The intracellular distribution of the WHIRLY1 protein and its functions in early barley leaf development

Barbara Karpinska[1], Nurhayati Razak[2], Euan James[3], Jenny Morris[3], Susan Verrall[3], Pete Hedley[3], Robert Hancock[3], and Christine Foyer[1]

[1]University of Birmingham
[2]University of Leeds
[3]The James Hutton Institute

April 05, 2024

Abstract

The WHIRLY (WHY) DNA/RNA binding proteins fulfil multiple but poorly characterised functions in leaf development. WHY1 transcript levels were highest in the bases of 7-day old barley leaves. Immunogold labelling revealed that the WHY1 protein was more abundant in the nuclei than the proplastids of the leaf bases. Transcript and metabolite profiling analysis of barley lines (W1-1 and W1-7) lacking WHY1, which show delayed greening compared to the wild type. While the transcript profile of leaf development was largely unchanged in W1-1 and W1-7 leaves, there were differences in levels of several transcripts encoding transcription factors associated with chloroplast development. These include a barley homologue of the Arabidopsis GATA transcription factor that regulates stomatal development, greening and chloroplast development, NAC1, two transcripts with similarity to Arabidopsis GLK1 and two transcripts encoding ARF transcriptions factors with functions in leaf morphogenesis and development. Chloroplast proteins were less abundant in the W1-1 and W1-7 leaves than the wildtype. The levels of TCA cycle metabolites and GABA were significantly lower in WHY1 knockdown leaves than the wild type. We conclude that WHY1 functions in the nuclei of the cells in the leaf bases contributes to the regulation of chloroplast development.

The intracellular distribution of the WHIRLY1 protein and its functions in early barley leaf development

Barbara Karpinska[1*], Nurhayati Razak[1*], Euan K. James[2], Jenny A. Morris[3], Susan R. Verrall[2], Pete E. Hedley[3], Robert D. Hancock[3] and Christine H. Foyer[1]

[1] School of Biosciences**,** College of Life and Environmental Sciences, University of Birmingham, Edgbaston, B15 2TT, UK

[2] Environmental Sciences, The James Hutton Institute, Invergowrie, Dundee DD2 5DA. United Kingdom.

[3] Cell and Molecular Sciences, The James Hutton Institute, Invergowrie, Dundee, DD2 5DA. United Kingdom.

*These authors contributed equally to the work reported here

Corresponding author: C.H.Foyer@bham.ac.uk

Running Head: WHIRLY1 functions in barley leaf development

Abstract

The WHIRLY (WHY) DNA/RNA binding proteins fulfil multiple but poorly characterised functions in leaf development. WHY1 transcript levels were highest in the bases of 7-day old barley leaves. Immunogold

labelling revealed that the WHY1 protein was more abundant in the nuclei than the proplastids of the leaf bases. Transcript and metabolite profiling analysis of barley lines (W1-1 and W1-7) lacking WHY1, which show delayed greening compared to the wild type. While the transcript profile of leaf development was largely unchanged in W1-1 and W1-7 leaves, there were differences in levels of several transcripts encoding transcription factors associated with chloroplast development. These include a barley homologue of the Arabidopsis GATA transcription factor that regulates stomatal development, greening and chloroplast development, NAC1, two transcripts with similarity to Arabidopsis GLK1 and two transcripts encoding ARF transcriptions factors with functions in leaf morphogenesis and development. Chloroplast proteins were less abundant in the W1-1 and W1-7 leaves than the wildtype. The levels of TCA cycle metabolites and GABA were significantly lower in WHY1 knockdown leaves than the wild type. We conclude that WHY1 functions in the nuclei of the cells in the leaf bases contributes to the regulation of chloroplast development.

Key words: D1 protein, chloroplast biogenesis, GENOMES UNCOUPLED; Photosynthesis, plastid-encoded RNA polymerase, RuBiSCO.

Abbreviations

Large subunit of RuBiSCO (RBCL); light harvesting chlorophyll a/b binding complex (LHC*);* mitochondrial dysfunction stimulon (MDS); nuclear (n); nuclear encoded RNA polymerase (NEP); photosynthesis-associated nuclear genes (*PhANGs*); plastid (pt); plastid-encoded RNA polymerase (PEP); RADICAL-INDUCED CELL DEATH1 (RCD1); ribulose-1, 5-bisphoshate carboxylase (RuBiSCO); small RuBiSCO subunit of (RBCS); RNA-binding proteins (RBPs)

Background

Cellular compartmentalization is essential for the regulation of metabolism and gene expression (Harrington, Feliu, Wiuf, & Stumpf, 2013). Reciprocal communication between the mitochondria, chloroplasts and nuclei is not only vital for the efficient functions of these compartments, but it also ensures the rapid adjustment of their protein content and composition to changing environmental conditions. Mitochondria-derived and plastid-derived retrograde signals are therefore important components in the regulation of nuclear gene expression (Diaz et al., 2018; Grubler et al., 2017; Kindgreen et al., 2012; Pogson, Woo, Forster, & Small, 2008). Retrograde signalling between the chloroplast and nucleus are not easily distinguished from those that operate during the proplastid-to-chloroplast transition in leaf development because of crossover between the biogenic and operational control of chloroplast functions (Pogson et al., 2008). In plastids, transcription is under the control of two types of RNA polymerases, a unique eubacterial-type plastid-encoded polymerase (PEP) and phage-type nucleus-encoded polymerases (NEPs). These RNA-polymerases specifically regulate the transcription of different subsets of genes but can also co-regulate a portion of the plastidial genes. The formation of chloroplasts from proplastids requires the establishment of the PEP complex. The PEP complex, which 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, & Pfannschmidt, 2011). PEP status/activity provides positive retrograde signals from the chloroplasts that convey essential information to the nucleus to promote *PhANG* expression.

The WHY family of proteins, which are specific to the plant kingdom (Desveaux et al., 2004) have a putative KGKAAL DNA binding domain that allows binding to ssDNA molecules of differing nucleotide sequence (Grabowski, Miao, Mulisch, & Krupinska, 2008) which may allow them to function as PAPs allowing the possibility of a functional interaction between these proteins (Días et al., 2017).

Mitochondria to nucleus signalling, which involves two key transcription factors: *ANAC013* and *ANAC017* , is also linked to plastid to nucleus signalling (Shapiguzov et al., 2019). The *ANAC013* and*ANAC017* transcription factors are released from the endoplasmic reticulum upon perception of appropriate signals and translocated to the nucleus, where they activate the expression of a specific set of genes called mitochondrial dysfunction stimulon (MDS) genes that include the alternative oxidases, *SOT12* and *ANAC013* (De Clercq et al., 2013; Safrany et al., 2008). The enhanced expression of *ANAC013*provides positive feedback regulation

of the signalling pathway. The nuclear-localised RADICAL-INDUCED CELL DEATH1 (RCD1) protein suppresses ANAC013 and ANAC017 functions (Shapiguzov et al., 2019). In addition, SOT12 belongs to the group of MDs genes that overlap with the genes induced by the SAL1, 3'-phosphoadenosine 5'-phosphate (PAP) chloroplast retrograde signalling pathway (Van Aken, & Pogson, 2017).

All plants have two *WHY* genes (*WHY1* and *WHY2*). *WHY1* encodes a protein that is located in chloroplasts and nuclei while *WHY2* encodes a mitochondria-targeted protein (Desveaux, Maréchal, & Brisson, 2005). WHY1 protein interacts with thylakoid membrane proteins and with the chloroplast nucleoids (Krupinska et al., 2014; Melonek et al., 2010). Unlike many other species, Arabidopsis has a third WHY gene, *AtWHY3* that is targeted to plastids (Krause et al., 2005). However, the intracellular localization of the WHY proteins appears to be flexible and determined by developmental and environmental signals. For example, the WHY2 protein that is primarily associated with mitochondrial nucleoids, was found in mitochondria, chloroplasts and nuclei during leaf senescence (Huang et al., 2020). Moreover, it appears that WHY3 can compensate for WHY2 in the Arabidopsis *why 2-1* mutant because WHY3 can be targeted to both chloroplasts and mitochondria (Golin et al., 2020). The expression of WHY2 in Arabidopsis decreased the expression of genes encoded by the chondriome (Maréchal et al., 2008). Similarly, the expression of the tomato SlWHY2 in transgenic tobacco plants led to mitochondrial gene transcription and stabilization of mitochondrial functions (Zhao et al., 2018).

Barley leaves deficient in the WHY1 protein have higher levels of chlorophyll than the wild type with an enhanced abundance of plastome-encoded transcripts (Comadira et al., 2015; Krupinska *et al.* , 2019). In contrast, the leaves of the Arabidopsis *why1* mutant and *why1why3* double mutants are phenotypically similar to the wild type. However, a *why1why3polIb-1* triple mutant defective in WHY1, WHY3, and the DNA polymerase 1B (Pol1B) exhibited a severe yellow-variegated phenotype (Lepage, Zampini, & Brisson, 2013). WHY1, WHY3 and RECA1 are associated with the chloroplast RNase H1 AtRNH1C protein and work together to maintain chloroplast genome integrity (Wang *et al* ., 2021). Maize transposon insertion lines in *WHY1* (Zmwhy1-1) have equivalent amounts of chloroplast DNA (cpDNA) to the wild type but are deficient in plastid ribosomes resulting in an albino phenotype (Prikryl, Watkins, Friso, van Wijk, & Barkam, 2008).

We have previously characterized the phenotypes, and metabolite and transcriptome profiles of three RNAi-knockdown barley lines (W1-1, W1-7 and W1-9) that have very low levels of *HvWHY1* expression (Comadira et al., 2015). The formation of plastid ribosomes and the establishment of photosynthesis was delayed in the RNAi-knockdown barley lines (Krupinska et al., 2019). Given the growing number of reports showing that WHY1 functions as a transcription factor in the nucleus, regulating the expression of genes involved in a wide range of processes including phytohormone synthesis, development and defence, we posed the hypothesis that WHY1 in the nuclei of developing leaves could also control chloroplast development. We therefore investigated the intracellular distribution of WHY1 between proplastids and nuclei in the bases of developing wild type barley leaves. We also characterized the transcript and metabolite profiles of barley lines (W1-1 and W1-7) lacking WHY1. We discuss the data indicating that the observed delay in plastid development in barley lines lacking WHY1 results from functions of the protein in the nuclei as well as the plastids.

Materials and Methods

Plant Material and Growth Conditions

Seeds of two independent transgenic barley (*Hordeum vulgare* L. cv. Golden Promise) lines (W1-1 and W1-7) with RNAi knockdown of the *WHIRLY1* gene and wild type controls were produced as described previously (Krupinska et al., 2002). Barley seeds (1 per pot) were sown in pots in compost (SHL professional potting compost) in controlled environment chambers with a 16h light/8h dark photoperiod, with an irradiance of 250 μmol $m^{-2}s^{-1}$, 20ºC/16ºC day/night temperature regime and 60% relative humidity. The primary leaves were harvested after 7 days.

Leaf Pigments

Leaf pigments were extracted from barley leaves with 80% (v/v) acetone. Absorbance of chlorophylls was measured at 663 and 646 nm and concentrations were calculated using the formula given by Lichtenthaler (1987).

qPCR

Reverse transcription of 1 µg of RNA into cDNA was performed using the QuantiTect Reverse Transcription Kit (Qiagen). The qPCR was performed using QuantiFast SYBR Green PCR kit (Qiagen) in the presence of 0.5 µM primers in a CFX96 thermocycler (Biorad, Hercules, CA, USA) following the manufacturer's instructions. PCR conditions were as follows: incubation at 95 °C for 5 min, 45 cycles 10 s 95°C and 30 s 60°C. Additionally melting curve analysis was performed at the end of each run to ensure specificity of the products. The same master mix without cDNA was used as negative control. The following primers were used: WHY1 (AK365452.1) Fwd 5´-GATGGGAATGGTCGCTTTTT -3´, Rev 5´-CCATGATGTGCGGTATGATG -3´), ACTIN 11 (AY145451) Fwd 5'- CGACAATGGAACCGGAATG-3', Rev 5'-CCCTTGGCGCATCATCTC-3') and Elongation factor 1- a (Z50789) Fwd 5´-TTGGTGGCATTGGAACTGTG -3´Rev 5'-CAAACCCACGCTTGAGATCC-3´.

Microarray processing and analysis

Microarray processing and analysis was performed on leaf RNA extracts from three biological replicates per treatment using a custom designed barley Agilent microarray (A-MEXP-2357; www.ebi.ac.uk/array-express) as previously described (Comadira et al 2015). Raw data can be accessed via the array express website (www.ebi.ac.uk/array-express) using accession number (E-MTAB-9882).

Western Blots

Total proteins were extracted with protein extraction buffer (Agrisera) supplemented with 5 mM DTT and the cocktail of protease inhibitors to prevent protein degradation. 10 µg of proteins were separated on 15% acrylamide SDS-PAGE and transferred to 0.45 -µm nitrocellulose membrane (Amersham 10600003). 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.

Immunogold Labeling

Young barley leaves were cut transversely with a fresh razorblade into 1 mm diameter strips. The samples were then prepared for immunogold labelling (IGL) and transmission electron microscopy (TEM) according to Rubio et al. (2019). Briefly, they were fixed immediately in 4% paraformaldehyde + 0.5% glutaraldehyde in 0.05 M sodium cacodylate (pH 7.0) overnight. After dehydration in ethanol the samples were infiltrated and embedded in LR White resin. The leaf samples were sectioned on a Leica UCT ultramicrotome and the sections (80 nm) collected on Ni grids coated with pyroxylin. After 1 h blocking in IGL buffer (Rubio et al. 2019) the grids were incubated for 2 h at room temperature in a polyclonal antibody (diluted 1:10 in IGL buffer) raised against barley WHIRLY1 in goat (Agrisera, Vannas, Sweden). After two washes in IGL buffer (10 min per wash) the grids were then incubated for 2 h in rabbit anti-goat IgG 10 nm gold (Aurion Immuno Gold Reagents & Accessories, Wageningen, Netherlands). The grids were finally washed twice in IGL buffer (5 min per wash) and 10 times in dH2O (30 s per wash). The immunogold labelled grids were viewed and photographed under a JEOL JEM1400 transmission electron microscope, and gold particles were counted on 8 - 10 representative images from each sample taken at the same magnification (x4000).

RESULTS

Expression and Localisation of WHY1 in Wild type and WHY knockdown Barley Leaves

Since monocotyledonous leaves show a gradient of development from base to tip, we sampled the base and

tip of the first leaves of 7-day old wild type seedlings to establish the intracellular localisation of WHY1. The levels of WHY1 transcripts were highest in the basal regions of 7-day old wild type leaves decreasing progressively from the middle to tip (Fig. 1). Immunogold labelling revealed that the WHY1 protein was found in developing plastids of the cells in the lowest region of the leaf bases that are not exposed to light (Fig. 2A), as well as in mature chloroplasts in the leaf tip (Fig. 2B). Furthermore, imaging of the nucleus in the leaf base revealed that gold labelling was abundant (Fig 2C, D) where it appeared to be mostly clustered with electron dense chromatin (Fig. S1). In the leaf base we observed an average density of 15.88 ± 2.04 (n = 8) gold particles in plastids and 26.30 ± 2.40 (n = 10) gold particles in nuclei however, due to the presence of large vacuoles, we were unable to obtain high quality nuclear images from the leaf tip. Labelling was largely absent from the cytosol (Fig. 2).

WHY1-deficient barley leaves develop in a similar manner to the wild type except that the greening of each leaf is delayed in the absence of WHY1 (Krupinska et al., 2020). To obtain an understanding of developmental delay we divided leaves into three sections; base, middle and tip. The basal sections of 7-day old wild type barley leaves had significantly less chlorophyll (Fig. S2) than the middle and tip sections. A similar chlorophyll gradient was observed in W1-1 and W1-7 leaves although the chlorophyll content in each section was lower than in the corresponding wild type leaves (Fig. S2).

Transcript Profiles of Leaf Development Indicate Co-ordinated Cell Development and Maturation

To obtain a broader understanding of developmental processes in wild type plants we conducted a transcriptomic comparison of the base, middle and tip regions of seven day old leaves. One-way analysis of variance identified 440 transcripts that were significantly differentially abundant ($P < 0.05$) in the different leaf regions. Transcripts were subjected to hierarchical clustering analysis which revealed five major clusters (Fig. 3). Cluster A comprised 69 transcripts that exhibited a gradient of abundance from low in the leaf base to high in the tip. This cluster included six transcripts encoding transcription factors homologous to Arabidopsis transcripts that have been shown to have roles in leaf development (Table S1). MLOC_74058.1 exhibits homology to an Arabidopsis transcription factor NGATHA3 (AT1G01030) involved in the control of leaf shape and expressed in leaf tips under the control of TCP (TEOSINTE BRANCHED 1, CYCLOIDEA and PROLIFERATING CELL FACTOR) transcription factors (Ballester et al., 2015). The latter family were represented by MLOC_14785.1 which exhibited homology to Arabidopsis TCP5 (AT5G60970). A gene (MLOC_70809.1) encoding a homologue of Arabidopsis GATA, NITRATE-INDUCIBLE, CARBON METABOLISM INVOLVED (GNC) transcription factor (AT5G56860) that regulates stomatal development, greening and chloroplast development (Bastakis, Hedtke, Klermund, Grimm, & Schwechheimer, 2018; Klermund et al., 2016; Zubo et al., 2018) was also present in cluster A.

Two of the transcription factors identified within cluster A (Fig. 3) were associated with the control of senescence in response to metabolic signals. AK373121 exhibits homology to an Arabidopsis zinc finger family protein METHYLENE BLUE SENSITIVITY 1 (MBS1; AT3G02790) responsible for acclimation or cell death in dose-dependent response to $^1O_2$ (Shumbe et al., 2017) while MLOC_64240.2 and MLOC_53744.1 both share homology to AT1G56010 encoding NAC1, a senescence associated transcription factor under the control of auxin (Kim et al., 2011). Further evidence for the upregulation of senescence-associated processes in the leaf tip was the increased abundance of transcripts (AK370424, MLOC_47161.1) encoding proteins with homology to AUXIN-INDUCED IN ROOT CULTURES 3 (AT2G04160) and SENESCENCE-ASSOCIATED GENE 12 (AT5G45890; SAG12), endopeptidases required for protein turnover (James et al., 2018; Neuteboom, Veth-Tello, Cludesdale, Hooykaas, & van der Zaal, 1999). Furthermore, several transcripts (MLOC_56129.2, MLOC_57630.1, AK374126) encoding proteins homologous to proteins required for ubiquitin mediated protein turnover exhibited greatest abundance in the leaf tip (Table S1).

Cluster B was the largest of the clusters comprising 187 transcripts that exhibited a gradient of abundance from high to low from the leaf base to the leaf tip (Fig. 3). Seventeen transcripts encoding transcription factors were identified, several of which exhibited homology to Arabidopsis transcripts with functions in photomorphogenesis and development. Two transcripts (AK364144, MLOC_73144.4) showed homology to Arabidopsis auxin response factors (AT4G30080, AT1G19220; ARF) with functions in leaf morphogenesis

and development (Liu, Jia, Wang, & He, 2011; Schuetz, Fidanza, & Mattsson, 2019). Similarly, AK376150 and AK365841 are homologues of Arabidopsis genes encoding INDETRMINATE DOMAIN 15 (AT2G01940) and GATA TRANSCRIPTION FACTOR 2 (AT2G45050), with functions in leaf morphogenesis and photomorphogenesis, respectively (Cui et al., 2013; Luo et al., 2010). As described below, a feature of cluster B were large numbers of transcripts associated with lipid and wax metabolism. Interestingly, we identified a transcript (AK364135) with homology to an Arabidopsis transcript encoding the class I TCP transcription factor TCP14 (AT3G47620). In Arabidopsis class I TCP transcription factors including TCP14 are master regulators of cuticle biosynthesis (Camoirano et al., 2020) and are required for the induction of genes involved in gibberellin biosynthesis and cell expansion in response to temperature (Ferrero, Viola, Ariel, & Gonzalez, 2019). Similarly, several transcripts in cluster B were associated with polyphenol metabolism and a transcription factor (AK361986) homologous to Arabidopsis MYB4 (AT4G38620) which functions in the control of flavonoid biosynthesis (Wang et al., 2020) was also identified in this cluster.

Consistent with the hypothesis that cells at the leaf base were undergoing division and expansion, 14 transcripts categorised as cell wall associated were identified in cluster B (Table S1). These included several transcripts (AK248822.1, AK356936, MLOC_36439.1, MLOC_43237.1, MLOC_12096.1, MLOC_73204.3) with homology to transcripts encoding Arabidopsis expansins with a well-established role in cell wall loosening, leaf initiation and subsequent growth (Marowa, Ding, & Kong, 2016). A further three transcripts (MLOC_61972.1, AK361522, AK361278) encoded xyloglucan endotransglycosylases (XTHs) that function in cell expansion by loosening cell walls (Rose, Braam, Fry, & Nishitani, 2002). Furthermore, transcripts encoding two pectin modifying enzymes, a methylesterase (MLOC_54267.1) and an acetylesterase (MLOC_55102.5) were highly abundant in the leaf base.

Transcripts associated with lipid metabolism were also highly represented within cluster B, consistent with the hypothesis that active cuticle biosynthesis is occurring in the basal portion of the leaf. For example, MLOC_67622.1 and MLOC_45058.1 both exhibited homology to Arabidopsis transcripts encoding 3-KETOACYL-COA SYNTHASE 6 (KCS6, AT1G68530). Plants carrying mutations in *KCS6* exhibited significant reductions in branched and unbranched long chain alkanes and alcohols in cuticular wax (Buster, & Jetter, 2017). Similarly, AK252678.1 shared homology with AT5G43760 encoding KCS20 a very long chain fatty acid synthase required for cuticular wax and root suberin biosynthesis (Lee et al., 2009). Two transcripts (MLOC_54056.1 and AK370579) shared homology with AT1G02205 encoding ECERIFERUM 1, mutants of which exhibit similar defects to *KCS6* mutant lines (Buster, & Jetter, 2017). Taken together these data suggest that the basal portion of the leaf comprises actively dividing/expanding cells.

Cluster C (Fig. 3) comprised 43 transcripts that were abundant at the leaf base, scarce in the middle section of the leaf with intermediate abundance in the leaf tip. Cluster D comprised 127 transcripts that exhibited a similar high to low abundance profile during leaf maturation as observed for cluster B. However, the expression gradient in cluster D was considerably greater than observed for cluster B. Like cluster B, cluster D contained several transcripts encoding proteins associated with cell wall metabolism including expansins, XTHs and pectin modifying enzymes (Table S1). Cluster D additionally contained transcripts encoding proteins required for cellulose biosynthesis where two transcripts (MLOC_66568.3, MLOC_68431.4) exhibited homology to Arabidopsis cellulose synthases (AT5G44030, AT5G17420) and a further two transcripts (MLOC_7722.1, AK370617) exhibited homology to an Arabidopsis transcript encoding the membrane anchored COBRA-LIKE 4 (AT5G15630) which plays a key function in cellulose deposition (Brown, Zeef, Ellis, Goodacre, & Turner, 2005).

Furthermore, cluster D contained transcripts associated cell expansion, cell polarity, organ patterning and development. MLOC_53132.1 exhibited significant homology to Arabidopsis transcripts (AT4G08950) encoding EXORDIUM, a brassinosteroid responsive gene that acts upstream of wall-associated kinases and expansins to promote cell expansion (Schröder, Lisso, Lange, & Müssig, 2009). Indeed, a transcript (AK364262) with homology to the Arabidopsis transcript encoding WALL-ASSOCIATED KINASE 2 (AT1G21270) required for turgor driven cell expansion was also identified (Kohorn et al., 2006). Several transcripts associated with vascular development and patterning were present. MLOC_58644.1 and AK359559 exhibited homology to

Arabidopsis AT2G34710 and AT5G62880 encoding PHABULOSA and ROP11, respectively which play roles in xylem patterning during early cellular differentiation (Müller et al., 2016; Nagashima et al., 2018). Furthermore, MLOC_69397.2 homologous to an Arabidopsis transcript (AT1G79430) encoding ALTERED PHLOEM DEVLOPMENT with a role promoting phloem development was present (Bonke, Thitamadee, Mähönen, hauser, & Helariutta, 2003). Finally, a transcript encoding a basic-helix-loop-helix transcription factor (MLOC_55768.1) with similarity to an Arabidopsis transcript encoding bHLH93 (AT5G65640) was also present. This transcription factor interacts with FAMA which controls differentiation of guard cells in the leaf epidermis (Ohashi-Ito and Bergmann, 2006). Taken together these results imply active cellular expansion and differentiation in the base of the leaf and indicate that these processes are complete in the more mature leaf regions. Transcripts in cluster E displayed a similar pattern of abundance to those in cluster C with minimum abundance in the mid region of the leaf.

Loss of WHY1 has Specific Effects on Leaf Transcript Abundance

Knockdown of the WHIRLY protein clearly delayed greening of emerging barley leaves (Fig. S2). To determine whether the WHIRLY protein influences other aspects of leaf development, a comparative transcriptome analysis of basal, mid and tip sections of leaves of wild type and two whirly knockdown lines was conducted. Significant differences in transcript abundance ($P<0.05$) based upon leaf position, genotype or the interaction of the two factors was determined using 2-way analysis of variance which revealed 1732 transcripts dependent on line, 2240 transcripts dependent on leaf region and only 23 which exhibited a distribution in abundance based on an interaction between the two factors (Table S2).

Of the 23 transcripts exhibiting a genotype by leaf region interaction in their patterns of abundance, three transcripts (AK370975, MLOC_56051.1, MLOC_56052.1) that exhibited homology to chlorophyll binding proteins had a low abundance in all sections of wild type leaves and although at higher abundance in W1-1 leaves both genotypes exhibited a pattern of reducing abundance from base to tip. In contrast, W1-7 leaves exhibited a reverse pattern of abundance increasing from base to tip. A similar pattern of abundance was observed for a transcript encoding a thylakoid luminal protein (AK370198). These data are consistent with delayed assembly of the photosystems in WHY-deficient leaves. Two other transcripts encoding proteins with functions in plastid biogenesis and metabolism, respectively also exhibited a genotype by leaf region dependent expression pattern. MLOC_9203.2 encodes a tubulin-like protein with homology to Arabidopsis AT2G36250 encoding an FtsZ protein essential for chloroplast division (Osteryoung, Stokes, Rutherford, Percival, & Lee, 1998) while MLOC_69205.1 encodes an oxaloacetate/malate antiporter acting as a malate valve to balance NADPH/ATP ratios in the plastid (Selinski, & Schiebe, 2019).

We then compared the transcriptome profiles of the three developmental regions in the W1-1 and W1-7 leaves with those of the wild type leaves focusing on the key transcripts discussed above in the wild type developmental pattern (Fig. 4, Table S3). Of these transcripts, most exhibited similar patterns of abundance in the developmental profiles of all genotypes. Notable exceptions were MLOC_70809.1 encoding a homologue of Arabidopsis GATA, NITRATE-INDUCIBLE, CARBON METABOLISM INVOLVED (GNC) transcription factor (AT5G56860) that regulates stomatal development, greening and chloroplast development, MLOC_53744.1 encoding NAC1, a senescence associated transcription factor under the control of auxin, SAG12 (MLOC_47161.1) and two transcripts encoding ARF transcriptions factors (AK364144, MLOC_73144.4) with functions in leaf morphogenesis and development. The patterns of abundance of these transcripts were perturbed in the absence of WHIRLY1 relative to wild type. While SAG12 transcripts were more abundant in all the sections of the W1-7 line than the other genotypes, the ARF and NAC1 transcripts were less abundant in all sections relative to the other genotypes (Fig. 4). These differences are indicative of divergent developmental programmes in leaves deficient in the WHY1 protein.

As a key phenotype of WHY1 knockdown was reduced and delayed greening (Fig. S2) we next sought to examine the shift in abundance of transcripts associated with light signalling and plastid development in leaf regions of different age. Twenty-nine transcripts associated with light dependent plastid biogenesis were identified in the developmental profile of the wild type, of which 22 exhibited a high to low gradient of abundance from the leaf base to leaf tip (Fig. 5). Six transcripts encoded chloroplastic ribosomal proteins and

a further six had putative functions in plastid gene expression. Three transcripts (MLOC53063.1, AK3635024 and AK363292) exhibited similarity to AT2G03200 which encodes a putative chloroplast nucleoid DNA binding protein while a further two transcripts (MLOC_65040.1 and MLOC_69013.1) exhibited similarity to AT5G10770 containing similar domain structures (Table S4). The plastid nucleoids comprise multiple copies of DNA, RNA and a range of proteins with functions in replication, gene expression and DNA binding that are believed to control plastid gene expression in a way analogous to chromatin (Powikrowska, Oetke, Jensen, & Krupinska, 2014). A further transcript MLOC_9558.1 encoded a protein with similarity to plastid encoded RNA polymerase alpha subunit, essential for light dependent plastid biogenesis (Yoo et al., 2019). The levels of this transcript declined from base to tip in wild type and WHY 1-1 but the levels of this transcript were high in the base and tip regions of the WHY1-7 leaves.

Transcript Analysis of the Basal Section of Leaves Suggests Delayed Plastid Development and Perturbation of Plastid Gene Expression in WHY1 Knockdown Lines

In wild type leaves, transcripts associated with key developmental processes were highly abundant in the basal section of leaves and declined in the middle and tip sections. We therefore conducted one-way ANOVA to determine significantly differentially abundant transcripts in the base of wild type, W1-1 and W1-7 leaves. This identified 1267 transcripts that were differentially abundant dependent on genotype of which 540 were not assigned a function (Table S5). The remaining 727 transcripts were assigned to a range of functions associated with primary and secondary metabolism, RNA, DNA and protein processing, cellular organisation and development (Table S5).

Seventy-six transcripts encoding proteins associated with plastid biogenesis and development were differentially expressed, (Table S6; Fig 6) the majority of which were more abundant in the base of 7-day old leaves of WHY1 knockdown plants than the wild type. These included transcripts encoding elements of the photosynthetic electron transport chains and Calvin cycle proteins including Rubisco (Table S6). Furthermore, several transcripts (AK248405.1, MLOC_39198.3, MLOC_52167.2 and MLOC_63408.1) homologous to Arabidopsis transcripts encoding proteins required for FeS cluster assembly (Bastow, Bych, Crack, Le Brun, & Balk, 2016; Hu, Kato, Sumida, Tanaka, & Tanaka, 2017; Roland et al., 2020; Xu, Adams, Chua, & Møller, 2005;) were less abundant in the wild type than WHY1. Many transcripts encoded proteins associated with the targeting and translocation of proteins to the chloroplast as well as several proteins associated with protein folding such as AK250352.1 and AK361117 which both exhibit homology to AT3G60370 encoding an immunophillin with protein folding activity required for the assembly of PSII (Lima et al., 2006) and AK251420.1 encoding a chaperonin homologous to AT2G28000 required for plastid division (Suzuki et al., 2009). Transcripts homologous to Arabidopsis transcripts encoding several plastid localised Clp proteases were also present including several (MLOC_32972.1, MLOC_4257.1, MLOC_64141.1, MLOC_68297.2, MLOC_861.2) that were previously demonstrated to play a role in plastid biogenesis (Kim et al., 2009).

A small number of transcripts exhibited a pattern of low abundance in W1-7. These included two transcripts (MLOC_56051.1, MLOC_56052.1) similar to AT2G34420 encoding a chlorophyll a-b binding protein, MLOC_61567.1 similar to plastid encoded ATCG01010 encoding an NADH dehydrogenase subunit. Interestingly, the levels of two transcripts (MLOC_65876.1, AK353571) with similarity to an Arabidopsis transcript (AT2G20570; GLK1) encoding a transcription factor required for the expression of nuclear encoded photosynthetic genes (Waters et al., 2009) were lower in all sections of the WHY1-7 leaves.

Leaf Developmental Stage Influences Primary Metabolite Profiles

To identify shifts in metabolism associated with leaf development, untargeted GC/MS analysis of a range of polar and non-polar compounds was undertaken. A total of 107 chromatographic features were resolved representing 36 non-polar and 71 polar components. Thirty of the features were unidentified with the remaining 77 identified based on their relative retention times and mass spectra. Only 26 components exhibited significant differences in abundance dependent on leaf section of which 19 were present in the polar fraction and 7 in the non-polar fraction (Table S7). Twenty of these compounds were identified comprising sugars, organic acids, amino acids, fatty acids and alcohols (Fig. 7). A minor unoximated fructose peak exhibited a

decline from leaf base to tip however, this represented less than 10% of the total fructose pool and neither of the major oximated peaks exhibited significant change (Table S7). Mannitol exhibited a similar change in abundance to the unoximated fructose peak while galactose was least abundant in the leaf base but higher in the mid and tip regions (Fig. 7). The TCA cycle organic acids fumarate and succinate were most highly abundant in the base and mid regions of the leaf but declined dramatically in the leaf tip perhaps due to a redirection of flux into amino acids, the majority of which exhibited an increase as the leaf aged (Fig.7). Significantly, although both glycine and serine increased as the leaf aged, glycine increased to a much greater extent meaning a higher glycine/serine ratio indicative of increased photorespiration (Novitskaya, Trevanion, Driscoll, Foyer, & Noctor, 2002) from leaf base to tip. Of the four lipophilic compounds that exhibited significant changes in abundance, all increased in the leaf tip relative to the leaf basal region.

Loss of WHY1 has Specific Effects on C/N Metabolism

Several transcripts associated with primary metabolic pathways were differentially abundant in the basal portion of wild type and WHY1 knockdown leaves (Fig. 8A). Significant differences were observed in transcripts encoding enzymes associated with the Calvin cycle, starch and sugar metabolism, glycolysis, the TCA cycle and amino acid metabolism. Many of these transcripts were more abundant in WHY1 knockdown than wild type leaves although transcripts encoding hexokinase (MLOC_54094.1), β-amylase (AK368826) a methionine S-methyltransferase (AK368357), an O-acetylcysteine thiol-lyase (AK248898.1) and an arginase (MLOC_65968.1) were consistently less abundant in WHY1 knockdown lines.

Differences in transcript abundance were reflected by significant differences in primary metabolic profiles (Fig. 8B). Twenty-two of 71 polar compounds analysed by GC/MS were significantly differentially abundant in wild type and WHY1 knockdown leaf basal regions. Eight of these compounds were unidentified with the remainder comprising primarily organic and amino acids. All of the TCA cycle components detected were significantly lower in WHY1 knockdown leaves than in wild type leaves as was the non-proteinaceous amino acid γ-amino butyric acid (GABA) that functions as a cytosolic bypass of specific steps (Sweetlove, Beard, Nunes-Nesi, Fernie, & Ratcliffe, 2010). Similarly, serine and aspartate were present at lower concentrations in WHY1 knockdown leaves while glycine, valine, leucine, threonine and isoleucine were present at higher concentrations, particularly in W1-7 (Fig. 8B).

Delayed Accumulation of Photosynthetic Proteins in the WHY Knockdown Lines

The levels of LHCBII protein were slightly higher in all regions of the leaves of 7-day old W1-1 and W1-7 seedlings than the wild type (Fig. 9). In contrast, the D1 protein, the RPS1 and the RBCS and RBCL proteins were much less abundant in the leaves of 7-day old W1-1 and W1-7 seedlings than the wild type (Fig. 9).

Discussion

The concept that the distribution of WHY1 between the nuclei and chloroplasts plays a key role in the regulation of plant development has arisen in recent years (Ren et al., 2017; Lin e al., 2019). The phosphorylation of the WHY1 protein favours partitioning to nuclei, a process that increase with leaf age (Guan *et al* ., 2018). However, little information is available on the distribution of WHY1 between the nuclei and proplastids of developing leaves. The immunogold labelling analyses reported here revealed that over 60% of the WHY1 protein in the basal sections of the leaves was located in the nuclei (Fig. 2). Hence, it is likely that WHY1 fulfils its functions in the nuclei of developing leaves as well as the plastids. Chloroplast development is delayed in the absence of WHY1, a process that has largely been ascribed to the functions of WHY1 in chloroplasts (Prikryl, Watkins, Friso, van Wijk, & Barkam, 2008; Kurpinska et al., 2019). However, WHY1 functions as a transcriptional activator in the nucleus, binding to the AT-rich region of the kinesin gene promoter to activate kinesin gene expression (Xiong *et al.,* 2009), and to the GTCAAT motif of the *S40* promoter (Krupinska *et al.,* 2014), and to a combination motif of GTNNNAAT and AT-rich motif of downstream target genes, such as *WRKY53* , *WRKY33* , *SPO11* , and *PR1* to regulate leaf senescence and other processes in*A. thaliana*(Miao*et al.,*2013; Ren*et al.,* 2017; Huang *et al.* , 2018). Our transcriptome profiles of the three developmental regions in the W1-1 and W1-7 leaves compared to the wild type leaves provide

some insights into how the developmental pattern of transcripts is changed in the absence of WHY1. While most transcripts exhibited similar patterns of abundance in the developmental profiles of all genotypes there were some clear exceptions such as a GNC transcription factor regulating stomatal development, greening and chloroplast development and NAC1, an auxin-regulated senescence-associated transcription factor, two transcripts encoding ARF transcriptions factors with functions in leaf morphogenesis and development and two transcripts with similarity to Arabidopsis GLK1 that encodes a transcription factor required for the expression of nuclear encoded photosynthetic genes (Waters et al., 2009). Moreover, SAG12 transcripts were more abundant in all the sections of the W1-7 line than the other genotypes. These differences are indicative of divergent chloroplast development programmes in leaves deficient in the WHY1 protein.

The WHY1 protein binds to ERF-binding cis elements in the promoter regions of genes such as ERF109 (REDOX RESPONSIVE TRANSCRIPTION FACTOR 1, RRTF1) (Miao et al., 2013. ERF109 is involved in plant stress responses and participates in reactive oxygen species (ROS) signalling and the regulation of developmental programs, such as jasmonate-dependent initiation of lateral root development (Huang*et al.* , 2019). WHY proteins have previously been reported to be involved in the regulation of shoot and root development. For example, WHY2 was shown to be a major regulator of root apical meristem development (McCoy et al., 2021). Similarly, the expression of WHY1 (nWHY1) in the nucleus of Arabidopsis *why1* mutants led to changes in the levels of transcripts associated with plant development during early growth, whereas expression of WHY1 in plastids increased the abundance of transcripts associated with salicylic acid synthesis (Lin et al., 2020). The binding of WHY proteins to the PB element of the*9-cis epoxycarotenoid dioxygenase (NCED)1* gene in cassava activated expression leading to increased abscisic acid levels (Yan,*et al.,* 2020). Hence, the presence of WHY proteins in the nucleus clearly influences the expression of genes involved in the synthesis of phytohormones that control plant growth and defence. The primary cause of the delay in greening observed in the barley leaves lacking WHY1 may therefore result from the absence of WHY1-dependent regulation of nuclear gene expression.

The action WHY1 as a transcription factor in the nucleus also regulates the expression of genes associated with photosynthesis and carbon metabolism. For example, WHY1 binds to the promoter of the rbcS gene that encodes the small subunit of the potato ribulose-1, 5-carboxylase, oxygenase under cold stress (Zhuang *et al.* , 2020), while WHY2 binds to the promoters of the *SWEET11/15* genes that encode sucrose transporters, leading to the modulation of starch allocation and silique development (Huang *et al* ., 2020). Here we report that the absence of WHY1 has a significant impact on the levels of transcripts encoding enzymes associated with the Calvin cycle, starch and sugar metabolism, glycolysis, the TCA cycle and amino acid metabolism, many of which were more abundant in WHY1-deficient leaves than the wild type. However, transcripts encoding enzymes such as β-amylase were less abundant in WHY1 knockdown lines. WHY1 is known to bind to the ERE-like element of the *AMY3-L* promoter, activating the expression of amylase and starch degradation. WHY1 also binds to the ERE element of the *ISA2* promoter to inhibit isoamylase-mediated starch-synthesis (Zhuang *et al.* , 2019). The absence of WHY1 from the nuclei of developing barley leaves could therefore lead to the observed changes in primary metabolites reported here. For example, all the metabolites of the TCA cycle that were detected were significantly lower in WHY1 knockdown leaves than the wild type, as were GABA. Other amino acids such as glycine, valine, leucine, threonine and isoleucine were higher in WHY1-deficient leaves than the wild type. It may be that WHY1 can bind to promotors of a wide range of housekeeping genes in the nucleus, to modulate their expression in response to developmental and environmental signals.

The WHY1-deficient barley leaves showed delayed chloroplast ribosome formation and acquisition of photosynthetic activity suggesting that WHY1 contributes to the coordination of nuclear and plastome gene expression (Krupinska *et al.* , 2019). The role of the WHY proteins in organelle to the nucleus retrograde signalling has long been suspected (Foyer et al., 2014) but remains to be established. Similarly, the factors that determine the partitioning of WHY1 between the chloroplasts and nuclei remain to be fully characterised. The compartmentation of WHY1 between the plastids and nuclei is influenced by protein phosphorylation, particularly as the leaves enter senescence (Ren*et al* ., 2017). Phosphorylation of the WHY1 protein by CIPK14 kinase or oxidation by the addition of hydrogen peroxide causes a re-distribution of WHY1 from the

plastids to the nucleus (Ren et al., 2017; Lin et al., 2019). However, little information is available about the regulation of WHY1 partitioning in other situations such as the bases of developing leaves. Earlier evidence indicated that WHY1 can move from the plastids to the nucleus (Isemer et al., 2012). Direct transfer of WHY1 from the plastids to the nuclei through contact sites or stromules (Hanson et al, 2018) is possible but remains to be proven. Moreover, the plastid-localized WHY1 affects miRNA biogenesis in the nucleus (Swida-Barteczka et al., 2018) suggesting that WHY1 influences chloroplast to nucleus signalling.

In summary, our understanding of WHY protein functions have greatly increased in recent years, as has our knowledge of the flexibility of their localisation and overlap of functions. The data presented here provides new insights into the intracellular localization of the WHY1 in developing leaves, highlighting how WHY1 in the nucleus might control chloroplast development.

ACKNOWLEDGEMENTS

We thank BBSRC (UK) for financial support (BB/N004914/1). The James Hutton Institute is supported by the Scottish Government Rural and Environmental Science and Analytical Services division. The authors thank Professor Karin Krupinska for supplying the seeds of the transgenic barley lines used in the experiments.

References

Ballester P., Navarrete-Gómez M., Carbonero P., Oñate-Sánchez L. & Ferrándiz C. (2015). Leaf expansion in Arabidopsis is controlled by a TCP-NGA regulatory module likely conserved in distantly related species. Physiologia Plantarum, 155(1), 21-32.

Bastakis E., Hedtke B., Klermund C., Grimm B. & Schwechheimer C. (2018). LLM-Domain B-GATA transcription factors play multifaceted toles in controlling greening in Arabidopsis. The Plant Cell, 30, 582-599.

Bastow E.L., Bych K., Crack J.C., Le Brun N.E. & Balk J. (2017). NBP35 interacts with DRE2 in the maturation of cytosolic iron-sulphur proteins in *Arabidopsis thaliana* . The Plant Journal, 89, 590-600.

Bonke M., Thitamadee S., Mähönen A.P., Hauser M.-T. & Helariutta Y. (2003) APL regulates vascular tissue identity in Arabidopsis. Nature, 426, 181-186.

Brown D.M., Zeef L.A.H., Ellis J., Goodacre R. & Turner S.R. (2005). Identification of novel genes in Arabidopsis involved in secondary cell wall formation using expression profiling and reverse genetics. The Plant Cell, 17, 2281-2295.

Camoirano A., Arce A.L., Ariel F.D., Alem A.L., Gonzalez D.H. & Viola I.L. (2020). Class I TCP transcription factors regulate trichome branching and cuticle development in Arabidopsis. Journal of Experimental Botany, in press, doi:10.1093/jxb/eraa257.

Comadira G., Rasool B., Karpinska B., Morris J., Verrall S.R., Hedley P.E., Foyer C.H. & Hancock R.D. (2015). Nitrogen deficiency in barley (*Horedeum vukgare*) seedlings induces molecular and metabolic adjustments that trigger aphid resistance. Journal of Experimental Botany, 66(12), 3639-3655.

Comadira G., Rasool B., Karpinska B., Márquez García B., Morris J., Verrall S.R., Bayer M., Hedley P.E., Hancock R.D., & Foyer, C.H. (2015). WHIRLY1 functions in the control of responses to N-deficiency but not aphid infestation in barley (*Hordeum vulgare*). Plant Physiology, 168, 1140-1151.

De Clercq I., Vermeirssen V., Van Aken O., Vandepoele K., Murcha M.W., Law S.R., Inzé D., Ng S., Ivanova A., Rombaut D., van de Cotte B., Jaspers P., Van de Peer Y., Kangasjärvi J., Whelan J., & Van Breusegem F. (2013). The membrane-bound NAC transcription factor ANAC013 functions in mitochondrial retrograde regulation of the oxidative stress response in Arabidopsis. The Plant Cell, 25, 3472–3490.

Desveaux D., Subramaniam R., Després C., Mess J.N., Lévesque C., Fobert P.R., Dangl J.L., & Brisson N (2004). A "Whirly" transcription factor is required for salicylic acid dependent disease resistance in Arabidopsis. Developmental Cell, 6, 229-240.

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

Diaz M.G., Hernandez-Verdeja T., Kremnev D., Crawford T., Dubreuil C., & Strand A. (2018). Redox regulation of PEP activity during seedling establishment in Arabidopsis thaliana. Nature Communications, 9, 50.

Dietz K., & Pfannschmidt T. (2011) Novel regulators in photosynthetic redox control of plant metabolism and gene expression. Plant Physiology, 155, 1477-1485.

Foyer C.H., Karpinska B.. & 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, 369, 20130226.

Golin, S., Negroni, Y. L., Bennewitz, B., Klösgen, R. B., Mulisch, M., La Rocca, N., Cantele, F., Vigani, G., Lo Schiavo, F., Krupinska, K., & Zottini, M. (2020). WHIRLY2 plays a key role in mitochondria morphology, dynamics, and functionality in *Arabidopsis thaliana* . Plants Direct, 4, e00229. https://doi.org/10.1002/pld3.229

Grabowski E., Miao Y., Mulisch M., & 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, 1800–1804.

Grubler B., Merendino L., Twardziok S.O., Mininno M., Allorent G., Chevalier F., Liebers M., Blanvillain R., Mayer K.F.X., Lerbs-Mache S., Ravanel S., & Pfannschmidt T. (2017). Light and plastid signals regulate different sets of genes in the albino mutant Pap7-1. Plant Physiology, 175, 1203-1219.

Guan Z., Wang W., Yu X., Lin L., & Miao. (2018). Comparative proteomic analysis of coregulation of CIPK14 and WHIRLY1/3 mediated pale yellowing of leaves in Arabidopsis. International Journal of Molecular Sciences, 19, 2231.

Harrington H.A., Feliu E., Wiuf C., & Stumpf M.P.H. (2013). Cellular compartments cause multistability and allow cells to process more information. Biophysical Journal, 104, 1824-1831.

Hu X., Kato Y., Sumida A., Tanaka A. & Tanaka R. (2017). The $SUFBC_2D$ complex is required for the biogenesis of all major classes of plastid Fe-S proteins. The Plant Journal, 90, 235-248.

Huang D., Lin W., Deng B., Ren Y., & 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, 2352.

Huang H., Ullah F., Zhou D.-X., Yi M. & Zhao Y. (2019) Mechanisms of ROS Regulation of Plant Development and Stress Responses. Frontiers in Plant Science 10, 800

Huang C., Yu J., Cai Q., Chen Y., Li Y., Ren Y., & Miao Y. (2020). Triple-localized WHIRLY2 influences leaf senescence and silique development via carbon allocation. Plant Physiology, 184, 1348–1362.

Isemer R, Mulisch M, Schäfer A, Kirchner S, Koop HU, Krupinska K (2012) Recombinant Whirly1 translocates from transplastomic chloroplasts to the nucleus. FEBS Lett 586: 85–88

James M., Poret M., Masclaux-Daubresse C., Marmagne A., Coquet L., Jouenne T., Chan P., Trouverie J. & Etienne P. (2018). SAG12, a major cysteine protease involved in nitrogen allocation during senescence for seed production in *Arabidopsis thaliana.* Plant and Cell Physiology, 59(10), 2052-2063.

Kim J., Rudella A., Rodriguez V.R., Zybailov B. Olinares P.D.B. & van Wijk K.J. (2009). Subunits of the plastidClpPR protease complex have differential contributions to embryogenesis, plastid biogenesis, and plant development in *Arabidopsis* . The Plant Cell, 21, 1669-1692.

Kim J.I., Murphy A.S., Baek D., Lee S.-W., Yun D.-J., Bressan R.A. & Narasimhan M.L. (2011) YUCCA6 over-expression demonstrates auxin function in delaying leaf senescence in *Arabidopsis thaliana* . Journal of Experimental Botany, 62(11), 3981-3992.

Kindgren P., Kremnev D., Blanco N.E., de Dios Barajas Lopez J., Fernandez A.P., Tellgren-Roth C., Kleine T., Small I., & Strand A. (2012) The plastid redox insensitive 2 mutant of Arabidopsis is impaired in PEP activity and high light-dependent plastid redox signalling to the nucleus. Plant Journal, 70, 279-291.

Kohorn B.D., Kobayashi M., Johansen S., Riese J., Huang L.-F., Koch K., Fu S., Dotson A. & Byers N. (2006) An Arabidopsis cell wall-associated kinase required for invertase activity and cell growth. The Plant Journal, 46(2), 307-316.

Klermund C., Ranftl Q.L., Diener J., Bastakis E., Richter R. & Schwechheimer, C. (2016). LLM-Domain B-GATA transcription factors promote stomatal development downstream of light signaling pathways in*Arabidopsis thaliana* hypocotyls. The Plant Cell, 28, 646-660.

Krause K., Kilbienskia I., Mulischb M., Rödigerc A., Schäfera A., & Krupinska, K. (2005). DNA-binding proteins of the Whirly family in Arabidopsis thaliana are targeted to the organelles. FEBS Letters, 579, 3707-3712.

Krupinska K, Haussühl K, Schäfer A, van der Kooij TA, Leckband G, Lörz H, Falk J (2002) A novel nucleus-targeted protein is expressed in barley leaves during senescence and pathogen infection. Plant Physiology, 130, 1172–1180.

Krupinska K., Oetke S., Desel C., Mulisch M., Schäfer A., Hollmann J., Kumlehn J., & Hensel G. (2014). WHIRLY1 is a major organizer of chloroplast nucleoids. Fronteirs in Plant Science, 5, 432.

Krupinska K., Braun S., Nia M.S., Schäfer A., Hensel G., & Bilger W. (2019). The nucleoid-associated protein WHIRLY1 is required for the coordinate assembly of plastid and nucleus-encoded proteins during chloroplast development. Planta, 249, 1337–1347.

Lepage E., Zampini E., & Brisson N. (2013). Plastid genome instability leads to reactive oxygen species production and plastid-to-nucleus retrograde signaling in Arabidopsis. Plant Physiology, 163, 867–881.

Li X., Henry R., Yuan J., Cline K. & Hoffman N.E. (1995). A chloroplast homologue of the signal recognition particle subunit SRP54 is involved in the posttranslational integration of a protein into thylakoid membranes. Proceedings of the National Academy of Sciences USA, 92, 3789-3793.

Lichtenthaler H.K. (1987). Chlorophylls and carotenoids - pigments of photosynthetic biomembranes. Methods in Enzymology, 148, 350-382.

Lima A., Lima S., Wong J.H., Phillips R.S., Buchanan B.B. & Luan S. (2006) A redox-active FKBP-type immunophilin functions in accumuylation of the photosystem II supercomplex in *Arabidiopsis thaliana* . Proceedings of the National Academy of Sciences USA, 103, 12631-12636.

Lin W, Huang D, Shi X, Deng B, Ren Y, Lin W, Miao Y (2019) H2O2 as a feedback signal on dual-located WHIRLY1 associates with leaf senescence in Arabidopsis. Cells 8: 1585.

Liu Z., Jia L., Wang H. & He Y. (2011) HYL1 regulates the balance between adaxial and abaxial identity for leaf flattening via miRNA-mediated pathways. Journal of Experimental Botany, 62(12), 4367-4381.

Luo X.-M., Lin W.-H., Zhu S., Zhu J.-Y., Sun Y., Fan X.-Y., Cheng M., Hao Y., Oh E., Tian M., Liu L., Zhang M., Xie Q., Chong K. & Wang Z.-Y. (2010). Integration of light- and brassinosteroid-signaling pathways by a GATA transcription factor in *Arabidopsis* . Developmental Cell, 19, 872-883.

Marechal A., Parent J.-S., Sabar M., Veronneau-Lafortune F., Abou-Rached C., & Brisson N. (2008). Over-expression of mtDNA-associated AtWhy2 compromises mitochondrial function. BMC Plant Biology, 8, 42.

Marechal A., Parent J.S., Veronneau-Lafortune F., Joyeux A., Lang B.F., & Brisson N. (2009). Whirly proteins maintain plastid genome stability in Arabidopsis. Proceedings of the National Academy of Sciences USA, 106, 14693–14698.

Marowa P., Ding A. & Kong Y. (2016). Expansins: roles in plant growth and potential applications in crop improvement. Plant Cell Reports, 35, 949-965.

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

Morley S.A., & Nielsen B.L. (2016). Chloroplast DNA copy number changes during plant development in organelle DNA polymerase mutants. Frontiers in Plant Science, 7, 57.

Motchoulski A. & Liscum E. (1999). *Arabidopsis* NPH3: A NPH1 photoreceptor-Interacting protein essential for phototropism. Science, 286, 961-964.

Muller C.J., Valdes A.E., Wang G., Ramachandran P., Beste L., Uddenberg D. & Carlsbecker A. (2016) PHABULOSA mediates an auxin signaling loop to regulate vascular patterning in Arabidopsis. Plant Physiology, 170, 956-970.

Nagashima Y., Tsugawa S., Mochizuki A., Sasaki T., Fukuda F. & Oda Y. (2018) A Rho-based reaction-difusion system governs cell wall patterning in metaxylem vessels. Scientific Reports , 8, 11542.

Neuteboom L.W., Veth-Tello L.M., Cludesdale O.R., Hooykaas P.J.J. & van der Zaal B.J. (1999). A novel subtilisin-like protease gene from*Arabidopsis thaliana* is expressed at sites of lateral root emergence. DNA Research, 6(1), 13-19.

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

Ohashi-Ito K. & Bergmann D.C. (2006) *Arabidopsis* FAMA controls the final proliferation/

differentiation switch during stomatal development. The Plant Cell, 18, 2493-2505.

Osteryoung K.W., Stokes K.D., Rutherford S.M/, Percival A.L. & Lee W.Y. (1998). Chloroplast division in higher plants requires members of two functionally divergent gene families with homology to bacterial*ftsZ* . The Plant Cell, 10, 1991-2004.

Pogson B.J., Woo N.S., Forster B., & Small I.D. (2008). Plastid signalling to the nucleus and beyond. Trends in Plant Science, 13, 602-609.

Powikrowska M., Oetke S., Jensen P.E. & Krupinska K. (2014) Dynamic composition, shaping and organization of plastid nucleoids. Frontiers in Plant Science, 5, 424.

Prikryl J., Watkins K.P., Friso G., van Wijk K.J., & 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, 5152–5165.

Richter R., Behringer C., Zourelidou M. & Schwechheimer C. (2013). Convergence of auxin and gibberellin signaling on the regulation of the GATA transcription factors GNC and GNL in *Arabidopsis thaliana* . Proceedings of the National Academy of Sciences USA, 110, 13192-13197.

Roland M.m Przybyla-Toscano J., Vignols F., Berger N., Azam T., Christ L., Santoni V., Wu H.-C., Dhalleine T., Johnson M.K., Dubos C., Couturier J. & Rouhier N. (2020). The plastidial *Arabidopsis thaliana* NFU1 protein binds and delivers [4Fe-4S] clusters to specific client proteins. The Journal of Biological Chemistry, 295, 1727-1742.

Rose J.K.C., Braam J., Fry S.C. & Nishitani K. (2002) The XTH family of enzymes involved in xyloglucan endotransglucosylation and endohydrolysis: Current perspectives and a new unifying nomenclature. Plant and Cell Physiology, 43(12), 1421-1435.

Rubio M.C., Calco-Begueria L., Diaz-Mendoza M., Elhiti M., Moore M., Matamoros M.A., James E.K., Diaz I., Perez-Rontome C., Villar I., Sein-Echaluce V.C., Hebelstrup K.H., Dietz K.-J. & Becana M. (2019). Phytoglobins in the nuclei, cytoplasm and chloroplasts modulate nitric oxide signaling and interact with abscisic acid. The Plant Journal, 100, 38-54.

Safrany J., Haasz V., Mate Z., Ciolfi A., Feher B., Oravecz A., Stec A., Dallmann G., Morelli G., Ulm R., & Nagy F. (2008). Identification of a novel cis-regulatory element for UV-B-induced transcription in Arabidopsis. The Plant Journal, 54, 402–414.

Schroder F., Lisso J., Lange P. & Mussig C. (2009) The extracellular EXO protein mediates cell expansion in Arabidopsis leaves. BMC Plant Biology, 9, 20.

Schuetz M., Fidanza M. & Mattsson J. (2019). Identification of auxin response factor-encoding genes expressed in distinct phases of leaf vein development and with overlapping functions in leaf formation. Plants, 8, 242.

Selinski J. & Scheibe R. (2019). Malate valves: old shuttles with new perspectives. Plant Biology, 21 (Suppl 1), 21-30.

Shapiguzov A., Vainonen J.P., Hunter K., Tossavainen H., Tiwari A., Jarvi S., Hellman M., Aarabi F., Alseekh S., Wybouw B., Van Der Kelen K., Nikkanen L., Krasensky-Wrzaczek J., Sipari N., Keinanen M., Tyystjarvi E., Rintamaki E., De Rybel B., Salojarvi J., Van Breusegem F., Fernie A.R., Brosche M., Permi P., Aro E.-M., Wrzaczek M., & Kangasjarvi J. (2019). Arabidopsis RCD1 coordinates chloroplast and mitochondrial functions through interaction with ANAC transcription factors. eLife, 8, e43284.

Shumbe L., D'Alessandro S., Shao N., Chevalier A., Ksas B., Bock R. & Havaux M. (2017). METHYLENE BLUE SENSITIVITY 1 (MBS1) is required for acclimation of Arabidopsis to singlet oxygen and acts downstream of β-cyclocitral. Plant, Cell and Environment, 40, 216-226.

Streatfield S.J., Weber A. Kinsman E.A., Häusler R.E., Li J., Post-Beittenmiller D., Kaiser W.M., Pyke K.A., Flügge U.-I. & Chory J. (1999). The phosphoenolpyruvate/phosphate translocator is required for phenolic metabolism, palisade cell development, and plastid-dependent nuclear gene expression. The Plant Cell, 11, 1609-1621.

Sweetlove L.J. Beard K.F.M., Nunes-Nesi A., Fernie A.R. & Ratcliffe R.G. (2010). Not just a circle: flux modes in the plant TCA cycle. Trends in Plant Science, 15, 462-470.

Suetsugu N., Takemiya A., Kong S.-G., Higa T., Komatsu A., Shimazaki K., Kohchi T. & Wada M. (2016). RPT2/NCH1 subfamily of NPH3-like proteins is essential for the chloroplast accumulation response in land plants. Proceedings of the National Academy of Sciences USA, 113, 10424-10429.

Sun A.Z., & Guo F.Q. (2016). Chloroplast retrograde regulation of heat stress responses in plants. Frontiers in Plant Science, 7, 398.

Suzuki K., Nakanishi H., Bower J., Yoder D.W., Osteryoung K.W. & Miyagishima S. (2009) Plastid chaperonin proteins Cpn60 and Cpn60are required for plastid division in *Arabidopsis thaliana* . BMC Plant Biology, 9, 38.

Swida-Barteczka A, Krieger-Liszkay A, Bilger W, Voigt U, Hensel G, Szweykowska-Kulinska Z, Krupinska K (2018) The plastid-nucleus located DNA/RNA binding protein WHIRLY1 regulates microRNA levels during stress in barley (Hordeum vulgare L.). RNA Biology 15, 886–891

Usadel B., Nagel A., Steinhauser D., Gibon Y., Blasing O.E., Redestig H., Sreenivasulu N., Krall L., Hannah M.A., Poree F., Fernie A.R. & Stitt M. (2006). PageMan: an interactive ontology tool to generate, display, and annotate overview graphs for profiling experiments. BMC Bioinformatics, 7, 535.

Van Aken O., & Pogson B.J. (2017). Convergence of mitochondrial and chloroplastic ANAC017/PAP-dependent retrograde signalling pathways and suppression of programmed cell death. Cell Death and Dif-

ferentiation, 24, 955–960.

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

Wang X.-C., Wu J., Guan M.-L., Zhao C.-H., Geng P. & Zhao Q. (2020).*Arabidopsis* MYB4 plays dual roles in flavonoid biosynthesis. The Plant Journal, 101, 637-652.

Waters M.T., Wang P., Korkaric M., Capper R.G., Saunders N.J. & Langdale J.A. (2009). GLK transcription factors coordinate expression of the photosynthetic apparatus in *Arabidopsis* . The Plant Cell, 21, 1109-1128.

Wu G.Z., Meyer E.H., Richter A.S., Schuster M., Ling Q., Schottler M.A., Walther D., Zoschke R., Grimm B., Jarvis R.P., & Bock R. (2019). Control of retrograde signalling by protein import and cytosolic folding stress. Nature Plants, 5, 525-538.

Xu X.M., Adams S., Chua N.-H. & Moller S.G. (2005). AtNAP1 represents an atypical SufB protein in *Arabidopsis* plastids. The Journal of Biological Chemistry, 280, 6648-6654.

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

Yoshimura K., Mori T., Yokoyama K., Koike Y., Tanabe N., Sato N., Takahashi H., Maruta T. & Shigeoka S. (2011). Identification of alternative splicing events regulated by an Arabidopsis serine/arginine-like protein, atSR45a, in response to high-light stress using a tiling array. Plant and Cell Physiology, 52, 1786-1805.

Yoo, C.Y., Pasoreck E.K., Wang H., Cao J, Blaha G.M., Weigel D. & Chen M. (2019) Phytochrome activates the plastid-encoded RNA polymerase for chloroplast biogenesis via nucleus-to-plastid signaling. Nature Communications, 10, 2629.

Yoo H.H., Kwon C., Lee M.M., & Chang I.K. (2007). Single-stranded DNA binding factor AtWHY1 modulates telomere length homeostasis in *Arabidopsis* . The Plant Journal, 49, 442–451.

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

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

Figure Legends

Figure 1: The relative abundance of WHY1 transcripts in the base, middle and tip sections of the first leaf of 7-day old wild type seedlings estimated by qRT-PCR (A) and Western blot (B). Data in panel A are presented as means +- SE (n = 3). Different letters represent statistical differences assessed by One-way ANOVA followed by Tukey's *post hoc* test.

Figure 2: Distribution of WHY1 in barley leaf determined by immunogold labelling. Leaf sections were prepared for immunogold labelling and the presence of WHY1 detected by ploy clonol antibody as described. Images are representative of plastids (p) in the leaf base (A); chloroplasts (c) in the leaf tip (B); nucleus with nucleolus (nu) in the leaf base (m, mitochondrion) (C) and a detail of the nucleus in the leaf base (D). Example gold particles are indicated by white arrows. Scale bars representing 0.5 (A, B, D) or 1 μm (C) are provided in individual panels.

Figure 3: Cluster analysis comparison of abundance of transcripts that differ significantly in the base, middle and tip regions of wild type 7-day old barley leaves. Relative transcript abundance is represented

according to the legend shown. Transcripts were grouped in clusters (A-E) as indicated and are ordered as listed in table S1.

Figure 4: Cluster analysis comparison of abundance of key transcripts that show a developmental pattern of expression in the base, middle and tip regions of wild type, W1-1 and W1-7 barley leaves. Relative transcript abundance is represented according to the legend shown. Transcripts mentioned in the text are indicated and other transcripts are ordered as indicated in table S3.

Figure 5: Cluster analysis comparison of abundance of transcripts associated with light signalling and plastid development exhibiting significant differences in abundance in base, middle and tip regions of wild type 7 day old barley leaves. Relative transcript abundance is represented according to the legend shown. Gene accession numbers are indicated to the right of each row and further details are provided in table S4.

Figure 6: Cluster analysis comparison of abundance of transcripts associated with plastid biogenesis and development exhibiting significant differences in abundance in the base of wild type, W1-1 and W1-7 7 day old barley leaves. Relative transcript abundance is represented according to the legend shown. Transcripts are ordered as indicated in Table S6.

Figure 7: Polar and non-polar metabolites exhibiting significant differences in abundance in base, middle and tip regions of wild type 7 day old barley leaves. Abundance is shown relative to basal regions for each compound where bars represent mean and lines SE (n=4). , base; , middle; , tip. C14:0, tetradecanoic acid; C17:0, heptacanoic acid; C22 alc, docosanol; C24 alc, tetracosanol.

Figure 8: Relative abundance of transcripts associated with primary metabolism (A) and primary metabolites (B) exhibiting significant differences in abundance in the basal region of 7-day old wild type and WHY1 knockdown barley leaves.

Figure 9: Western blots of selected chloroplast proteins in the base, middle and tip sections of the first leaves of wild type (WT) and W1-1 and W1-7 of seedlings 7 days after germination. Proteins detected are chlorophyll a/b-binding proteins: LHCB1 and LHCB2, the small subunit of RUBISCO (RBCS), chloroplast ribosomal protein S1 (RPS1), , the large subunit of RUBISCO (RBCL), the photosystem II protein (D1) and WHIRLY1 (WHY1).

Figure S1: Representative transmission electron micrographs of nucleus in base of WT barley leaves indicating the association of WHY1 protein (red arrows) with electron dense chromatin (Ch).

Figure S2: A comparisons of leaf chlorophyll content in the base, middle and tip sections of the first leaves of 7-day old wild type (WT), W1-1 and W1-7 seedlings. Data are presented as mean ± SE (n = 6). Different letters represent statistical differences assessed by one-way ANOVA followed by Tukey's *post hoc* test.

Table Legends

Table S1: Transcripts significantly differentially abundant in the base, middle and tip of wild type 7day old barley leaves. Cluster indicates the gene cluster in which the transcripts appear according to fig. 3.

Table S2: Transcripts in the base, middle and tip of 7-day wild type, W1-1 and W1-7 barley leaves that exhibit significant changes in abundance dependent on genotype, leaf position or interaction of the two factors.

Table S3: Transcripts associated with cell wall metabolism that exhibit differential abundance in the base, middle and tip of 7-day old wild type barley leaves.

Table S4: Transcripts associated with light signalling and plastid biogenesis that exhibit differential abundance in the base, middle and tip of 7-day old wild type barley leaves. Transcripts are ordered in accordance with data shown in Fig. 5

Table S5: Transcripts significantly differentially abundant in the basal region of wild type, W1-1 and W1-7 barley leaves.

Table S6: Transcripts associated with plastid biogenesis and development significantly differentially abundant in the basal region of wild type, W1-1 and W1-7 barley leaves. Transcripts are ordered in accordance with Fig. 6.

Table S7: Metabolites significantly differentially abundant in the base, middle and tip of 7-day old wild type barley leaves. Relative abundance is described as mean and standard error. P values represent significance and FDR represents the calculated rate of false positive. Unknown compounds identified in the non-polar (NP) or polar (P) fractions of leaf extracts are labelled according to the MS scan number of their peak apex, fatty acids are labelled according the convention Cx:y where x is the number of carbon atoms and y is the number of unsaturated bonds, fatty alcohols are labelled Cx alc where c is the number of carbon atoms, amino acids are labelled according to standard three letter abbreviations.