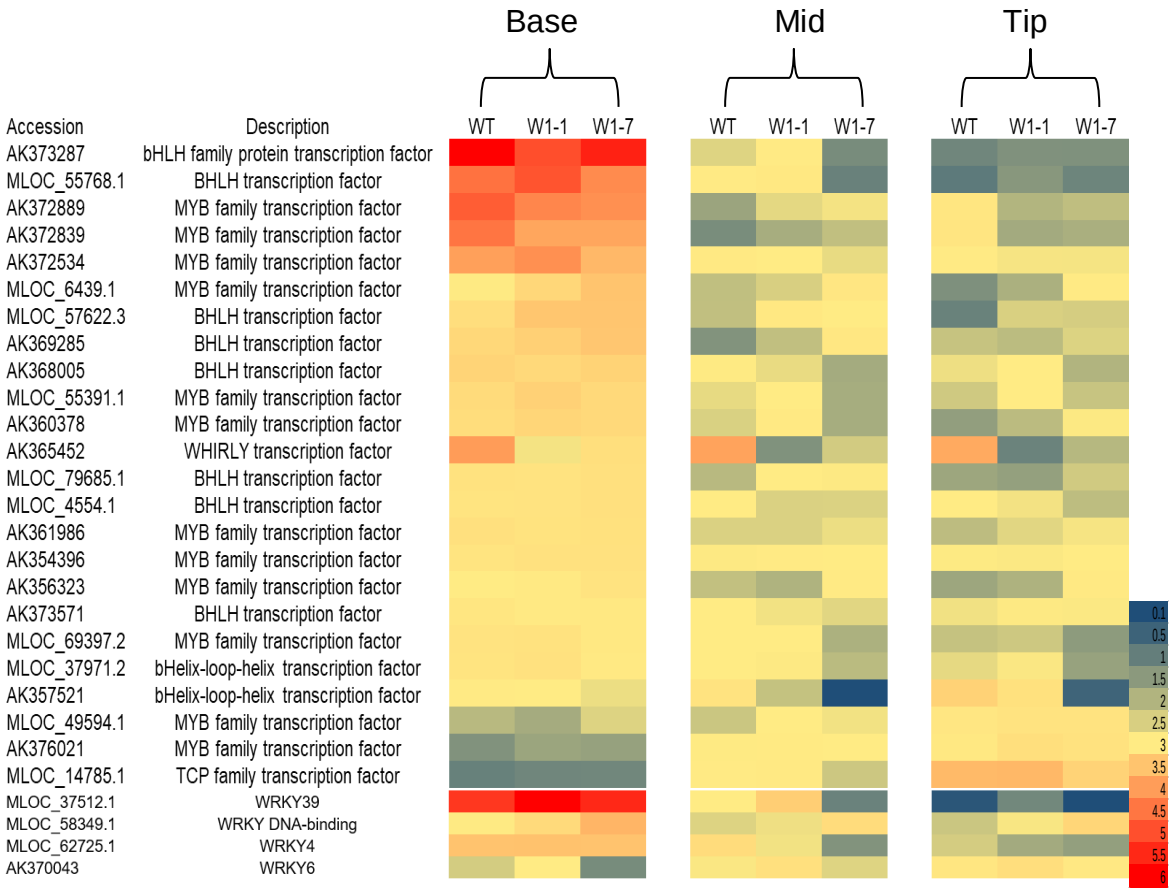


Figure 5.19: Heat map of transcript abundance of key transcripts associated with transcription factors at 7 days old.

The differentially expressed transcripts were analysed using two-way analysis of variance (ANOVA), with the genotypes (WT, W1-1, W1-7) as factors, and a Bonferroni multiple-testing correction at a *p*-value of ≤0.05, FC>2 (Genespring12, Aligent Technologies). Each column showed transcript abundance in the WT, W1-1 and W1-7 in all the leaf sections. Accession numbers were indicated together with a brief description of the gene lists. The relative abundance of transcripts associated with transcription factors is illustrated according to the scale bar shown.

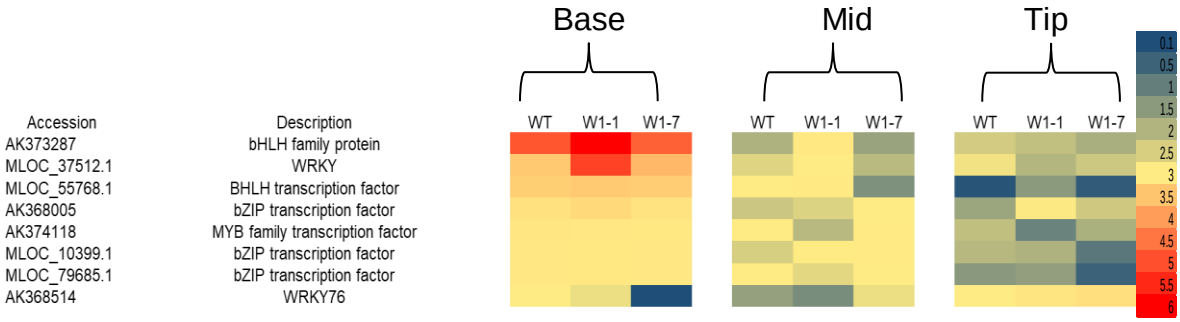


Figure 5.20: Heat map of transcript abundance of key transcripts associated with transcription factors at 14 days old.

The differentially expressed transcripts were analysed using two-way analysis of variance (ANOVA), with the genotypes (WT, W1-1, W1-7) as factors, and a Bonferroni multiple-testing correction at a *p*-value of ≤0.05, FC>2 (Genespring12, Aligent Technologies). Each column showed transcript abundance in the WT, W1-1 and W1-7 in all the leaf sections. Accession numbers were indicated together with a brief description of the gene lists. The relative abundance of transcripts associated with transcription factors is illustrated according to the scale bar shown.

5.2.3 Transcript changes in WHY1-deficient barley leaves dependent on genotype and leaf region

The levels of several transcripts were changed in the WHY1-deficient barley seedlings, dependent on genotype and leaf region, 7 days after germination (Figure 5.21). Transcripts encoding valyl-tRNA synthetase were highest in the base sections of the W1-7 leaves relative to W1-1 and the wild type Figure 5.21. However, the levels of valyl-tRNA synthetase transcripts were similar to those in the middle and tip sections of the leaves of the wild type and WHY1-deficient barley (Figure 5.21). Analysis of *wp1* mutants of *Oryza sativa* showed that the valyl-tRNA synthetase (OsValRS2) protein was targeted to the mitochondria and chloroplasts (Wang et al., 2016). This *wp1* mutants had pale and albino phenotypes at the seedling stage with impaired chloroplast ribosomes biogenesis in the leaves, suggesting that OsValRS2 plays a role in chloroplast development and ribosome biogenesis (Wang et al., 2016). The levels of transcripts encoding DNA-directed RNA polymerase subunit alpha were higher in the base, middle and tip sections of the W1-7 leaves 7 days after germination (Figure 5.21). The abundance of a transcript encoding PPR (MLOC_67650.1) was higher in the base sections of the W1-7 leaves compared to these sections in the W1-1 and wild type leaves (Figure 5.21). However, the levels of this transcript were similar in the middle and tip sections of the leaves of all genotypes (Figure 5.21). The levels of transcripts encoding a basic helix-loop-helix (AK357521) protein were significantly lower in the middle and tip sections of the W1-7 leaves compared to W1-1 and the wild type (Figure 5.21).

The abundance of a redox-associated transcript (MLOC_21848.2) was lower in the base sections of the W1-7 leaves relative to W1-1 and the wild type (Figure 5.21). The levels of transcripts encoding this redox protein were similar in the middle and tip sections of the leaves of all the genotypes (Figure 5.21). A number of transcripts associated with photosynthesis were significantly lower in the base sections of the W1-7 leaves compared to W1-1 and wild type (Figure 5.21). The levels of a transcript encoding a chlorophyll A-B binding protein were similar in

the middle sections of the leaves of all the genotypes (Figure 5.21). However, the levels of this transcript were higher in the tip sections of the W1-7 leaves compared to W1-1 and wild type (Figure 5.21).

Transcripts encoding a late embryogenesis abundant protein, related to the group 3 classification of these proteins, were changed in the WHY1-deficient barley leaves, dependent on both genotype and leaf region, at 14 days after germination.

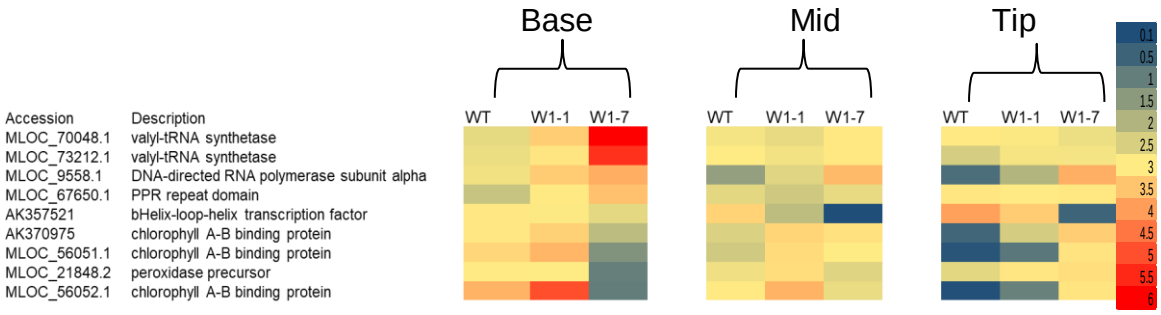


Figure 5.21: Heat map of transcript abundance of key transcripts changes in the WHY1-deficient barley, dependent on genotype and leaf region at 7 days old.

The differentially expressed transcripts were analysed using two-way analysis of variance (ANOVA), with the genotypes (WT, W1-1, W1-7) as factors, and a Bonferroni multiple-testing correction at a *p*-value of ≤0.05, FC>2 (Genespring12, Aligent Technologies). Each column showed transcript abundance in the WT, W1-1 and W1-7 in all the leaf sections. Accession numbers were indicated together with a brief description of the gene lists. The relative abundance of transcripts is illustrated according to the scale bar shown.

5.2.4 Differentially expressed transcripts in the wild type and WHY1-deficient barley

A one-way ANOVA was used to analyse the differentially expressed genes in each of the genotypes, independently, across leaf development to understand better the differences in the leaf transcript profile between the WHY-deficient line and the wild type at 7 days after germination. In total, only 441 transcripts changed in the wild type (Figure 5.22). However, many more transcripts were changed in abundance with respect to the stage of leaf development i.e. the section of the leaf in the W1-1 and W1-7 mutants compared to the wild type (Figure 5.23).

The total number of transcripts changed in W1-1 were 5555, with 6809 transcripts changed in W1-7 (Figure 5.22) relative to the wild type. Of these, 17 transcripts were unique to the wild type, while 36 were shared with the W1-1 barley leaves (Figure 5.22). A total of 1984 transcripts changed in W1-1, while 3166 were shared with W1-7 (Figure 5.22). A total of 3255 transcripts were changed exclusively in W1-7, with only 19 transcripts shared with wild type (Figure 5.22). In total, 369 transcripts were common to the wild type, W1-1 and W1-7 leaves (Figure 5.22).

Transcriptome profile patterns that were characteristic of the base, middle and tip sections of the wild type leaves were identified at 7 days after germination (Figure 5.23 A). In total, 441 transcripts were changed across the base, middle and tip sections of the wild type leaves (Figure 5.23 A). In contrast, 5555 transcripts were changed in the base, middle and tip sections of the W1-1 leaves, at 7 days after germination (Figure 5.23 B) while 6809 transcripts were changed across these sections in the W1-7 leaves (Figure 5.23 C). Highly distinctive patterns of transcripts were identified in the cluster analysis comparisons of the wild type, W1-1 and W1-7 leaves (Figure 5.23). However, the developmental patterns of

the transcript profiles in the W1-1 and W1-7 leaves were similar relative to the wild type (Figure 5.23).

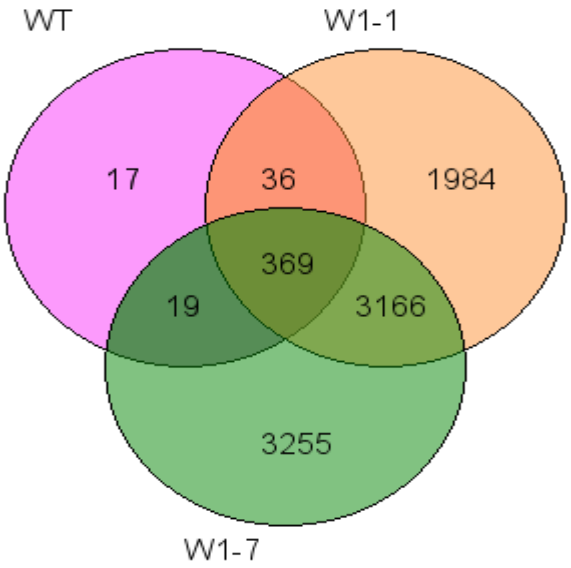


Figure 5.22: Transcript profile comparison of wild type, W1-1 and W1-7 barley leaves during leaf development, 7 days after germination.

Venn diagram illustrating the total number of differentially expressed transcripts in the wild type, W1-1 and W1-7, at 7 days after germination. The differentially expressed genes across the leaf development were analysed independently under each genotype, using one-way ANOVA. Significant differences were analysed using a moderated Student's *t*-test, with a Benjamini–Hochberg multiple testing correction ($p<0.05$; fold change >2; Genespring 12, Agilent Technologies).

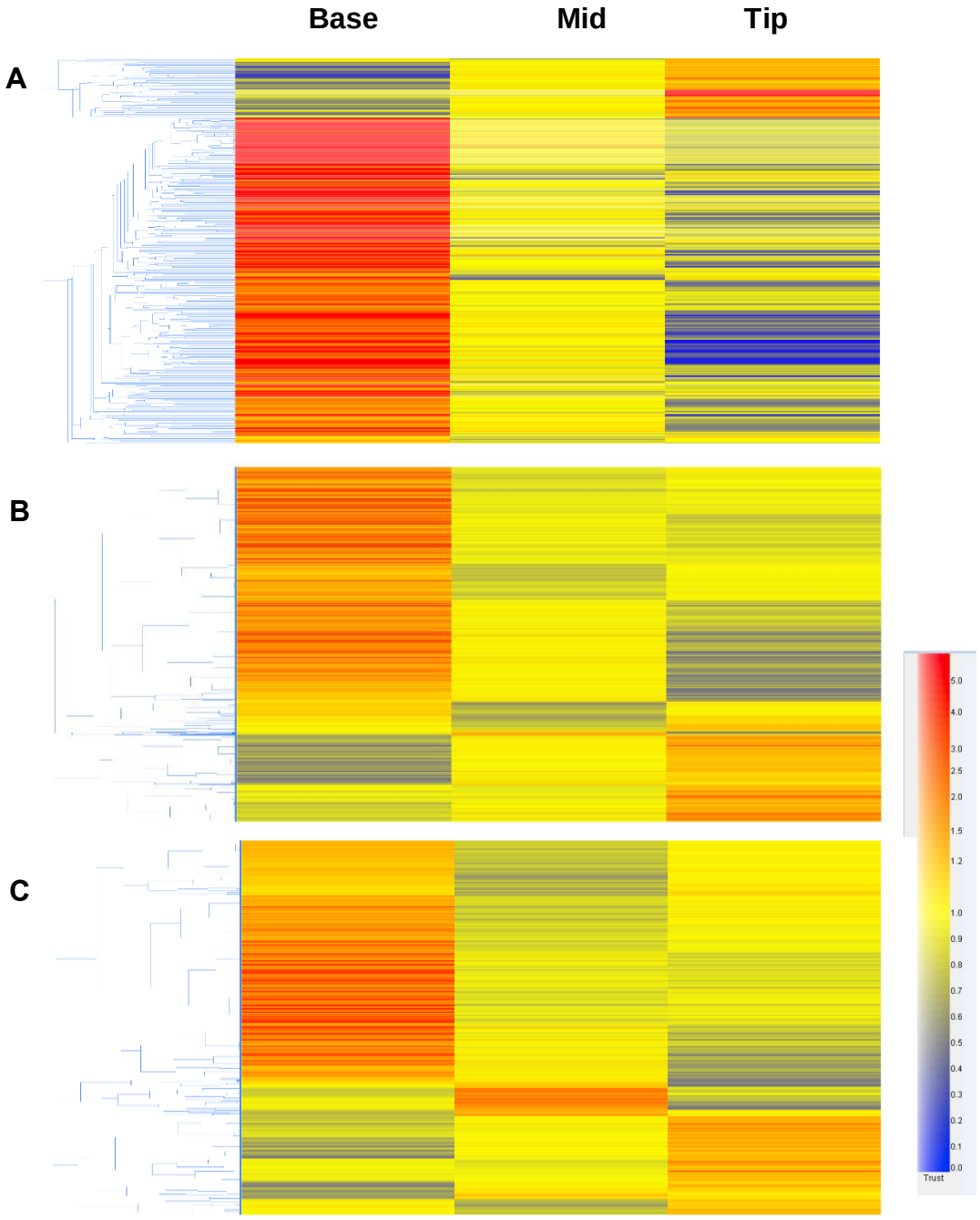


Figure 5.23: Comparison of the hierarchical clustering of differentially expressed transcripts in the (A) wild type, (B) W1-1 and (C) W1-7, in 7-day-old seedlings.

The differentially expressed genes in the base, middle and tip sections of the first leaves were analysed independently under each genotype, using one-way ANOVA. Significant differences were analysed using a moderated Student's *t*-test, with a Benjamini–Hochberg multiple testing correction ($p<0.05$; fold change >2; Genespring 12, Agilent Technologies). Each column showed transcript abundance in the WT, W1-1 and W1-7 in all the leaf sections. The relative abundance of transcipts is illustrated according to the scale bar shown.

5.2.5 An overview of transcript changes in the wild type, W1-1 and W1-7 independently

Developmental changes in the transcript profiles of the wild type, W1-1 and W1-7 leaves were analysed using a one-way ANOVA. The levels of cell-wall-associated transcripts were highest in the base sections of the wild type, W1-1 and W1-7 leaves, as shown by over-representation analysis (ORA) and Wilcoxon analysis (Appendix D, Appendix E, Appendix F, Appendix G and Appendix H). In contrast, the levels of cell-wall-associated transcripts were lowest in the tip sections of the wild type, W1-1 and W1-7 leaves, as illustrated by ORA and Wilcoxon analysis (Appendix D, Appendix E, Appendix F, Appendix G and Appendix H).

5.2.6 Transcript changes in the W1-7 lines relative to the wild type

Following the above analysis, particular focus was placed on characterising the changes in transcript abundance in the W1-7 leaves relative to the wild type because the differences in this line were much greater than those detected in W1-1 leaves. Transcript changes in the base, middle and tip sections of W1-7 leaves relative to the wild type were analysed independently using a volcano plot (Figure 5.24 and Figure 5.25).

5.2.6.1 Number of differentially expressed transcripts in W1-7 relative to the wild type

In total, 325 the levels of transcripts were changed in the base sections of the W1-7 leaves relative to the wild type (Figure 5.24). Moreover, the levels of 3189 transcripts were changed in the middle sections of the W1-7 leaves relative to the wild type, (Figure 5.24) with only 245 transcripts changed in abundance in the tip sections of the W1-7 leaves relative to the wild type (Figure 5.24). Many transcripts were expressed in a developmental manner in the base, middle and tip sections of the leaves of W1-7 relative to the wild type (Figure 5.25). In contrast, there were no differences in the base, middle and tip sections in W1-1 leaves relative to the wild type, using a volcano plot (data in Appendix I).

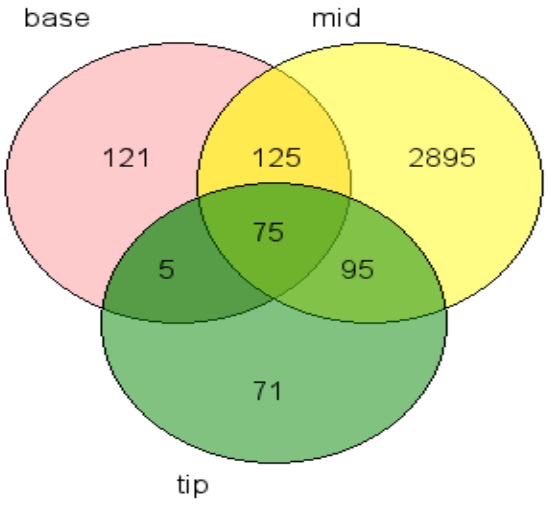


Figure 5.24: Transcript profile comparison in the W1-7 lines relative to the wild-type barley leaves during leaf development, 7 days after germination.

Venn diagram illustrating the total number of differentially abundant transcripts that changed in the base, middle and tip sections of W1-7 leaves relative to the wild type at 7 days old. Pairwise comparisons between the wild type and W1-7, for each leaf position (base, middle and tip), were performed using volcano plots, with a Benjamini–Hochberg multiple testing correction (p<0.05; fold change >2; Genespring 12, Agilent Technologies).

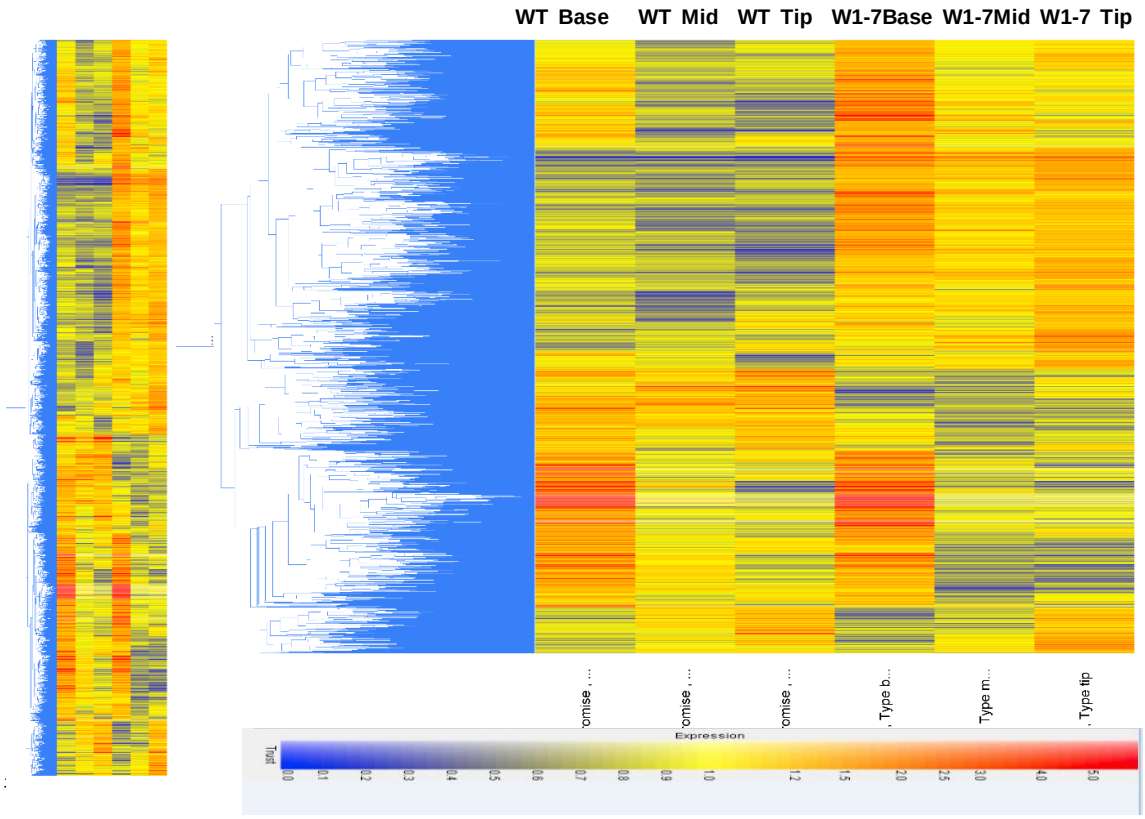


Figure 5.25: Comparison of hierarchical clustering of differentially expressed transcripts in W1-7 relative to the wild type in 7-day-old seedlings.

The differentially expressed genes were analysed independently for each genotype, using one-way ANOVA. Significant differences were analysed using a moderated Student's *t*-test, with a Benjamini–Hochberg multiple testing correction (p<0.05; fold change >2; Genespring 12, Agilent Technologies). Each column showed transcript abundance in the WT, W1-1 and W1-7 in all the leaf sections. The relative abundance of transcripts is illustrated according to the scale bar shown.

5.2.6.2 Transcripts associated with chloroplast ribosomal proteins

Several of the transcripts associated with chloroplast ribosomal proteins were increased in abundance in the base sections of the W1-7 leaves compared to the wild type (Figure 5.26). The levels of transcripts encoding chloroplast ribosomal proteins were decreased in the tip sections of the wild type leaves compared to W1-7 leaves (Figure 5.26). Clear differences in the profiles of chloroplast ribosomal proteins were observed in the W1-7 leaves relative to the wild type (Figure 5.26).

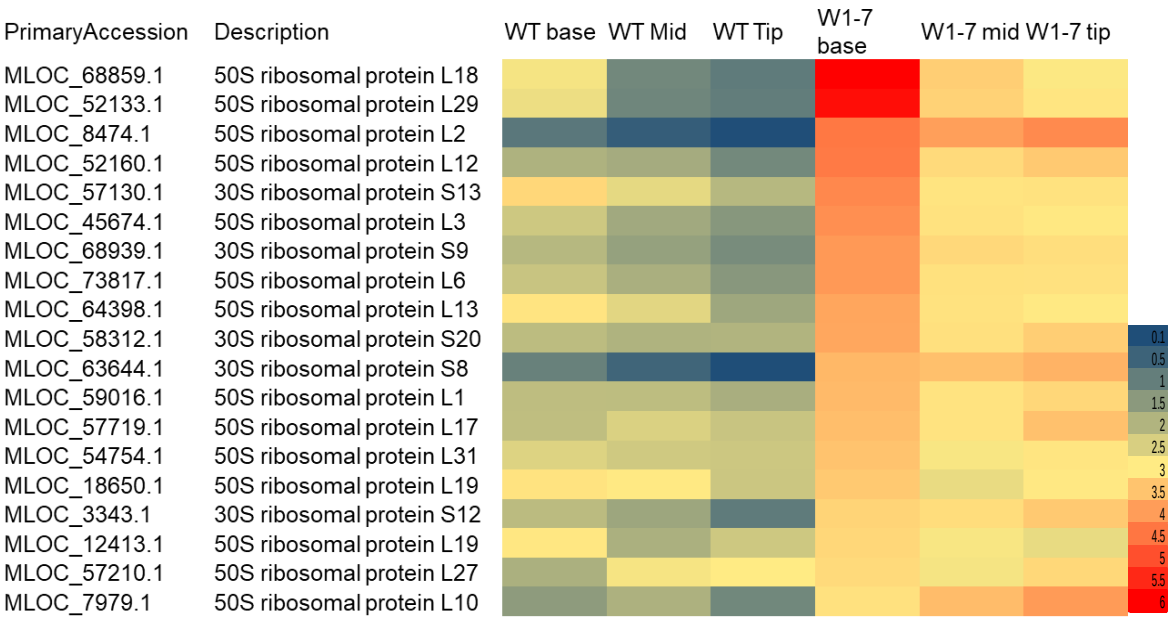


Figure 5.26: Comparison of heat maps of transcript abundance of key transcripts associated with the chloroplast ribosomal protein in W1-7 relative to the wild type in 7-day-old seedlings.

The differentially expressed transcripts were analysed using two-way analysis of variance (ANOVA), with the genotypes (WT, W1-1, W1-7) as factors, and a Bonferroni multiple-testing correction at a *p*-value of ≤0.05, FC>2 (Genespring12, Aligent Technologies). Each column showed transcript abundance in the WT, W1-1 and W1-7 in all the leaf sections. Accession numbers were indicated together with a brief description of the gene lists. The relative abundance of transcripts associated with chloroplast ribosomal proteins is illustrated according to the scale bar shown.

5.2.6.3 Transcripts associated with photosynthesis

The abundance of many plastome-encoded transcripts, such as those encoding NAD(P)H, oxidoreductases, constituents of PSI and PSII and the ATP synthase was decreased in the wild type relative to W1-7 leaves (Figure 5.27 A). Interestingly, these transcripts were significantly higher in all leaf sections of the W1-7 leaves compared to the wild type with one exception (MLOC_34266.1) (Figure 5.27 A). In contrast, many of the photosynthesis-associated transcripts encoded by nuclear genes were lower in abundance in the tip sections of the wild type leaves compared to the W1-7 leaves (Figure 5.27 B). The level of transcripts associated with chloroplast-targeted proteins that are encoded by the nucleus were also significantly higher in all the sections of the W1-7 leaves compared to the wild type with one exception (MLOC_37052.1) (Figure 5.27 B).

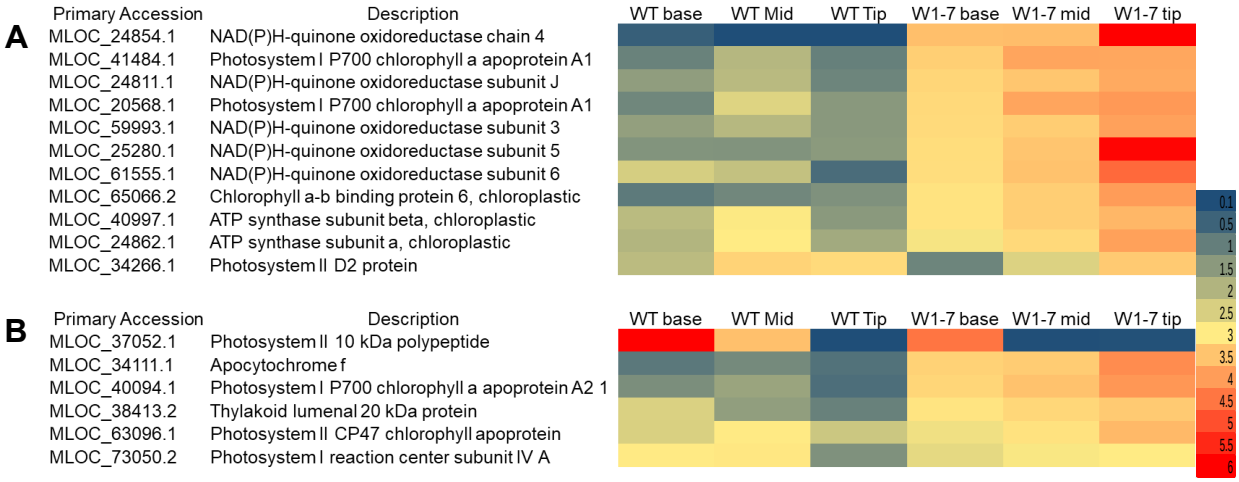


Figure 5.27: Comparison of heat maps of transcript abundance of key transcripts associated with photosynthesis that are (A) chloroplast encoded and (B) nucleus encoded in W1-7 relative to the wild type in 7-day-old seedlings.

The differentially expressed transcripts were analysed using two-way analysis of variance (ANOVA), with the genotypes (WT, W1-1, W1-7) as factors, and a Bonferroni multiple-testing correction at a *p*-value of ≤0.05, FC>2 (Genespring12, Aligent Technologies). Each column showed transcript abundance in the WT, W1-1 and W1-7 in all the leaf sections. Accession numbers were indicated together with a brief description of the gene lists. The relative abundance of transcripts associated with photosynthesis is illustrated according to the scale bar shown.

5.2.6.4 Transcripts associated with RNA processing

Transcripts encoding plastid transcriptionally active 6, which is involved in plastid gene expression, were significantly higher in the base sections of the W1-7 leaves than the wild type (Figure 5.28). The levels of DNA-directed RNA polymerase subunit beta were significantly higher in all sections of the W1-7 leaves compared to the wild type, the values being highest in the base sections of the W1-7 leaves (Figure 5.28).

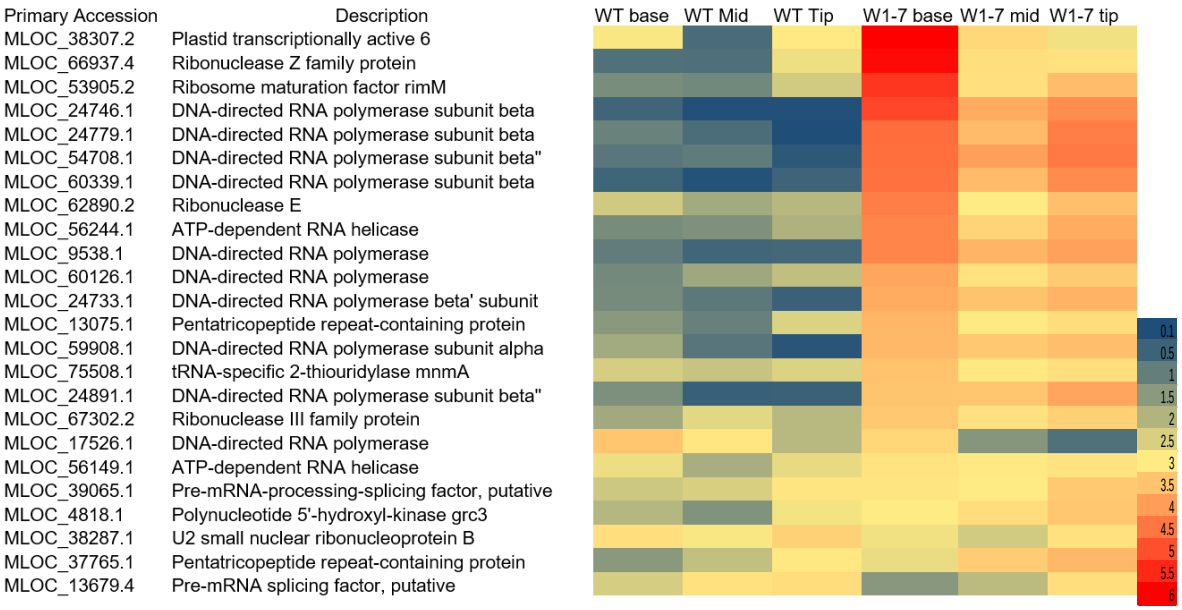


Figure 5.28: Comparison of heat maps of transcript abundance of key transcripts associated with RNA processing in W1-7 relative to the wild type in 7-day old seedlings.

The differentially expressed transcripts were analysed using two-way analysis of variance (ANOVA), with the genotypes (WT, W1-1, W1-7) as factors, and a Bonferroni multiple-testing correction at a *p*-value of ≤0.05, FC>2 (Genespring12, Aligent Technologies). Each column showed transcript abundance in the WT, W1-1 and W1-7 in all the leaf sections. Accession numbers were indicated together with a brief description of the gene lists. The relative abundance of transcripts associated with RNA processing is illustrated according to the scale bar shown.

5.2.6.5 Hormone-associated transcripts

Large numbers of transcripts encoding hormone-associated proteins were significantly more abundant in the basal sections of the wild type and W1-7 leaves compared to other leaf regions (Figure 5.29). In contrast to the levels of transcripts encoding hormone-associated proteins in the basal sections in the wild type and W1-7, the levels of transcripts encoding auxin-related proteins were lowest in the tip sections of the wild type and W1-7 leaves (Figure 5.29).

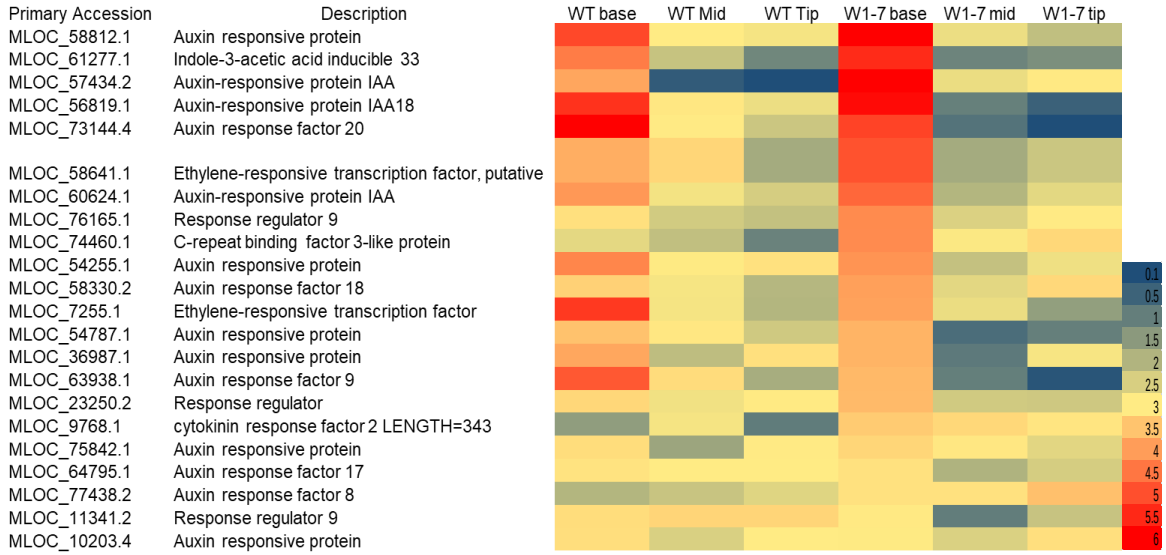


Figure 5.29: Comparison of heat maps of transcript abundance of key transcripts associated with hormones in W1-7 relative to the wild type in 7-day-old seedlings.

The differentially expressed transcripts were analysed using two-way analysis of variance (ANOVA), with the genotypes (WT, W1-1, W1-7) as factors, and a Bonferroni multiple-testing correction at a *p*-value of ≤0.05, FC>2 (Genespring12, Aligent Technologies). Each column showed transcript abundance in the WT, W1-1 and W1-7 in all the leaf sections. Accession numbers were indicated together with a brief description of the gene lists. The relative abundance of transcripts associated with hormones is illustrated according to the scale bar shown.

5.2.6.6 Light signalling and FAR-RED IMPAIRED RESPONSE 1 (FAR1)-associated transcripts

The levels of transcripts encoding a number of light signalling (MLOC_11221.1, MLOC_38189.2, MLOC_59573.1 and MLOC_56584.3) proteins were highest in the basal sections of the wild type and W1-7 leaves (Figure 5.30 A). The levels of these transcripts were significantly lower in the middle and tip sections of the W1-7 leaves than the wild type (Figure 5.30 A). Transcripts encoding an early light-induced protein (MLOC_78997.1) were significantly lower in all the leaf sections in the wild type compared to the W1-7 leaves (Figure 5.30 A). Interestingly, lower levels of light-associated transcripts (MLOC_65070.2 and MLOC_60978.1) were found in the basal sections of the wild type and W1-7 leaves than other sections (Figure 5.30 A). The levels of these transcripts was highest in the middle and tip sections of the leaves of the wild type and W1-7 leaves than the leaf bases (Figure 5.30 A). A transcript (MLOC_33258.3) encoding FAR1 was lower in all the sections of the wild type leaves than the W1-7 leaves (Figure 5.30 B). Similarly, a transcript (MLOC_53882.4) encoding FAR1 was also lower in all the leaf sections in the wild type than the W1-7 leaves (Figure 5.30 B).

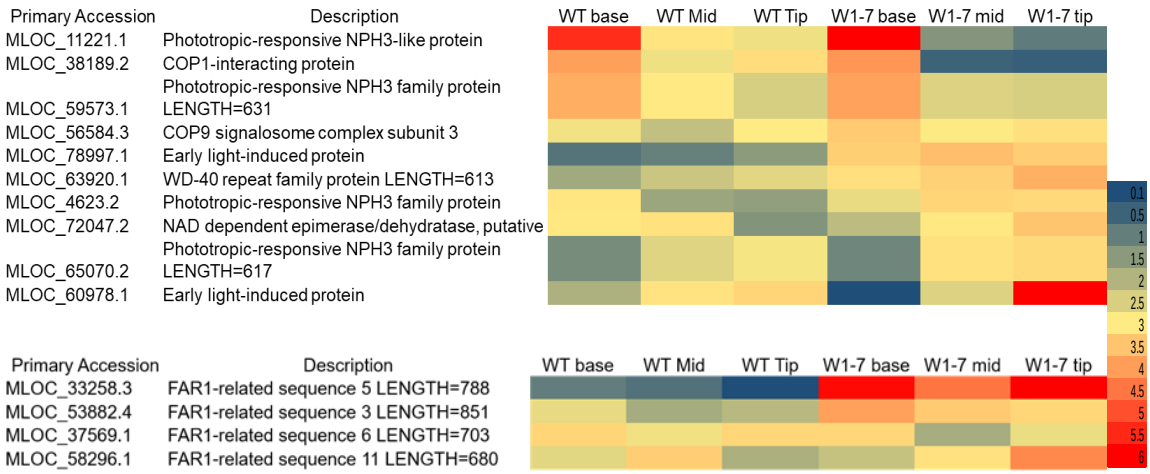


Figure 5.30: Comparison of heat maps of transcript abundance of key transcripts associated with a (A) light signalling and (B) a FAR1-like protein in W1-7 relative to the wild type in 7-day-old seedlings.

The differentially expressed transcripts were analysed using two-way analysis of variance (ANOVA), with the genotypes (WT, W1-1, W1-7) as factors, and a Bonferroni multiple-testing correction at a *p*-value of ≤0.05, FC>2 (Genespring12, Aligent Technologies). Each column showed transcript abundance in the WT, W1-1 and W1-7 in all the leaf sections. Accession numbers were indicated together with a brief description of the gene lists. The relative abundance of transcripts associated with light signalling and a FAR1-like protein is illustrated according to the scale bar shown.

5.2.6.7 Transcripts encoding transcription factors

A large number of transcripts encoding transcription factors were changed in abundance in the W1-7 leaves relative to the wild type. For example, transcripts encoding the bZIP transcription factor MLOC_51623.1, the RNA polymerase sigma factor MLOC_59299.1 and the WRKY transcription factor 21 (MLOC_44455.1) were significantly lower in the wild type than W1-7 leaves.

Figure 5.31: Comparison of heat maps of transcript abundance of key transcripts associated with transcription factors in W1-7 relative to the wild type in 7-day old seedlings.

The differentially expressed transcripts were analysed using two-way analysis of variance (ANOVA), with the genotypes (WT, W1-1, W1-7) as factors, and a Bonferroni multiple-testing correction at a *p*-value of ≤0.05, FC>2 (Genespring12, Aligent Technologies). Each column showed transcript abundance in the WT, W1-1 and W1-7 in all the leaf sections. Accession numbers were indicated together with a brief description of the gene lists. The relative abundance of transcripts associated with transcription factors is illustrated according to the scale bar shown.

5.3 Discussion

The studies reported in this chapter revealed significant differences in the transcript profiles of the developing leaves of 7 and 14 day old WHY1-deficient barley seedlings relative to the wild type. These differences were most pronounced in the W1-7 leaves compared to the wild type. This analysis provides new insights into the functions of WHY1 in chloroplast development.

The levels of transcripts encoding chloroplast 50S and 30S ribosomal subunits were higher in the basal sections of the W1-7 leaves than the wild type (Figure 5.1 & Figure 5.12). Moreover, the expression of the plastid-encoded ribosomal transcripts was greatly increased in the WHY1-deficient lines. This finding suggests that chloroplast ribosome biogenesis is significantly changed in the basal regions of W1-7 leaves compared to the wild type. These data are also in agreement with the results presented in Chapter 4 showing that the splicing of plastid rRNAs of the 23S and 4.5 rRNAs was aberrant in the WHY1-deficient lines. Taken together, these findings demonstrate that in the absence of WHY1, the plastid-encoded ribosomal transcripts are not effectively translated into ribosomal subunits. The developing WHY1-deficient leaves lack the chloroplast ribosomes required to translate plastid-encoded transcripts such as those encoding photosynthetic proteins. There is also a lack of coordination in the expression of ribosomal genes in the nucleus and chloroplasts.

Unlike the expression of plastome genes encoding photosynthetic proteins (see Chapter 4), the levels of transcripts associated with photosynthesis that are encoded by nuclear genes were generally higher in the WHY1-deficient lines than the wild type, especially during early chloroplast biogenesis (Figure 5.27). The greening of the chloroplasts was significantly delayed in the WHY1-deficient lines compared to the wild type. Hence, chloroplast biogenesis requires a functional WHY1. The high plastome copy number and increased abundance of nuclear transcripts encoding photosynthetic proteins were not sufficient to compensate

for the loss of WHY1. The levels of plastome-encoded transcripts were significantly lower in the WHY1-deficient lines than the wild type, with much lower levels of the encoded proteins, as discussed in Chapter 4. The lack of plastome-encoded transcripts/proteins explains the delayed greening of the WHY1-deficient leaves, despite the high expression of nuclear-encoded photosynthetic genes (Chapter 4).

Moreover, the levels of many transcripts involved in RNA processing such as pentatricopeptide repeat (PPR) were increased in the W1-7 leaves relative to the wild type (Figure 5.28). PPR proteins are family of RNA-binding proteins that function in RNA processing in chloroplast gene expression (Stern et al., 2010, Shikanai and Fujii, 2013). Similarly, RNA helicases are involved in processes such as RNA synthesis, modification, cleavage and degradation as well as in ribosome biogenesis and translation initiation (Banroques et al., 2011, Linder and Jankowsky, 2011). The high levels of transcripts encoding such proteins in the WHY1-deficient leaves suggests that there is enhanced expression of all the genes associated RNA processing as well as those encoding chloroplast ribosomal subunits. The enhanced expression may be an attempt to compensate for low levels of chloroplast ribosomal proteins/translation.

The levels of transcripts encoding protein kinases were increased in the WHY1-deficient lines compared to the wild type (Figure 5.6 and Figure 5.15). For example, the levels of transcripts encoding protein kinase AK356512 were higher in the WHY1-deficient leaves than the wild type. This transcript shows homology to *Arabidopsis* (AT3G54090.1), which encodes a fructokinase-like protein. Members of the pfkB-carbohydrate kinase family are potential plastidial thioredoxin targets (Arsova et al., 2010). *Arabidopsis* knockout mutants displayed an albino phenotype and had impaired in chloroplast development, with defects in PEP-dependent transcription (Arsova et al., 2010). These data suggest that post-translational regulation via protein kinases may be important in the control of chloroplast development.

In addition, the levels of transcripts encoding FAR1-like proteins (MLOC_33258.3) were increased in the W1-7 relative to the wild type, particularly at 7 days after germination. The abundance of this transcript was lower in the embryos of the dry seeds of the WHY1-deficient than the wild type (Appendix A.1). The expression of FAR1-like proteins is important during the establishment of photosynthesis but it may not be required in the embryos. The levels of transcripts encoding FAR1-like proteins were also found to be increased in an earlier study of W1-7 leaves at 21 days old (Comadira et al., 2015). The FAR1 gene encodes a transposase-derived transcription factor that activates the expression of *FAR-RED ELONGATED HYPOCOTYL1* and *FHY1-LIKE* and modulate phytochrome A to promote chlorophyll biosynthesis and chloroplast division (Tang et al., 2012). The FAR1 transcription factor is also a positive regulator of ABA signalling in *Arabidopsis* (Tang et al., 2013). An increase in ABA sensitivity was observed during seed germination in lines expressing WHY1 in the plastids (Isemer et al., 2012a). The absence of WHY1 may result in an increase in light signalling pathways mediated by FAR1 that integrate light and ABA signalling (Tang et al., 2013) in the W1-7 leaves compared to the wild type (Figure 5.30 A).

The leaf transcript profiling analysis also revealed that transcripts encoding a Val-tRNA synthetase were significantly more abundant in the W1-7 leaves than the wild type (Figure 5.21). Previous studies have shown that rice Val-tRNA synthetase mutants called white panicle 1 (*wp1*) have pale or albino phenotypes, suggesting that the Val-tRNA synthetase plays important role in chloroplast development and chloroplast ribosome biogenesis (Wang et al., 2016). Furthermore, many others transcripts were changed in the WHY1-deficient leaves at earliest stages of leaf development. Of these, transcripts associated with redox processes such as thioredoxin, peroxidases and hormones metabolism were significantly more abundant in the WHY1-deficient lines than the wild type during early stages of leaf development. Similarly, transcripts encoding transcription factors of the MYB, bHLH and WRKY families were increased in the WHY1-deficient lines compared to the wild type. However, there was also a number of transcripts encoding MYB, bHLH and a few WRKY family

transcription factors that were decreased in the WHY1-deficient lines compared to the wild type. The observed changes in transcription factors and proteins involved redox processes and also transcriptions factors as well as proteins involved in defence in hormone signalling, metabolism and developmental processes demonstrate that a significant readjustment of gene expression is required in leaves lacking WHY1 during early leaf development and chloroplast biogenesis.

Chapter 6 . Metabolic leaf profile of the WHY1-deficient lines during chloroplast development

6.1 Introduction

During plant development, young leaves are metabolic sinks that require the import of nutrients to sustain their metabolism, growth and development. The leaf undergoes a transition from a metabolic sink to a source during development. Plants are photoautotrophs and maintain growth and development through photosynthesis, via the delivery of reduced carbon (C) compounds. The reduced C compounds produced by photosynthesis are reoxidised in respiration, producing energy and C skeletons for processes such as the incorporation of inorganic N into amino acids (Foyer et al., 2003). Signal transduction and the regulation of gene expression are required to regulate the transition of the emerging leaf from the sink state to a source with functional photosynthesis during development. This process also involves changes in the accumulation of metabolites which are connected to biochemical phenotypes (Satou et al., 2014).

Metabolite profiling analysis can provide an understanding of gene functions. The metabolites produced in complex networks of biochemical pathways are classified either as primary metabolites (e.g. amino acids, sugars, sugar phosphates, and organic acids) or secondary metabolites (e.g. phenylpropanoids) that often have important functions in specific environmental conditions (Satou et al., 2014). Metabolite profiling provides a snap-shot of the metabolic status of an organ (Weckwerth, 2003, Saito and Matsuda, 2010). The biosynthesis of most amino acids begins in the chloroplasts, while most of the proteins involved in these metabolic pathways are encoded by nuclear genes and transported into the chloroplast (Satou et al., 2014). Leaf metabolic profiling approaches have been used in comparisons of albino and pale green *Arabidopsis*

mutants that showed increases in several amino acids such as glutamine, glutamate and asparagine as a result of altered levels of nuclear-encoded chloroplast proteins (Satou et al., 2014).

Leaf metabolite profiles have previously been analysed in 17-day old WHY1-deficient barley (W1-7) seedlings and the wild type (WT) the under optimal and limiting nitrogen nutrition (Comadira et al., 2015). The metabolite profiles of the W1-7 leaves were shown to be similar to the wild type but there was a significant difference in the levels of leaf sucrose. In addition, there were low levels of reducing sugars and tricarboxylic acid cycle intermediates but no differences in the leaf amino acid pools (Comadira et al., 2015). The aim of the studies reported in this chapter was to extend this analysis by examining the metabolite profiles of different leaf sections early in leaf development in the WHIRLY1-deficient barley (W1-1, W1-7) seedlings and the wild type (WT). Furthermore, these studies will allow an understanding of the changes in the metabolite pools that occur during chloroplast development. Therefore, WHY1 and wild type seedlings were grown for 7 and 14 days after germination. The metabolite profiles were determined in the base, middle (Mid) and tip sections of the first leaves.

6.2 Results

6.2.1 Sample variation

Principal component analysis (PCA) of the metabolic profile of the first leaves of WHY1-deficient lines showed a definitive separation from the wild type when comparing lines and leaf positions (Figure 6.1). When score 1 was plotted against score 2 in terms of genotype, the W1-7 had a clear separation from the W1-1 and the wild type (Figure 6.1). There was a clear separation in the leaf positions (base, middle and tip) of the first leaves of W1-7 compared to the wild type (Figure 6.1 B). However, this was less apparent in the W1-1 compared to the wild type. Metabolites were also separated in leaf positions which was most clearly illustrated when score 2 was plotted against score 5 (Figure 6.1 B).

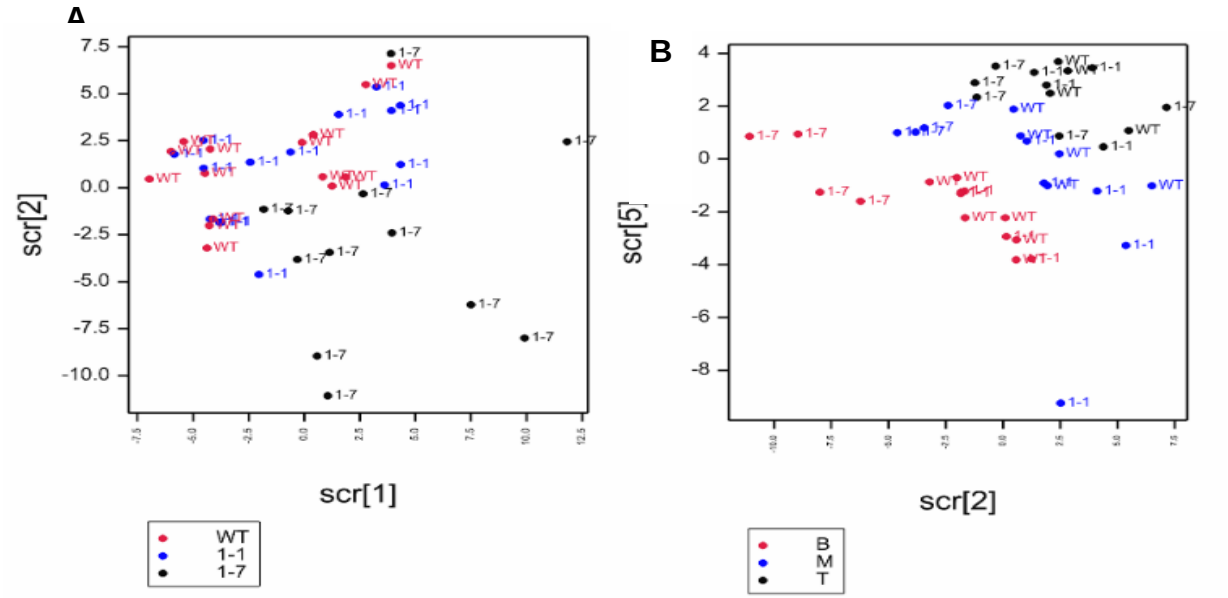


Figure 6.1: Principal components analysis (PCA) of the metabolic profiles of the first leaves of W1-1, W1-1 and wild type (WT), 7 days after germination.

The PCA analysis has been plotted against (A) genotypes and (B) leaf regions (base, middle and tip). (A) The line plot was separated according to genotypes (red: WT, blue: W1-1 and black: W1-7) and (B) the line plot that was separated according to leaf sections (red: base, blue: middle and black: tip). Scr: score. B: base, M: mid, T: tip.

6.2.2 Metabolite changes in the WHY1-deficient line relative to the wild type during leaf development at 7 days old.

In total, there were 79 metabolites that were changed in young leaves of WHY1-deficient lines both at 7- and 14-day old. These metabolites have been categorised in the following categories: amino acids, TCA cycle intermediates, sugars and fatty acids. The analysis illustrates an overview of changes in leaf metabolite profiles that that illustrating amino acids, TCA cycle intermediates, sugars and carboxylic acids (Figure 6.2).

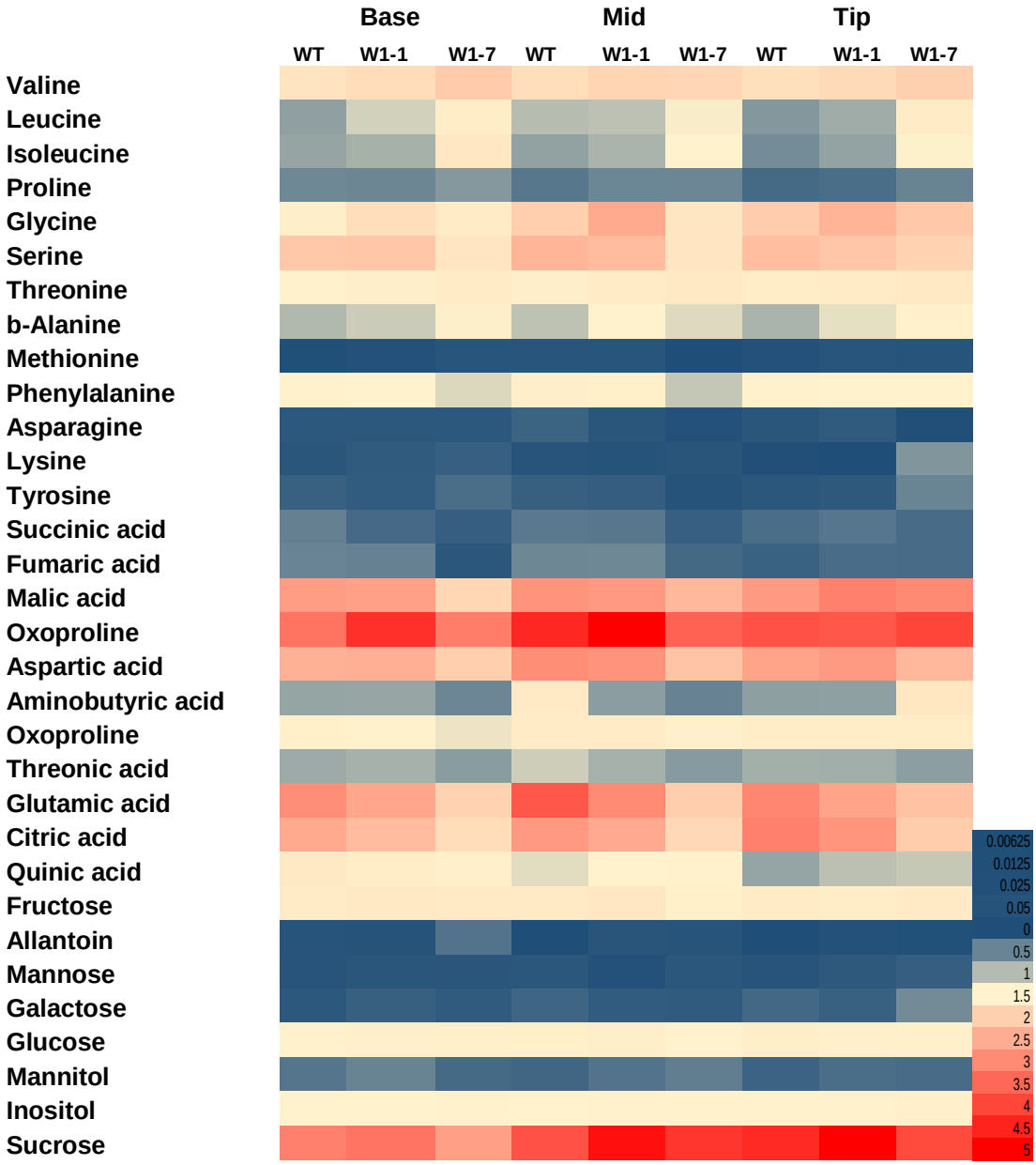


Figure 6.2: Heat map on the content of metabolites in the WHY1-deficient lines compared to the wild type, 7 days after germination.

The values shown on the heat map represent the relative concentration of the mean of each compound (n=4), which is estimated as response ratios calculated by peaks areas normalised to the internal standards areas (ribitol and nonadecanoic acid methyl ester) for each metabolite acquired. Statistical analysis for metabolite data was determined using 1-way ANOVA and was set to $p<0.05$ to take into account of the greater variability within the metabolite dataset.

6.2.2.1 Amino acids

The loss of WHY1 caused a large change in the amino acid profiles of the first leaves in WHY1-deficient leaves compared to the wild type. The levels of serine were significantly lower in the base and middle of the first leaves of W1-7 (Figure 6.3 B). However, there were no differences in the levels of serine in the tip leaves of the W1-7 compared to the wild type and W1-1, 7 days after germination (Figure 6.3 B). In contrast to the level of glycine, only the middle section of the first leaves of W1-1 had significantly higher levels of glycine compared to the W1-7 and wild type (Figure 6.3 A). The levels of glycine were all similar in the base and tip sections of the first leaves in all lines (Figure 6.3 A).The ratio of Gly/Ser was higher in the base, middle and tip sections of the first leaves in the W1-1 and W1-7 7 leaves compared to the wild type at 7 days after germination (Figure 6.3 C).

The levels of asparagine were similar in all leaves sections in the wild type, W1-1 and W1-7 barley leaves (Figure 6.4 A). However, the levels of aspartate were lower in the base and middle sections of the W1-7 leaves compared to the wild type and W1-1 (Figure 6.4 B). In contrast, the levels of aspartate were not different in the tip sections in all the barley (Figure 6.4). The Asn/Asp ratio was 2 times higher in the base section of the W1-7 leaves compared to the wild type and W1-1 (Figure 6.4).

The levels of isoleucine, leucine and valine were significantly increased in the base sections of the first leaves of W1-7 compared to wild type and W1-1, 7 days after germination (Figure 6.5). The levels of valine were significantly increased in the tip leaves of W1-7 compared to the wild type (Figure 6.5 C). The levels of isoleucine and leucine were similar in the middle and tip sections of the first leaves of the W1-7 compared to wild type and W1-1, 7 days after germination (Figure 6.5 A & B). The levels of glutamate and phenylalanine were significantly lower in the middle section of the first leaves of the W1-7 compared to the wild type (Appendix J). In contrast, the levels of glutamate and phenylalanine were all similar in the base and tip in all lines (Appendix J). There were no differences in all leaf regions in all the lines in the levels of lysine, threonic acid, proline, β-alanine and tyrosine (Appendix J).

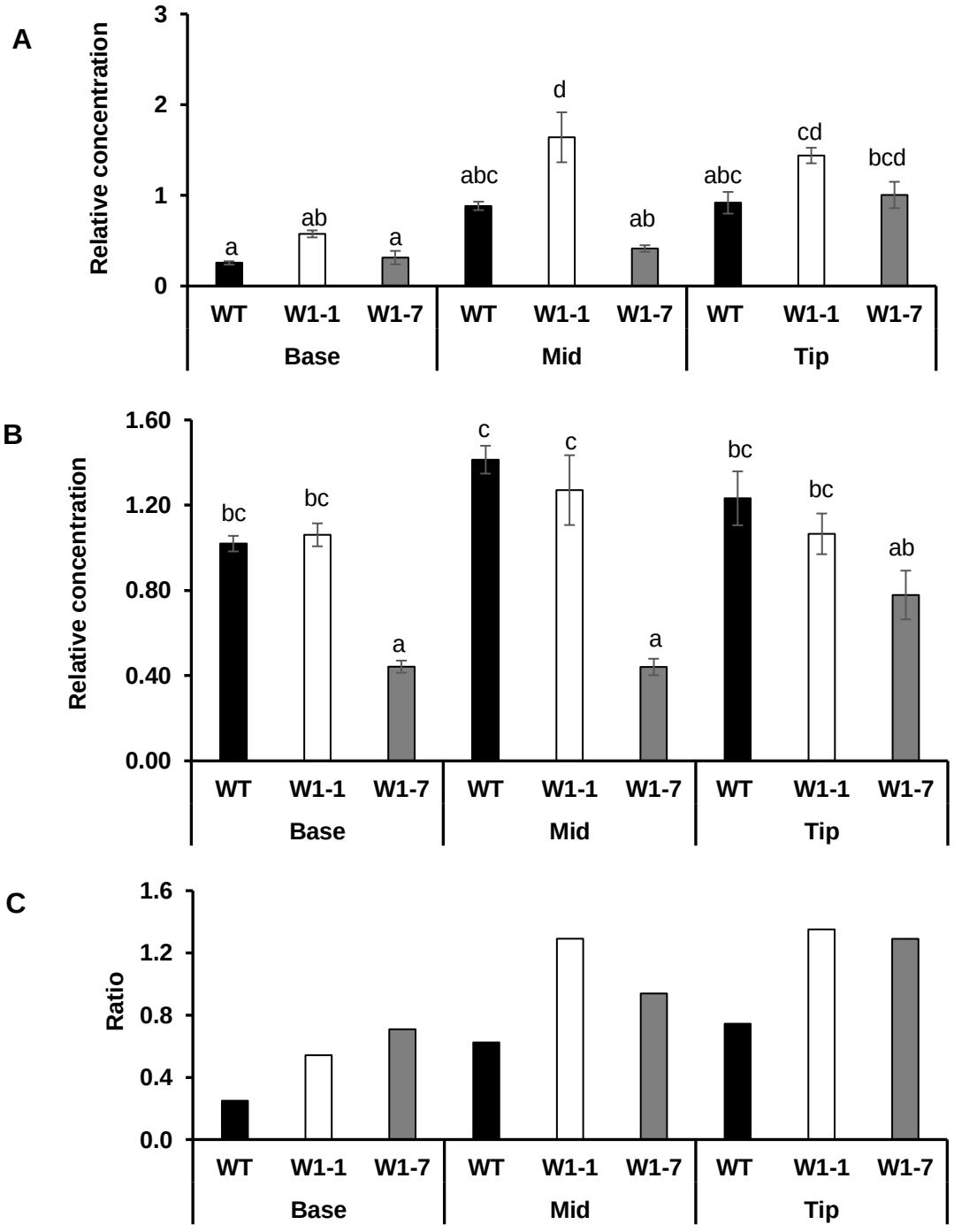


Figure 6.3: The levels of amino acids (A) glycine, (B) serine and (C) glycine to serine ratio in the base, middle (Mid) and tip sections of the first leaves of wild type (WT) and W1-1 and W1-7 seedlings, 7 days after germination.

Relative concentration was the mean compound (n=4) and normalised to the internal standards (n=4). Significant differences (letters) were analysed using Tukey's HSD test 1-way ANOVA (p<0.05).

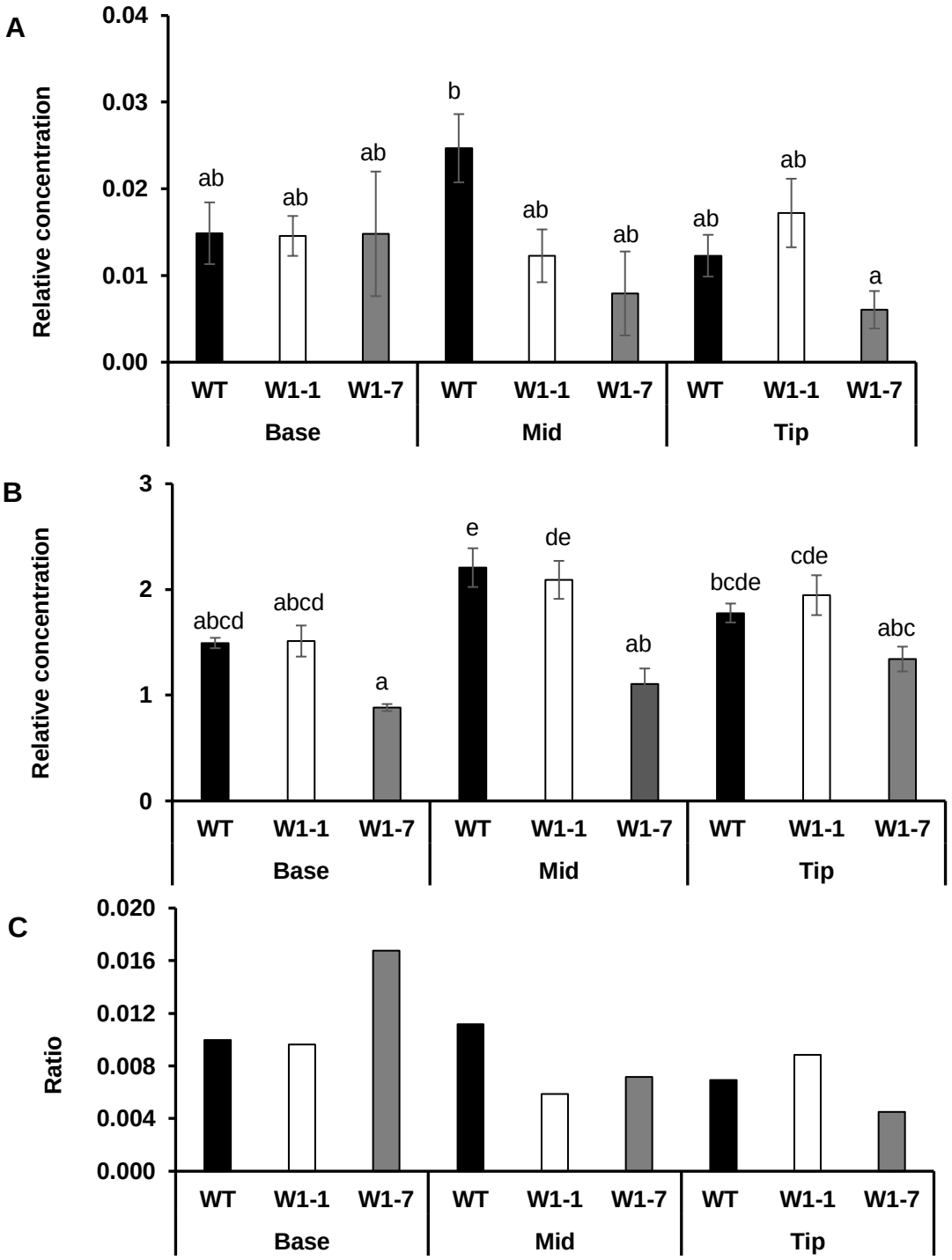


Figure 6.4: The levels of amino acids (A) asparagine, (B) aspartate and (C) the ratio of asparagine to aspartate in the base, middle (Mid) and tip sections of the first leaves of wild type (WT) and W1-1 and W1-7 seedlings, 7 days after germination.

Relative concentration was the mean compound (n=4) and normalised to the internal standards. Significant differences (letters) were analysed using Tukey's HSD test 1-way ANOVA (p<0.05).

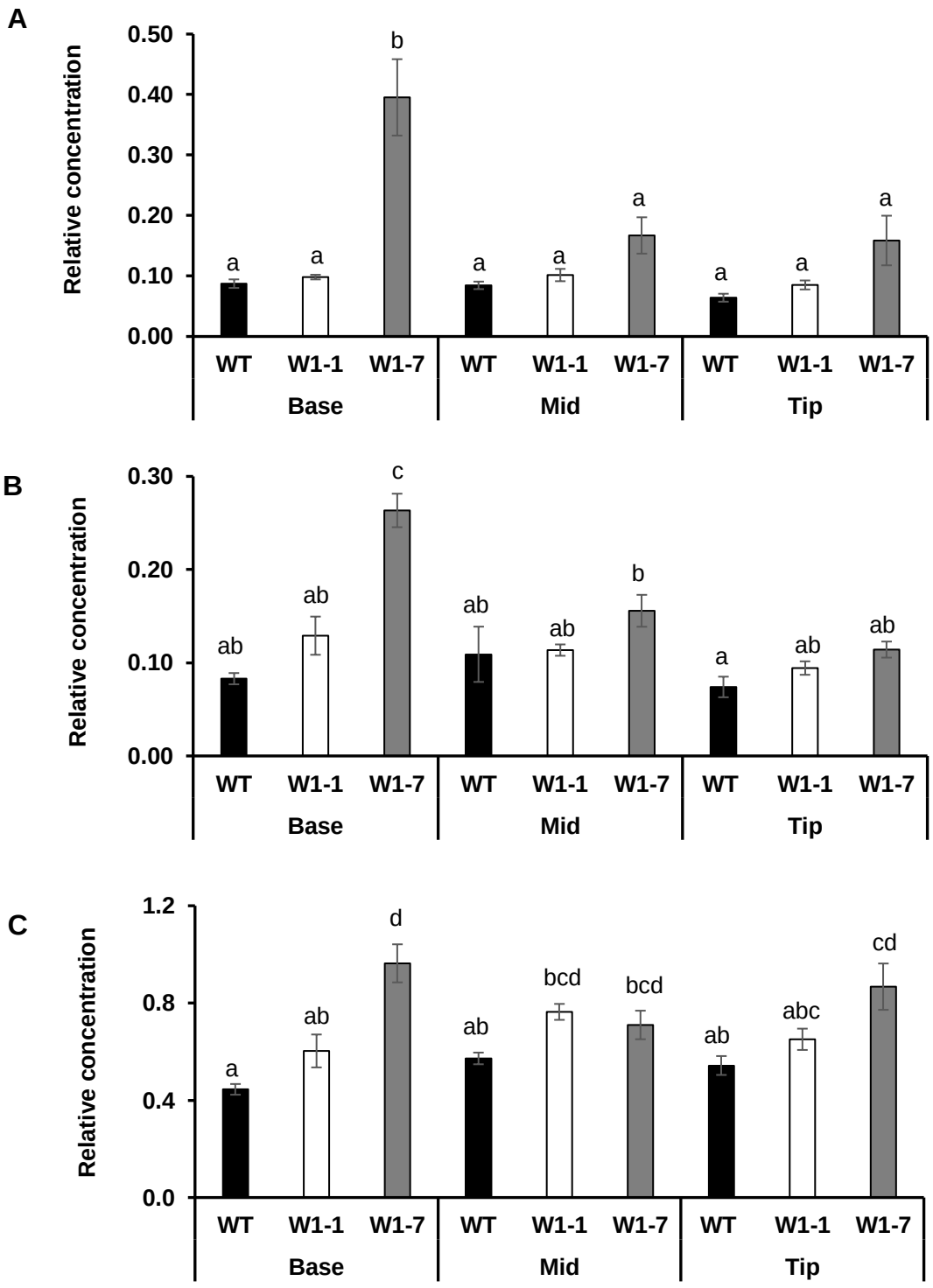


Figure 6.5: The levels of amino acids (A) isoleucine (B) leucine and (C) valine in the base, middle (Mid) and tip sections of the first leaves of wild type (WT) and W1-1 and W1-7 seedlings, 7 days after germination.

Relative concentration was the mean compound (n=4) and normalised to the internal standards. Significant differences (letters) were analysed using Tukey's HSD test 1-way ANOVA (p<0.05).

6.2.2.2 Carbohydrates

The levels of fructose and sucrose were all similar in the first leaves of all leaf sections in all lines (Figure 6.6 A & B). The levels of glucose were also similar in the first leaves of all leaf sections in all lines (Figure 6.6 C). However, the level of glucose in the tip section of the first leaves of W1-7 was significantly lower compared to the wild type, 7 days after germination (Figure 6.6 C). In addition, the levels of galactose and inositol were all similar in the first leaves of all leaf sections in all lines, 7 days after germination (Appendix J).

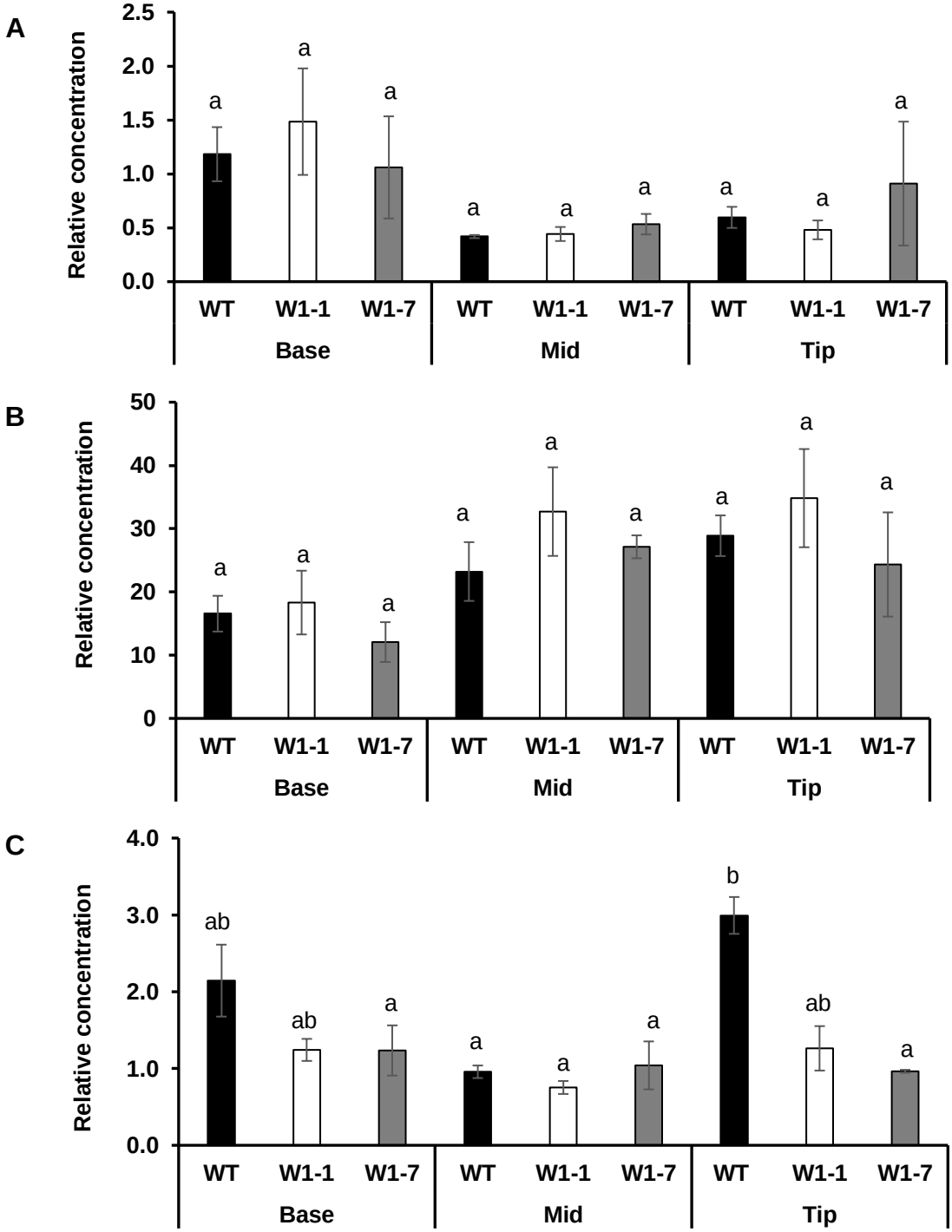


Figure 6.6: The levels of amino acids (A) fructose, (B) sucrose and (C) glucose in the base, middle (Mid) and tip sections of the first leaves of wild type (WT) and W1-1 and W1-7 seedlings, 7 days after germination.

Relative concentration was the mean compound (n=4) and normalised to the internal standards. Significant differences (letters) were analysed using Tukey's HSD test 1-way ANOVA (p<0.05).

6.2.2.3 TCA cycle intermediates

The levels of a large number of TCA cycle intermediates such as malate and fumarate were significantly lower in the base sections of the first leaves of W1-7 compared to the wild type and W1-1, 7 days after germination (Figure 6.7 A & B). The levels of malate and fumarate were lower in the middle sections of the first leaves of W1-7 compared to wild type and W1-1, 7 days after germination (Figure 6.7 A& B). The level of succinate was significantly lower in the base section of the first leaves of W1-7 compared to the wild type and had no difference with W1-1, 7 days after germination (Figure 6.7 C). The level of succinate was lower in the middle section of the first leaves of W1-7 compared to the wild type and W1-1, 7 days after germination (Figure 6.7 C). Interestingly, the levels of malate, fumarate and succinate in the tip of WHY1-deficient leaves were higher than the wild type, 7 days after germination (Figure 6.7). In contrast, the levels of citrate were all similar in all leaves sections in all lines, 7 days after germination (Appendix J).

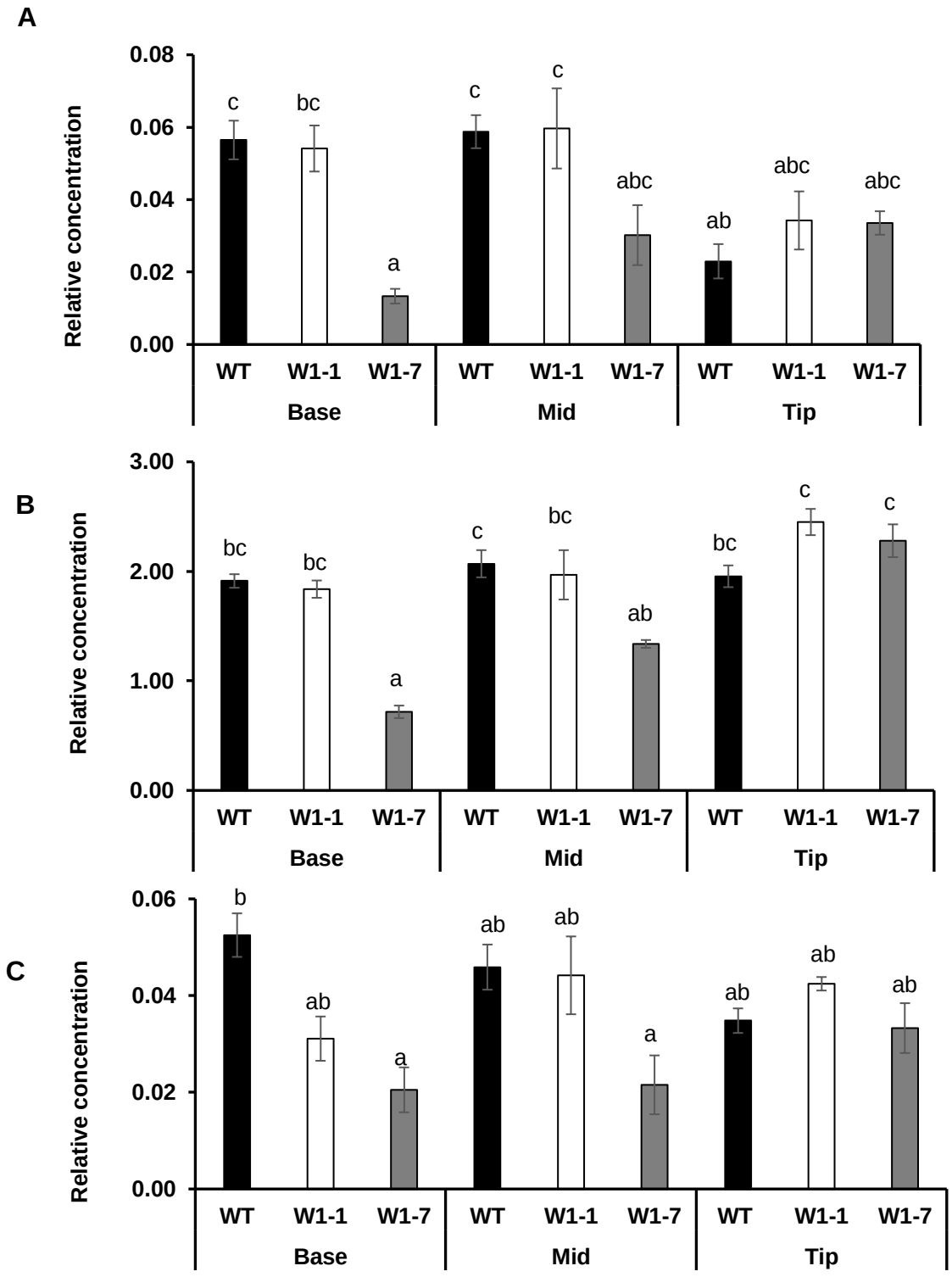


Figure 6.7: The levels of (A) fumarate, (B) malate and (C) succinate in the base, middle (Mid) and tip sections of the first leaves of wild type (WT) and W1-1 and W1-7 seedlings, 7 days after germination.

Relative concentration was the mean compound (n=4) and normalised to the internal standards. Significant differences (letters) were analysed using Tukey's HSD test 1-way ANOVA (p<0.05).

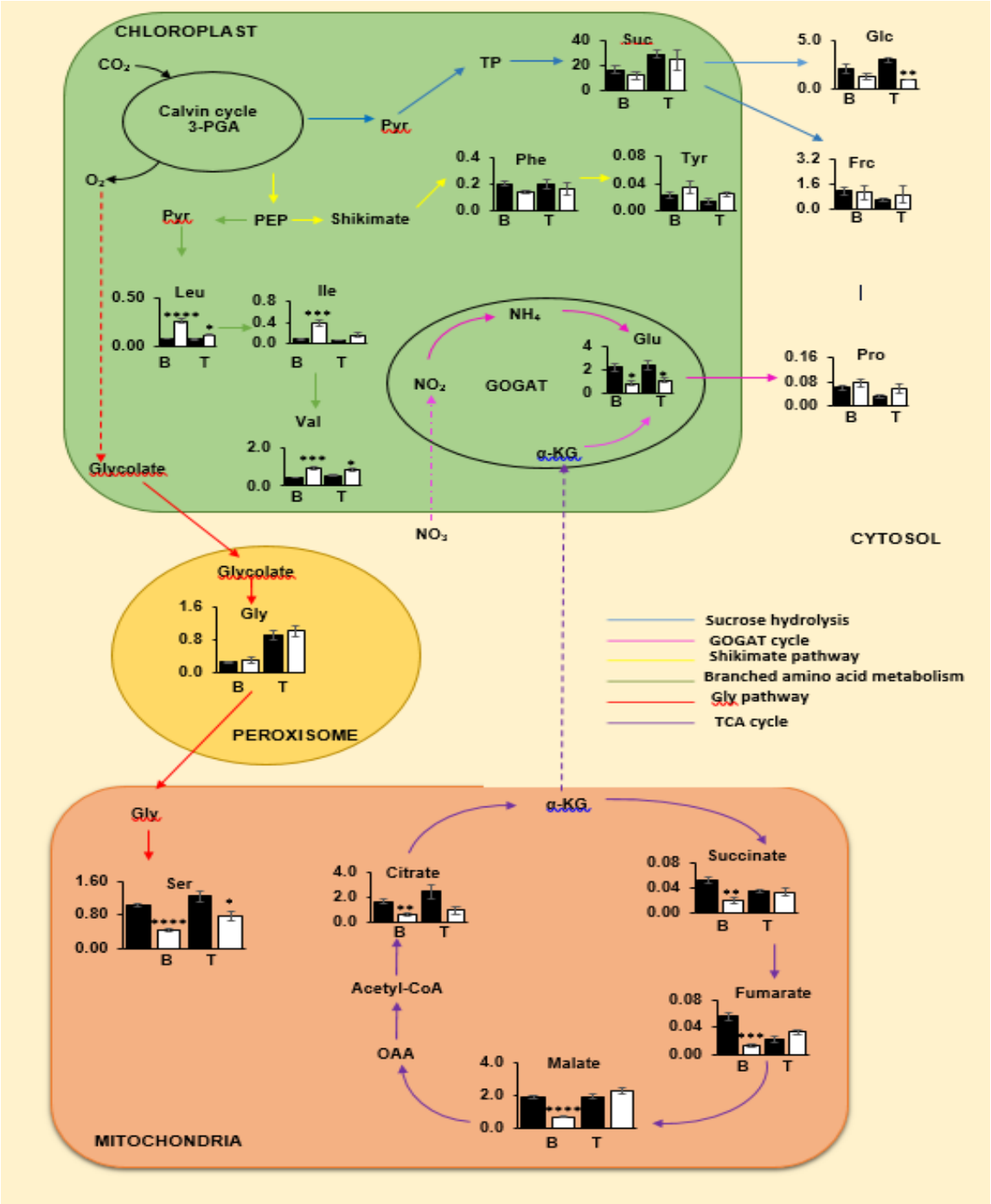


Figure 6.8: A comparison of the leaf metabolite profiles in the first leaves of the W1-7 compared to the wild type at 7-day old analysis, shown as a schematic of key metabolic pathways.

The bar chart represents the relative concentration of each metabolite in the W1-7 and wild type (black bar= WT and white = W1-7 base). Relative concentration was the mean compound (n=4) and normalised to the internal standards. Significant differences (letters) were analysed using the Student's *t*-test (p<0.05).

6.2.2.4 Fatty acids

In general, the fatty acids contents of the first leaves in the W1-1 and W1-7 had no significant differences in all leaf sections compared to the wild type, 7 days after germination.

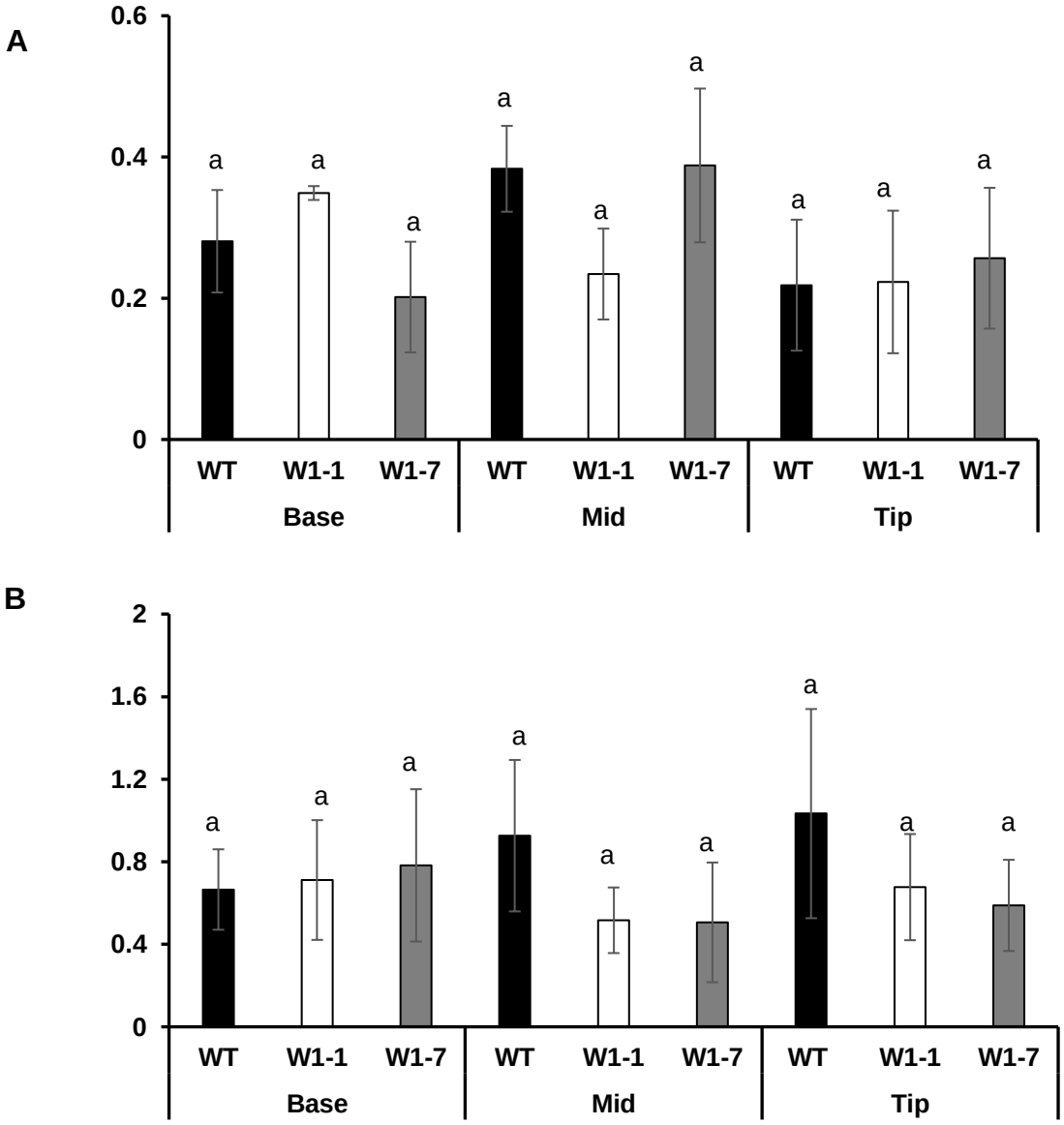


Figure 6.9: The levels of (A) pentadecanoic and (B) octadecenoic in the base, middle (Mid) and tip sections of the first leaves of wild type (WT) and W1-1 and W1-7 seedlings, 7 days after germination.

Relative concentration was the mean compound (n=4) and normalised to the internal standards. Significant differences (letters) were analysed using Tukey's HSD test 1-way ANOVA (p<0.05).

6.2.3 Metabolite changes in the WHY1-deficient line relative to the wild type during leaf development at 14 days old.

The following experiments were performed to determine the levels of metabolite profiles of different leaf sections in the WHY1-deficient (W1-1, W1-7) seedlings and the wild type (WT) at 14 days old. However, only three biological replicates were used in this analysis as many samples for 14 days old analysis were degraded during the processing.

6.2.3.1 Amino acids

The level of glycine was significantly higher in the tip leaves of the first leaves of W1-7 compared to the wild type with no significant changes in other leaf sections in all lines, 14 days after germination (Figure 6.10). The ratio of Gly/Ser was also higher in the tip of W1-7 compared to wild type, 14 days after germination (Figure 6.10). There were no significant differences in the levels of serine, isoleucine, leucine and valine in all the leaf sections of all lines, 14 days after germination (Figure 6.11).

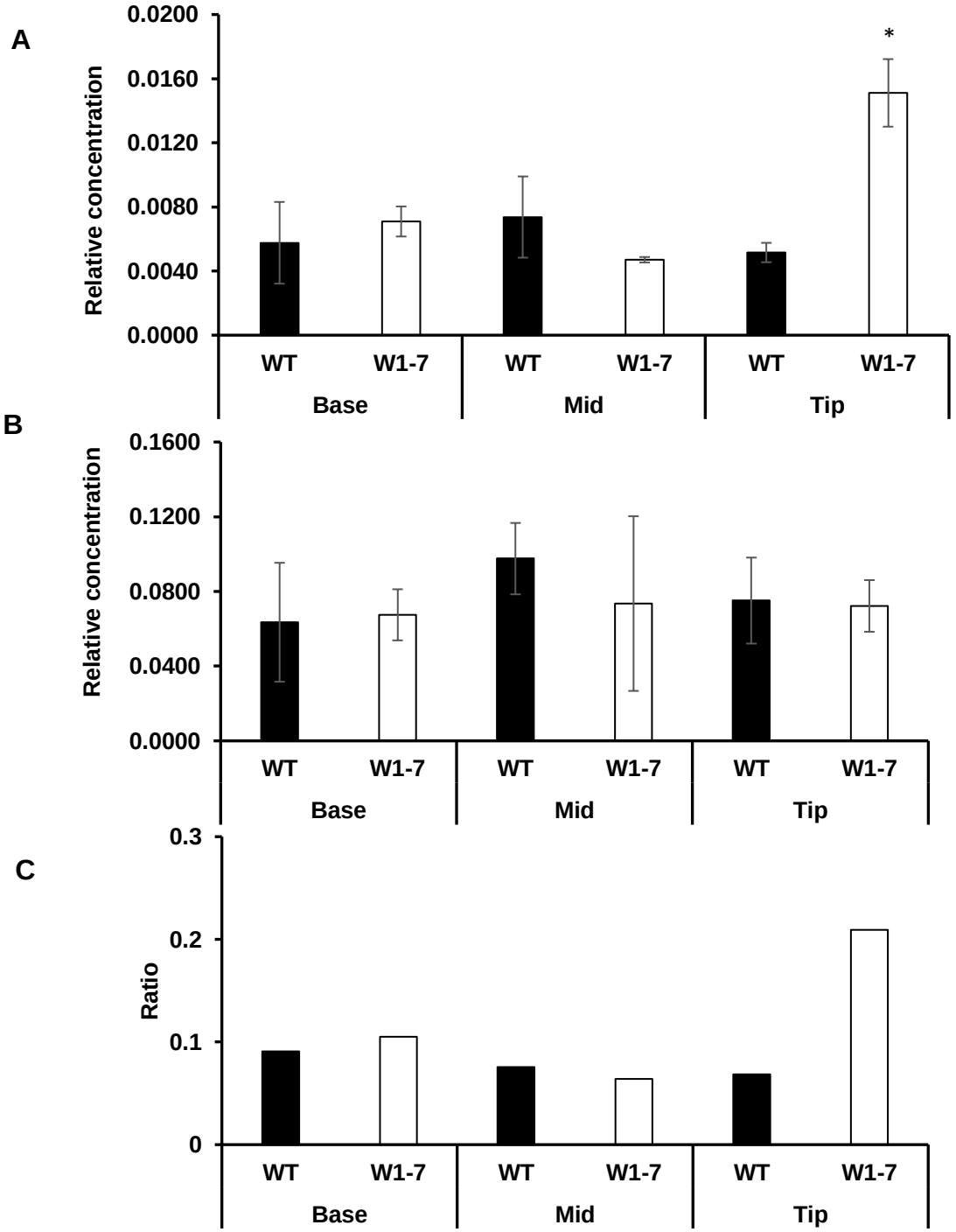


Figure 6.10: The levels of (A) glycine, (B) serine and (C) ratio of gly/ser in the base, middle (Mid) and tip sections of the first leaves of wild type (WT) and W1-7 seedlings, 14 days after germination.

Relative concentration was the mean compound (n=3) and normalised to the internal standards. Significant differences were analysed using Student's *t*-test ($p<0.05$).

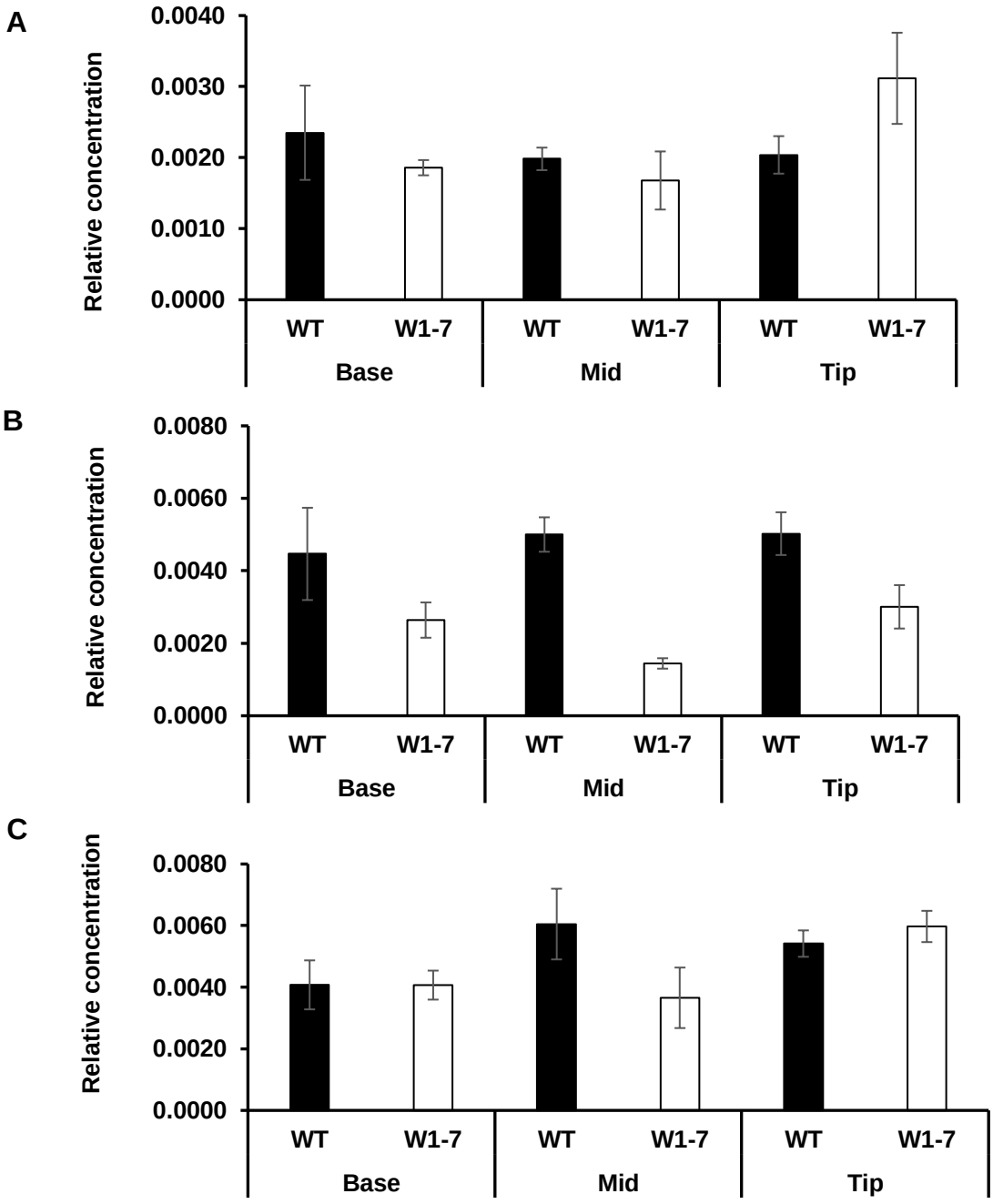


Figure 6.11: The levels of (A) isoleucine, (B) leucine and (C) valine in the base, middle (Mid) and tip sections of the first leaves of wild type (WT) and W1-7 seedlings, 14 days after germination.

Relative concentration was the mean compound (n=3) and normalised to the internal standards. Significant differences were analysed using Student's *t*-test ($p<0.05$).

6.2.3.2 Carbohydrates

The levels of sucrose were similar in the base and tip sections of the first leaves of W1-7 but the middle section of the first leaves of W1-7 had lower level of sucrose compared to the wild type, 14 days after germination (Figure 6.12 A). However, the levels of fructose and glucose were lower in the base sections of the first leaves in the W1-7 compared to the wild type, 14 days after germination (Figure 6.12 B & C). Similar levels of fructose and glucose were observed in the middle sections of the first leaves both in W1-7 and wild type, 14 days after germination (Figure 6.12 B & C).However, the levels of fructose and glucose were higher in the tip section of the first leaves W1-7 compared to wild type, 14 days after germination (Figure 6.12 B & C).

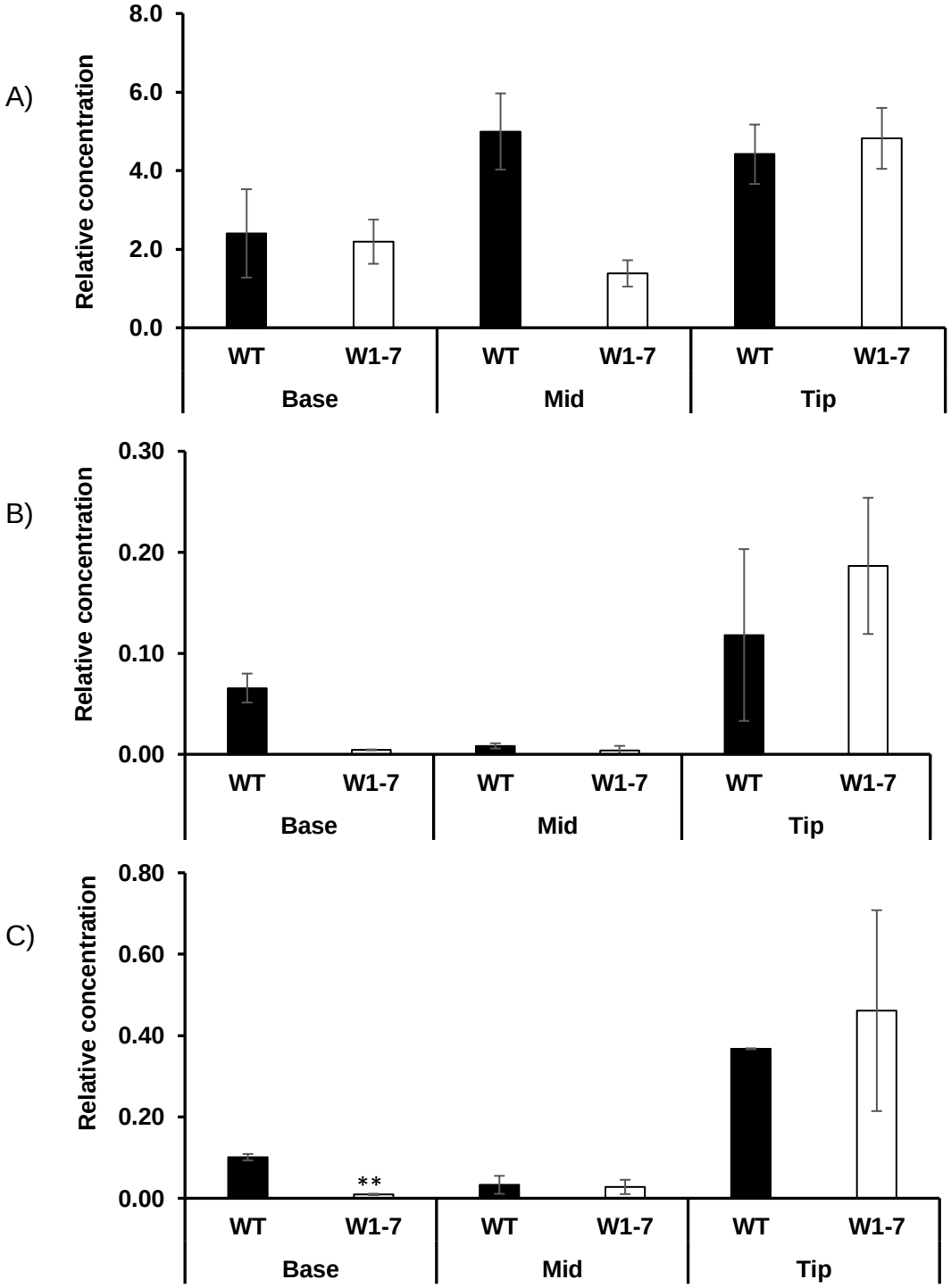


Figure 6.12: The levels of amino acids (A) sucrose, (B) fructose and (C) glucose in the base, middle (Mid) and tip sections of the first leaves of wild type (WT) and W1-7 seedlings, 14 days after germination.

Relative concentration was the mean compound (n=3) and normalised to the internal standards. Significant differences (letters) were analysed using Student's *t*-test ($p<0.05$).

6.2.3.3 TCA cycle intermediate

There were no differences in the levels of TCA cycle intermediates, such as fumarate, malate and succinates in all the leaf sections in all lines, 14 days after germination (Figure 6.13).

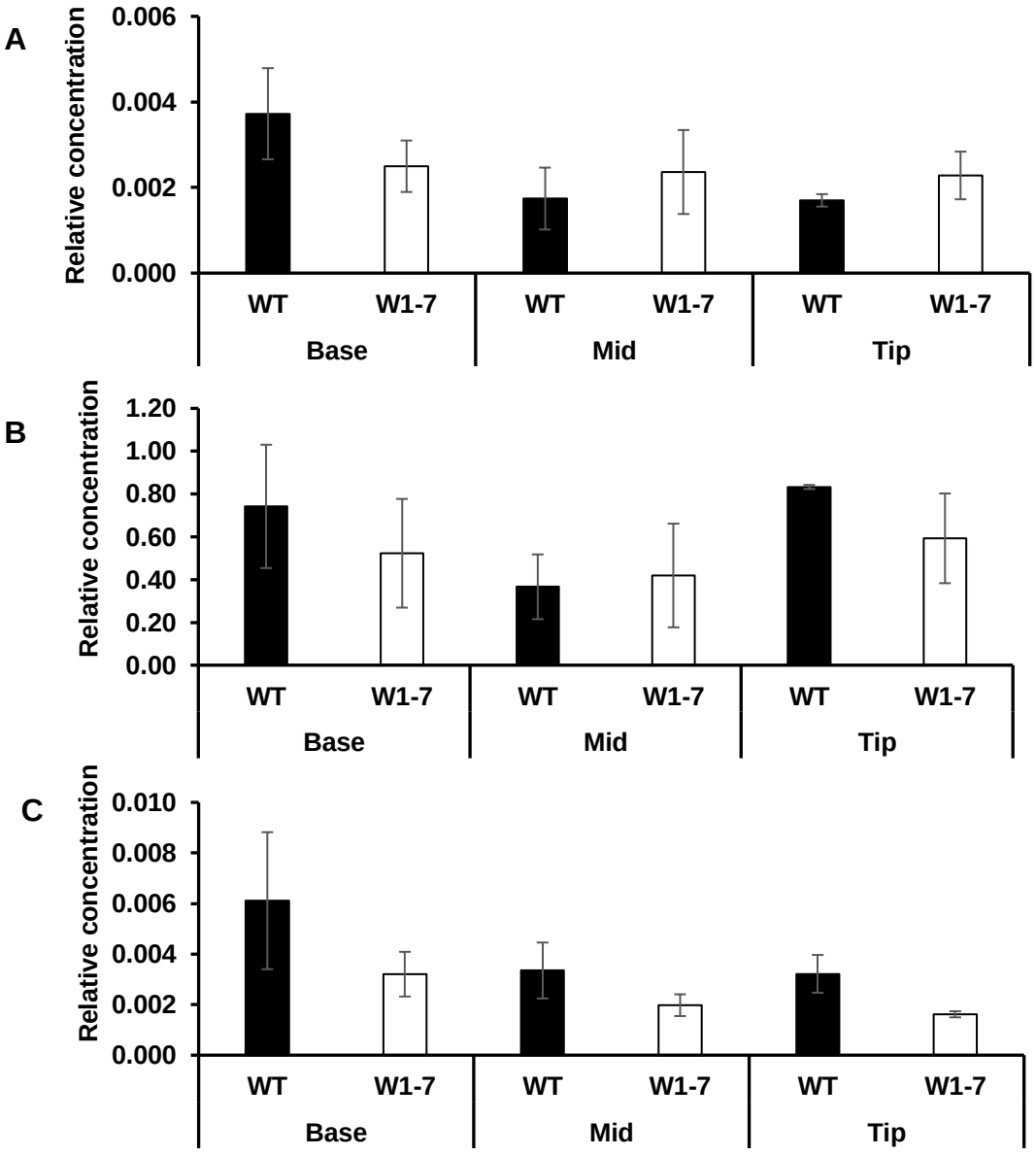


Figure 6.13: The levels of (A) fumarate, (B) malate and (C) succinate in the base, middle (Mid) and tip sections of the first leaves of wild type (WT) and W1-7 seedlings, 14 days after germination.

Relative concentration was the mean compound (n=3) and normalised to the internal standards. Significant differences (letters) were analysed using Student's *t*-test ($p<0.05$).

6.3 Discussion

The studies reported here have characterised the metabolite changes that occurred during barley leaf development, together with the effects of the loss of WHY1 on this process. The PCA analysis showed a clear separation in the data obtained from the W1-7 and wild type leaves (Figure 6.1). This analysis also showed that there were differences in the metabolite profiles of each of the leaf regions (base, middle and tip). However, these differences were less apparent in the W1-1 leaves than the W1-7 leaves compared to the wild type (Figure 6.1). These findings show that the loss of WHY1 protein has a marked effect on the developmental profile of leaf metabolites.

The ratio of Gly/Ser can be used as a marker of photorespiration (Novitskaya et al., 2002). The data presented here showed that Gly/Ser ratios were higher in the first leaves of W1-1 and W1-7 than the wild type in 7-day old seedlings (Figure 6.3 C). This suggests that photorespiration was increased by the loss of the WHY1 protein. An increase in photorespiration may favour ROS production in the peroxisomes in the leaves lacking of WHY1 protein. This is consistent with the results obtained in the transcriptomic profiling analysis (Chapter 5), which showed that transcripts associated with redox signalling were greatly increased in the WHY1-deficient lines. This may be related to the increase in mRNAs associated with photorespiration in the barley leaves lacking WHY1 protein. Moreover, the ratios of Gly/Ser and Asn/Asp have been used as markers of carbon and nitrogen status related to photosynthesis and respiration, respectively (Novitskaya et al., 2002). The Asn/Asp ratios were increased in the base sections of the W1-7 leaves of the 7-day old seedlings compared to the wild type (Figure 6.4 C). Conversely, the Asn/Asp ratios were lower in the middle and tip sections of the first leaves in both of the W1-1 and W1-7 than the wild type (Figure 6.4 C). These results were also consistent with findings of the transcript profiling analysis, which showed that many transcripts associated with photosynthesis and photorespiration were higher in the WHY1-deficient leaves, particularly in the basal sections (Chapter 5). These findings also suggest that nitrogen-rich amino

acids accumulate in the basal regions of the W1-7 leaves, 7 days after germination. Furthermore, the basal sections of the W1-7 leaves have increased levels of branched-chain amino acids such as isoleucine, leucine and valine compared to the wild type in 7-day old seedlings (Figure 6.5).

The levels of cycle intermediates such as fumarate, malate and succinate were significantly lower in the basal sections of the W1-7 leaves compared to the wild type in 7-day old seedlings (Figure 6.7). This finding is consistent with the results in Chapter 4 (Figure 4.6), in which the onset of efficient photosynthesis was delayed in W1-7 plants compared to the wild type during leaf development. An increased flux through the TCA cycle may be required to sustain leaf growth in the absence of functional photosynthesis. The loss of the WHY1 protein, therefore, has a strong effect on the leaf metabolite pools of the developing barley leaves, particularly regarding metabolites associated with respiratory metabolism and primary nitrogen assimilation.

The lower levels of sucrose measured in the basal sections of the W1-7 leaves compared to the wild type in 7-day old seedlings (Figure 6.6 B) is consistent with a delayed onset of photosynthesis in the WHY1-deficient leaves (Figure 4.6). High sucrose levels are associated with the establishment of photosynthesis even in the basal sections of the wild type leaves. Lower levels of glucose were also measured in the basal sections of the W1-7 leaves (Figure 6.6 C). Glucose is generally produced by the invertase-mediated breakdown of sucrose to glucose and fructose. The increased levels of hexoses are probably required to sustain the increased flux through the TCA cycle particularly in the basal sections of the W1-7 leaves.

There was also a marked difference in the levels of fructose and glucose in the tip sections of the wild type and W1-7 at 14 days after germination (Figure 6.12). The levels of fructose and glucose were higher both in both the wild type and W1-7 leaves at 14 days compared to 7 days after germination. The greater hydrolysis of sucrose in the tips of the leaves may be due to senescence in the leaf tip (Pourtau et al., 2006). There were no significant differences in the levels of fatty

acids such as octadecenoic acid and pentadecanoic acid in any of the leaf regions in any line 7 days after germination. Similarly, there were no marked changes in the fatty acid levels of the W1-7 leaves compared to the wild type at 14-day old. Thus, the absence of a functional WHY1 protein had less effects on the leaf metabolite profiles of 14 days leaves that the 7 leaves of 7-day old seedlings. This suggests that the leaves of all lines reached reach a similar metabolic status at 14 days. This finding is consistent with the measurements of other parameters that serve as markers for chloroplast development shown in Chapter 4.

Chapter 7 . LEA5 and WHIRLY1 interactions with DEA (D/H)-box RNA 22 in *Arabidopsis*

7.1 Introduction

Plants produce ROS as a product of primary metabolism and as specific signals that regulate plant growth and defence under optimal and stress conditions (Foyer and Noctor, 2009, Suzuki et al., 2012). Late embryogenesis abundant (LEA)-like protein 5, (LEA5 also called LEA38: At4g02380) is one of a number of proteins, whose expression is changed in plants following exposure to environmental stresses, phytohormones and ROS. The LEA5 protein, which is a member of the LEA-3 group of protein family, is unusual because other members of this family are not regulated in response to oxidative stress (Mowla et al., 2006). *LEA5* is constitutively expressed in roots but not in seeds. In leaves, *LEA5* transcripts exhibit a diurnal pattern of regulation, being low in the light and abundant in the dark. However, the expression of *LEA5* is induced in leaves in the light following exposure to abiotic and biotic stresses, as well as oxidants and phytohormones (Salleh et al., 2012, Mowla et al., 2006).

Constitutive overexpression of *LEA5* increased root growth and shoot biomass, whereas both of these processes were decreased in anti-sense lines or *lea5* knockout mutants. In contrast, photosynthesis was more sensitive to drought in the overexpressing lines, suggesting that LEA5 protein modulates photosynthesis in plants exposed to stress (Mowla et al., 2006). While *LEA5-YFP* expression studies have shown that the LEA5 protein is localised in mitochondria (Salleh et al., 2012), tandem affinity purification (TAP) tagging studies suggested that the LEA5 protein is able to interact with chloroplast proteins (collaboration with Geert De Jaeger in the Department of Plant Systems Biology, VIB, Gent). The TAP analysis of protein-protein interactions showed that LEA5 interacted the

RH22 in the dark and in the light in all the samples analysed. This finding is interesting because LEA5 is predicted to contain a putative chloroplast transit peptide (Hundertmark and Hincha, 2008).

RH22 is also called heavy seed (HS) 3 (At1g59990), which is localised to plastids, is involved in RNA processing and metabolism (Kanai et al., 2013). The loss of RH22 function is lethal (Chi et al., 2012). Moreover, the *rh22* mutants exhibit defects in young seedlings. Knock-down mutants have a pale phenotype with defects in plastid development (Chi et al., 2012, Kanai et al., 2013). The *rh22* knock-down mutant phenotype is caused by aberrant processing of rRNAs (23S rRNA and 4.5S rRNA) for the large subunit (50S) of the plastid ribosome (Chi et al., 2012). These findings show that the RH22 is required for chloroplast ribosome assembly in *Arabidopsis*, with effects on the chloroplast, seed development and seedling growth (Chi et al., 2012, Kanai et al., 2013).

In the following experiments, a split-yellow fluorescent protein (YFP) assay system was used to analyse protein-protein interactions. In split-YFP assays, the two non-fluorescent halves of the YFP are attached to target proteins and expressed in living cells. If the target proteins interact then the YFP halves come together to produce the fluorescent protein and fluorescence signal. As well as its applications in the study of protein-protein interactions, YFP methods can also be used to study the intracellular localisation of proteins (Horstman et al., 2014). Hence, split-YFP techniques have been widely used to study plant processes (Citovsky et al., 2006, Citovsky et al., 2008, Ohad et al., 2007, Waterworth et al., 2015).

The WHY1-deficient barley seedlings show a similar pattern of aberrant processing of the plastid rRNAs (23S rRNA and 4.5S rRNA) to that of mutants lacking a functional RH22. The WHY family are single-stranded DNA binding proteins (Desveaux et al., 2000). In *Arabidopsis*, WHY1 is targeted to the nucleus and chloroplasts, while WHY2 is located in the mitochondria (Krause et al., 2005). Similarly, in barley, WHY1 is localised in the nucleus and chloroplasts and WHY2

is found in the mitochondria (Melonek et al., 2010). This chapter reports the results of experiments designed to explore the binding of the LEA5 and WHY1 proteins to RH22 in *Arabidopsis*.

The following studies were undertaken to explore the intracellular localisation of the LEA5 protein using several different approaches. Firstly, the subcellular localisation of LEA5 was studied using transgenic *Arabidopsis* lines that had been transformed to express a LEA5-YFP fusion protein under the control of the 35S promoter. Secondly, transient expression of a LEA5-YFP protein in *Arabidopsis* leaf protoplasts was used to analyse the intracellular localisation of this protein. Thirdly, *in vivo* interactions between the RH22, LEA5 and WHY1 were characterised using the split-YFP approach. The specific aims of the study were as follows:

1) To determine whether LEA5 is localised in chloroplasts using stable transgenic lines transformed expressing a LEA5-YFP fusion protein, together with transient expression of LEA5-YFP in *Arabidopsis* protoplasts.

2) To confirm the interactions between LEA5 and RH22 proteins in chloroplasts.

3) To examine interactions between WHY1 and RH22 in *Arabidopsis* chloroplasts.

7.2 Results

7.2.1 Intracellular localisation of LEA5.

Several approaches were used to elucidate the function of LEA5 in *Arabidopsis*. Firstly, various web-based prediction tools (TargetP, MitoProt and iPSORT) were used to determine the potential intracellular localisation of LEA5. The results of these studies are summarised in Table 7.1.

Table 7.1 : Predicted intracellular localisation of LEA5 using web-based prediction tools.

Values indicate the probability of mitochondrial and plastid localisation using different software programs (TargetP, MitoProt and iPSORT).

Program	Predicted target organelle	
	Mitochondrion	Plastid
TargetP	0.089	0.481
MitoProt	0.9595	-
iPSORT	-	Plastid signal peptide

TargetP predicts the subcellular location of proteins based on the predicted presence of the N-terminal pre-sequences: chloroplast transit peptide (cTP), mitochondrial targeting peptide (mTP) or secretory pathway signal peptide (SP). In a similar manner, iPSORT predicts the subcellular localisation of N-terminal sorting signals, based on the presence of mitochondrial targeting peptide (mTP) or chloroplast transit peptide (cTP) sequences. In contrast, MitoProt calculates the N-terminal protein region that can support a mitochondrial targeting sequence and cleavage site. In all cases, the data is provided as probabilities. The location with the highest score is the most likely according to the software applied. The relationship between the scores for any one program may also indicate how certain the prediction is. For example, as shown in Table 7.1, TargetP value for plastids was 0.481 but only 0.089 for mitochondria. This data suggests that there is a high probability of targeting to plastids with a lower probability of targeting to

mitochondria. However, the MitoProt software predicted a high probability for mitochondrial localisation for LEA5 (0.9595). Meanwhile, iPSORT predicted subcellular localisation of LEA5 to be only in plastids. Taken together, these data suggest that LEA5 is targeted both to mitochondria and to plastids. However, the prediction programs only provide preliminary evidence of the potential intracellular localisation of LEA5. Therefore, transgenic *A. thaliana* lines expressing a *LEA5*-YFP fusion protein under the control of the 35S CaMV promoter were used to localise the LEA protein using confocal laser scanning microscopy (Mohd Salleh et al., 2011).

7.2.2 Studies on intact leaves from LEA5-YFP-expressing plants

In the previous studies, a stable expression of LEA5-YFP showed that LEA5 localises to the mitochondria (Mohd Salleh et al., 2012). In this experiment, the LEA5-YFP fusion protein was detected in the attached youngest leaves of 5-day old seedlings of the transgenic *A. thaliana* lines using a Zeiss LSM 700 inverted confocal microscope (Figure 7.1). The seedlings were collected from controlled environment chambers during the light period immediately prior to measurement. The chloroplast auto-fluorescence signal is shown in Figure 7.1a and the YFP signal, detected using an excitation wavelength of 488 nm and an emission wavelength of 505-530 nm, is shown in Figure 7.1b. The merged YFP and chlorophyll fluorescence signals shown in Figure 7.1 demonstrate that there is a high degree of overlap between the YFP and chlorophyll fluorescence signals, as indicated by the yellow areas in the Figure 7.1. However, the *Arabidopsis* seeds of LEA5 should be genotyped in the next experiment.

a) Chlorophyll fluorescence b) YFP emission c) Overlay

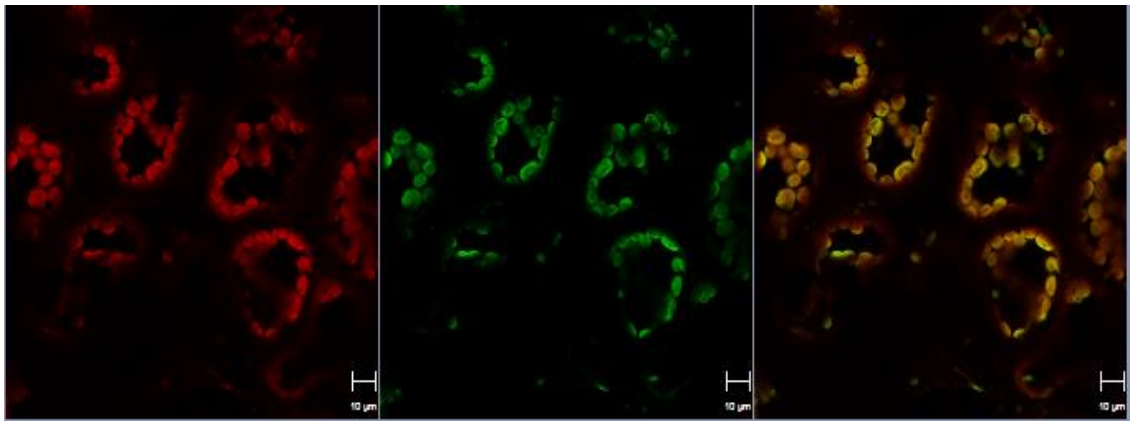


Figure 7.1: Intracellular localisation of LEA5.

The chlorophyll fluorescence signal (a, red), the YFP signal (b, green) and the merged chlorophyll fluorescence and YFP signals (c; yellow). Scale bar = 10 µm.

7.2.3 Studies on intact protoplasts from LEA5-YFP-expressing leaves

The localisation of the YFP tagged LEA protein was also determined in intact protoplasts, which were prepared from leaves of 3 week- old transgenic *Arabidopsis* plants expressing the *LEA5*-YFP construct. The YFP signal detected in protoplasts using a LSM510 META confocal microscope is shown in Figure 7.2a. The corresponding chloroplast auto-fluorescence signal is shown in Figure 7.2b and the merged YFP and chlorophyll fluorescence signals are shown in Figure 7.2c. The data shown in Figure 7.2 demonstrates an overlap between the YFP and chlorophyll fluorescence signals, as indicated by the yellow areas.

a) YFP emission b) Chlorophyll fluorescence c) Overlay

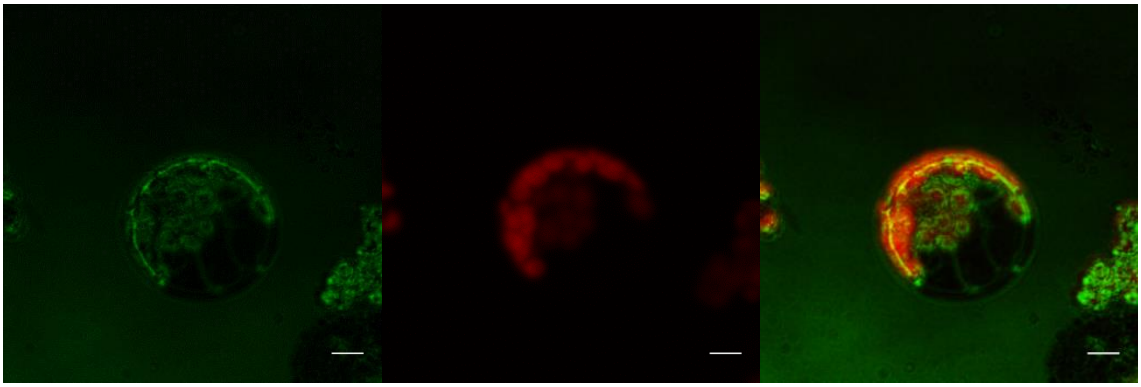


Figure 7.2: Confocal microscope images of intact protoplasts isolated from LEA5-YFP expressing leaves.

The YFP signal (a, green), the chlorophyll florescent signal (b, red) and the merged YFP and chlorophyll fluorescent signals (c; yellow). Scale bar = 10 µm.

7.2.4 LEA5 localisation in the chloroplasts and mitochondria

The intracellular localisation of the YFP-tagged protein was investigated in intact protoplasts, which were prepared from the full length of 35S-LEA5-YFP construct. In these experiments the mitochondria were identified by staining with Mitotracker Red CMXRos. The YFP signal of the protoplasts is shown in Figure 7.3a, the mitochondria-tracker signal is shown in Figure 7.3b and the corresponding chloroplast auto-fluorescence signal is shown Figure 7.3c. The merged YFP, mito-tracker and chlorophyll fluorescence signals are shown in Figure 7.3d. The data shown in Figure 7.3 demonstrate that LEA5 localises to both the mitochondria and chloroplasts.

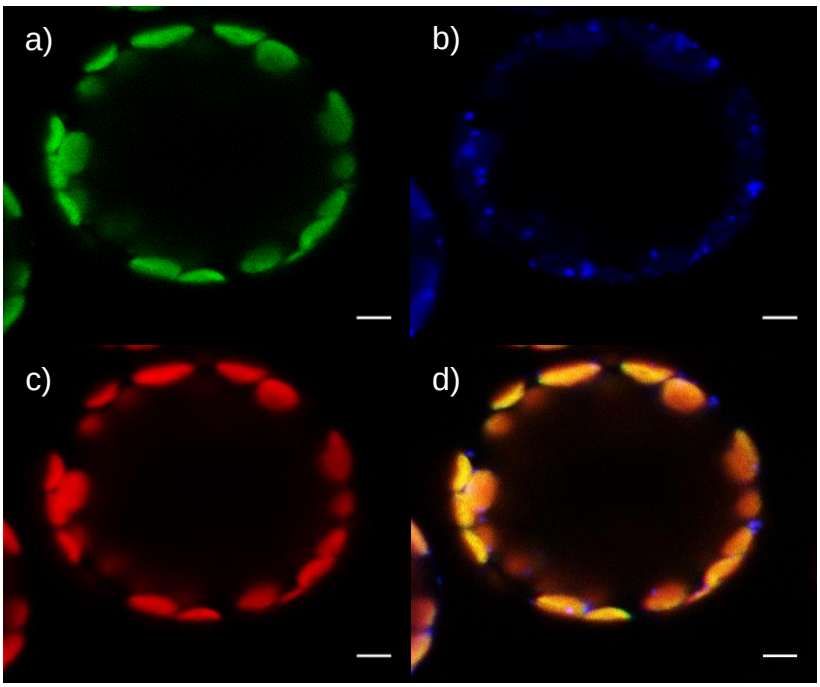


Figure 7.3: Confocal microscopy images of intact *A. thaliana* mesophyll protoplasts expressing the full length 35S-LEA5-YFP.

The YFP (a,green), the mito-tracker staining (b, blue), the chlorophyll fluorescence (c, red) and the overlay (a, green plus b, blue, plus c, red) in the same cell. Scale bar = 5 µm.

The dual chloroplast and mitochondria localisation of the LEA5 protein was confirmed in intact protoplasts, which transiently expressed the 35S-LEA5-YFP construct, together with the targeting domain of the alternative oxidase1 (AOX1) bound to a red fluorescent protein (RFP) marker system. For the chloroplast localisation, the 35S-LEA5-YFP construct was expressed together with the targeting domain of the small subunit (SSU) of ribulose-1, 5-bisphosphate carboxylase/oxygenase (RuBisCO).

The YFP signal of the protoplasts is shown in Figure 7.4a, the signal from mitochondria is shown in Figure 7.4b and the merged YFP and mitochondria targeted signal are shown in Figure 7.4c. These data demonstrate that the full length of 35S-LEA5-YFP and mitochondria-targeted AOX-RFP co-localise to the mitochondria.

a) YFP emission b) RFP emission c) Overlay

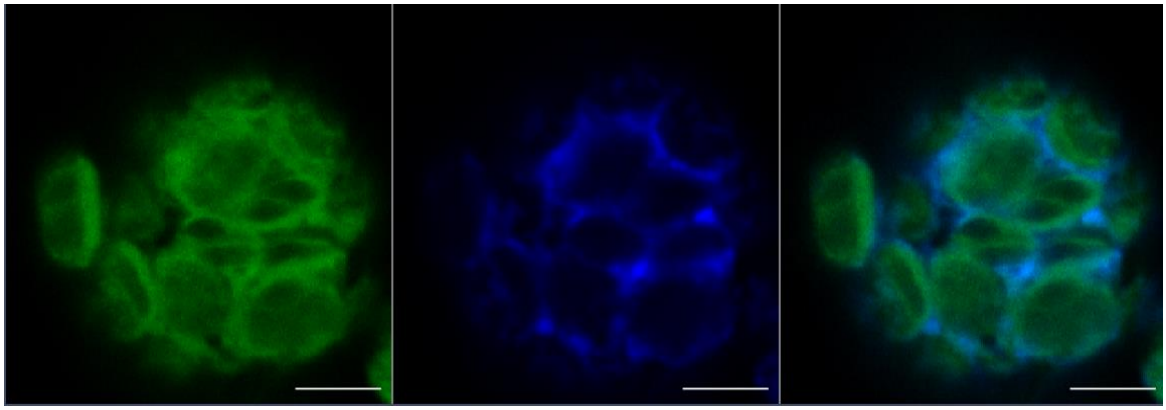


Figure 7.4: Confocal microscopy images of intact *A. thaliana* mesophyll protoplasts transiently expressing 35S-LEA5-YFP and AOX-RFP.

The YFP (a, green), the RFP (b, blue) and the overlay (a, red plus b, green) in the same cell at 40 lens. Scale bar = 10 µm.

Similarly, when the 35S-LEA5-YFP construct was expressed together with the SSU-RFP marker system, the 35S-LEA5-YFP and SSU-RFP co-localise to the chloroplast, as shown in Figure 7.5. The YFP signal of the protoplasts is shown in Figure 7.5a. The signal from the chloroplast is shown in Figure 7.5b and the merged YFP and chloroplast-targeted signals are shown in Figure 7.5c.

a) YFP emission b) RFP emission c) Overlay

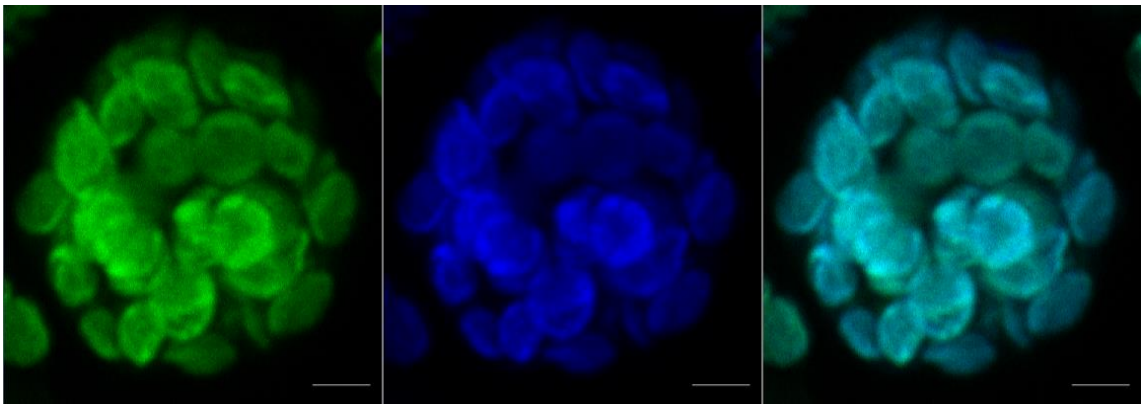


Figure 7.5: Confocal microscopy images of intact *A. thaliana* mesophyll protoplasts transiently expressing 35S-LEA5-YFP and SSU-RFP.

The YFP (a, green), the RFP (b, green) and the overlay (a, red plus b, green) in the same cell at 40 lens. Scale bar = 10 µm.

7.2.5 Expression of the LEA5 protein in protoplasts

The interactions between LEA5 and the RH22 were investigated using split-YFP probes in a transient expression system. Two other tagged proteins, amino peptidase P2 (APP2) GFP, which localises to chloroplasts, and amino peptidase P1 (APP1)-GFP, which localises to the cytosol, were used as controls. The *LEA5* and *RH22* genes were first cloned into the pDH51-GW-YFPn and pDH51-GW-YFPc vectors. Transient expression of the fusion protein using light microscope was monitored in intact *A. thaliana* protoplasts (Figure 7.6).

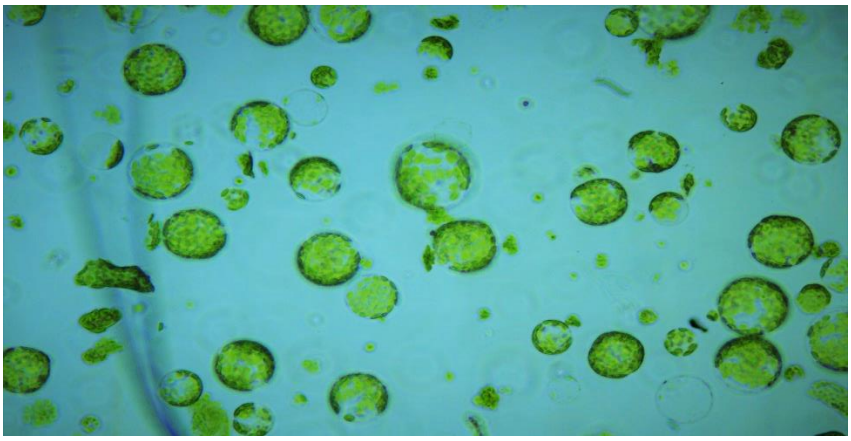


Figure 7.6: Light microscope image of a typical leaf mesophyll protoplast preparation made from the leaves of 3 week old *Arabidopsis* seedlings.

The chlorophyll auto-fluorescence signal from the intact *A. thaliana* protoplasts is shown in Figure 7.7a left panel and the mito-tracker signal is shown Figure 7.7b middle panel. The merged chlorophyll auto fluorescence signal and mito-tracker signal is shown in Figure 7.7c right panel. The markers for the chloroplasts (chlorophyll fluorescence) and the mitochondria (mito-tracker) are distinct (Figure 7.7c).

a) Chlorophyll fluorescence b) Mito-tracker c) Overlay

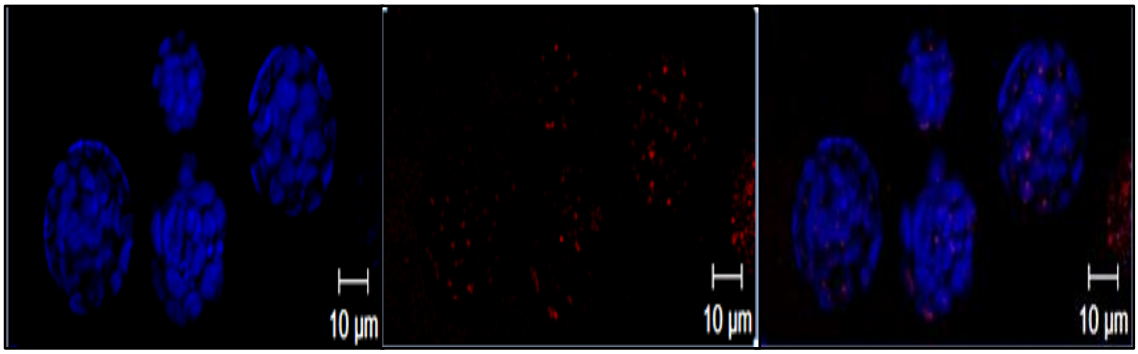


Figure 7.7: Confocal microscopy images of intact *A. thaliana* mesophyll protoplasts stained with Mitotracker Red CMXRos.

The chlorophyll fluorescence (a, blue), the mito-tracker staining (b, red) and the overlay (c) in the same cell. Scale bar = 10 µm.

The chlorophyll fluorescence signal of protoplasts expressing the pDH51-GW-YFPn and pDH51-GW-YFPc vectors are shown in Figure 7.8a. The YFP signal generated by the expression of LEA5-YFPn and LEA5-YFPc is shown Figure 7.8. The merged YFP and chlorophyll fluorescence signal generated by the expression of LEA5-YFPn and LEA5-YFPc is shown in Figure 7.8. There is no clear overlap between the YFP and chlorophyll fluorescence signals.

a) Chlorophyll fluorescence b) YFP emission c) Overlay

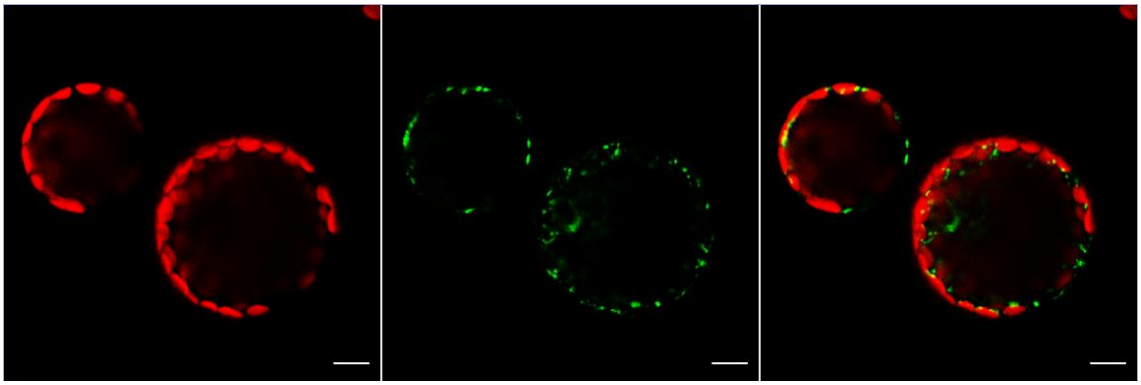


Figure 7.8: Confocal microscopy images of intact *A. thaliana* mesophyll protoplasts transiently expressing a) LEA5-YFPn with LEA5-YFPc.

The chlorophyll fluorescence (a, red), the YFP (b, green) and the overlay (c) in the same cell. Scale bar = 10 µm.

7.2.6 Expression of the DEA (D/H)-box RNA helicase 22 protein in protoplasts

The RH22 cDNA was cloned into the pDH51-GW-YFPn and pDH51-GW-YFPc vectors. Transient expression of the fusion protein in intact *A. thaliana* protoplasts was monitored using the Zeiss LSM700 inverted confocal microscopy. The chlorophyll fluorescence signal is shown in Figure 7.9a. The YFP signal generated by the expression of RH22-YFPn and RH22-YFPc is shown (Figure 7.9b). The merged YFP and chlorophyll fluorescence signals are shown in Figure 7.9c. There is a clear overlap between the YFP and chlorophyll fluorescence signals (Figure 7.9).

a) Chlorophyll fluorescence b) YFP emission c) Overlay

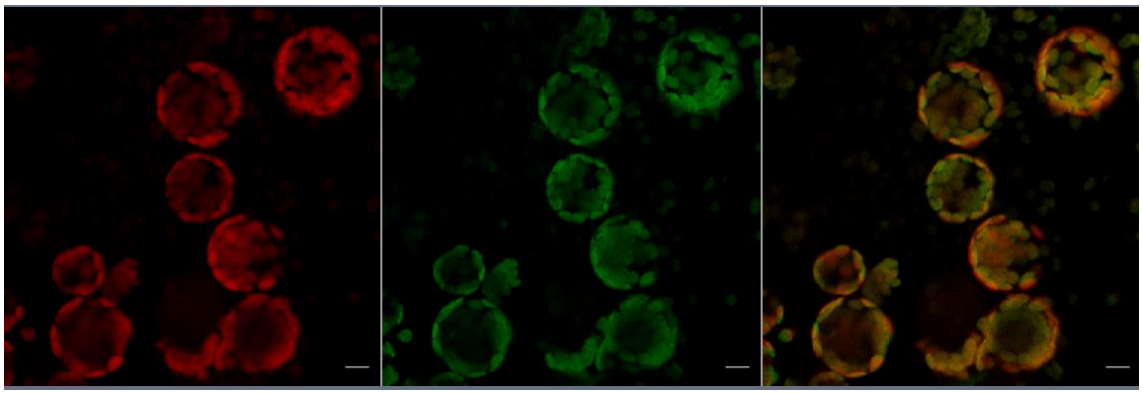


Figure 7.9: Confocal microscopy images of intact *A. thaliana* mesophyll protoplasts transiently expressing DEA (D/H)-box RNA helicase 22-YFPn and DEA (D/H)-box RNA helicase 22-YFPc.

The chlorophyll fluorescence (a, red), the YFP (b, green) and the overlay (c) in the same cell. Scale bar = 10 µm.

7.2.7 Expression of the chloroplast (APP2-GFP) and cytosolic (APP1-GFP) marker proteins in *A. thaliana* mesophyll protoplasts

The marker protein, APP2-GFP, which localises to the chloroplasts, was used as a control in these studies. The chlorophyll fluorescence signal from protoplasts expressing APP2-GFP is shown in Figure 7.10a. The APP2-GFP fluorescence is shown in Figure 7.10b. The merged GFP and chlorophyll fluorescence signal is shown in Figure 7.10c. There is an overlap between the GFP and chlorophyll fluorescence signals (Figure 7.10c). This is clearly visible at the periphery of the protoplasts. However, there are some yellow areas that indicate an overlap of the signals, even though it is known that APP2 localises to chloroplasts.

a) Chlorophyll Fluorescence b) GFP emission c) Overlay

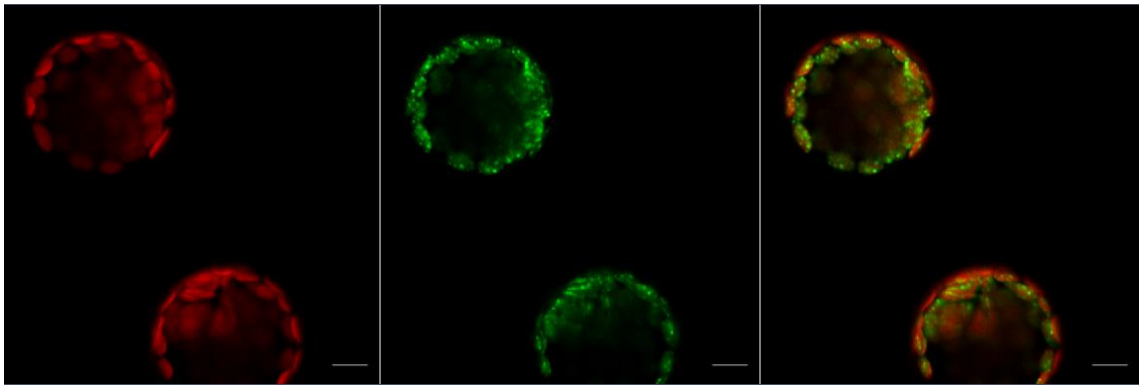


Figure 7.10: Confocal microscopy images of intact *A. thaliana* mesophyll protoplasts transiently expressing amino peptidase P2 (APP2; At3g05350). The chlorophyll fluorescence (a, red), the GFP (b, green) and the overlay (c) in the same cell. Scale bar = 10 µm.

The second marker protein, APP1-GFP, which localises to the cytosol, was used as a second control in these studies. The chlorophyll fluorescence signal for protoplasts expressing APP1-GFP is shown Figure 7.11a. The APP2-GFP fluorescence signal is shown Figure 7.11b. The merged GFP and chlorophyll fluorescence signals are shown in Figure 7.11c. The data in Figure 7.11c shows that there is no overlap between the GFP and chlorophyll fluorescence signals. This finding was in line with our predictions because APP1 localises to the cytosol and not the chloroplasts.

a) Chlorophyll fluorescence b) GFP emission c) Overlay

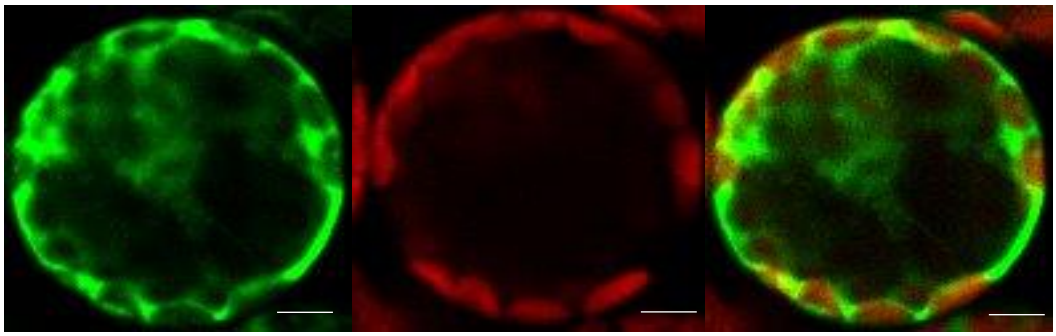
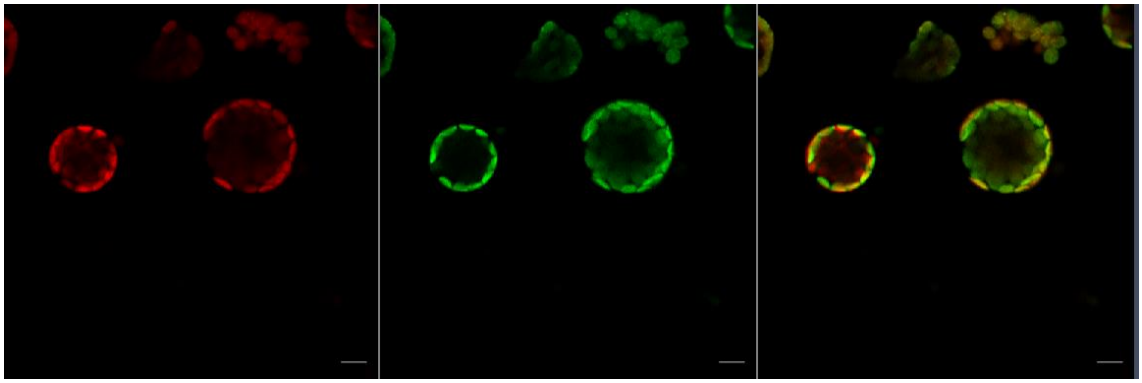


Figure 7.11: Confocal microscopy images of intact *A. thaliana* mesophyll protoplasts transiently expressing amino peptidase P1 (APP1; At4g36760). The chlorophyll fluorescence (a, red), GFP (b, green) and overlay (c) in the same cell. Scale bar = 5 µm.

7.2.8 Interactions of the LEA5 protein with RH22

Interaction between LEA5 and chloroplast RH22 was investigated using the split-YFP transient expression system. The chlorophyll fluorescence signal from intact *A. thaliana* mesophyll protoplasts transiently expressing LEA5-YFPn, together with RH22-YFPc is shown Figure 7.12a (upper panel). YFP signal generated by these protoplasts is shown in Figure 7.12b (upper panel). The data in Figure 7.12c (upper panel) reveal an overlap between the YFP and chlorophyll fluorescence signals. The yellow signal observed in Figure 7.12c indicates that both LEA5 and RH22 interact in the chloroplasts. The detail of this interaction is also seen in Figure 7.12 (lower panel), which shows the same images at higher magnification.

a) Chlorophyll fluorescence b) YFP emission c) Overlay



a) Chlorophyll fluorescence b) YFP emission c) Overlay

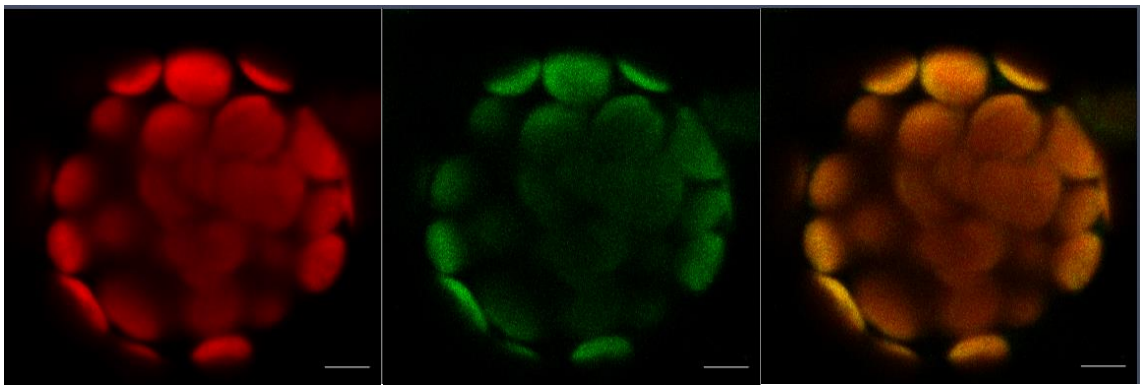


Figure 7.12: Confocal microscopy images of intact *A. thaliana* mesophyll protoplasts transiently expressing LEA5-YFPn with DEA (D/H)-box RNA helicase 22-YFPc.

Upper panel showed interactions at 20x magnification (upper panel). Scale bar = 10 µm. Lower panel showed interactions at 40 lens indicating chlorophyll fluorescence (a, red), YFP (b, green) and the overlay (c) in the same cell. Scale bar = 5 µm.

To confirm the interaction between LEA5 and RH22 in the chloroplasts, transient expression assays were performed using the LEA5-YFPc and RH22-YFPn probes (Figure 7.13). The chlorophyll fluorescence signal shown in Figure 7.13a and the YFP signal generated by the expression of LEA5-YFPc and RH22-YFPn shown in Figure 7.13b show a high degree of overlap (Figure 7.13c). This overlap can be seen more clearly in single protoplast images from this experiment (Figure 7.14)

a) Chlorophyll fluorescence b) YFP emission c) Overlay

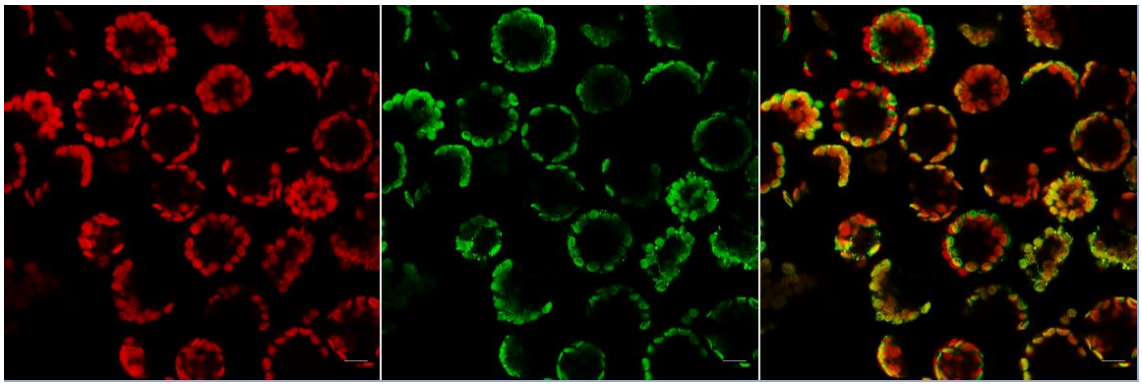


Figure 7.13: Confocal microscopy images of intact *A. thaliana* mesophyll protoplasts transiently expressing LEA5-YFPc with DEA (D/H)-box RNA helicase 22-YFPn using a 40 lens.

The chlorophyll fluorescence (a, red), the YFP (b, green) and the overlay (c) in the same cell. Scale bar = 10 µm.

a) Chlorophyll fluorescence b) YFP emission c) Overlay

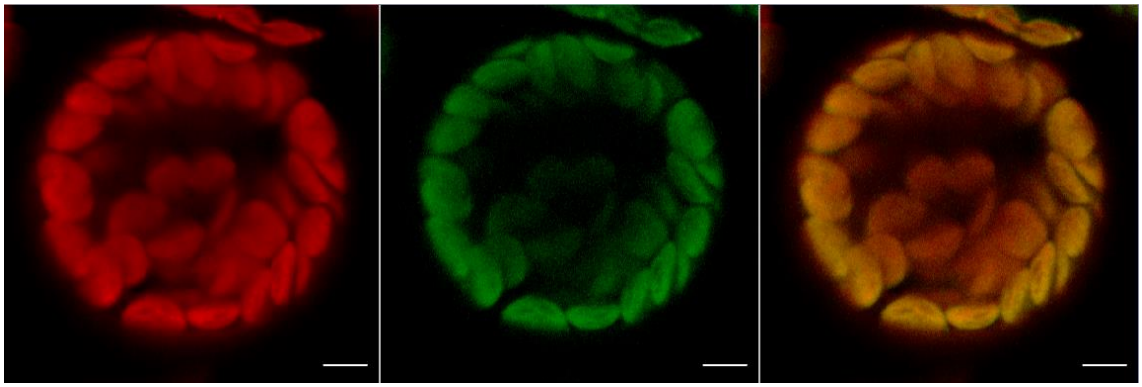


Figure 7.14: Confocal microscopy images of a single intact *A. thaliana* mesophyll protoplast transiently expressing LEA5-YFPc with DEA (D/H)-box RNA helicase 22-YFPn using 20 lens magnifications.

Chlorophyll fluorescence (a, red), YFP (b, green) and the overlay (c). Scale bar = 5 µm.

7.2.9 Interactions of the WHY1 protein with RH22

WHY1 was cloned into the pDH51-GW-YFPn and pDH51-GW-YFPc vectors. Transient expression of the fusion protein in intact protoplasts was monitored using confocal microscopy. The chlorophyll florescence signal from the intact *A. thaliana* mesophyll protoplasts transiently expressing WHY1-YFPn, together with RH22-YFPc is shown in Figure 7.15a (upper panel). The YFP signal generated in these protoplasts is shown in Figure 7.15b (upper panel). The data in Figure 7.15c (upper panel) reveals an overlap between the YFP and chlorophyll fluorescence signals. The data shown in Figure 7.15 (upper panel) indicate that both LEA5 and RH22 interact in the chloroplasts. The detail of this overlap is further confirmed in Figure 7.15 (lower panel), which shows the interaction of WHY1-YFPc together with RH22-YFPn. The chlorophyll florescence signal is shown in Figure 7.15a (lower panel). The YFP signal generated by these protoplasts is shown in Figure 7.15b (lower panel). The data in Figure 7.15c (lower panel) reveal an overlap between the YFP and chlorophyll fluorescence signals in the chloroplasts.

a) Chlorophyll fluorescence b) YFP emission c) Overlay

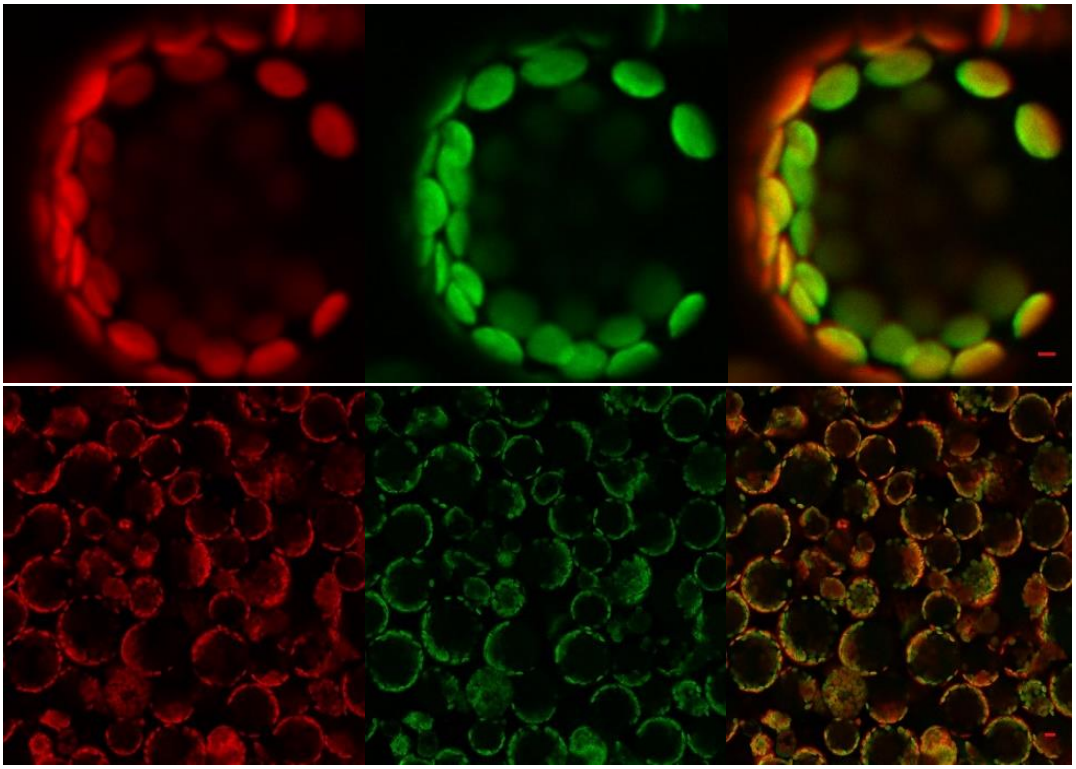


Figure 7.15: Confocal microscopy images of a single intact *A. thaliana* mesophyll protoplast transiently expressing AtWHY1YFPn with DEA (D/H)-box RNA helicase 22-YFPc.

Upper panel showed interactions at 40x magnification (upper panel). Scale bar = 10 µm. Lower panel showed interactions at 20 lens indicating chlorophyll fluorescence (a, red), YFP (b, green) and the overlay (c) in the same cell. Scale bar = 5 µm.

As a negative control, LEA5-YFPn was expressed in the same protoplasts as RH22-YFPn**.** In this case, no YFP signal was generated in the intact protoplasts (Figure 7.16a) although the chlorophyll fluorescence signal was clearly visible (Figure 7.16b).

a) Chlorophyll fluorescence b) YFP emission c) Overlay

Figure 7.16: Confocal microscopy images of intact *A. thaliana* mesophyll protoplasts transiently expressing LEA5-YFPn with DEA (D/H)-box RNA helicase 22-YFPn.

The chlorophyll fluorescence (a, red), YFP (b, no signal) and the overlay (c) in the same cell. Scale bar = 10 µm.

7.3 Discussion

LEA5 was first identified in a screening for novel genes involved in oxidative stress tolerance (Mowla et al., 2006). This protein was shown to localise to the mitochondria of transformed *Arabidopsis* plants expressing a *SAG21*-YFP fusion protein driven by the CaMV 35S promoter (Salleh et al., 2012). In agreement with this study, LEA5 also named LEA38 by Candat et al. 2014 was one of three LEA3 family proteins (LEA37, LEA38, and LEA41) that was found to be targeted to the mitochondria.

The data obtained from available web-based prediction tools (Table 7.1) suggest the possibility that the LEA5 protein could also be targeted to plastids, but with a lower probability than targeting to mitochondria. Moreover, a previous PhD student, Daniel Shaw, undertook a tandem affinity purification (TAP) tagging study to identify proteins that interacted with LEA5 in the Geert De Jaeger lab, Department of Plant Systems Biology, VIB, Gent, Belgium. These studies consistently showed that LEA5 interacted with chloroplast proteins. The chloroplast RH22 was identified as a significant LEA5 binding partner in all experiments.

RNA helicases are required for the rearrangement of ribonucleoproteins and the regulation of gene expression. The RH22 is required for the accumulation of mRNAs from plastid genes. The *hs3-1* mutants that lack a functional of RH22 protein exhibit defects in plastid development (Kanai et al., 2013). Two RNA polymerases, the plastid-encoded RNA polymerase (PEP) and the nuclear-encoded RNA polymerase (NEP) are required for the transcription of chloroplast genes. The transcripts that were changed in abundance in *hs3-1* mutants included both NEP- and PEP-dependent genes, suggesting that this helicase is involved in posttranscriptional regulation, that are distinct from NEP- or PEP-dependent transcription (Kanai et al., 2013). Thus, if LEA5 binds to RH22, then it may regulate or influence the accumulation of chloroplast mRNAs.

In the study reported here, the intracellular location of LEA5 was first studied in leaves protoplasts from the same *LEA5*-YFP-expressing lines that had been used in an earlier study (Salleh et al., 2012). The images presented in Figure 7.1 and Figure 7.2 indicate that the LEA5-YFP fluorescence signal overlaps with the chlorophyll auto-fluorescence signal, suggesting that LEA5 can localise to plastids. Furthermore, the data presented in Figure 7.3 show that LEA5-YFP localises to both of the chloroplast and mitochondria. Further confirmation of the chloroplast localisation of LEA5 was carried out using 35S-LEA5-YFP plus SSU-RFP. The data shown in Figure 7.5 indicate that 35S-LEA5-YFP localises in the chloroplasts together with SSU-RFP. In addition, the data shown in Figure 7.4 indicate that 35S-LEA5-YFP and the mitochondria targeted alternative oxidase1 (AOX-RFP) co-localise in the mitochondria. However, there is as yet no simple explanation for why the plastid localisation was not observed in earlier studies (Candat et al., 2014, Salleh et al., 2012). Previous studies showed that stable expression of LEA5-YFP localises only to the mitochondria (Salleh et al., 2012). However, in this study, LEA5 was shown to be localised in the chloroplasts of isolated protoplasts. Thus, LEA5 expression in chloroplasts may only be transient. The stability of the LEA5 protein in chloroplasts may be determined by many factors including the availability of binding partners. In contrast, we may suppose that the LEA5 protein is much more stable in mitochondria as shown in a previous publication (Salleh et al., 2012). Therefore, further experiments, involving the uptake of LEA5 protein into chloroplasts and mitochondria should be performed.

The split-YFP assay system can be used to study the intracellular location proteins. Experiments were performed using a split-YFP assay system to determine whether the LEA5 protein localises to the chloroplasts of protoplasts transiently expressing the two halves of the protein in pDH51-GW-YFPn and pDH51-GW-YFPc vectors (Figure 7.8). In this case, the merged YFP and chlorophyll fluorescence signals did not show any marked areas of overlap (Figure 7.8). However, a similar situation was observed when the marker protein APP2-GFP which localises to the chloroplasts, was used as a control. In this case, the chlorophyll fluorescence signal in protoplasts expressing APP2-GFP

showed little overlap with the GFP signal (Figure 7.10). In contrast, however, the YFP signal generated by the expression of RH 22-YFPn and RH22-YFPc showed good overlap with the chlorophyll fluorescence signal (Figure 7.9), as would be predicted from the known chloroplast location of this protein.

The split-YFP assay system was also used to study the interactions between proteins *in situ* because both interacting partners have to be localised in the same compartment in order for interactions to occur. In this way, the interaction between LEA5 and chloroplast RH22 was demonstrated (Figure 7.12) using YFP fluorescence emission. There was an overlap between the YFP and chlorophyll fluorescence signals, showing that this interaction occurred in the chloroplasts (Figure 7.13 and Figure 7.14). However, as discussed above neither the LEA5 protein nor the chloroplast APP2-GFP protein was found to localise to the chloroplasts in comparable experiments (Figure 7.13, Figure 7.14) suggesting that self-assembly between two halves of YFP is unlikely. However, further controls should be performed to confirm this point. For example, additional negative controls might involve a mutated version (Miller et al., 2015) of either the LEA protein or the RH22 protein in which the amino acids involved in binding have been mutated. Using such controls, it will be possible to estimate the appropriate signal-to- noise ratio in these split-YFP experiments and quantify the signals. Taken together, the results presented here show for the first time, the interaction between LEA5 and RH22 in the plastids This binding may play a crucial role in post-transcriptional regulation and gene expression under stress conditions.

The binding of LEA5 to RH22 is interesting and merits further exploration. For example, the presence of the LEA5 protein in chloroplasts might be a response to stress, the subsequent interaction between LEA5 with RH22 leading to altered stability of chloroplast mRNAs. It is therefore important to establish whether the activity of the helicase, which is involved in plastid ribosome assembly and other aspects of RNA metabolism, is altered by LEA5-binding. A key question concerns the role of this interaction in plant response to stress. An interaction between the

chloroplast RH22 and WHY1 was demonstrated using the same split-YFP transient expression system (Figure 7.15). This interaction occurred in the chloroplasts (Figure 7.15). In addition, young *hs3-1* seedlings exhibited a pale greening phenotype but older seedlings (20-day old) had a comparable phenotype to the wild type (Chi et al., 2012). This phenotype was very similar to the phenotype of WHY1-deficient barley seedlings. The leaves of 7-day old WHY1-deficient barley seedlings (W1-1, W1-7) showed a “delayed greening” phenotype (Figure 4.1), but at 14 days the developing leaves are similar to the wild type. The delay in greening phenotype at a young age both in mutants lacking the RH22 and WHY1-deficient barley seedlings is caused by the aberrant splicing of rRNAs (23S rRNA + 4.5S rRNA) for the large subunit (50S) of the plastid ribosome (Chi et al., 2012). The RH22 is required for the accumulation of plastid mRNAs during seed development and seedling growth, as well as ensuring seed oil biosynthesis by maintaining plastid mRNA levels (Kanai et al., 2013). WHY1 may interact with RH22 to regulate its functions in RNA processing.

Chapter 8 . General discussion

Crop production worldwide is inherently dependent on photosynthesis and the acclimation of the photosynthetic processes to changing environmental conditions. In this thesis, the development of photosynthesis in barley leaves was explored by focussing on the roles of WHY1 and LEA5 in the chloroplast. Using RNA-seq and microarrays, these studies have characterised the transcript profiles of the embryos of dry seeds and young leaves in the WHY1-deficient barley lines in comparison to the wild type. Moreover, the changes in gene expression profiles observed in the developing leaves were related to changes in the metabolite profiles and the abundance of pigments and specific proteins to gain deeper insights into the role of WHY1 in chloroplast development and chloroplast to nucleus signalling. In addition, the role of WHY1 in seed production was characterised. The binding of WHY1 to the RH22 was also characterised as was the binding of RH22 to the LEA5 protein. These studies allow deeper insights into the functions of these two proteins in chloroplasts. Several conclusions can be drawn from these studies, as discussed in detail below:

8.1 WHY1 plays a key role in chloroplast development in barley leaves

Transcripts encoding photosynthetic proteins were highly expressed even in the embryos of the dry seeds of the W1-7 lines. While leaf growth was similar to the wild type in the WHY1-deficient lines, WHY1-deficient barley leaves show a “delayed greening” phenotype in a strictly developmental manner. The delay in the establishment of photosynthesis was characterised by the slow accumulation of photosynthetic pigments, as well as a delayed accumulation of transcripts encoding photosynthetic proteins and in the levels of specific photosynthetic proteins. The delay in the establishment of photosynthesis was marked in 7-day

old leaves but less pronounced in leaves at 14 days. Therefore, the processes that require WHY1 occur early in chloroplast biogenesis but are less important at the later stages. These observations allow an analysis of WHY1 functions in the chloroplast. It is well-known that this protein binds to single-stranded DNA (Desveaux et al., 2000) and that it is required for the production of functional chloroplast ribosomes in maize leaves (Prikryl et al., 2008). The data presented here show that WHY1 is also required for the accumulation of chloroplast ribosomal proteins in barley leaves at the early stages of development. Moreover, the data presented show that WHY1 is required for the splicing of 23S rRNA and 4.5S rRNA, which is necessary for the formation of 50S large subunit of the plastid ribosomes in barley as it is in maize. However, in contrast to the situation in maize, where the leaves never accumulate chlorophyll or functional chloroplasts, barley leaves are merely delayed in greening. Furthermore, the maize phenotype could be due at least in part to the nature of the transposon insertion. In contrast, the WHY1-deficient lines produced by RNAi are relatively free from such complications. In conclusion, these data show that WHY1 is required for the translation of transcripts produced in the plastids.

Moreover, WHY1 is required for the transcription of chloroplast-encoded genes. Specifically, transcripts encoding key photosynthetic proteins such as D1 and the large subunit of RuBiSCO were much lower in the WHY1–deficient lines than the wild type. A key question concerns how WHY1 is able to exert such a strong influence over plastid transcription and translation. One possible answer is the observed binding of WHY1 to the RH22 protein in the chloroplasts. The *Arabidopsis rh22* mutants exhibit a pale-green phenotype similar to the developing leaves in the WHY1-deficient lines (Kanai et al., 2013). Moreover, like the WHY1-deficient lines, the mature leaves of the *rh22* mutants *Arabidopsis* were similar to the WT plants. Knock-down of *RH22* leads to a delayed-greening phenotype with defects in plastid development resulting in the slow growth in seed development (Kanai et al., 2013). The *rh22* phenotype is caused by aberrant splicing of 23S rRNA and 4.5S (Chi et al., 2012). Crucially, the data identified that WHY1 is a new protein that influences leaf development and chloroplast biogenesis in barley leaves, thus, it may function to integrate gene expression in

the chloroplasts and nucleus. Besides that, it can be a great tool to increase crop productivity as WHY1-deficient lines produced more seed yields in the T4 generation despite deficient in photosynthesis as well as in photosynthesis. However, further experiments will be required to understand why WHY1-deficient lines produce more seed. For accuracy, seed yields should be measured through successive generations to make sure that the traits are accurately related to the presence of transgene and not to other factors that are associated with the transformation procedures, tissue culture etc that will alter fertility. The metabolomics profiling analysis revealed that many metabolites associated with sugar and respiratory metabolism, and nitrogen assimilation were changed in the absence of WHY1, particularly in the basal regions of the leaves. WHY1 is, therefore, an important regulator of metabolism during chloroplast development. Moreover, many chloroplast-targeted proteins, such as ribosomal proteins were affected by the loss of WHY1. The chloroplast ribosomal proteins are essential in the production key proteins required for the photosynthetic apparatus during early chloroplast development. Hence, the increased abundance of transcripts encoding these proteins may be an attempt to compensate for the failure to establish the photosynthetic apparatus in the absence of WHY1.

It can be concluded that WHY has multiple functions in the chloroplasts and plays diverse roles in chloroplast transcription, translation, the regulation of DNA copy number and post-translational mechanisms, as well as RNA processing. The findings also suggest that the chloroplast development in the WHY1-deficient barley seedlings is postponed/delayed rather than prevented because mature chloroplasts gradually form as the WHY1-deficient leaves undergo development. It has also been demonstrated in this thesis that WHY1 plays a role in the coordination of nuclear and plastome gene expression related to the production of photosynthetic proteins during chloroplast biogenesis.

In the future, it would be interesting to characterise the structure of the young leaves of the WHY1-deficient lines in comparison to the wild type. For example, leaf ultrastructure may be studied using transmission electron micrographs. It is

also important to study the number of plastids/chloroplasts per cell and the plastid/chloroplast size in the young leaves of the WHY1-deficient lines in comparison to the wild type. Moreover, it is crucial to characterise the proteins that interact with WHY1. Knowledge of the WHY1 binding partners and the sequences involved will greatly help understanding the functions of WHY1 bindings in the chloroplast. This will be particularly interesting to study in the future. It would be also interesting to study how WHY1 binding to RH22 alters its processing activity and whether WHY1 barley protein interacts with barley RH22 in the same way as in *Arabidopsis*. In addition, it will be interesting to characterise the high light responses of WHY1-deficient seedlings. WHY1 has been shown here to have a role in chloroplast development. Thus, it would be interesting to study how the responses to other factors that play key roles in chloroplast biogenesis are changed in the WHY1-deficient seedlings. Lincomycin (LINC) and other inhibitors such as norflurazon (NF) can be used to study chloroplast-to-nucleus signalling pathways in the different lines. These inhibitors induce strong oxidation of the cytosol and nuclei, then the repression of nuclear gene expression form the basis of a screen, which revealed the plastid-localised GENOMES UNCOUPLED (GUN) pathway (Karpinska et al., 2017). The *gun* mutants retain nuclear gene expression of chloroplast-associated genes after the treatment with NF (Susek et al., 1993). The expression of nuclear-encoded *PHOTOSYNTHESIS-ASSOCIATED NUCLEAR GENES* (*PhANGs*) such as genes encoding components of the photosynthetic electron transfer chain (*LIGHT-HARVESTING CHLOROPHYLL A/B BINDING COMPLEX* (*LHC*) or is down-regulated when chloroplast biogenesis is disturbed (Hess et al., 1994, Ruckle et al., 2007, Susek et al., 1993). The loss of GUN1 protein functions during protein import in chloroplast cause uncouple *PhANG* expression from the developmental state of the plastids (Wu et al., 2019).

In summary, the findings presented in this thesis have characterised chloroplast development in barley leaves and how it is controlled by WHY1. These results have increased our understanding of chloroplast development and the role of WHY1 in the coordination of nuclear and plastid-encoded photosynthetic genes. The new knowledge on WHY1 functions in the establishment of photosynthesis

in barley hence lays the foundations for further experiments seeking to explore chloroplast to nucleus retrograde signalling mechanisms under optimal and stress conditions. These findings also have wider implications because they will ultimately be useful to researchers seeking to improve crop production and resilience, as well as food security worldwide.

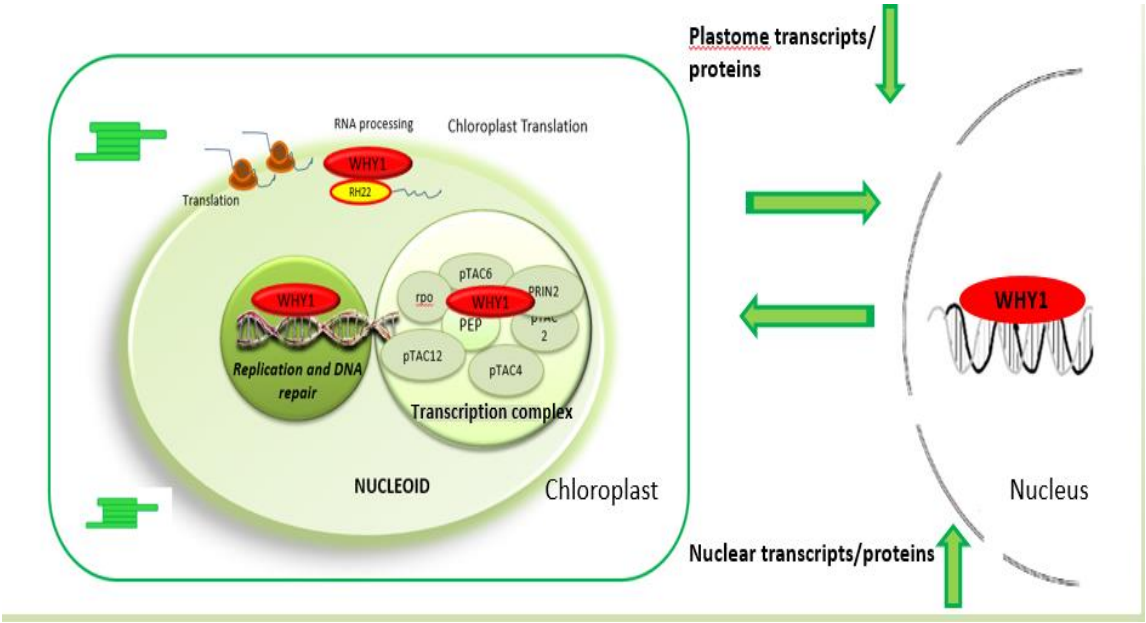


Figure 8.1: Summary of key findings that were achieved in the role of WHY1 in chloroplasts.

The WHY family is single-stranded DNA binding proteins and dually localised both to nucleus and chloroplast (Desveaux et al., 2000, Grabowski et al., 2008). In this thesis, it has been shown that WHY1 is required for chloroplast development in barley leaves. Moreover, WHY1 is required only at the early stages of developing leaf. The findings also demonstrated that WHY1 is necessary for a transition from proplastids to mature chloroplasts regulating nuclear and plastid-encoded gene expression and protein synthesis, which the levels of plastid transcripts and proteins were lower in the WHY1-deficient lines. In contrast, the levels of nuclear-encoded chloroplasts transcripts and proteins were significantly higher in the WHY1-deficient leaves. In addition to WHY binding to the DNA in chloroplast nucleoids (Krupinska et al., 2014), this finding demonstrates that WHY1 protein interacts with RH22 and may regulate its functions in RNA processing. RH22 is required for the splicing of chloroplast ribosomal proteins (Chi et al., 2012). The WHY1-deficient barley seedlings showed aberrant splicing of 23S rRNA and 4.5S rRNA, hence, impairs chloroplast ribosome functions as well as the transcription of plastome genes. It also can be concluded that the loss of WHY1 functions disrupts chloroplast to nucleus signalling during barley leaf development.

8.2 The role of LEA5 in chloroplast development

The aim of the studies on LEA5 was to identify interacting proteins. LEA5 was first identified as a protein induced by and confer tolerance to oxidative stress (Mowla et al., 2006). *AtLEA5* is constitutively expressed in roots but not in leaves and seeds (Mowla et al., 2006). Interestingly, *AtLEA5* transcripts exhibit a diurnal pattern of regulation, where transcripts abundant in the dark are repressed in the light. Expression of *AtLEA5* in leaves is induced in leaves in the light by exposure to abiotic and biotic stresses, as well as by oxidants and phytohormones (Mowla et al., 2006, Salleh et al., 2012). Constitutive overexpression of *AtLEA5* increased root growth and shoot biomass, whereas both of these processes were impaired in antisense lines or *atlea5* knockout mutants (Salleh et al., 2012). While *LEA5*-YFP expression studies localised LEA5 only to mitochondria (Mohd Salleh et al., 2012), the analysis of protein-protein interactions using tandem affinity purification (TAP) showed that LEA5 interacts with proteins that are predominantly localised in chloroplasts, particularly the RH22, which is essential for RNA processing in plastids. RH22 is localised to plastids and is required for chloroplast development (Chi et al., 2012, Kanai et al., 2013). Using a split YFP approach, LEA5 was shown to interact with RH22 in chloroplasts, suggesting this protein has functions in chloroplasts, particularly in plants exposed to stress. The binding of LEA5 with RH22 may result in alterations of RH22 functions. This would only occur in the light when exposed to stress. Thus, LEA binding to RH22 may regulate translation in the chloroplasts resulting in the downregulation of photosynthesis. This hypothesis should be tested in future work, for example, in the *Arabidopsis* lines constitutively overexpressing *LEA5*. Hence, LEA5 may play a specific role in the downregulation of photosynthesis in response to oxidative stress or drought stress (Mowla et al., 2006). This concept is consistent with the finding that under stress conditions, decreases in photosynthesis are often observed. Further studies are required to answer this question and also whether LEA5 is imported into chloroplasts. Moreover, the binding of LEA5 to RH22 requires further exploration and characterisation, particularly with regards to the

effects of LEA5 on RNA processing in the chloroplasts and on translation in mitochondria and chloroplasts.

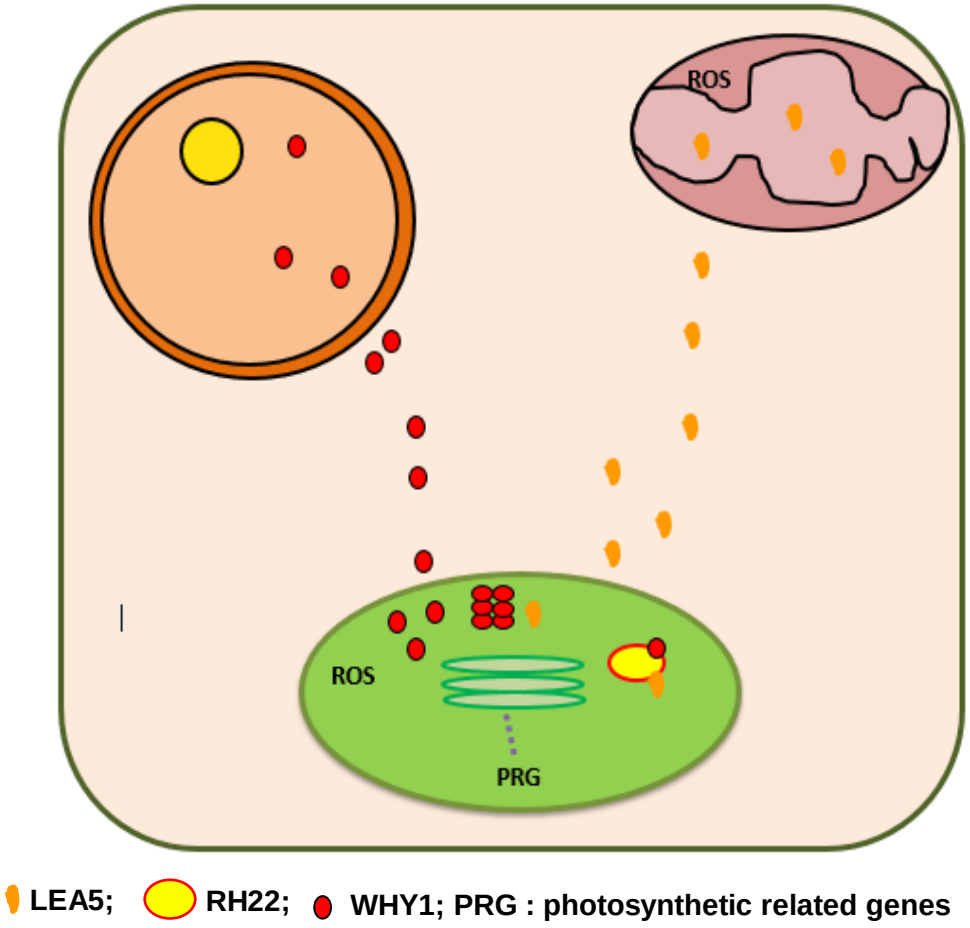


Figure 8.1: Summary of key findings that were achieved in the role of LEA5 in chloroplasts.

The WHY1 protein is a dual location protein targeted in the plastids and nucleus of the same cell (Grabowski et al., 2008) and the RH22 protein is located in plastids (Kanai et al., 2013). The LEA5 protein is located in the mitochondria (Mohd Salleh et al., 2012). The data presented here show that by using split-YFP approach, the LEA5 interacted with RH22 in chloroplasts. This binding influences translation in the chloroplasts resulting in the downregulation of photosynthesis during stress. In addition, the *Arabidopsis* WHY1 protein was also shown to interact with RH22 using split-YFP transient expression system thus may regulates its functions in RNA processing.

Appendix A

A.1 Transcripts that were involved RNA and DNA binding activities in the WHIRLY deficient embryos (W1-7) relative to the wild type.

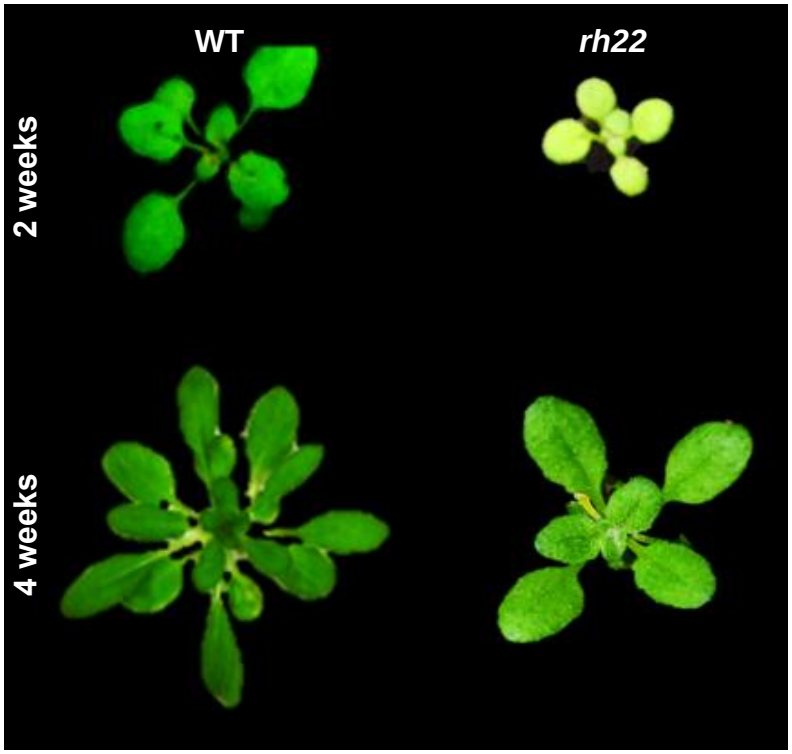
Description	Accession	Relative expression (log2FC)
bHLH transcription factor	AK368273	4.4
Aquaporin 7	MLOC_13871.1	4.3
Gnk2-homologous domain	AK376143	4.5
Leucine-rich repeat	MLOC_71648.1	4.5
DNA-binding WRKY	AK367216	1.6
WRKY transcription factor 32	MLOC_66134.2	3.1
WRKY transcription factor 40	MLOC_10687.2	-1.7
Myb transcription factor	AK367954	4.8
Myb domain protein	MLOC_7426.1	4.0
Myb domain protein	MLOC_52439.6	2.3
MYB-related transcription factor	MLOC_7981.1	2.1
Myb family transcription factor	MLOC_8187.2	1.7
Myb family transcription factor	AK356219	6.6
Myb transcription factor	MLOC_9835.2	-1.9
Single myb histone 6	MLOC_34636.1	-1.9
Myb family transcription factor	MLOC_53628.1	-2.5

Description	Accession	Relative expression (log2FC)
Zinc finger protein	MLOC_78744.1	6.2
Zinc finger protein	AK370626	6.5
Zinc finger protein	AK366628	5.5
Zinc finger protein	AK358990	4.6
Zinc finger protein	AK376150	2.2
Zinc finger protein	AK355626	2.1
Zinc finger	MLOC_2875.1	2.1
Zinc finger	MLOC_53961.1	2
Zinc finger	MLOC_57429.1	2
Zinc finger	AK369046	1.9
Zinc finger	MLOC_19593.3	1.8
Zinc finger	MLOC_55252.1	1.8
Zinc finger protein	MLOC_50966.5	1.7
Zinc finger	AK362870	1.7
Zinc finger	MLOC_9929.1	1.7
Zinc finger	MLOC_74185.1	1.7
Zinc finger	AK355820	1.7
Zinc finger	AK362022	1.6
Zinc finger	AK365791	1.6
Ring finger family protein	AK366156	1.6
Zinc finger protein	AK372198	1.6
Zinc finger	MLOC_11609.3	1.6
Zinc finger	AK359310	1.5
Zinc finger	MLOC_61611.1	-1.7
Zinc finger	MLOC_57307.2	-1.7
FAR1-related sequence	MLOC_33258.3	-2.8
FAR1-related sequence	MLOC_30557.1	-4.9
Zinc finger	AK367522	5
Zinc finger protein	AK250895.1	4.7

Zinc finger protein	AK376680	4.7
Zinc finger protein	AK358990	4.7
Ring finger family protein	MLOC_6942.1	4.3
Zinc finger	MLOC_61836.6	4
Zinc finger	MLOC_65371.1	3.7
Zinc finger	AK363655	3.7
Zinc finger	AK370596	3.5
Zinc finger	MLOC_75538.2	3.2
Zinc finger	MLOC_40031.2	3.1
GATA transcription factor 29	MLOC_15510.1	3
Zinc finger	MLOC_58410.4	3
Zinc finger	AK364624	3
Ring finger family protein	MLOC_7401.1	3
Zinc finger	MLOC_55203.2	2.9
Ring finger family protein	MLOC_34316.2	2.9
Zinc finger	MLOC_2619.1	2.8
Zinc finger	AK375119	2.8
Ring finger family protein	MLOC_54705.1	2.7
Zinc finger	AK375174	2.7
Zinc finger	MLOC_63032.1	2.5
Zinc finger	MLOC_66027.1	2.4
Zinc finger	MLOC_65351.2	2.4
Zinc finger	MLOC_17636.1	2.4
Zinc finger	MLOC_54674.1	2.4
Zinc finger	MLOC_75092.2	2.4
Ring finger family protein	MLOC_10290.2	2.3
Ring finger family protein	MLOC_2019.1	2.3
Ring finger family protein	MLOC_32712.1	2.1
BZIP transcription factor family protein	MLOC_70302.1	5.1
BZIP transcription factor family protein	MLOC_72368.4	2.8
BZIP transcription factor family protein	AK370288	2.3
BZIP transcription factor family protein	AK373425	2
Basic-leucine zipper (bZIP)	MLOC_51930.1	1.7
Basic-leucine zipper (bZIP)	MLOC_71642.4	1.7

Basic-leucine zipper (bZIP)	MLOC_10399.1	1.6
Basic-leucine zipper (bZIP)	MLOC_63436.1	2
RNA splicing & processing		
Pre-mRNA-splicing factor 38	AK356654	1.8
Pre-mRNA splicing factor ATP-dependent RNA helicase-like protein	MLOC_16173.1	4
Telemore binding, DNA repair, chromatin organisation		
Telomere-binding protein	MLOC_57410.3	2.1
Transcription factor BIM1	MLOC_62335.1	2.2
Telomere-binding protein 1	AK376011	2
base-excision DNA repair	AK356205	1.5
Replication protein A 32 kDa subunit	MLOC_81884.2	1.8
DNA replication	MLOC_10314.1	1.6
DNA replication	MLOC_67256.3	1.5
DNA mismatch repair protein	MLOC_50820.2	-2.5

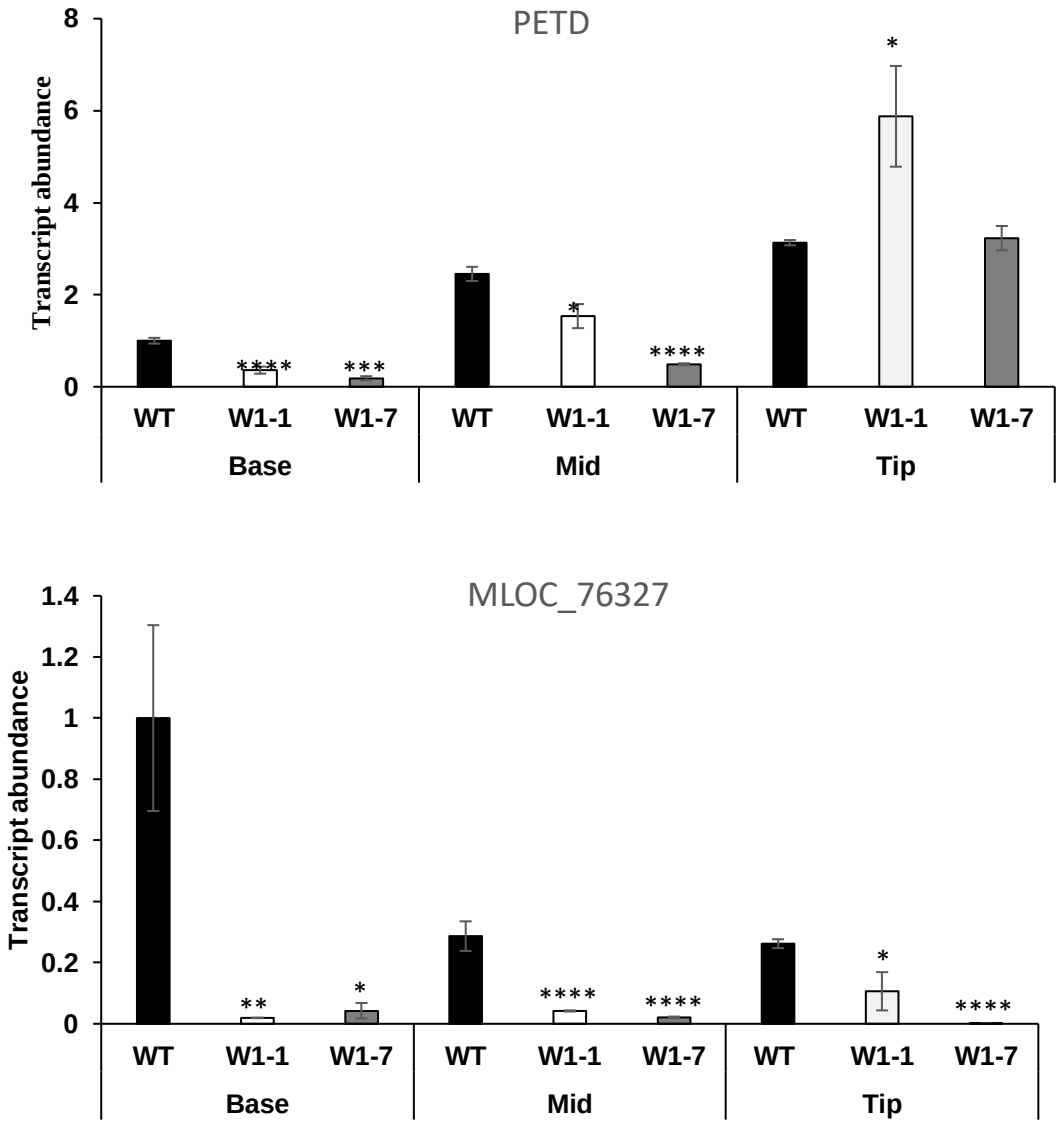
Appendix B



Comparison of shoot phenotypes of the wild type *Arabidopsis* and *rh22* at 7 and 14-day old.

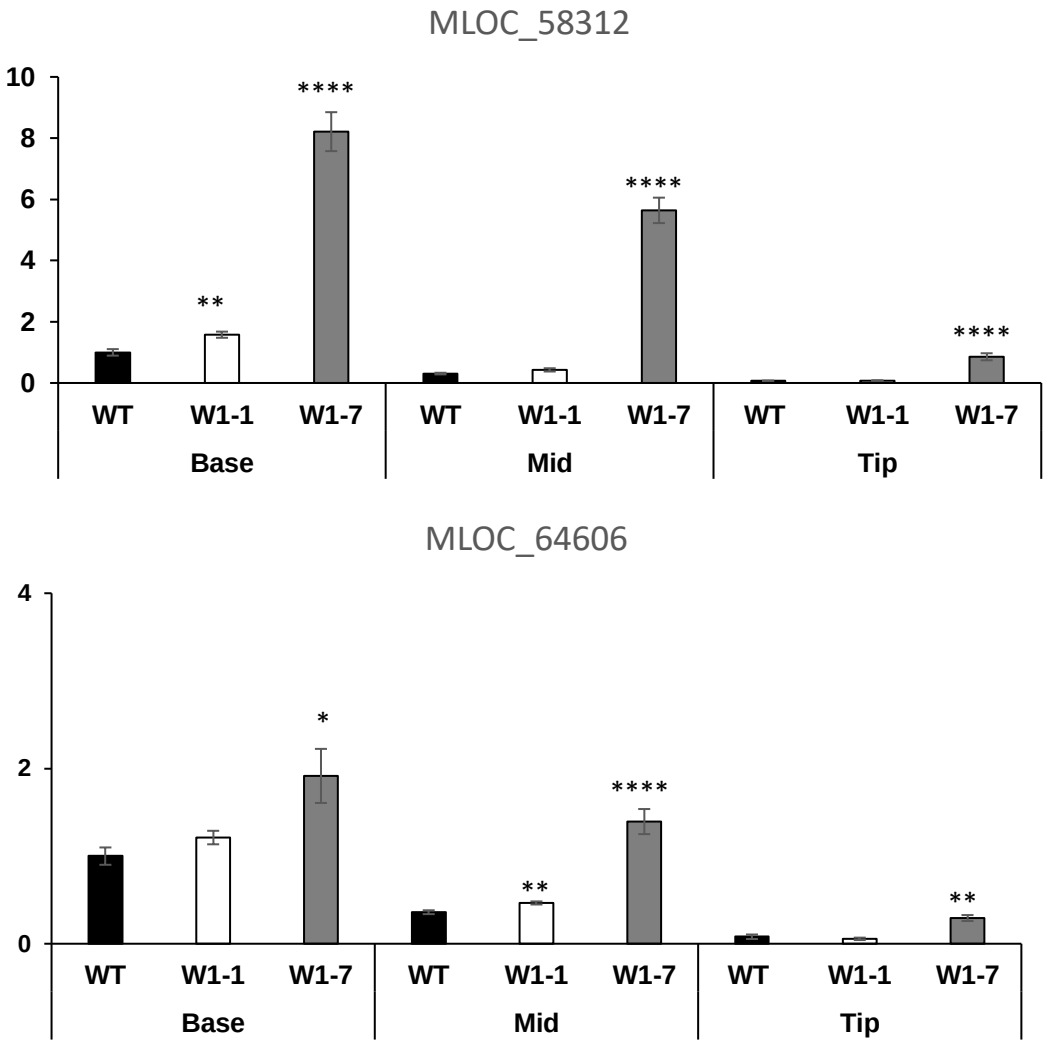
Seeds were kept at 4ºC for 3 days before the seedlings were sown in pots in soil in controlled environment chambers with a 16h light/ 8h dark photoperiod, irradiance of 200 µmol $m^{-2}s^{-1}$, 20ºC/16ºC day/night temperature regime and 60% relative humidity. (Scale bar =1 cm).

B.1 Plastid-encoded



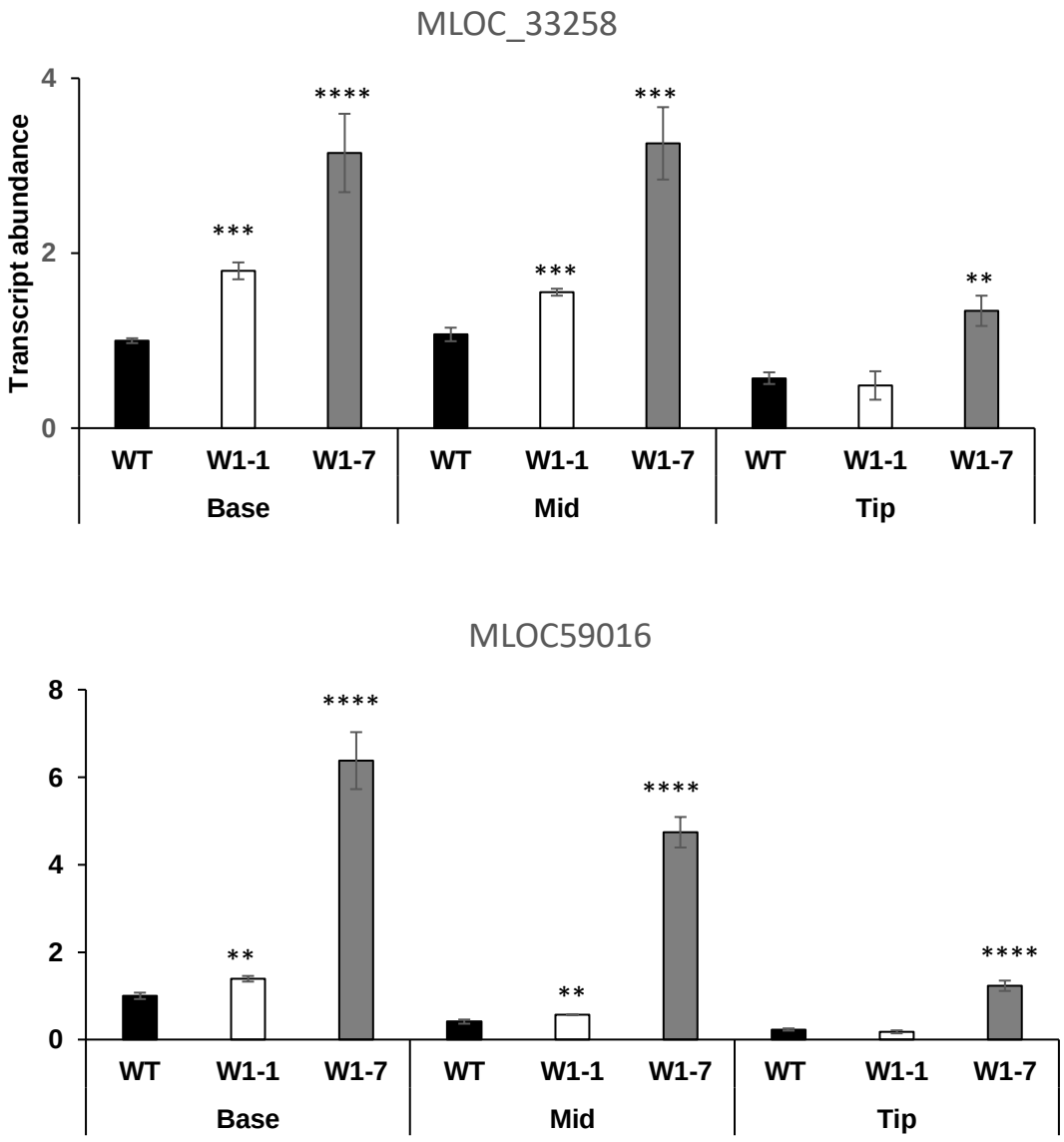
Levels of transcripts encoded by plastid genes; PETD and MLOC_76327 in the base, middle (Mid) and tip sections of the first leaves of wild type (WT), W1-1 and W1-7 seedlings 7 days after germination. Data was normalised to the 16S. Data was set to 1 and W1-1 and W1-7 were compared to the wild type. Values are represented as mean ± SE (n=6). Asterisks indicate significant differences between WHY1- deficient and wild type plants as estimated by the Student's t-test (*p<0.05; **p<0.01; ***p<0.001 and ****p<0.0001).

B.2 Nuclear-encoded transcripts



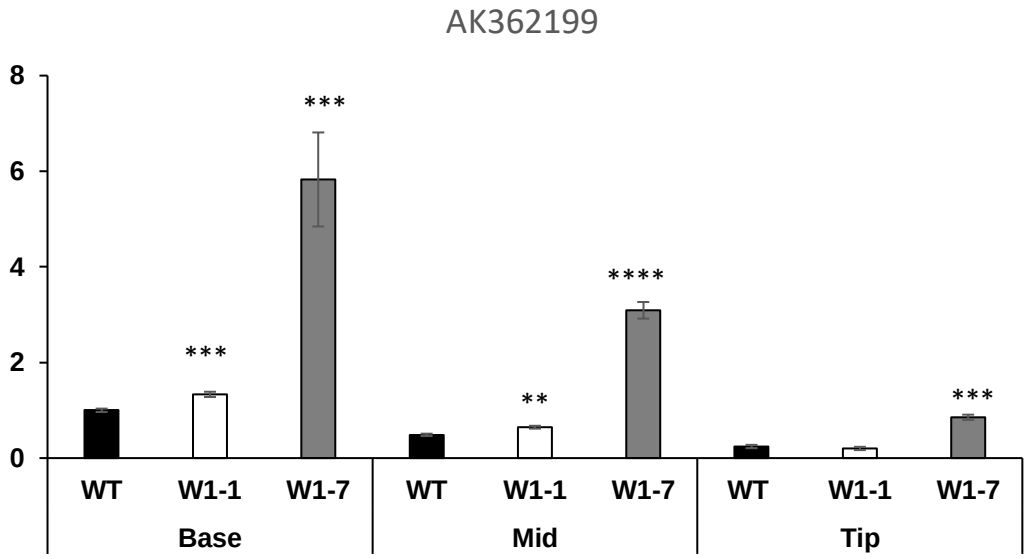
Levels of chloroplast-targeted transcripts encoded by nuclear genes; MLOC_58312 and MLOC_64606 in the base, middle (Mid) and tip sections of the first leaves of wild type (WT), W1-1 and W1-7 seedlings 7 days after germination. Data was set to 1, and W1-1 and W1-7 were compared to the wild type. Values are represented as mean ± SE (n=6). Asterisks indicate significant differences between WHY1-deficient and wild type plants as estimated by the Student's *t*-test (*****p<0.05; **p<0.01; ***p<0.001 and ********p<0.0001).

B.2 Nuclear- encoded transcripts



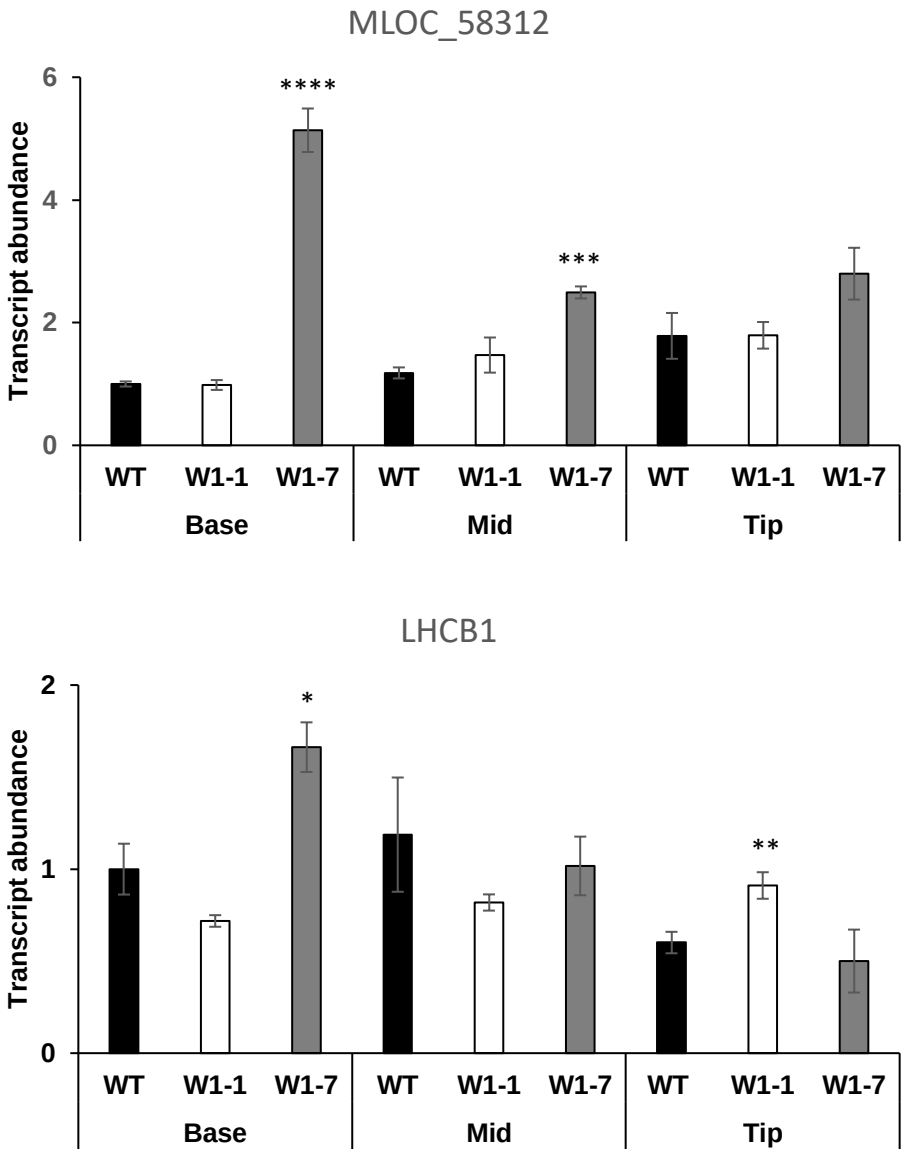
Levels of chloroplast-targeted transcripts encoded by nuclear genes; MLOC_33258 and MLOC_59016 in the base, middle (Mid) and tip sections of the first leaves of wild type (WT), W1-1 and W1-7 seedlings 7 days after germination. Data was set to 1, and W1-1 and W1-7 were compared to the wild type. Values are represented as mean ± SE (n=6). Asterisks indicate significant differences between WHY1-deficient and wild type plants as estimated by the Student's *t*-test (*****p<0.05; **p<0.01; ***p<0.001 and ********p<0.0001).

B.2 Nuclear- encoded transcripts



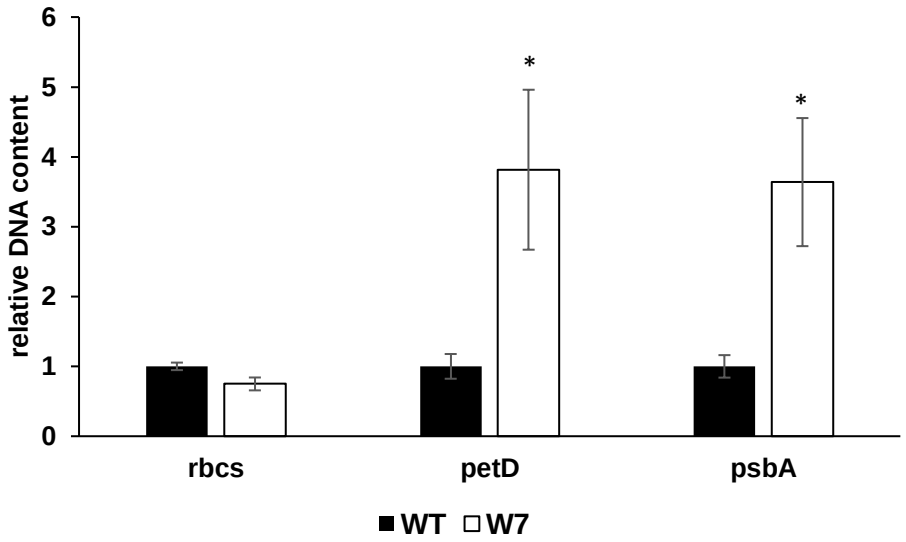
Levels of chloroplast-targeted transcripts encoded by nuclear genes; AK362199 in the base, middle (Mid) and tip sections of the first leaves of wild type (WT), W1-1 and W1-7 seedlings 7 days after germination. Data was set to 1, and W1-1 and W1-7 were compared to the wild type. Values are represented as mean ± SE (n=6). Asterisks indicate significant differences between WHY1-deficient and wild type plants as estimated by the Student's *t*-test (*****p<0.05; **p<0.01; ***p<0.001 and ********p<0.0001).

B.2 Nuclear- encoded transcripts



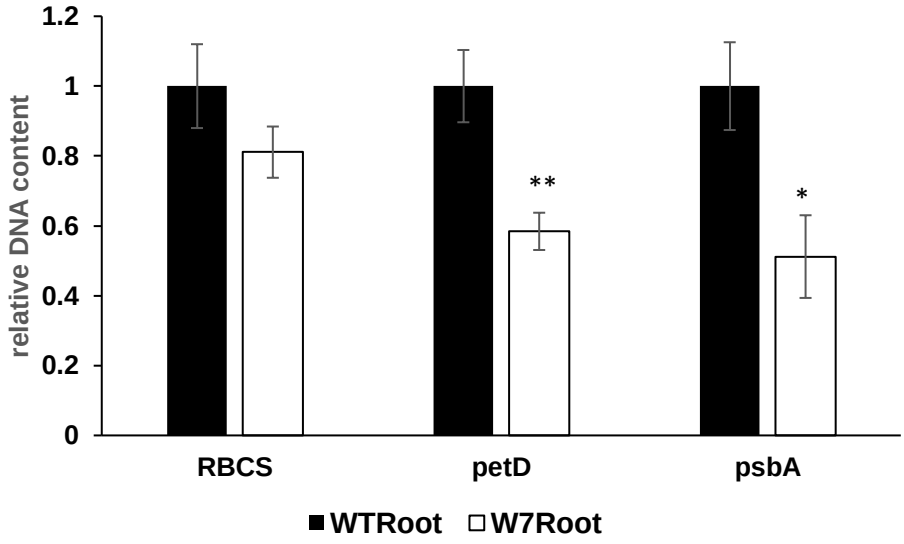
Levels of chloroplast-targeted transcripts encoded by nuclear genes; MLOC_58312 and LHCB1 in the base, middle (Mid) and tip sections of the first leaves of wild type (WT), W1-1 and W1-7 seedlings 14 days after germination. Data was set to 1, and W1-1 and W1-7 were compared to the wild type. Values are represented as mean ± SE (n=6). Asterisks indicate significant differences between WHY1-deficient and wild type plants as estimated by the Student's *t*-test (**p*<0.05; ***p*<0.01; ****p*<0.001 and *****p*<0.0001).

B.3 Plastid DNA content in mature leaves



The ratios of plastid (pt) DNA levels to nuclear (n) levels (ptDNA/nDNA ratios) in the mature leaves of wild type, W1-1 and W1-7 seedlings at 21 days after germination. Ratios were measured using specific primers to the plastome targets *petD* and *psbA*, with *rbcS* as a reference for the nuclear genome. Data was normalised to the *18S* rDNA gene and values for the WT were set to 1. Values are represented as mean ± SE (n=6). Asterisks indicate significant differences between WHY1-deficient and wild type plants as estimated by the Student's *t*-test (**p*<0.05; ***p*<0.01; ****p*<0.001 and *****p*<0.0001).

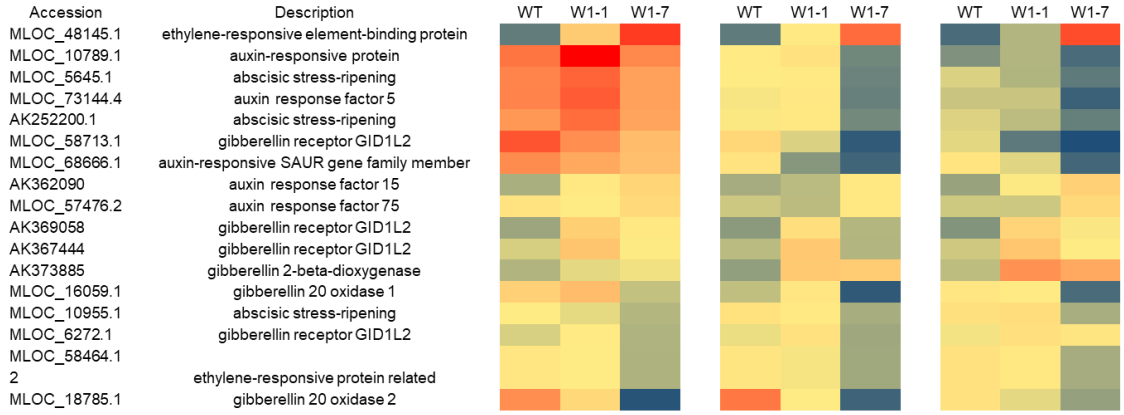
B.3.1 Plastid DNA content in root



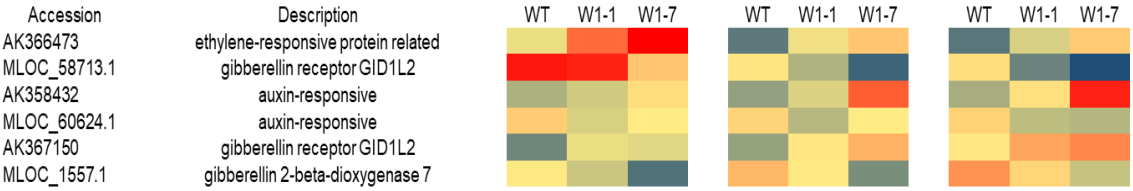
The ratios of plastid (pt) DNA levels to nuclear (n) levels (ptDNA/nDNA ratios) in the root of wild type, W1-1 and W1-7 seedlings at 7days after germination. Ratios were measured using specific primers to the plastome targets *petD* and *psbA*, with *rbcS* as a reference for the nuclear genome. Data was normalised to the 18S rDNA gene and values for the WT were set to 1. Values are represented as mean ± SE (n=6). Asterisks indicate significant differences between WHY1-deficient and wild type plants as estimated by the Student's *t*-test (**p*<0.05; ***p*<0.01; ****p*<0.001 and *****p*<0.0001).

Appendix C

C.1 Transcripts associated with hormones according to genotype-dependent transcripts



Heat map of transcript abundance of key transcripts associated with hormones at 7 days old according to genotype-dependent transcripts. The differentially expressed transcripts were analysed using two-way ANOVA, with the leaf sections (base, middle and tip) as factors, and a Bonferroni multiple-testing correction at a *p*-value of ≤0.05, FC>2 (Genespring12, Aligent Technologies).

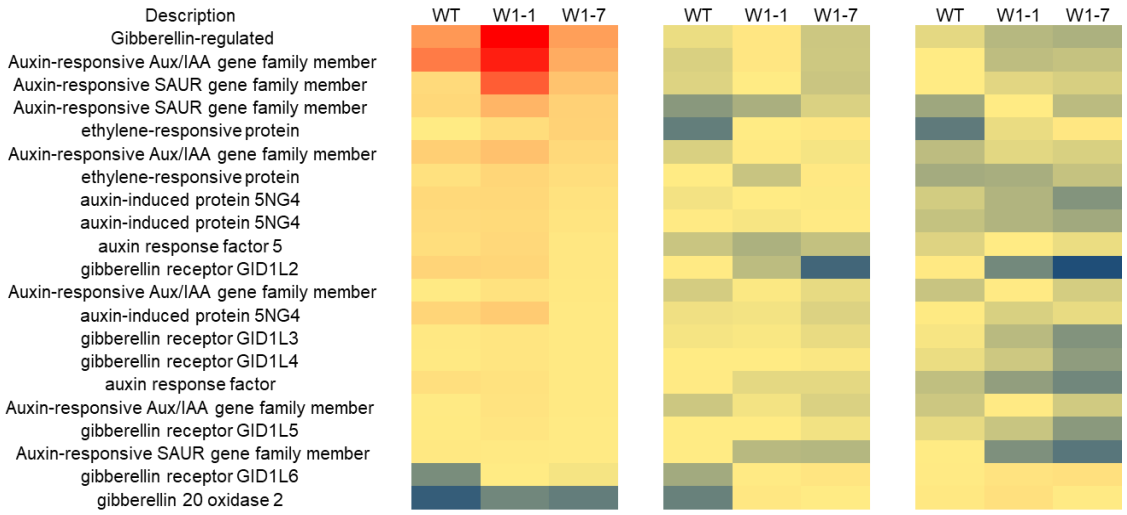


Heat map of transcript abundance of key transcripts associated with hormones at 14 days old according to genotype-dependent transcripts. The differentially expressed transcripts were analysed using two-way ANOVA, with the leaf sections (base, middle and tip) as factors, and a Bonferroni multiple-testing correction at a *p*-value of ≤0.05, FC>2 (Genespring12, Aligent Technologies).

C.2 Transcripts associated with hormones dependent on leaf regions

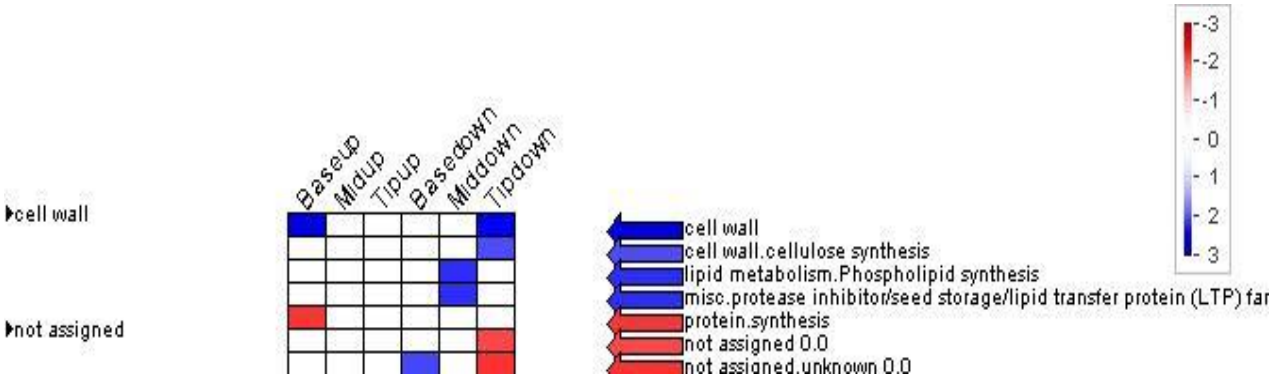
Heat map of transcript abundance of key transcripts associated with hormones at 7 days old dependent on leaf region. The differentially expressed transcripts were analysed using two-way ANOVA, with the leaf sections (base, middle and tip) as factors, and a Bonferroni multiple-testing correction at a *p*-value of ≤0.05, FC>2 (Genespring12, Aligent Technologies).

C.2.1 Transcripts associated with hormones dependent on leaf regions

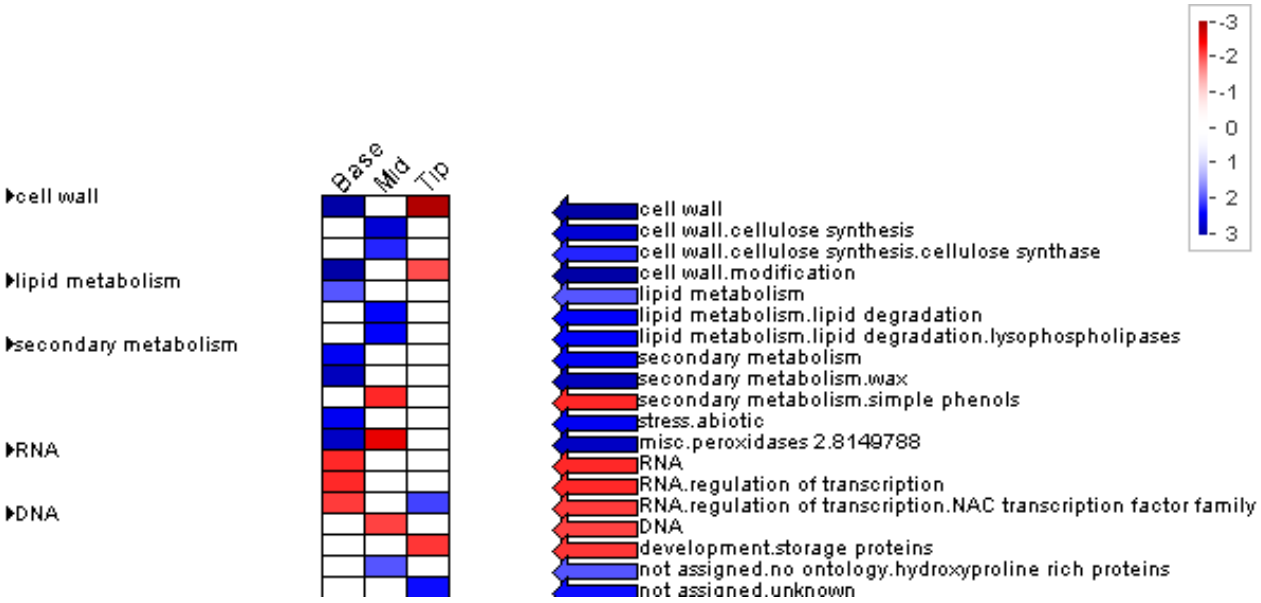


Heat map of transcript abundance of key transcripts associated with hormones at 14 days old dependent on leaf region. The differentially expressed transcripts were analysed using two-way ANOVA, with the leaf sections (base, middle and tip) as factors, and a Bonferroni multiple-testing correction at a *p*-value of ≤0.05, FC>2 (Genespring12, Aligent Technologies).

Appendix D

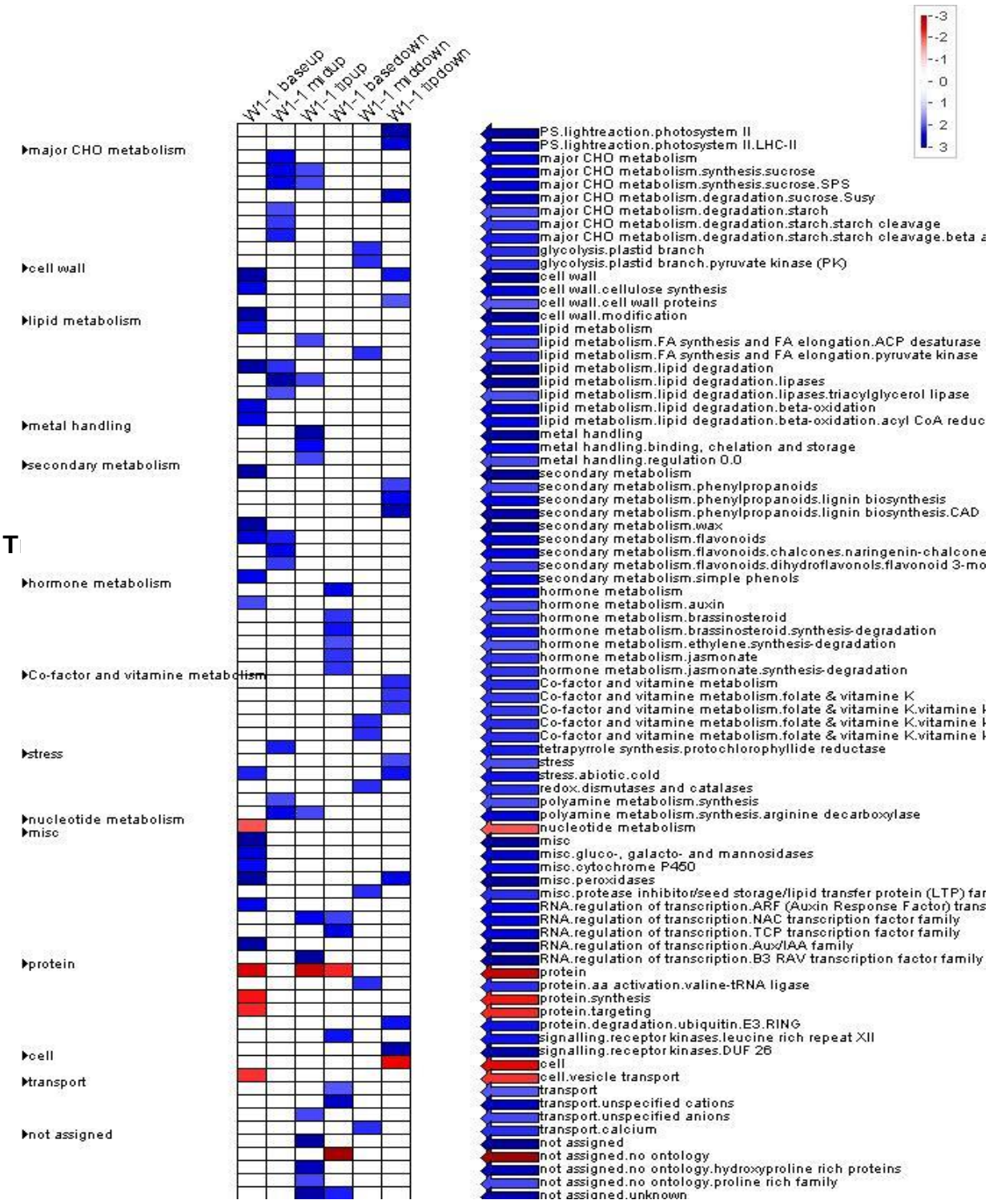


Transcript changes in the base, middle and tip leaves of the wild type on MapMan according to ORA analysis.



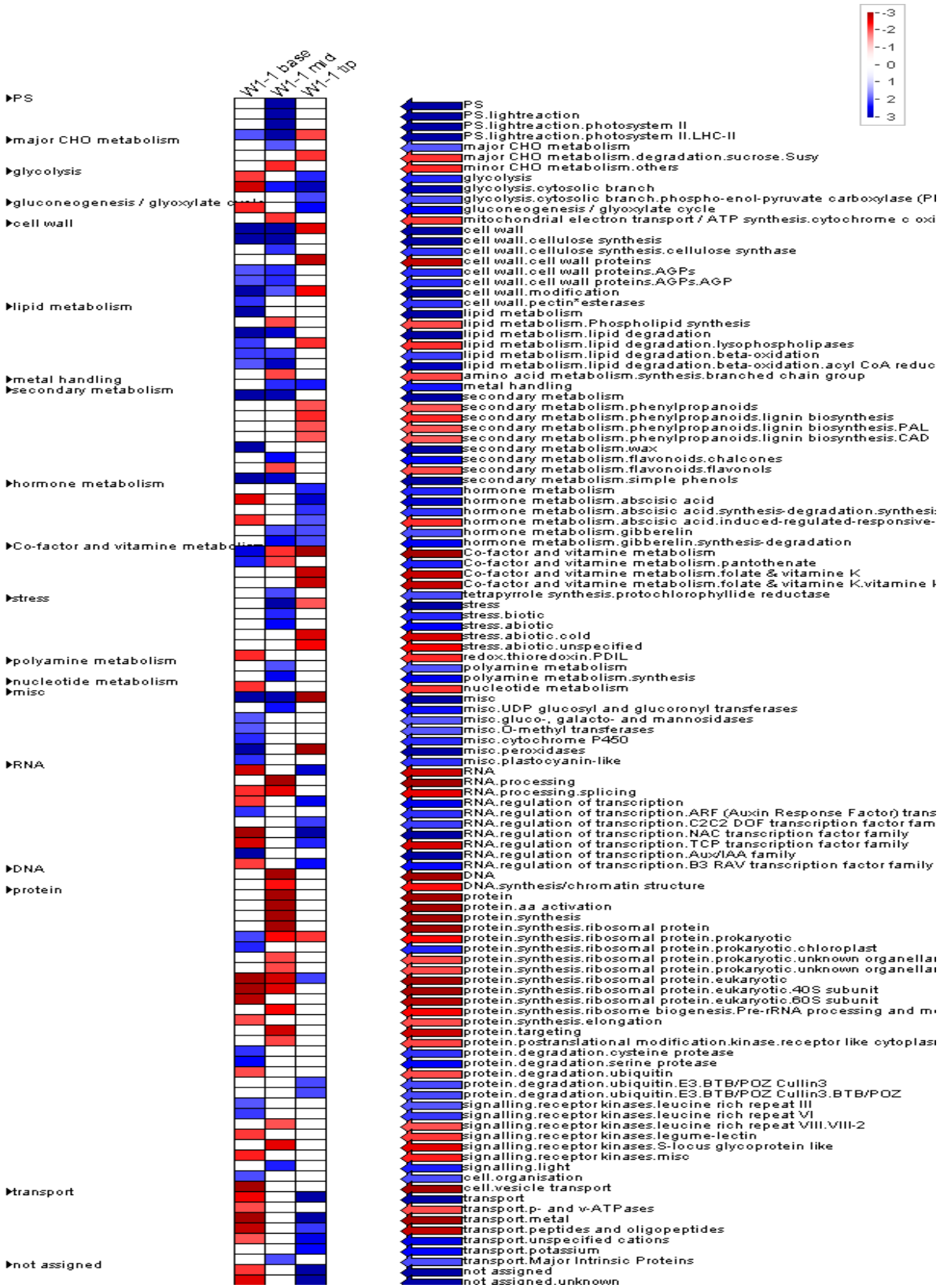
Transcript changes in the base, middle and tip leaves of the wild type at 7 day old analysis on MapMan according to Wilcoxon analysis.

Appendix E



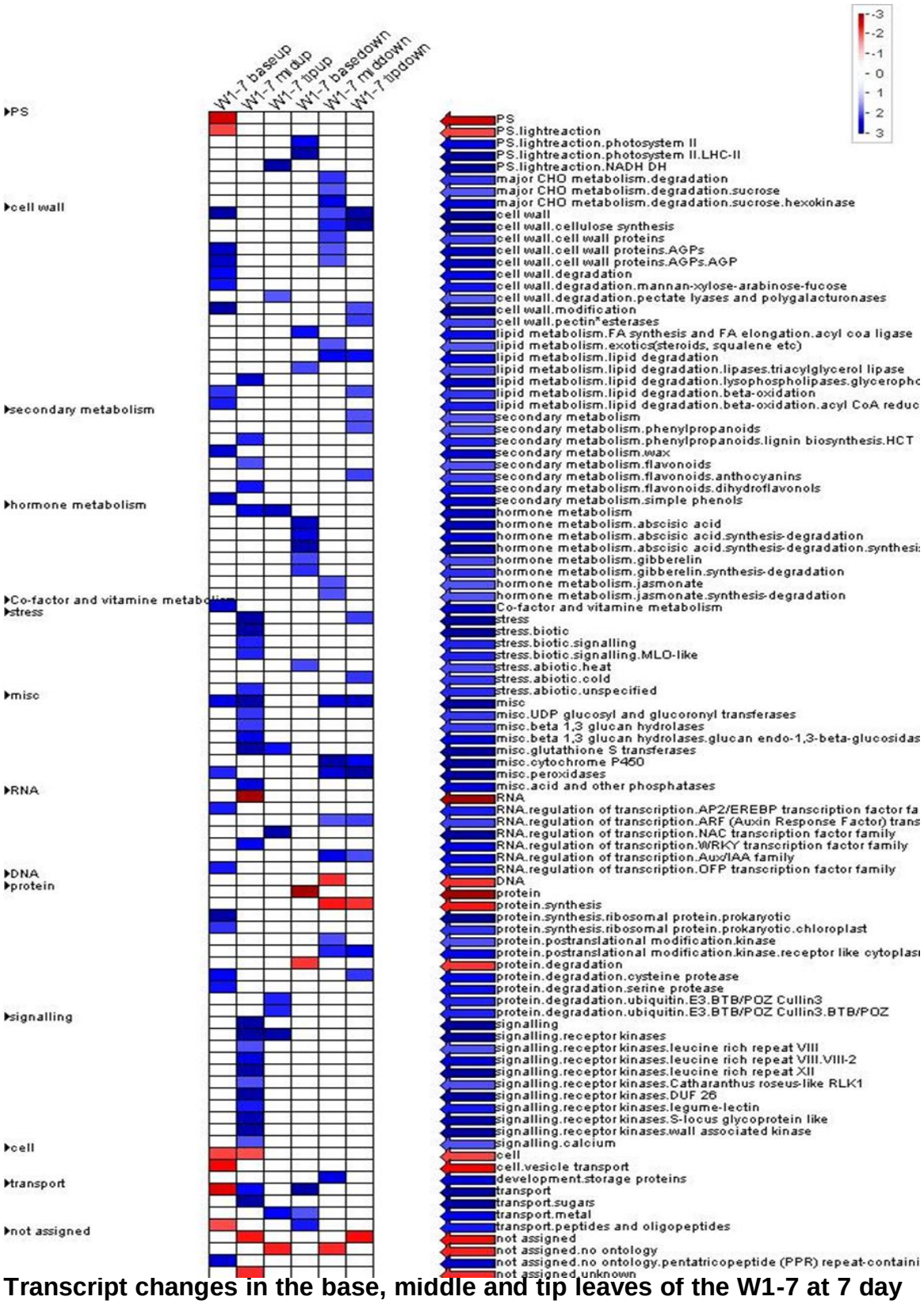
Transcript changes in the base, middle and tip leaves of the W1-1 at 7 day old analysis on MapMan according to ORA analysis.

Appendix F



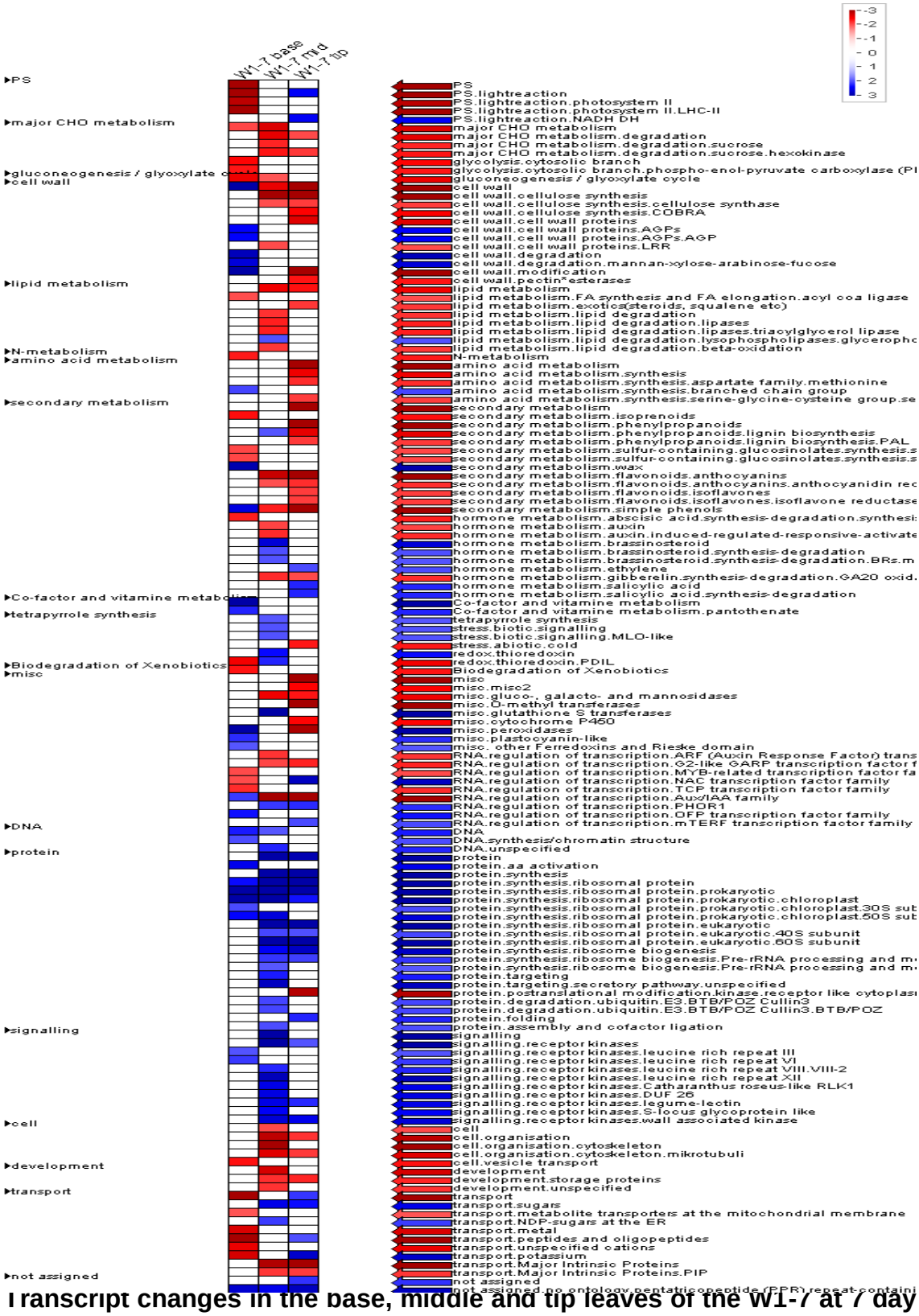
Transcript changes in the base, middle and tip leaves of the W1-1 at 7 day old analysis on MapMan according to Wilcoxon analysis.

Appendix G



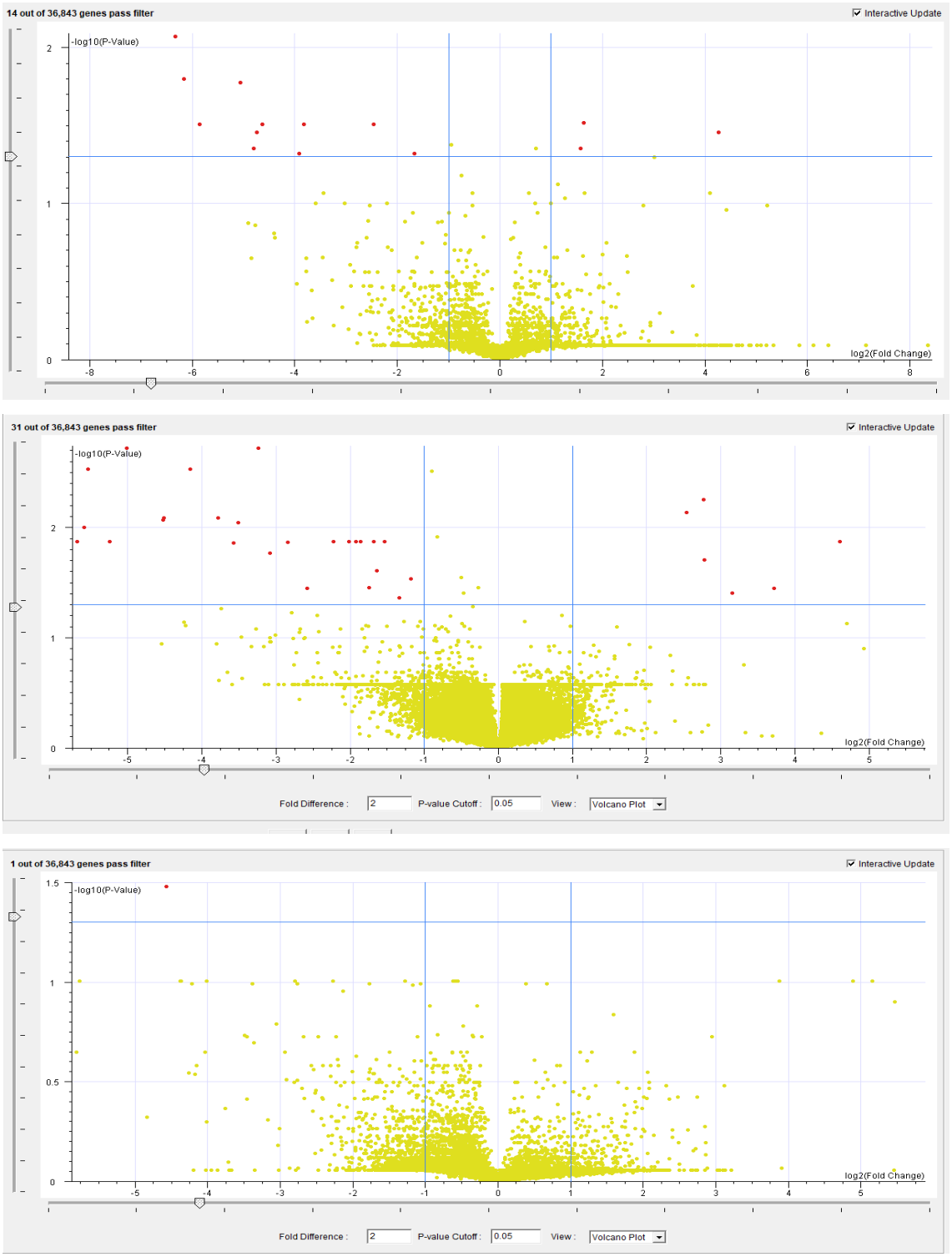
Transcript changes in the base, middle and tip leaves of the W1-7 at 7 day old analysis on MapMan according to ORA analysis.

Appendix H



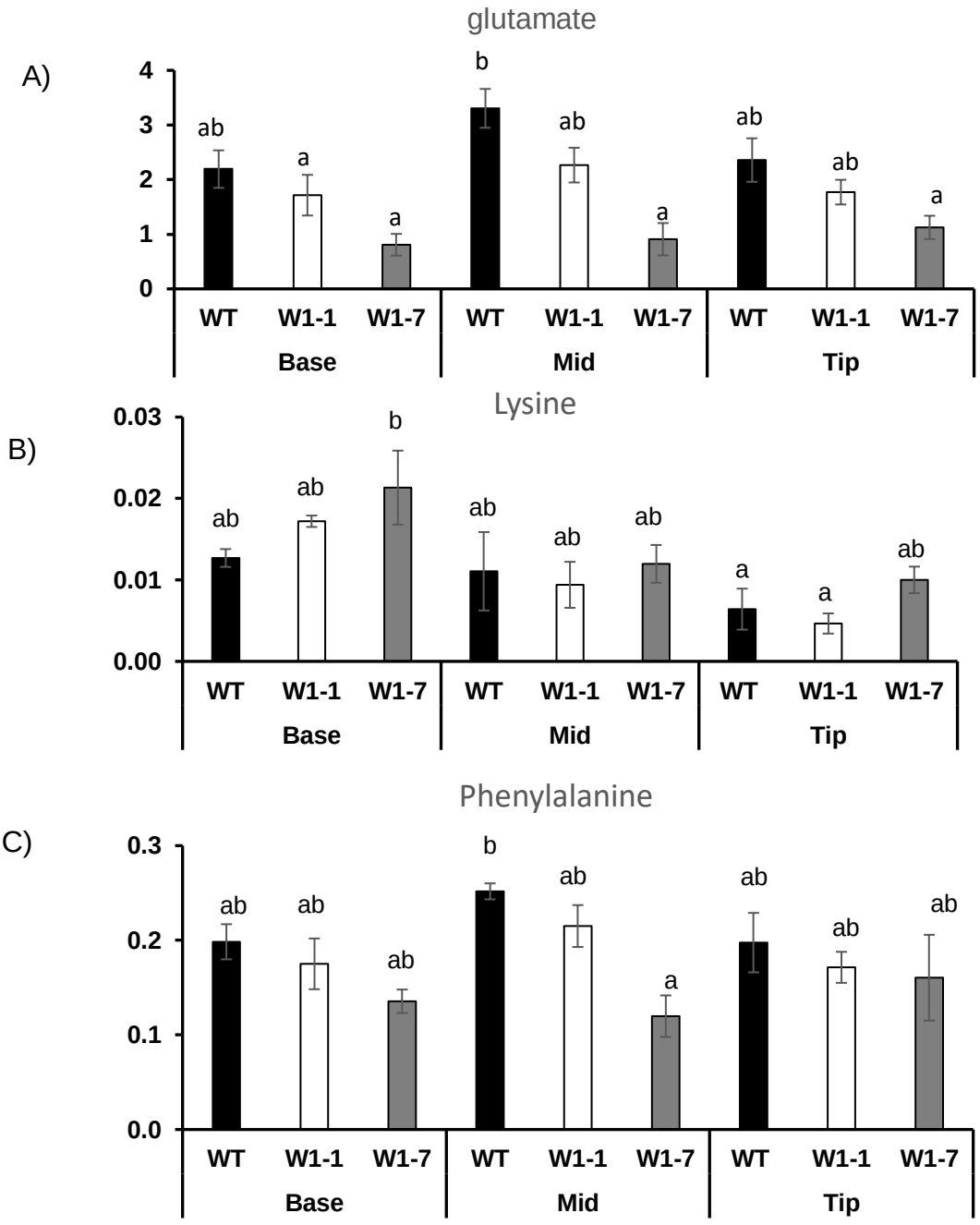
Transcript changes in the base, middle and tip leaves of the W1-7 at 7 day old analysis on MapMan according to Wilcoxon analysis.

Appendix I

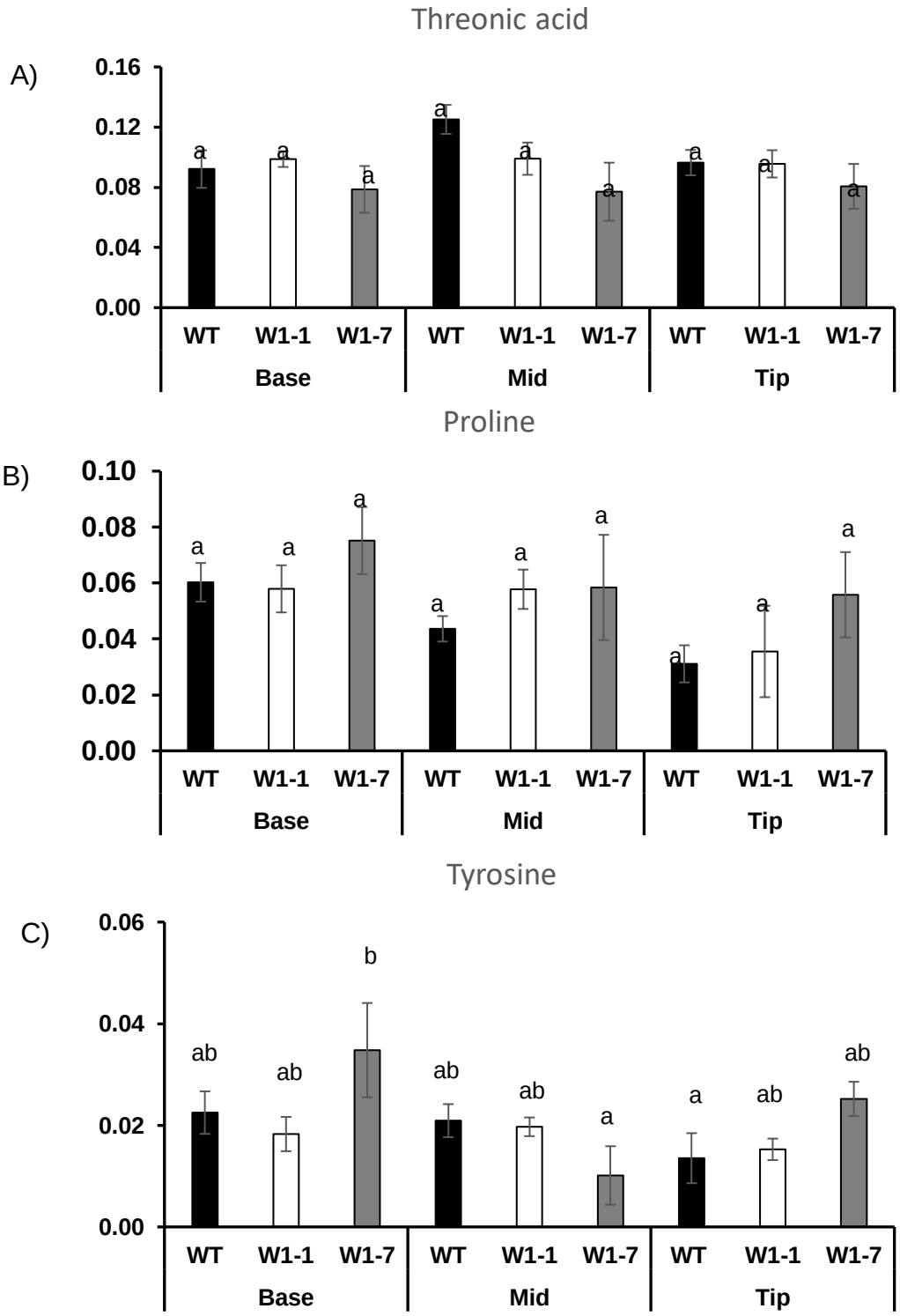


Number of differentially expressed transcripts in W1-1 relative to the wild type at 7 day old analysis was analysed using volcano plot.

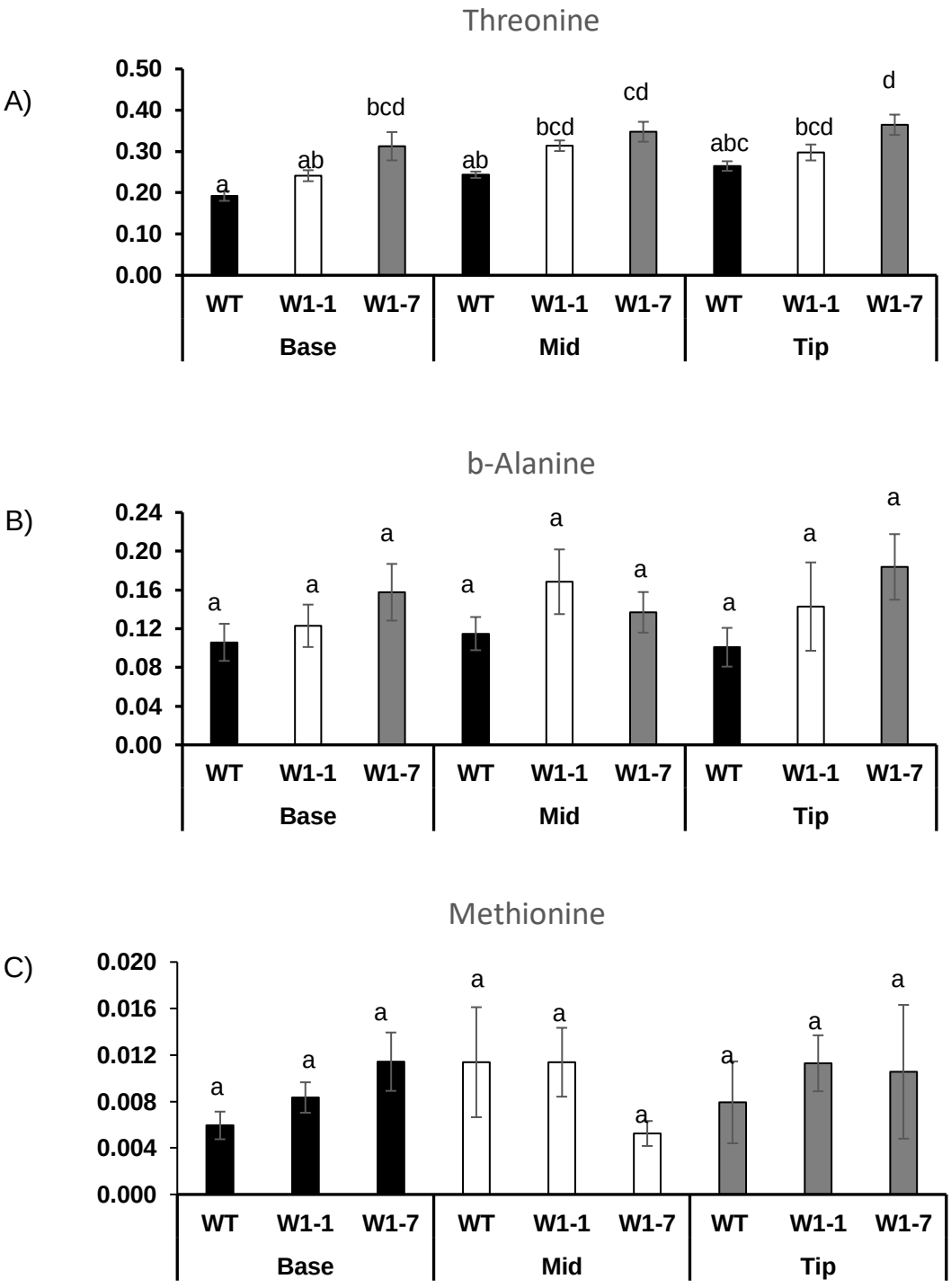
Appendix J



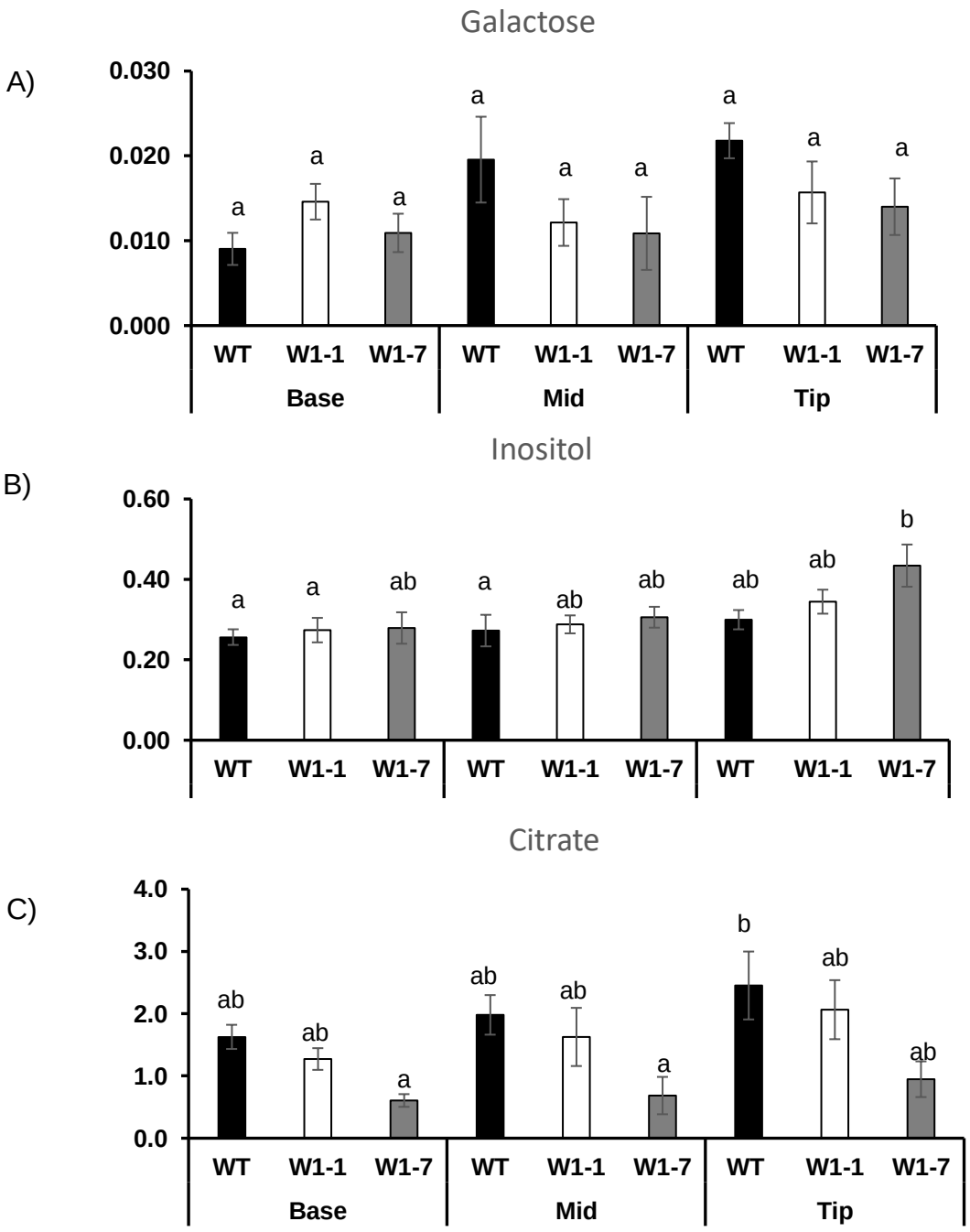
The levels of amino acids (A) glutamate, (B) lysine and (C) phenylalanine in the base, middle (Mid) and tip sections of the first leaves of wild type (WT) and W1-1 and W1-7 seedlings, 7 days after germination. Relative concentration was the mean compound (n=4) and normalised to the internal standard (n=4). Significant differences (letters) were analysed using Tukey's HSD test 1-way ANOVA (p<0.05).



The levels of amino acids (A) threonic acid, (B) proline and (C) tyrosine in the base, middle (Mid) and tip sections of the first leaves of wild type (WT) and W1-1 and W1-7 seedlings, 7 days after germination. Relative concentration was the mean compound (n=4) and normalised to the internal standard (n=4). Significant differences (letters) were analysed using Tukey's HSD test 1-way ANOVA (p<0.05).



The levels of amino acids (A) threonine, (B) b-alanine and (C) methionine in the base, middle (Mid) and tip sections of the first leaves of wild type (WT) and W1-1 and W1-7 seedlings, 7 days after germination. Relative concentration was the mean compound (n=4) and normalised to the internal standard (n=4). Significant differences (letters) were analysed using Tukey's HSD test 1-way ANOVA (p<0.05).



The levels of amino acids (A)galactose, (B) inositol and (C) citrate in the base, middle (Mid) and tip sections of the first leaves of wild type (WT) and W1-1 and W1-7 seedlings, 7 days after germination. Relative concentration was the mean compound (n=4) and normalised to the internal standard (n=4). Significant differences (letters) were analysed using Tukey's HSD test 1-way ANOVA (p<0.05

List of References

Alexa, A. and Rahnenfuhrer, J. 2018. Gene set enrichment analysis with topGO. *R package version 2.35.0.*

Allen, J. F., de Paula, W. B., Puthiyaveetil, S. and Nield, J. 2011. A structural phylogenetic map for chloroplast photosynthesis. *Trends in Plant Science,* **16**(12), pp.645-655.

Allison, L. A., Simon, L. D. and Maliga, P. 1996. Deletion of rpoB reveals a second distinct transcription system in plastids of higher plants. *The EMBO Journal,* **15**(11), pp.2802-2809.

Arsova, B., Hoja, U., Wimmelbacher, M., Greiner, E., Üstün, Ş., Melzer, M., Petersen, K., Lein, W. and Börnke, F. 2010. Plastidial Thioredoxin *z* Interacts with Two Fructokinase-Like Proteins in a Thiol-Dependent Manner: Evidence for an Essential Role in Chloroplast Development in *Arabidopsis* and *Nicotiana benthamiana*. *The Plant Cell,* **22**(5), pp.1498-1515.

Banroques, J., Cordin, O., Doere, M., Linder, P. and Tanner, N. K. 2011. Analyses of the functional regions of DEAD-box RNA "helicases" with deletion and chimera constructs tested in vivo and in vitro. *Journal of Molecular Biology,* **413**(2), pp.451-472.

Barajas-López, J. d. D., Blanco, N. E. and Strand, Å. 2013. Plastid-to-nucleus communication, signals controlling the running of the plant cell. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research,* **1833**(2), pp.425-437.

Barkan, A. 2011. Expression of Plastid Genes: Organelle-Specific Elaborations on a Prokaryotic Scaffold. *Plant Physiology,* **155**(4), pp.1520-1532.

Baumgartner, B. J., Rapp, J. C. and Mullet, J. E. 1989. Plastid Transcription Activity and DNA Copy Number Increase Early in Barley Chloroplast Development. *Plant Physiology,* **89**(3), pp.1011-1018.

Bellafiore, S., Barneche, F., Peltier, G. and Rochaix, J.-D. 2005. State transitions and light adaptation require chloroplast thylakoid protein kinase STN7. *Nature,* **433**, pp.892.

Bewley, J. D. 1997. Seed Germination and Dormancy. *Plant Cell,* **9**(7), pp.1055–1066.

Bi, Y. M., Kant, S., Clarke, J., Gidda, S., Ming, F., Xu, J., Rochon, A., Shelp, B. J., Hao, L., Zhao, R., Mullen, R. T., Zhu, T. and Rothstein, S. J. 2009. Increased nitrogen-use efficiency in transgenic rice plants over-expressing a nitrogen-responsive early nodulin gene identified from rice expression profiling. *Plant,Cell & Environment,* **32**(12), pp.1749-1760.

Blanco, N. E., Guinea-Diaz, M., Whelan, J. and Strand, A. 2014. Interaction between plastid and mitochondrial retrograde signalling pathways during changes to plastid redox status. *Philosophical Transactions of the Royal Society B: Biological Sciences,* **369**(1640), p20130231.

Bock, R. and Timmis, J. N. 2008. Reconstructing evolution: gene transfer from plastids to the nucleus. *Bioessays,* **30**(6), pp.556-566.

Börner, T., Aleynikova, A. Y., Zubo, Y. O. and Kusnetsov, V. V. 2015. Chloroplast RNA polymerases: Role in chloroplast biogenesis. *Biochimica et Biophysica Acta (BBA) - Bioenergetics,* **1847**(9), pp.761-769.

Bruce, B. D. 2000. Chloroplast transit peptides: structure, function and evolution. *Trends in Cell Biology,* **10**(10), pp.440-447.

Cahoon, E. B., Shockey, J. M., Dietrich, C. R., Gidda, S. K., Mullen, R. T. and Dyer, J. M. 2007. Engineering oilseeds for sustainable production of industrial and nutritional feedstocks: solving bottlenecks in fatty acid flux. *Current Opinion in Plant Biology,* **10**(3), pp.236-244.

Candat, A., Paszkiewicz, G., Neveu, M., Gautier, R., Logan, D. C., Avelange-Macherel, M.-H. and Macherel, D. 2014. The Ubiquitous Distribution of

Late Embryogenesis Abundant Proteins across Cell Compartments in Arabidopsis Offers Tailored Protection against Abiotic Stress. *The Plant Cell,* **26**(7), pp.3148-3166.

Cappadocia, L., Maréchal, A., Parent, J.-S., Lepage, É., 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., Zampini, É., Lepage, É., 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.

Carrie, C., Giraud, E. and Whelan, J. 2009a. Protein transport in organelles: Dual targeting of proteins to mitochondria and chloroplasts. *THe FEBS Journal,* **276**(5), pp.1187-1195.

Carrie, C., Kuhn, K., Murcha, M. W., Duncan, O., Small, I. D., O'Toole, N. and Whelan, J. 2009b. Approaches to defining dual-targeted proteins in Arabidopsis. *The Plant Journal,* **57**(6), pp.1128-1139.

Carrie, C. and Small, I. 2013. A reevaluation of dual-targeting of proteins to mitochondria and chloroplasts. *Biochimica et Biophysica Acta* **1833**(2), pp.253-259.

Caverzan, A., Passaia, G., Rosa, S. B., Ribeiro, C. W., Lazzarotto, F., & Margis-Pinheiro, M. 2012. Plant responses to stresses: Role of ascorbate peroxidase in the antioxidant protection. *Genetics and Molecular Biology,* **35**(4), pp.1011-1019.

Cavrak, V. V., Lettner, N., Jamge, S., Kosarewicz, A., Bayer, L. M. and Mittelsten Scheid, O. 2014. How a retrotransposon exploits the plant's heat stress response for its activation. *PLoS Genetics,* **10**(1), p1004115.

Chacinska, A., Koehler, C. M., Milenkovic, D., Lithgow, T. and Pfanner, N. 2009. Importing Mitochondrial Proteins: Machineries and Mechanisms. *Cell,* **138**(4), pp.628-644.

Chan, K. X., Crisp, P. A., Estavillo, G. M. and Pogson, B. J. 2010. Chloroplast-to-nucleus communication: current knowledge, experimental strategies and relationship to drought stress signaling. *Plant Signaling and Behavior,* **5**(12), pp.1575-1582.

Chen, M., Liu, X., Jiang, S., Wen, B., Yang, C., Xiao, W., Fu, X., Li, D., Chen, X., Gao, D. and Li, L. 2018. Transcriptomic and Functional Analyses Reveal That PpGLK1 Regulates Chloroplast Development in Peach (*Prunus persica*). *Frontiers in Plant Science,* **9**, p34.

Chi, W., He, B., Mao, J., Li, Q., Ma, J., Ji, D., Zou, M. and Zhang, L. 2012. The Function of RH22, a DEAD RNA Helicase, in the Biogenesis of the 50S Ribosomal Subunits of *Arabidopsis* Chloroplasts. *Plant Physiology,* **158**(2), pp.693-707.

Citovsky, V., Gafni, Y. and Tzfira, T. 2008. Localizing protein–protein interactions by bimolecular fluorescence complementation in planta. *Methods,* **45**(3), pp.196-206.

Citovsky, V., Lee, L.-Y., Vyas, S., Glick, E., Chen, M.-H., Vainstein, A., Gafni, Y., Gelvin, S. B. and Tzfira, T. 2006. Subcellular Localization of Interacting Proteins by Bimolecular Fluorescence Complementation in Planta. *Journal of Molecular Biology,* **362**(5), pp.1120-1131.

Comadira, G., Rasool, B., Kaprinska, B., Garcia, 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.

Creissen, G., Reynolds, H., Xue, Y. and Mullineaux, P. 1995. Simultaneous targeting of pea glutathione reductase and of a bacterial fusion protein to chloroplasts and mitochondria in transgenic tobacco. *Plant Journal,* **8**(2), pp.167-175.

Cruz, J. l., DieterKressler and PatrickLinder. 1999. Unwinding RNA in *Saccharomyces cerevisiae*: DEAD-box proteins and related families. *Trends in Biochemical Sciences,* **24**(5), pp.192-198.

Davidson, S. E., Reid, J. B. and Helliwell, C. A. 2006. Cytochromes P450 in gibberellin biosynthesis. *Phytochemistry Reviews,* **5**(2-3), pp.405-419.

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. *Plant Cell,* **7**(5), pp.589-598.

Desveaux, Allard, Brisson and Sygusch. 2002. A new family of plant transcription factors displays a novel ssDNA-binding surface. *Nature Structural and Molecular 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 *PR10-a* Gene Activation in Potato. **12**(8), pp.1477-1489.

Desveaux, D., Marechal, 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., Despres, C., Mess, J. N., Levesque, C., Fobert, P. R., Dangl, J. L. and Brisson, N. 2004. A "Whirly" transcription factor is required for salicylic acid-dependent disease resistance in Arabidopsis. *Dev Cell,* **6**(2), pp.229-40.

Di Cola, A., Klostermann, E. and Robinson, C. 2005. The complexity of pathways for protein import into thylakoids: it's not easy being green. *Biochemical Society Transactions,* **33**(5), pp.1024-1027.

Díaz, M. G., Hernández-Verdeja, T., Kremnev, D., Crawford, T., Dubreuil, C. and Strand, Å. 2018. Redox regulation of PEP activity during seedling establishment in *Arabidopsis thaliana*. *Nature Communications,* **9**(1), p50.

Dolezal, P., Likic, V., Tachezy, J. and Lithgow, T. 2006. Evolution of the molecular machines for protein import into mitochondria. *Science,* **313**(5785), pp.314-318.

Dubreuil, C., Jin, X., Barajas-Lopez, J. D., Hewitt, T. C., Tanz, S. K., Dobrenel, T., Schroder, W. P., Hanson, J., Pesquet, E., Gronlund, A., Small, I. and Strand, A. 2018. Establishment of Photosynthesis through

Chloroplast Development Is Controlled by Two Distinct Regulatory Phases. *Plant Physiology,* **176**(2), pp.1199-1214.

Dudek, J., Rehling, P. and van der Laan, M. 2013. Mitochondrial protein import: common principles and physiological networks. *Biochimica et Biophysica Acta,* **1833**(2), pp.274-285.

Estavillo, G. M., Crisp, P. A., Pornsiriwong, W., Wirtz, M., Collinge, D., Carrie, C., Giraud, E., Whelan, J., David, P., Javot, H., Brearley, C., Hell, R., Marin, E. and Pogson, B. J. 2011. Evidence for a SAL1-PAP chloroplast retrograde pathway that functions in drought and high light signaling in *Arabidopsis. The Plant Cell,* **23**(11), pp.3992-4012.

Fernández, A. P. and Strand, Å. 2008. Retrograde signaling and plant stress: plastid signals initiate cellular stress responses. *Current Opinion in Plant Biology,* **11**(5), pp.509-513.

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 B: Biological Sciences,* **369**(1640), p20130226.

Foyer, C. H. and Noctor, G. 2009. Redox regulation in photosynthetic organisms: signaling, acclimation, and practical implications. *Antioxidant and Redox Signaling,* **11**(4), pp.861-905.

Foyer, C. H., Parry, M. and Noctor, G. 2003. Markers and signals associated with nitrogen assimilation in higher plants. *Journal of Experimental Botany,* **54**(382), pp.585-593.

Gagliardi, D., Kuhn, J., Spadinger, U., Brennicke, A., Leaver, C. J. and Binder, S. 1999. An RNA helicase (AtSUV3) is present in *Arabidopsis thaliana* mitochondria. *FEBS Letters,* **458**(3), pp.337-342.

Garay-Arroyo, A., Colmenero-Flores, J. M., Garciarrubio, A. and Covarrubias, A. A. 2000. Highly hydrophilic proteins in prokaryotes and eukaryotes are common during conditions of water deficit. *The Journal of Biological Chemistry,* **275**(8), pp.5668-5674.

Gentle, I., Gabriel, K., Beech, P., Waller, R. and Lithgow, T. 2004. The Omp85 family of proteins is essential for outer membrane biogenesis in mitochondria and bacteria. *The Journal of Cell Biology,* **164**(1), pp.19-24.

Gill, S. S. and Tuteja, N. 2010. Reactive oxygen species and antioxidant machinery in abiotic stress tolerance in crop plants. *Plant Physiology and Biochemistry,* **48**(12), pp.909-930.

Goyal, K., Walton, L. J. and Tunnacliffe, A. 2005. LEA proteins prevent protein aggregation due to water stress. *The Biochemical Journal,* **388**(1), pp.151-157.

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B. W., Nusbaum, C., Lindblad-Toh, K., Friedman, N. and Regev, A. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology,* **29**(7), pp.644-652.

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.

Green, B. R. 2011. Chloroplast genomes of photosynthetic eukaryotes. *The Plant Journal,* **66**(1), pp.34-44.

Grevich, J. J. and Daniell, H. 2005. Chloroplast Genetic Engineering: Recent Advances and Future Perspectives. *Critical Reviews in Plant Sciences,* **24**(2), pp.83-107.

Grzelczak, Z. F., Sattolo, M. H., Hanley-Bowdoin, L. K., Kennedy, T. D. and Lane, B. G. 1982. Synthesis and turnover of proteins and mRNA in germinating wheat embryos. *Canadian Journal of Biochemistry and Physiology,* **60**(3), pp.389-397.

Gubler, F., R Kalla, J. K. R. and Jacobsen, J. V. 1995. Gibberellin-regulated expression of a myb gene in barley aleurone cells: evidence for Myb

transactivation of a high-pI alpha-amylase gene promoter. *The Plant Cell,* **7**(11), pp.1879-1891.

Hajdukiewicz, P. T., Allison, L. A. and Maliga, P. 1997. The two RNA polymerases encoded by the nuclear and the plastid compartments transcribe distinct groups of genes in tobacco plastids. *The EMBO Journal,* **16**(13), pp.4041-4048.

Hanaoka, M., Kanamaru, K., Takahashi, H. and Tanaka, K. 2003. Molecular genetic analysis of chloroplast gene promoters dependent on SIG2, a nucleus-encoded sigma factor for the plastid-encoded RNA polymerase, in *Arabidopsis thaliana*. *Nucleic Acids Research,* **31**(24), pp.7090-7098.

Harpster, M. H., Mayfield, S. P. and Taylor, W. C. J. P. M. B. 1984. Effects of pigment-deficient mutants on the accumulation of photosynthetic proteins in maize. *Plant Molecular Biology,* **3**(2), pp.59-71.

Hedtke, B., Borner, T. and Weihe, A. 1997. Mitochondrial and chloroplast phage-type RNA polymerases in *Arabidopsis*. *Science,* **277**(5327), pp.809-811.

Hernández-Verdeja, T. and Strand, Å. 2018. Retrograde Signals Navigate the Path to Chloroplast Development. *Plant Physiology,* **176**(2), pp.967-976.

Hess, W. R., Müller, A., Nagy, F. and Börner, T. 1994. Ribosome-deficient plastids affect transcription of light-induced nuclear genes: genetic evidence for a plastid-derived signal. *Molecular and General Genetics* **242**(3), pp.305-312.

Hess, W. R., Prombona, A., Fieder, B., Subramanian, A. R. and Börner, T. 1993. Chloroplast rps15 and the rpoB/C1/C2 gene cluster are strongly transcribed in ribosome-deficient plastids: evidence for a functioning non-chloroplast-encoded RNA polymerase. *The EMBO Journal,* **12**(2), pp.563-571.

Ho, L. H. M., Giraud, E., Uggalla, V., Lister, R., Clifton, R., Glen, A., Thirkettle-Watts, D., Van Aken, O. and Whelan, J. 2008. Identification of Regulatory Pathways Controlling Gene Expression of Stress-

Responsive Mitochondrial Proteins in *Arabidopsis*. *Plant Physiology,* **147**(4), pp.1858-1873.

Horstman, A., Nougalli Tonaco, I. A., Boutilier, K. and Immink, R. G. H. 2014. A Cautionary Note on the Use of Split-YFP/BiFC in Plant Protein-Protein Interaction Studies. *International Journal of Molecular Sciences,* **15**(6), pp.9628-9643.

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), pp. 2352.

Huber, W., Carey, V. J., Gentleman, R., Anders, S., Carlson, M., Carvalho, B. S., Bravo, H. C., Davis, S., Gatto, L., Girke, T., Gottardo, R., Hahne, F., Hansen, K. D., Irizarry, R. A., Lawrence, M., Love, M. I., MacDonald, J., Obenchain, V., Oles, A. K., Pages, H., Reyes, A., Shannon, P., Smyth, G. K., Tenenbaum, D., Waldron, L. and Morgan, M. 2015. Orchestrating high-throughput genomic analysis with bioconductor. *Nature Methods,* **12**(2), pp.115-21.

Hundertmark, M. and Hincha, D. K. 2008. LEA (Late Embryogenesis Abundant) proteins and their encoding genes in *Arabidopsis thaliana*. *BMC Genomics,* **9**(1), pp. 118.

Isemer, R., Krause, K., Grabe, N., Kitahata, N., Asami, T. and Krupinska, K. 2012a. Plastid Located WHIRLY1 Enhances the Responsiveness of *Arabidopsis* Seedlings Toward Abscisic Acid. *Frontiers of Plant Science,* **3**, pp. 283.

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

Jacobsen, J. V., Pearce, D. W., Poole, A. T., Pharis, R. P. and Mander, L. N. 2002. Abscisic acid, phaseic acid and gibberellin contents associated with dormancy and germination in barley. *Physiologia Plantarum,* **115**(3), pp.428-441.

Jukanti, A. K., Heidlebaugh, N. M., Parrott, D. L., Fischer, I. A., McInnerney, K. and Fischer, A. M. 2008. Comparative transcriptome profiling of near-isogenic barley (*Hordeum vulgare*) lines differing in the allelic state of a major grain protein content locus identifies genes with possible roles in leaf senescence and nitrogen reallocation. *New Phytologist,* **177**(2), pp.333-349.

Jung, C., Lyou, S. H., Yeu, S., Kim, M. A., Rhee, S., Kim, M., Lee, J. S., Choi, Y. D. and Cheong, J. J. 2007. Microarray-based screening of jasmonate-responsive genes in Arabidopsis thaliana. *Plant Cell Reports,* **26**(7), pp.1053-63.

Kacprzak, S. M., Mochizuki, N., Naranjo, B., Xu, D., Leister, D., Kleine, T., Okamoto, H. and Terry, M. J. 2019. Plastid-to-Nucleus Retrograde Signalling during Chloroplast Biogenesis Does Not Require ABI4. *Plant Physiology,* **179**(1), pp.18-23.

Kanai, M., Hayashi, M., Kondo, M. and Nishimura, M. 2013. The plastidic DEAD-box RNA helicase 22, HS3, is essential for plastid functions both in seed development and in seedling growth. *Plant Cell Physiology,* **54**(9), pp.1431-1440.

Karpinska, B., Alomrani, S. O. and Foyer, C. H. 2017. Inhibitor-induced oxidation of the nucleus and cytosol in Arabidopsis thaliana: implications for organelle to nucleus retrograde signalling. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences,* **372**(1730), p20160392.

Ke, J., Wen, T.-N., Nikolau, B. J. and Wurtele, E. S. 2000. Coordinate Regulation of the Nuclear and Plastidic Genes Coding for the Subunits of the Heteromeric Acetyl-Coenzyme A Carboxylase. *Plant Physiology,* **122**(4), pp.1057-1072.

Kessler, F., Blobel, G., Patel, H. A. and Schnell, D. J. 1994. Identification of two GTP-binding proteins in the chloroplast protein import machinery. *Science,* **266**(5187), pp.1035-1039.

Kessler, F. and Schnell, D. 2009. Chloroplast biogenesis: diversity and regulation of the protein import apparatus. *Current Opinion in Cell Biology,* **21**(4), pp.494-500.

Kindgren, P. and Strand, Å. 2015. Chloroplast transcription, untangling the Gordian Knot. *New Phytologist,* **206**(3), pp.889-891.

Koehler, C. M. 2004. New developments in mitochondrial assembly. *Annual Review of Cell and Developmental Biology,* **20**, pp.309-335.

Kohler, D., Schmidt-Gattung, S. and Binder, S. 2010. The DEAD-box protein PMH2 is required for efficient group II intron splicing in mitochondria of Arabidopsis thaliana. *Plant Molecular Biology,* **72**(4-5), pp.459-467.

Konishi, T., Shinohara, K., Yamada, K. and Sasaki, Y. 1996. Acetyl-CoA carboxylase in higher plants: most plants other than gramineae have both the prokaryotic and the eukaryotic forms of this enzyme. *Plant & Cell Physiology,* **37**(2), pp.117-122.

Koussevitzky, S., Nott, A., Mockler, T. C., Hong, F., Sachetto-Martins, G., Surpin, M., Lim, J., Mittler, R. and Chory, J. 2007. Signals from Chloroplasts Converge to Regulate Nuclear Gene Expression. *Science,* **316**(5825), pp.715-719.

Kovács-Bogdán, E., Soll, J. and Bölter, B. 2010. Protein import into chloroplasts: The Tic complex and its regulation. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research,* **1803**(6), pp.740-747.

Kozjak, V., Wiedemann, N., Milenkovic, D., Lohaus, C., Meyer, H. E., Guiard, B., Meisinger, C. and Pfanner, N. 2003. An essential role of Sam50 in the protein sorting and assembly machinery of the mitochondrial outer membrane. *The Journal of Biological Chemistry,* **278**(49), pp.48520-48523.

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.

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

Krupinska, K., Haussuhl, K., Schafer, A., van der Kooij, T. A., Leckband, G., Lorz, H. and Falk, J. 2002. A novel nucleus-targeted protein is

expressed in barley leaves during senescence and pathogen infection. *Plant Physiology,* **130**(3), pp.1172-1180.

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**(432).

Lee, D. W., Kim, J. K., Lee, S., Choi, S., Kim, S. and Hwang, I. 2008. Arabidopsis Nuclear-Encoded Plastid Transit Peptides Contain Multiple Sequence Subgroups with Distinctive Chloroplast-Targeting Sequence Motifs. *The Plant Cell,* **20**(6), pp.1603-1622.

Leister, D. 2003. Chloroplast research in the genomic age. *Trends in Genetics,* **19**(1), pp.47-56.

Lepage, E., Zampini, E. 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.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R. and Subgroup, G. P. D. P. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics,* **25**(16), pp.2078-2079.

Lichtenthaler, H. K. 1987. Chlorophylls and carotenoids: pigments of photosynthetic biomembranes. *Methods in Enzymology.* Academic Press.

Linder, P. and Jankowsky, E. 2011. From unwinding to clamping - the DEAD box RNA helicase family. *Nature Reviews Molecular Cell Biology,* **12**(8), pp.505-516.

Liu, G., Kennedy, R., Greenshields, D. L., Peng, G., Forseille, L., Selvaraj, G. and Wei, Y. 2007. Detached and attached *Arabidopsis* leaf assays reveal distinctive defense responses against hemibiotrophic *Colletotrichum spp*. *Molecular Plant-Microbe Interactions,* **20**(10), pp.1308-1319.

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

Luk, E., Yang, M., Jensen, L. T., Bourbonnais, Y. and Culotta, V. C. 2005. Manganese activation of superoxide dismutase 2 in the mitochondria of *Saccharomyces cerevisiae*. *The Journal of Biological Chemistry,* **280**(24), pp.22715-227200.

Lyska, D., Meierhoff, K. and Westhoff, P. 2013. How to build functional thylakoid membranes: from plastid transcription to protein complex assembly. *Planta,* **237**(2), pp.413-428.

Malgorzata Grudkowska, B. Z. 2004. Multifunctional role of plant cysteine proteinases. *Acta Biochimica Polonica,* **51**, pp.609–624.

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.

Marechal, A., Parent, J. S., Veronneau-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.

Martin, W., Rujan, T., Richly, E., Hansen, A., Cornelsen, S., Lins, T., Leister, D., Stoebe, B., Hasegawa, M. and Penny, D. 2002. Evolutionary analysis of *Arabidopsis*, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus. *Proceedings of the National Academy of Sciences of the United States of America,* **99**(19), pp.12246-12251.

Mascher, M., Muehlbauer, G. J., Rokhsar, D. S., Chapman, J., Schmutz, J., Barry, K., Muñoz-Amatriaín, M., Close, T. J., Wise, R. P., Schulman, A. H., Himmelbach, A., Mayer, K. F. X., Scholz, U., Poland, J. A., Stein, N. and Waugh, R. 2013. Anchoring and ordering NGS contig assemblies by population sequencing (POPSEQ). **76**(4), pp.718-727.

Matthes, A., Schmidt-Gattung, S., Köhler, D., Forner, J., Wildum, S., Raabe, M., Urlaub, H. and Binder, S. 2007. Two DEAD-Box Proteins May Be Part of RNA-Dependent High-Molecular-Mass Protein Complexes in Arabidopsis Mitochondria. *Plant Physiology,* **145**(4), pp.1637-1646.

McHale, L., Tan, X., Koehl, P. and Michelmore, R. W. 2006. Plant NBS-LRR proteins: adaptable guards. *Genome Biology,* **7**(4), p212.

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.

Millar, A. A., Jacobsen, J. V., Ross, J. J., Helliwell, C. A., Poole, A. T., Scofield, G., Reid, J. B. and Gubler, F. 2006. Seed dormancy and ABA metabolism in *Arabidopsis* and barley: the role of ABA 8'-hydroxylase. *Plant Journal,* **45**(6), pp.942-954.

Miller, J. D., Arteca, R. N. and Pell, E. J. 1999. Senescence-Associated Gene Expression during Ozone-Induced Leaf Senescence in *Arabidopsis*. *Plant Physiology,* **120**(4), pp.1015-1024.

Miller, K. E., Kim, Y., Huh, W. K. and Park, H. O. 2015. Bimolecular Fluorescence Complementation (BiFC) Analysis: Advances and Recent Applications for Genome-Wide Interaction Studies. *Journal of Molecular Biology,* **427**(11), pp.2039-2055.

Mitschke, J., Fuss, J., Blum, T., Hoglund, A., Reski, R., Kohlbacher, O. and Rensing, S. A. 2009. Prediction of dual protein targeting to plant organelles. *New Phytology,* **183**(1), pp.224-235.

Moller, I. M. 2001. Plant mitochondria and oxidative stress: electron transport, NADPH turnover, and metabolism of reactive oxygen species. *Annual Review of Plant Physiology and Plant Molecular Biology,* **52**, pp.561-591.

Morgante, C. V., Rodrigues, R. A., Marbach, P. A., Borgonovi, C. M., Moura, D. S. and Silva-Filho, M. C. 2009. Conservation of dual-targeted proteins in *Arabidopsis* and rice points to a similar pattern of gene-

family evolution. *Molecular Genetics and Genomics,* **281**(5), pp.525-538.

Morley, S. A. and Nielsen, B. L. 2016. Chloroplast DNA Copy Number Changes during Plant Development in Organelle DNA Polymerase Mutants. **7**(57).

Mossmann, D., Meisinger, C. and Vogtle, F. N. 2012. Processing of mitochondrial presequences. *Biochimica et Biophysica Acta* **1819**(9-10), pp.1098-1106.

Mowla, S. B., Cuypers, A., Driscoll, S. P., Kiddle, G., Thomson, J., Foyer, C. H. and Theodoulou, F. L. 2006. Yeast complementation reveals a role for an *Arabidopsis thaliana* late embryogenesis abundant (LEA)-like protein in oxidative stress tolerance. *The Plant Journal,* **48**(5), pp.743-756.

Nelson, N. and Yocum, C. F. 2006. Structure and function of photosystems I and II. *Annual Review of Plant Biology,* **57**, pp.521-565.

Neuhaus, H. E. and Emes, M. J. 2000. NONPHOTOSYNTHETIC METABOLISM IN PLASTIDS. *Annual Review of Plant Physiology and Plant Molecular Biology,* **51**(1), pp.111-140.

Neupert, W. and Herrmann, J. M. 2007. Translocation of proteins into mitochondria. *Annual Review of Biochemistry,* **76**, pp.723-749.

Nishimura, K., Ashida, H., Ogawa, T. and Yokota, A. 2010. A DEAD box protein is required for formation of a hidden break in Arabidopsis chloroplast 23S rRNA. *Plant Journal,* **63**(5), pp.766-777.

Novitskaya, L., Trevanion, S. J., Driscoll, S., Foyer, C. H. and Noctor, G. 2002. How does photorespiration modulate leaf amino acid contents? A dual approach through modelling and metabolite analysis. *Plant, Cell & Environment,* **25**(7), pp.821-835.

Ohad, N., Shichrur, K. and Yalovsky, S. 2007. The Analysis of Protein-Protein Interactions in Plants by Bimolecular Fluorescence Complementation. *Plant Physiology,* **145**(4), pp.1090-1099.

Patro, R., Duggal, G., Love, M. I., Irizarry, R. A. and Kingsford, C. 2017. Salmon provides fast and bias-aware quantification of transcript expression. *Nature Methods,* **14**, p417.

Pogson, B. J., Ganguly, D. and Albrecht-Borth, V. 2015. Insights into chloroplast biogenesis and development. *Biochimica et Biophysica Acta* **1847**(9), pp.1017-1024.

Pospisil, P. 2009. Production of reactive oxygen species by photosystem II. *Biochimica et Biophysica Acta,* **1787**(10), pp.1151-1160.

Pospisil, P. 2016. Production of Reactive Oxygen Species by Photosystem II as a Response to Light and Temperature Stress. *Frontiers in Plant Science,* **7**, pp.1950-1950.

Pourtau, N., Jennings, R., Pelzer, E., Pallas, J. and Wingler, A. 2006. Effect of sugar-induced senescence on gene expression and implications for the regulation of senescence in *Arabidopsis*. *Planta,* **224**(3), pp.556-568.

Pribil, M., Labs, M. and Leister, D. 2014. Structure and dynamics of thylakoids in land plants. *Journal of Experimental Botany,* **65**(8), pp.1955-1972.

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.

Rehling, P., Model, K., Brandner, K., Kovermann, P., Sickmann, A., Meyer, H. E., Kuhlbrandt, W., Wagner, R., Truscott, K. N. and Pfanner, N. 2003. Protein insertion into the mitochondrial inner membrane by a twin-pore translocase. *Science,* **299**(5613), pp.1747-1751.

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.

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W. and Smyth, G. K. 2015. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res,* **43**(7).

Rolland, N., Curien, G., Finazzi, G., Kuntz, M., Marechal, E., Matringe, M., Ravanel, S. and Seigneurin-Berny, D. 2012. The biosynthetic capacities of the plastids and integration between cytoplasmic and chloroplast processes. *Annual Review of Genetics,* **46**, pp.233-264.

Ruban, A. V. 2014. Evolution under the sun: optimizing light harvesting in photosynthesis. *Journal of Experimental Botany,* **66**(1), pp.7-23.

Ruckle, M. E., DeMarco, S. M. and Larkin, R. M. 2007. Plastid Signals Remodel Light Signaling Networks and Are Essential for Efficient Chloroplast Biogenesis in Arabidopsis. *American Society of Plant Biologists,* **19**(12), pp.3944-3960.

Saito, K. and Matsuda, F. 2010. Metabolomics for functional genomics, systems biology, and biotechnology. *Annual Review of Plant Biology,* **61**, pp.463-489.

Sakamoto, W., Miyagishima, S.-y. and Jarvis, P. 2008. Chloroplast Biogenesis: Control of Plastid Development, Protein Import, Division and Inheritance. *The Arabidopsis Book,* **6**, p0110.

Salleh, F. M., Evans, K., Goodall, B., Machin, H., Mowla, S. B., Mur, L. A. J., Runions, J., Theodoulou, F. L., Foyer, C. H. and Rogers, H. J. 2012. A novel function for a redox-related LEA protein (SAG21/AtLEA5) in root development and biotic stress responses. *Plant, Cell & Environment,* **35**(2), pp.418-429.

Sasaki, Y. and Nagano, Y. 2004. Plant acetyl-CoA carboxylase: structure, biosynthesis, regulation, and gene manipulation for plant breeding. *Bioscience Biotechnology Biochemistry,* **68**(6), pp.1175-1184.

Satou, M., Enoki, H., Oikawa, A., Ohta, D., Saito, K., Hachiya, T., Sakakibara, H., Kusano, M., Fukushima, A., Saito, K., Kobayashi, M., Nagata, N., Myouga, F., Shinozaki, K. and Motohashi, R. 2014. Integrated analysis of transcriptome and metabolome of *Arabidopsis* albino or pale green mutants with disrupted nuclear-encoded chloroplast proteins. *Plant molecular biology,* **85**(4-5), pp.411-428.

Schleiff, E., Soll, J., Sveshnikova, N., Tien, R., Wright, S., Dabney-Smith, C., Subramanian, C. and Bruce, B. D. 2002. Structural and guanosine

triphosphate/diphosphate requirements for transit peptide recognition by the cytosolic domain of the chloroplast outer envelope receptor, Toc34. *Biochemistry,* **41**(6), pp.1934-1946.

Schmidt, O., Pfanner, N. and Meisinger, C. 2010. Mitochondrial protein import: from proteomics to functional mechanisms. *Nature Reviews Molecular Cell Biology,* **11**(9), pp.655-667.

Seki, M., Narusaka, M., Abe, H., Kasuga, M., Yamaguchi-Shinozaki, K., Carninci, P., Hayashizaki, Y. and Shinozaki, K. 2001. Monitoring the expression pattern of 1300 Arabidopsis genes under drought and cold stresses by using a full-length cDNA microarray. *Plant Cell,* **13**(1), pp.61-72.

Shahollari, B., Vadassery, J., Varma, A. and Oelmuller, R. 2007. A leucine-rich repeat protein is required for growth promotion and enhanced seed production mediated by the endophytic fungus *Piriformospora indica* in *Arabidopsis thaliana*. *Plant Journal,* **50**(1), pp.1-13.

Shapiguzov, A., Vainonen, J. P., Wrzaczek, M. and Kangasjarvi, J. 2012. ROS-talk - how the apoplast, the chloroplast, and the nucleus get the message through. *Frontiers in Plant Science,* **3**, p292.

Sharma, P., Jha, A. B., Dubey, R. S. and Pessarakli, M. 2012. Reactive Oxygen Species, Oxidative Damage, and Antioxidative Defense Mechanism in Plants under Stressful Conditions. *Journal of Botany,* **2012**, p26.

Shi, K., Gu, J., Guo, H., Zhao, L., Xie, Y., Xiong, H., Li, J., Zhao, S., Song, X. and Liu, L. 2017. Transcriptome and proteomic analyses reveal multiple differences associated with chloroplast development in the spaceflight-induced wheat albino mutant mta. *PLOS ONE,* **12**(5), p0177992.

Shikanai, T. and Fujii, S. 2013. Function of PPR proteins in plastid gene expression. *RNA Biology,* **10**(9), pp.1446-1456.

Soll, J. 2002. Protein import into chloroplasts. *Current Opinion in Plant Biology,* **5**(6), pp.529-535.

Soll, J. and Schleiff, E. 2004. Protein import into chloroplasts. *Nature Reviews Molecular Cell Biology,* **5**, p198.

Sreenivasulu, N., Usadel, B., Winter, A., Radchuk, V., Scholz, U., Stein, N., Weschke, W., Strickert, M., Close, T. J., Stitt, M., Graner, A. and Wobus, U. 2008. Barley grain maturation and germination: metabolic pathway and regulatory network commonalities and differences highlighted by new MapMan/PageMan profiling tools. *Plant Physiology,* **146**(4), pp.1738-1758.

Steiner, S., Schröter, Y., Pfalz, J. and Pfannschmidt, T. 2011. Identification of Essential Subunits in the Plastid-Encoded RNA Polymerase Complex Reveals Building Blocks for Proper Plastid Development. *Plant Physiology,* **157**(3), pp.1043-1055.

Stern, D. B., Goldschmidt-Clermont, M. and Hanson, M. R. 2010. Chloroplast RNA Metabolism. *Annual Review of Plant Biology,* **61**(1), pp.125-155.

Stonebloom, S., Burch-Smith, T., Kim, I., Meinke, D., Mindrinos, M. and Zambryski, P. 2009. Loss of the plant DEAD-box protein ISE1 leads to defective mitochondria and increased cell-to-cell transport via plasmodesmata. *Proceedings of the National Academy of Sciences of the United States of America,* **106**(40), pp.17229-17234.

Sun, Y. and Zerges, W. 2015. Translational regulation in chloroplasts for development and homeostasis. *Biochimica et Biophysica Acta (BBA) - Bioenergetics,* **1847**(9), pp.809-820.

Susek, R. E., Ausubel, F. M. and Chory, J. 1993. Signal transduction mutants of Arabidopsis uncouple nuclear CAB and RBCS gene expression from chloroplast development. *Cell,* **74**(5), pp.787-99.

Suzuki, N., Koussevitzky, S., Mittler, R. and Miller, G. 2012. ROS and redox signalling in the response of plants to abiotic stress. *Plant, Cell & Environment,* **35**(2), pp.259-270.

Sveshnikova, N., Soll, J. and Schleiff, E. 2000. Toc34 is a preprotein receptor regulated by GTP and phosphorylation. *Proceedings of the National Academy of Sciences of the United States of America,* **97**(9), pp.4973-4978.

Tan, B.-C., Joseph, L. M., Deng, W.-T., Liu, L., Li, Q.-B., Cline, K. and McCarty, D. R. 2003. Molecular characterization of the *Arabidopsis* 9-cis epoxycarotenoid dioxygenase gene family. *The Plant Journal,* **35**(1), pp.44-56.

Tang, W., Ji, Q., Huang, Y., Jiang, Z., Bao, M., Wang, H. and Lin, R. 2013. FAR-RED ELONGATED HYPOCOTYL3 and FAR-RED IMPAIRED RESPONSE1 transcription factors integrate light and abscisic acid signaling in *Arabidopsis*. *Plant Physiology,* **163**(2), pp.857-866.

Tang, W., Wang, W., Chen, D., Ji, Q., Jing, Y., Wang, H. and Lin, R. 2012. Transposase-derived proteins FHY3/FAR1 interact with PHYTOCHROME-INTERACTING FACTOR1 to regulate chlorophyll biosynthesis by modulating HEMB1 during deetiolation in *Arabidopsis*. *The Plant Cell,* **24**(5), pp.1984-2000.

Taylor, A. B., Smith, B. S., Kitada, S., Kojima, K., Miyaura, H., Otwinowski, Z., Ito, A. and Deisenhofer, J. 2001. Crystal structures of mitochondrial processing peptidase reveal the mode for specific cleavage of import signal sequences. *Structure,* **9**(7), pp.615-625.

Theis, J. and Schroda, M. 2016. Revisiting the photosystem II repair cycle. *Plant Signaling & Behavior,* **11**(9), p1218587.

Thomashow, M. F. 1999. Plant cold acclimation: Freezing Tolerance Genes and Regulatory Mechanisms. *Annual Review of Plant Physiology and Plant Molecular Biology,* **50**, pp.571-599.

Tiller, N. and Bock, R. 2014. The translational apparatus of plastids and its role in plant development. *Molecular Plant,* **7**(7), pp.1105-1120.

Tyystjarvi, E. 2013. Photoinhibition of photosystem II. *International Review of Cell and Molecular Biology,* **300**, pp.243-303.

Urao, T., Yamaguchi-Shinozaki, K. and Shinozaki, K. 2001. Plant Histidine Kinases: An Emerging Picture of Two-Component Signal Transduction in Hormone and Environmental Responses. *Science's STKE,* **2001**(109), pp.re18.

Van de Poel, B. and Van Der Straeten, D. 2014. 1-aminocyclopropane-1-carboxylic acid (ACC) in plants: more than just the precursor of ethylene. *Frontiers in Plant Science,* **5**, p640.

Vothknecht, U. C. and Westhoff, P. 2001. Biogenesis and origin of thylakoid membranes. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research,* **1541**(1), pp.91-101.

Wang, R., Guegler, K., LaBrie, S. T. and Crawford, N. M. 2000. Genomic analysis of a nutrient response in Arabidopsis reveals diverse expression patterns and novel metabolic and potential regulatory genes induced by nitrate. *Plant Cell,* **12**(8), pp.1491-1509.

Wang, Y., Wang, C., Zheng, M., Lyu, J., Xu, Y., Li, X., Niu, M., Long, W., Wang, D., Wang, H., Terzaghi, W., Wang, Y. and Wan, J. 2016. WHITE PANICLE1, a Val-tRNA Synthetase Regulating Chloroplast Ribosome Biogenesis in Rice, Is Essential for Early Chloroplast Development. *Plant Physiology,* **170**(4), pp.2110-2123.

Waters, M. T. and Langdale, J. A. 2009. The making of a chloroplast. *The EMBO Journal,* **28**(19), pp.2861-2873.

Waterworth, W. M., Drury, G. E., Blundell-Hunter, G. and West, C. E. 2015. Arabidopsis TAF1 is an MRE11-interacting protein required for resistance to genotoxic stress and viability of the male gametophyte. *The Plant Journal,* **84**(3), pp.545-557.

Weaver, L. M., Gan, S., Quirino, B. and Amasino, R. M. 1998. A comparison of the expression patterns of several senescence-associated genes in response to stress and hormone treatment. *Plant Molecular Biology,* **37**(3), pp.455-469.

Weckwerth, W. 2003. Metabolomics in systems biology. *Annual Review of Plant Biology,* **54**, pp.669-89.

Wiedemann, N., Frazier, A. E. and Pfanner, N. 2004. The protein import machinery oof mitochondria. *Journal of Biological Chemistry,* **279**(15), pp.14473-14476.

Woodson, J. D. and Chory, J. 2008. Coordination of gene expression between organellar and nuclear genomes. *Nature Reviews Genetics,* **9**, p383.

Woodson, J. D., Perez-Ruiz, J. M. and Chory, J. 2011. Heme synthesis by plastid ferrochelatase I regulates nuclear gene expression in plants. *Current Biology,* **21**(10), pp.897-903.

Woodson, J. D., Perez-Ruiz, J. M., Schmitz, R. J., Ecker, J. R. and Chory, J. 2013. Sigma factor-mediated plastid retrograde signals control nuclear gene expression. *The Plant journal : for cell and molecular biology,* **73**(1), pp.1-13.

Wu, F.-H., Shen, S.-C., Lee, L.-Y., Lee, S.-H., Chan, M.-T. and Lin, C.-S. 2009. Tape-Arabidopsis Sandwich - a simpler *Arabidopsis* protoplast isolation method. *Plant Methods,* **5**(1), p16.

Wu, G.-Z., Meyer, E. H., Richter, A. S., Schuster, M., Ling, Q., Schöttler, M. A., Walther, D., Zoschke, R., Grimm, B., Jarvis, R. P. and Bock, R. 2019. Control of retrograde signalling by protein import and cytosolic folding stress. *Nature Plants,* **5**(5), pp.525-538.

Xiao, W., Sheen, J. and Jang, J.-C. 2000. The role of hexokinase in plant sugar signal transduction and growth and development. *Plant Molecular Biology,* **44**(4), pp.451-461.

Xiao, Y., Savchenko, T., Baidoo, E. E., Chehab, W. E., Hayden, D. M., Tolstikov, V., Corwin, J. A., Kliebenstein, D. J., Keasling, J. D. and Dehesh, K. 2012. Retrograde signaling by the plastidial metabolite MEcPP regulates expression of nuclear stress-response genes. *Cell,* **149**(7), pp.1525-1535.

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), p437.

Yagi, Y. and Shiina, T. 2014. Recent advances in the study of chloroplast gene expression and its evolution. *Frontiers in Plant Science,* **5**(61).

Yamaguchi, K. and Subramanian, A. R. 2000. The plastid ribosomal proteins. Identification of all the proteins in the 50 S subunit of an organelle ribosome (chloroplast). *Journal of Biological Chemistry,* **275**(37), pp.28466-28482.

Yogev, O., Karniely, S. and Pines, O. 2007. Translation-coupled translocation of yeast fumarase into mitochondria in vivo. *Journal of Biological Chemistry,* **282**(40), pp.29222-29229.

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

Yoo, S.-D., Cho, Y.-H. and Sheen, J. 2007b. Arabidopsis mesophyll protoplasts: a versatile cell system for transient gene expression analysis. *Nature Protocols,* **2**, p1565.

Zhang, X. P. and Glaser, E. 2002. Interaction of plant mitochondrial and chloroplast signal peptides with the Hsp70 molecular chaperone. *Trends in Plant Science,* **7**(1), pp.14-21.

Zhelyazkova, P., Sharma, C. M., Förstner, K. U., Liere, K., Vogel, J. and Börner, T. 2012. The Primary Transcriptome of Barley Chloroplasts: Numerous Noncoding RNAs and the Dominating Role of the Plastid-Encoded RNA Polymerase. *The Plant Cell,* **24**(1), pp.123-136.

Zhou, X., Fei, Z., Thannhauser, T. W. and Li, L. 2011. Transcriptome analysis of ectopic chloroplast development in green curd cauliflower (*Brassica oleracea L. var. botrytis*). *BMC Plant Biology,* **11**, p169.

Zimmermann, P., Hirsch-Hoffmann, M., Hennig, L. and Gruissem, W. 2004. GENEVESTIGATOR. Arabidopsis microarray database and analysis toolbox. *Plant Physiology,* **136**(1), pp.2621-2632.

Zoschke, R. and Bock, R. 2018. Chloroplast Translation: Structural and Functional Organization, Operational Control, and Regulation. *The Plant Cell,* **30**(4), pp.745-770.

Zybailov, B., Rutschow, H., Friso, G., Rudella, A., Emanuelsson, O., Sun, Q. and van Wijk, K. J. 2008. Sorting signals, N-terminal modifications and abundance of the chloroplast proteome. *PLoS One,* **3**(4), p1994.