WHIRLY1 regulates aliphatic glucosinolate biosynthesis in early seedling development of Arabidopsis

Nguyen Thuy-Linh[1], Moutesidi Pinelopi[2], Glasneck Anike[3], Khosravi Solmaz[4], Abel Steffen[2], Hensel Götz[5], Krupinska Karin[3], Humbeck Klaus[1,*]

[1] Institute of Biology, Martin-Luther-University Halle-Wittenberg, 06120 Halle (Saale), Germany
[2] Department of Molecular Signal Processing, Leibniz Institute of Plant Biochemistry (IPB), 06120 Halle (Saale), Germany
[3] Institute of Botany, Christian-Albrechts-University (CAU), 24098 Kiel, Germany
[4] Department of Breeding Research, Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Gatersleben, 06466 Seeland, Germany
[5] Centre for Plant Genome Engineering, Institute of Plant Biochemistry, Heinrich-Heine-University Duesseldorf, 40255 Duesseldorf, Germany

** Author to whom correspondence should be addressed*

Abstract

WHIRLY1 belongs to a family of plant-specific transcription regulators capable of binding DNA or RNA in all three plant cell compartments containing genetic materials. In *Arabidopsis thaliana*, WHIRLY1 has been studied at the later stages of plant development, including flowering and leaf senescence, as well as in biotic and abiotic stress responses. In this study, novel *WHIRLY1* knock-out mutants of *A. thaliana* were prepared by CRISPR/Cas9 to investigate the role of AtWHIRLY1 during early seedling development. The loss-of-function of *WHIRLY1* in 5-day-old seedlings did not cause differences in the phenotype and the photosynthetic performance of the emerging cotyledons compared to the wild type. Nevertheless, RNA sequencing analysis revealed that the knock-out of *WHIRLY1* affected the expression of a small but specific set of genes during this critical phase of development. About 110 genes were found to be significantly deregulated in the knockout mutant. Among them, genes involved in the early steps of aliphatic glucosinolate biosynthesis happening in chloroplasts were suppressed in the knock-out mutants. The downregulation of these genes in the *WHIRLY1* knock-out line coincided with reduced aliphatic glucosinolate content. Hence, AtWHIRLY1 plays a role in regulating of aliphatic glucosinolate biosynthesis during early seedling development.

Introduction

The WHIRLY protein family comprises few plant-specific proteins sharing the eponymous single-stranded (ss)-DNA-binding domain Whirly (Desveaux et al., 2000, 2002). The members

of this family have the ability to bind preferentially to ssDNA but can also bind to dsDNA and RNA (Cappadocia et al., 2010; Desveaux et al., 2002). Each WHIRLY protein contains an N-terminal organelle target peptide (OTP), which targets them into different compartments. In most angiosperms, two proteins named WHIRLY1 and WHIRLY2 were identified as being targeted either to plastids or mitochondria, respectively (Krupinska et al., 2022). However, in some plant families, such as Brassicaceae, including *A. thaliana*, having one extra AtWHIRLY3 protein, subcellular localization of them is more complicated. Among them, two paralogous genes, *AtWHIRLY1* and *AtWHIRLY3*, share the ancestor origin with monocot *WHIRLY1* (Desveaux et al., 2005). AtWHIRLY3 was shown to be dually targeted to chloroplasts and mitochondria by *in organelle* import assay (Golin et al., 2020). Meanwhile, WHIRLY1 was shown to translocate from plastids to the nucleus in transplastomic tobacco plants synthesizing an HA-tagged AtWHIRLY1 inside chloroplasts (Isemer et al., 2012). Indeed, barley WHIRLY1 was also found in the nucleus (Grabowski et al., 2008). By its dual localization in chloroplasts and nucleus, WHIRLY1 is an excellent candidate for the communication between organelles and the nucleus during plant development and stress responses (Krupinska et al., 2022; Taylor et al., 2022).

The first discovered WHIRLY protein was potato WHIRLY1, which was shown to form a whirligig-like homotetramer structure called PBF-2. The binding of PBF-2 to the elicitor response element (ERE; TGACAnnnnTGTCA) on the promoter of the potato *StPR-10a* (pathogenesis-related gene 10a) induced its expression in response to pathogen attack (Desveaux et al., 2000, 2002). Similarly, WHIRLY1 can bind to the promoter of the *PR1* in Arabidopsis and thereby influence the plants' defense (Desveaux et al., 2004). In Arabidopsis *WHIRLY1* TILLING mutants, binding of the mutated WHIRLY1 protein to the *PR1* promoter was found to be reduced, coinciding with reduced resistance towards *P. parasitica* (Desveaux et al., 2004). The function of WHIRLY1 in response to biotic stresses was also observed in other species, such as maize (Kretschmer et al., 2017), casava (Liu et al., 2018), etc.

WHIRLY1 was reported to bind also to an ERE-like element overlapping with a W-box in the promoter region of *HvS40* in barley (Krupinska et al., 2014) and also to the GNNNAAATT sequence plus an AT-rich telomeric repeat-like sequence in the promoter of *AtWRKY53* in *A. thaliana* (Miao et al., 2013). Both genes function in leaf senescence (HUMBECK et al., 1996; Miao et al., 2004) and are proposed to be negatively regulated by WHIRLY1. In addition to functions in pathogen response and leaf senescence, WHIRLY proteins are involved in abiotic stress responses, e.g., drought, possibly acting via histone marker modulation at the promoter of stress-related genes including *HvNCED1* (Janack et al., 2016a; Manh et al., 2023).

The function of WHIRLY1 in early development was studied in monocots. *WHIRLY1* mutants of maize and rice were severely disturbed in chloroplast development and died at the three-to-four-leaf stage (Prikryl et al., 2008; Qiu et al., 2022). In contrast, neither T-DNA insertion mutants of Arabidopsis *WHIRLY1* nor *WHIRLY2* show apparent phenotypes. It has

been, therefore, discussed that AtWHIRLY3 might replace AtWHIRLY1 and AtWHIRLY2 (Krupinska et al., 2022). Analysis of the *why2* mutant showed that germination and early seedling development were compromised (Golin et al., 2020; Yuri L. Negroni et al., n.d.), whereas vegetative growth was unaffected. This result indicates that AtWHIRLY3 cannot replace all functions of AtWHIRLY2 during germination being in line with its low expression level at this stage of development (Golin et al., 2020).

So far, research on the biological functions of Arabidopsis WHIRLY1 mainly concerned senescence and flowering. Still, little is known about the involvement of WHIRLY1 in early seedling development, which plays a vital role in the growth cycle. This study, therefore, aims to investigate whether AtWHIRLY1, similarly to AtWHIRLY2, plays a role in germination and early seedling development. While most previous studies on the function of AtWHIRLY1 used T-DNA insertion mutant *why1-1* (SALK_023713), in which the T-DNA is integrated close to the start codon in the first exon (Huang et al., 2022; Isemer et al., 2012; Lin et al., 2019), in the present work, new *AtWHIRLY1* knock-out mutants prepared by site-directed mutagenesis employing CRISPR/Cas9 were used to study functions of WHIRLY1 at the early stage of development. In contrast to the *why2* mutant, the loss-of-function of *AtWHIRLY1* did not cause apparent changes in the phenotype of seedlings and photosynthetic performance. However, genome-wide gene expression analysis revealed a significant impact of *WHIRLY1* knockout on the expression of a small set of nuclear genes. These genes specifically encode enzymes involved in the chloroplast-located early steps of aliphatic glucosinolate (GSL) biosynthesis, leading to a lower aliphatic GSL abundance in the *AtWHIRLY1* knock-out mutant.

Results

CRISPR/Cas9-mediated knock-out mutants of AtWHIRLY1

So far, most studies about the AtWHIRLY1 function used the mutant *why1-1* with a T-DNA insertion in the first exon (SALK_023713), which had been first denominated KO-1 (Yoo et al., 2007). However, in this mutant and two other T-DNA insertion mutants, i.e., *why1-2* (SALK_147680; KO-2 in (Yoo et al., 2007) and *why1-3* (SALK_03900), mRNA level of *AtWHIRLY1* was relatively high (Supplemental Figure S1). Noticeably, the *why1-1* mutant still has a second in-frame ATG located downstream of the first ATG and it might serve as an alternative start codon for the translation of a shorter protein lacking the first 10 amino acids constituting the plastid transit peptide. Hence, this truncated form of WHIRLY1 might be synthesized in the *why1-1* mutant but mistargeted.

To preclude this possibility, novel *AtWHIRLY1* knock-out mutants were prepared by CRISPR/Cas9-mediated genome editing. Three independent Cas9-free homozygous mutant lines were generated and denotated as *why1-4*, *why1-5*, and *why1-6*. *why1-4 and why1-5* mutants contain point mutations, 25_27DelAA and 27_28InsT, respectively, shifting the reading

frame and altering the amino acid sequences of the residual proteins having 50 amino acids instead of 263 (Figure 1A). Meanwhile, the mutation in the line *why1*-6 resulted in a novel in-frame ATG codon after the 27_28InsA mutation which might produce a truncated WHIRLY1 protein. To avoid the effects of such an isoform, this mutant was not investigated in this study. The two mutant lines, *why1-4* and *why1-5*, were verified using PCR-coupled cleavage amplified polymorphism site analysis (Figure 1B), showing that they are homologous to the mutated alleles. Besides, the absence of a Cas9-coding sequence was checked by PCR with specific primers (Supplemental Table S3), showing that both lines are Cas9-free. The knock-out of *WHIRLY1* by CRISPR/Cas9 resulted in a substantial reduction of *WHIRLY1* transcript level compared to the wildtype (WT) (Figure 1C), whereas expression levels of the two others *WHIRLY* genes were similar as in the WT.

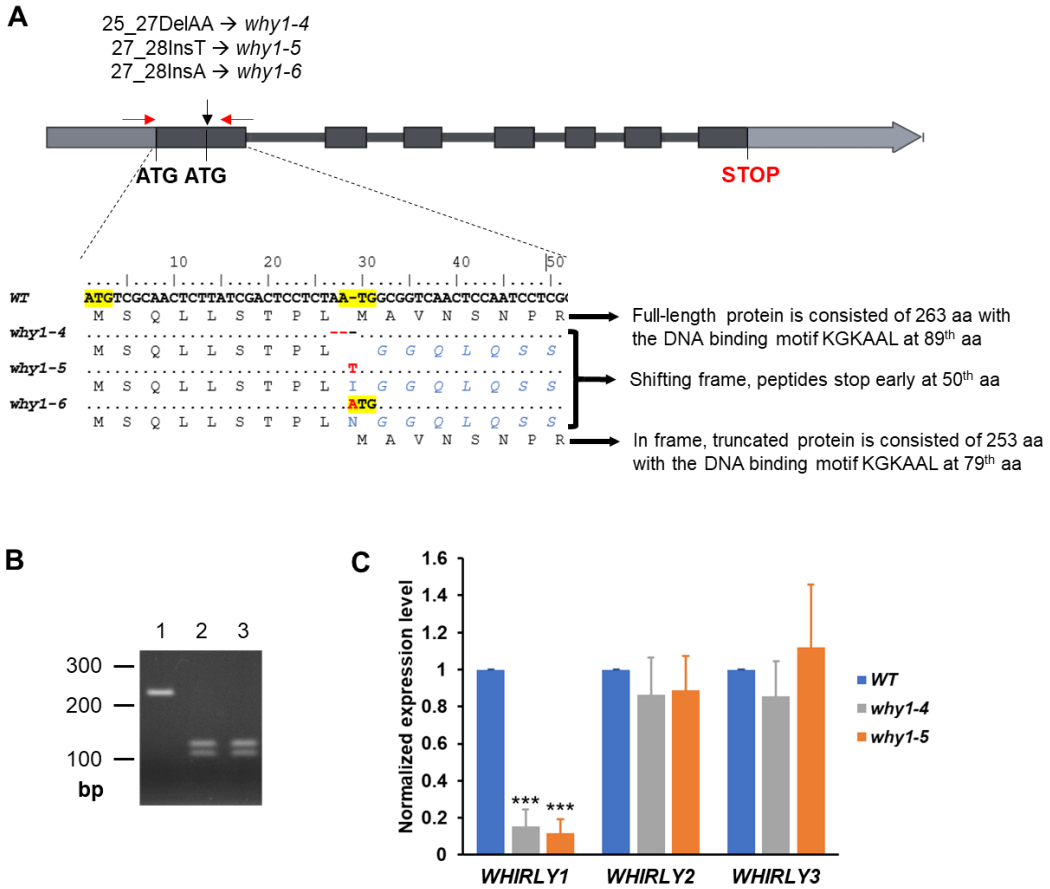

Figure 1. Genotyping of CRISPR/Cas9 mutants of *AtWHIRLY1*. (**A**) Scheme of three *WHIRLY1* knock-out mutants in the present work. The start codon (ATG) is highlighted and point mutations are depicted in red letter. Both *why1-4* and *why1-5* lines contain a frame-shift mutation, which changes the protein sequence starting at 10th aa. The mutation in w*hy1-6* might also cause frame shifting together with a novel ATG in frame within the WT allele. Size of expected mutant peptides encoded by these alleles is shown at the end of the alignment. (**B**) Two mutants, *why1-4* and *why1-5,* were verified using PCR-coupled CAPS. The WT allele gave a non-cut fragment of 231 bp while CRISPR/Cas9-mediated mutated alleles were digested into two products of 126 bp and 105 bp. Hence, a homologous mutant produces only two former bands. Lane 1: WT, lane 2: *why1-4*, lane 3: *why1-5*. (**C**) Normalized expression

level of three *WHIRLY* genes in three knock-out mutants by qRT-PCR compared to the WT. Bar chart represents average of mean and error bars indicate standard error of four independent biological replicates. Asterisks denote a statistically significant level of Student's t-test between qRT-PCR-based normalized expression level of genes in *WHIRLY1* mutants and the WT, (*) p-value < 0.05, (**) p-value < 0.01, (***) p-value < 0.001.

Loss-of-function of *WHIRLY1* did not affect morphology and photosystem II efficiency

To investigate the impact of WHIRLY1 on the development of seedlings, seeds of two CRISPR/Cas9 knock-out lines (*why1-4* and *why1-5*) and of the WT were grown for five days on medium. Seedlings developed roots and green cotyledons (Figure 2A). Morphologically, there was no apparent difference between all lines regarding root length, hypocotyl length, and fresh weight (Figure 2B). Photosystem II efficiency indicated by chlorophyll fluorescence (F_V/F_M) and chlorophyll content of cotyledons were similar between the wild type and mutants (Figure 2B). The results indicate that during early development under normal growth conditions, the loss-of-function of *WHIRLY1* has no evident influence on the vegetative growth and development of photosynthetic capacity.

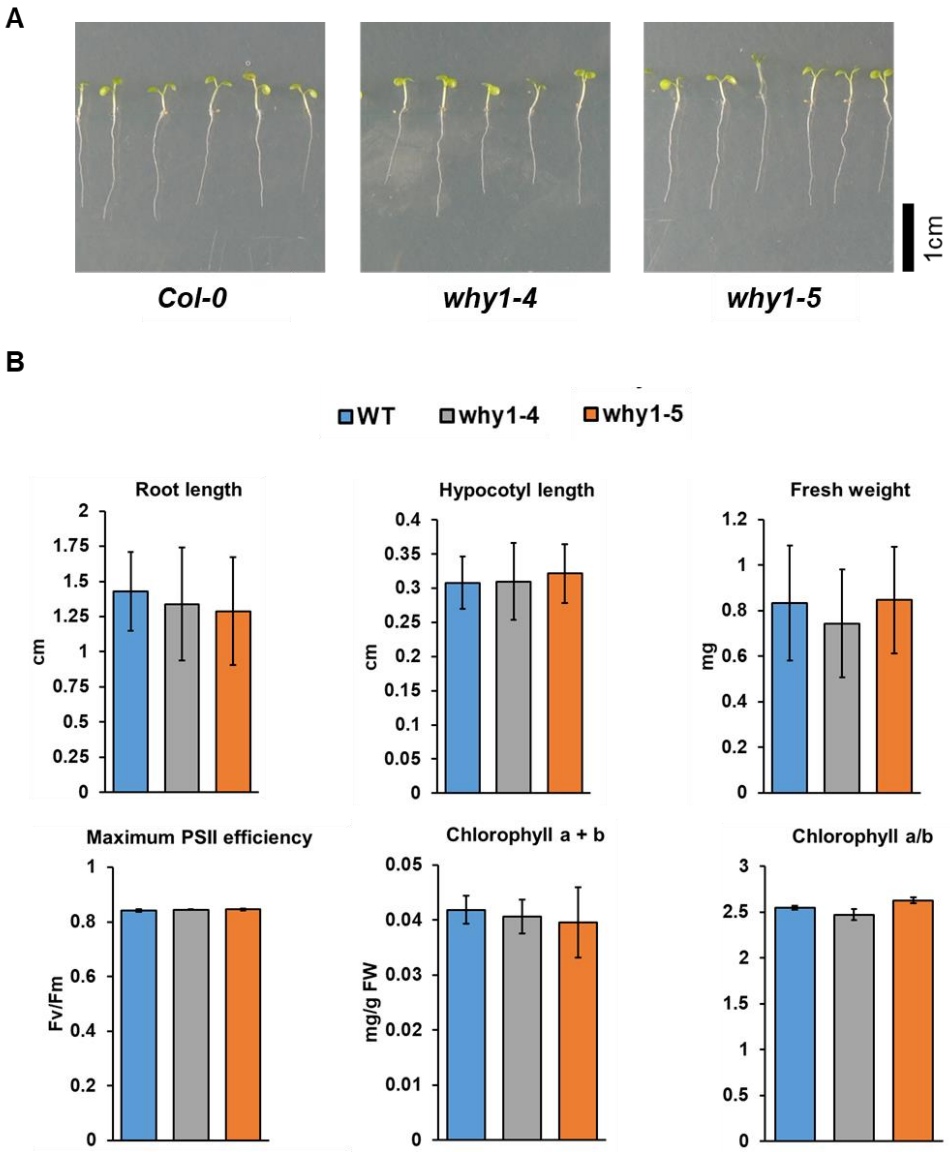

Figure 2. Phenotype of *WHIRLY1* knock-out mutant seelings. (**A**) Image of 5-day-old WT and two CRISPR/Cas9 knock-out mutants of *AtWHIRLY1*, *why1-4* and *why1-5*, grown vertically on half-strength MS plate under continuous light. (**B**) Root length, hypocotyl length, whole seedling fresh weight, maximum PSII efficiency (Fv/Fm), total chlorophyll a+b content, and chlorophyll a/b ratio of 5-day-old WT and *why1* seedlings. The bar charts show the average value and error bars show standard deviation of 4 biological replicates, each containing at least 20 seedlings. Asterisks indicate statistically significant level based on the Student's t-test, *** p-value < 0.001.

Loss-of-function of *WHIRLY1* specifically influenced the transcriptome of seedlings

To test the impact of loss-of-function of *AtWHIRLY1* on nuclear gene expression in seedlings at the 5-day stage, RNA sequencing (RNA-seq) was performed with three independent biological replicates of the *WHIRLY1* knock-out mutant *why1-5*, each consisting of a pool of around 30 seedlings grown in the same agar plate. RNA was isolated from whole

seedlings of the mutant and the WT. A summary of read numbers and read mapping rate is presented in Supplemental Table S1.

In each line, nearly 15 thousand genes were detected to be expressed (with an expression level above the threshold, i.e., fpkm value > 1). The DESeq2 program was used to identify the differences in gene expression between the two genotypes. Considering that mutant seedlings did not show appearent changes in morphology and photosystem II efficiency compard to the wild type, differentially expressed genes (DEGs) were calculated using a mild stringent cut-off of $|\log_2 \text{fold-change(FC)}| > 1$ and p-value < 0.01. As a result, 73 upregulated DEGs and 42 downregulated DEGs were identified in the knock-out mutant *why1-5* compared to the WT, indicating that WHIRLY1 regulates target genes both positively and negatively. The low number of DEGs (Figure 3A) suggests that, the loss-of-function of *WHIRLY1* during this early growth stage has no wide influence on the transcriptome, but has a rather specific impact on a small set of genes. Gene Ontology (GO) enrichment analysis failed to identify significant GO terms for upregulated genes in the line *why1-5*. On the other hand, a small set of 42 significantly downregulated DEGs in the knock-out line was enriched in genes related to the biosynthesis of sulfur-containing compounds. In particular, the glucosinolate (GSL) biosynthesis function group had a very high enrichment fold (Figure 3B). Besides GO enrichment analysis, a manual search for all upregulated and downregulated DEGs in databases (The Arabidopsis Information Resource (TAIR), www.arabidopsis.org and UniProt, https://www.uniprot.org/) revealed that in the *WHIRLY1* knockout, multiple DEGs are related to various other processes, including development and stress responses (Supplemental Table S2).

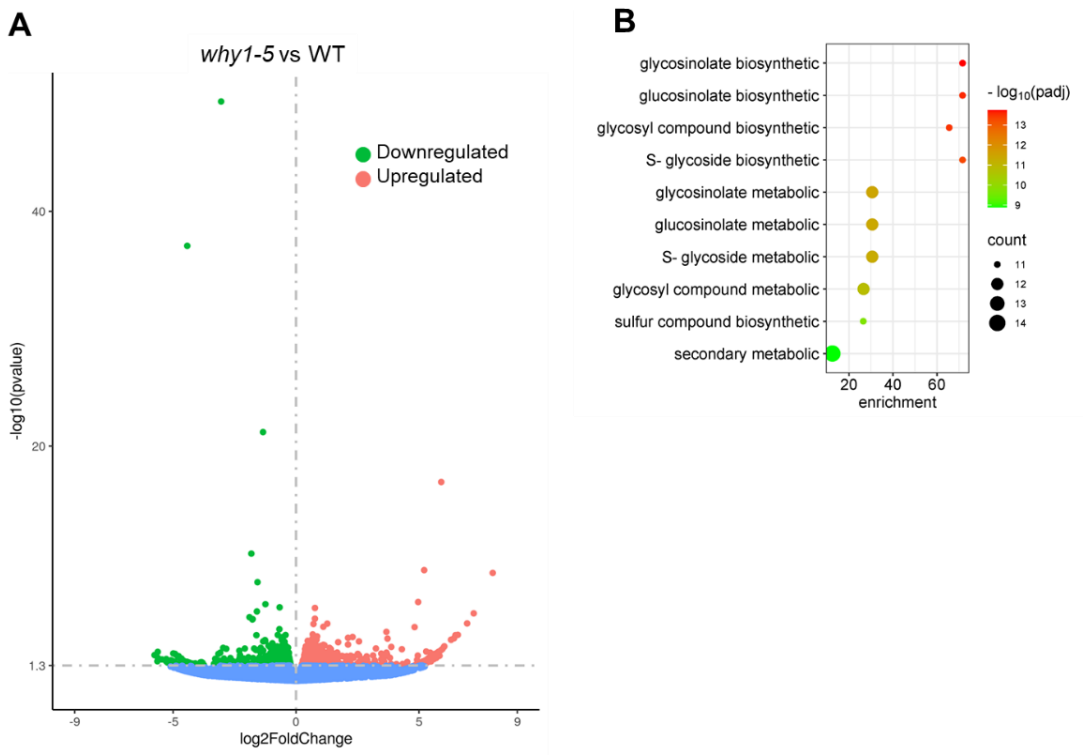

Figure 3. Transcriptome analysis of seedlings of the *why1-5* mutant and the WT. (**A**) The expression changes of DEGs in the mutant *why1-5* compared to the WT are shown by a Volcano plot. (**B**) GO enrichment analysis of down-regulated DEGs. The top 10 GO biological process terms according to p-adjusted value are shown.

Downregulation of genes encoding enzymes of the aliphatic glucosinolate pathway

A striking result of the RNA-seq analysis was that 10 downregulated DEGs are involved in glucosinolate (GSL) biosynthesis (Figure 3B). GSLs are secondary metabolites which are composed of a thioglucose, a sulfate group, and a variable side chain derived from sulfur-containing amino acids (Mérillon & Ramawat, 2017). Based on precursor amino acids, GSLs are categorized into two main groups: aliphatic (aGSL) or indole glucosinolates (iGSLs). As illustrated in Figure 4A, aliphatic GSL biosynthesis involves three main stages: side-chain elongation occurring in the plastid, core-structure formation and side-chain modification which both taking place in the cytosol.

According to the RNA-seq data, loss-of-function of *WHIRLY1* specifically affected the expression level of several genes encoding enzymes in the early steps of Methionine (Met)-derived aGSL biosynthesis, such as *BRANCHED-CHAIN AMINOTRANSFERASE 4* (*BCAT4*), *METHYLTHIOALKYLMALATE SYNTHASE 1* (*MAM1*), *ISOPROPYLMALATE ISOMERASE 1* (*IPMI1*), and *BILE ACID:SODIUM SYMPORTER 5* (*BASS5*), wherein most of them are plastid-located (Figure 4A). The downregulation of the aforementioned genes and additional genes encoding enzymes of GSL biosynthetic pathway in the knock-out line *why1-5* was validated by

qRT-PCR, showing a high correlation between RNA-seq and qRT-PCR data (Supplemental Figure S2). The expression level of numerous genes related to biosynthesis of aGSL in the *why1-5* seedling was substantially decreased to about a third of the WT level (Supplemental Figure S2). In contrast, the expression of most genes involved in iGSL biosynthesis was not affected in the *WHIRLY1* knock-out mutant (Supplemental Figure S2). Still, *SUPERROOT1* (*SUR1*), a gene shared between aGSL and iGSL pathways, had a slightly reduced expression level in the *why1*-5 mutant. Interestingly, in contrast to the knock-out mutant *why1*-5 prepared by genome editing, in the T-DNA insertion mutant *why1-1*, expression of GSL-related genes was unaltered (Figure 4B).

Biosynthesis of aGSL is known to be directly controlled by MYELOBLASTOSIS transcription factors (MYBs), including MYB28, MYB29, and MYB76, while iGSL biosynthesis is governed by MYB34, MYB51, and MYB122 (Mérillon & Ramawat, 2017). Therefore, the expression levels of the corresponding genes were also analyzed in seedlings of the knock-out mutant *why1-5*. The qRT-PCR revealed that the expression level of *MYB* genes was similar in the *WHIRLY1* knock-out mutant as in the WT (Figure 4B; Supplemental Figure S2). The findings suggest that WHIRLY1 does not act directly on the transcript abundance of the known upstream regulators of aGSL biosynthesis, i.e., *MYB28* and *MYB29*, but instead regulates expression of aGSL biosynthesis genes via another mechanism. The downregulation of aGSL biosynthetic genes was also confirmed in seedlings of the other knock-out line *why1-4* (Figure 4B).

The expression changes of GSL-related genes in the mutant *why1-5* were compared with the available RNA-array data of mature *MYB28* and *MYB29* knock-out plants (Sønderby et al., 2010). Interestingly, a similar expressional change of GSL biosynthesis genes was observed in the *myb28* and the *why1-5* lines (Figure 4A). In contrast to these two mutants, the *myb29* showed a different pattern in the expression of GSL-related genes. Indeed, the same genes involved in side-chain elongation steps were repressed due to the loss-of-function of either *WHIRLY1* or *MYB28*, but not *MYB29*. For instance, the expression level of *BCAT4*, which encodes the main enzyme initiating the side-chain elongation process, was significantly reduced in the knock-out lines *why1-5* and *myb28*. In contrast, the *BCAT4* transcript level did not differ in the *myb29* mutant. The same trend was observed with other aGSL-related genes, such as *MAM1, IPMI1,* and *BASS5*. On the contrary, genes encoding enzymes in the iGSL pathway were slightly but not significantly induced in the *why1-5* and *myb28* mutants, while some genes, such as *MYB34* and *CYP81F2*, were significantly deregulated in the *myb29* line. Interestingly, the effect of the loss-of-function of WHIRLY1 on aGSL-related genes was only observed in the seedling stage but not in the rosette of mature plants (Supplemental Figure S3).

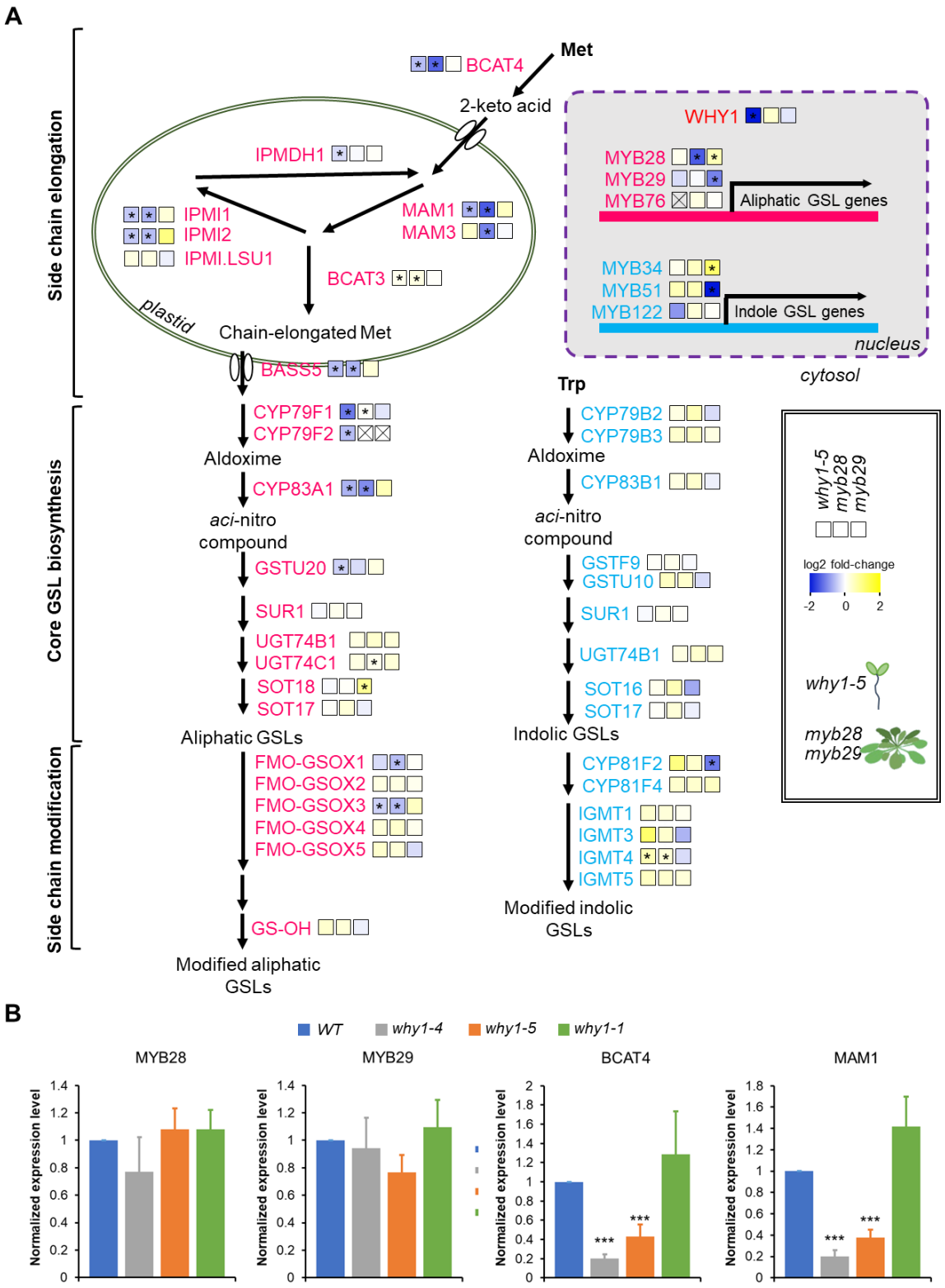

Figure 4. Loss-of-function of *WHIRLY1* affects GSL-related gene expression. (**A**) Heatmap showing changes in transcript levels of GSL-related genes due to the knock-out of *WHIRLY1* and due to knock-out of *MYB28* and *MYB29* (Sønderby et al., 2010). The $\log_2$FC is indicated by color bar, and asterisks (*) show statistical significant difference as result of RNA-seq/microarray. Scheme of two major GSL

biosynthesis pathways, aliphatic GSL and indole GSL was adapted from Augustine & Bisht, 2017. See here also for gene names. (B) Normalized expression level of selected GSL-related genes by qRT-PCR. Bar charts show the average and error bars show standard deviation of four independent biological replicates, each containing around 30 seedlings. Asterisks indicate statistically significant level of Student's t-test between *whirly1* knock-out mutants and WT; * p-value < 0.05, ** p-value < 0.01, *** pvalue < 0.001

Reduction of GSL content in the *WHIRLY1* knock-out mutant

To investigate whether the specific suppression of genes encoding enzymes in the early steps of aGSL biosynthesis has an influence on the aGSL content, whole seedlings were collected to quantify GSLs by LC-MS/MS. The amount of each GSL in each sample was calibrated using sinigrin, a compound not naturally found in *A. thaliana*. After that, GSL content was normalized to the samples fresh weights. GSL profile of the *WHIRLY1* knock-out mutant *why1-5* was in line with the transcriptomics data (Figure 5; Supplemental Table S3). In the *why1-5* mutant, the total GSL level in 5-day-old seedlings was decreased significantly compared to the WT. The lower GSL concentration in the mutant was due to a significant reduction of aGSL content but not of iGSL (Figure 5A).

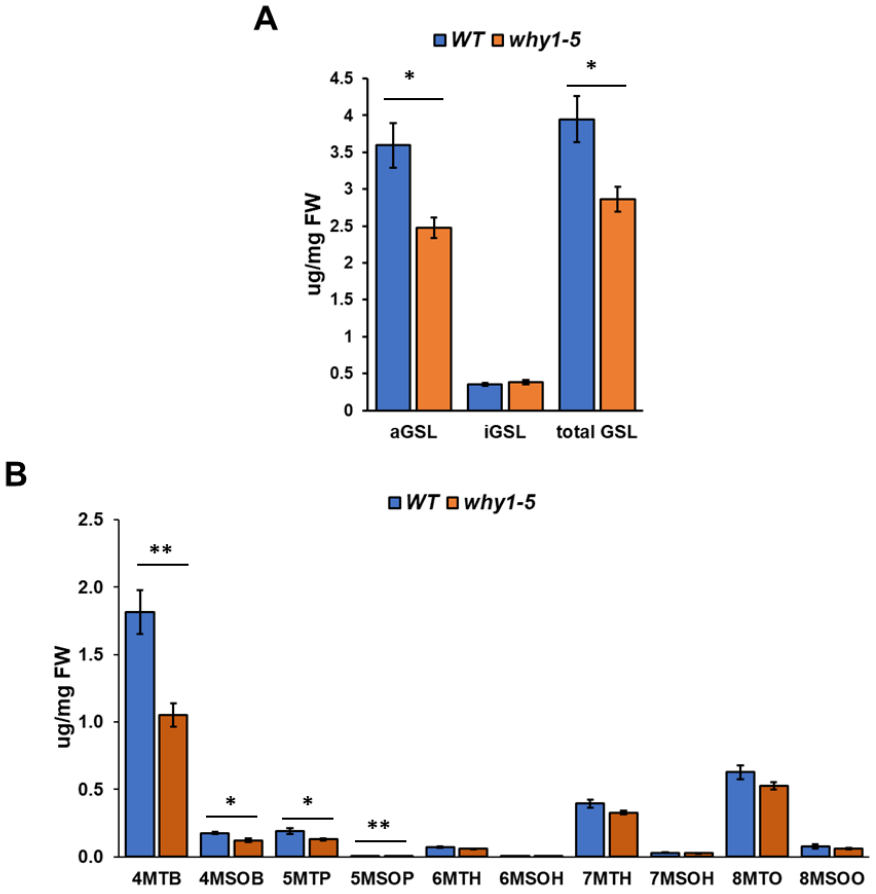

Figure 5. Glucosinolate contents in the *WHIRLY1* knock-out and the WT seedlings. (A) The concentration of aGSL, iGSL, and total GSL, and (B) individual aGSL compounds in the seedling *why1-*

5 and WT measured by LC-MS/MS. Bar charts show the average and error bars show standard deviation of three independent biological replicates, each containing four samples as pool of 3-4 seedlings. Asterisks indicate statistically significant level of Student's t-test between *why1-5* knock-out mutants and WT; * p-value < 0.05, ** p-value < 0.01, *** p-value < 0.001

To elucidate the impact of the knock-out of *WHIRLY1* on GSL biosynthesis, the concentration of seperated GSLs were identified. In *why1-5* mutant seedlings, the amount of each short-chain aGSL compound was significantly reduced (Figure 5B). Indeed, the WT contained almost twice as much 4MTB as the mutant (1.81 ug/mg FW *versus* 1.05 ug/mg FW). Moreover, 4MSOB and 5MTP concentrations in the *why1-5* mutant were reduced to about 60% of the WT level (Figure 5B). For comparison, long-chain aGSLs were only reduced mildly by less than 15% due to the knock-out of *WHIRLY1* (Figure 5B). In contrast, the iGSL content was slightly increased in the knock-out mutant *why1-5* compared to the WT (Supplemental Table S3).

Discussion

Knock-out of *WHIRLY1* has no impact on seedling development

WHIRLY1 has been described as a multifaceted transcription regulator that modulates various plant development and stress responses (Krupinska et al., 2022; Taylor et al., 2022). Previous studies on the function of Arabidopsis WHIRLY1 were primarily performed with T-DNA insertion mutants (Supplemental Figure S1). The T-DNA integrated at the very beginning of the coding sequence in both *why1-1* and *why1-2* mutants. However, a second in-frame start codon ATG downstream of the first one can potentially serve as an alternative start codon for translating of a truncated peptide of WHIRLY1. This shortened version lacks only 10 amino acids at the N-terminal of the plastid transit peptide, but its Whirly domain, including the DNA-binding motif KGKAAL, remains intact. This assumption aligns with the high level of WHIRLY1 expression level in these mutants (Supplemental Figure S1). Hence, despite being mentioned as knock-out mutants (Huang et al., 2022; Miao et al., 2013; Yoo et al., 2007), they likely contain the complete mature AtWHIRLY1. Nonetheless, the T-DNA insertion mutant *why1-1* has been reported to have an altered telomere length (Yoo et al., 2007) and to show earlier senescence (Miao et al., 2013) and flowering (Huang et al., 2022). These changes might be due to mistargeting of the N-terminally truncated WHIRLY1 protein starting from the second in-frame ATG codon.

To investigate the consequences of a complete knock-out of *AtWHIRLY1* on early seedling development, in this study, novel *WHIRLY1* mutants were generated using site-directed mutagenesis with CRISPR/Cas9. These knock-out lines have point mutations which cause a frame shift, changing the protein sequence starting from the 11th amino acid. Therefore, the mutated peptides are relatively short (50-amino-acid long) and have no functionalities (Figure 1A).

Despite the true knock-out of *AtWHIRLY1*, these novel mutants, even during seedling development, did not show an obvious phenotype, suggesting that the presence of WHIRLY1 seems unnecessary for average seedling growth. It has already been described in dicots that *WHIRLY1* single mutants and *WHIRLIES* double/triple mutants showed no appearing phenotype under control conditions (Golin et al., 2020; Maréchal et al., 2009; Zhuang et al., 2019). The lack of an apparent phenotype can be explained by an unknown mechanism compensating WHIRLY1-knock-out effects on chloroplast performance and development of seedlings. This complementation in Arabidopsis *WHIRLY1* knock-out mutants could be done preferably by AtWHIRLY3, which was shown to be dually targeted into both chloroplast and mitochondria (Golin et al., 2020). Recently, it has been observed that the expression level of *AtWHIRLY3* was enhanced at shoots of *whirly2* knock-out mutant at the 2- to 4-true-leaf stage, indicating a compensation of WHIRLY2 deficiency by enhanced abundance of WHIRLY3 (Golin et al., 2020). Therefore, it is likely that AtWHIRLY3 might take over AtWHIRLY1 function in seedling development in *WHIRLY1* knock-out mutants. Indeed, AtWHIRLY1 and AtWHIRLY3 share 77% similarity in protein sequence and might have similar functions under certain conditions (Desveaux et al., 2004). For comparison, monocot plants with WHIRLY1 deficiency show obvious phenotypes likely due to lack of a third WHIRLY protein. The RNAi-mediated knockdown *HvWHIRLY1* barley showed delayed formation of green tissue as well as suppressed development of photosynthetic function (Krupinska et al., 2019; Saeid Nia et al., 2023), while lacking ZmWHIRLY1 in maize or OsWHIRLY1 in rice caused abnormal embryos and albino plants (Prikryl et al., 2008; Qiu et al., 2022).

Despite the novel *WHIRLY1* knock-out mutants lacking an apparent phenotype, RNA-seq analysis of the *why1-5* mutant seedlings revealed distinct alterations in nuclear gene expression, indicating that the regulation of specific genes cannot be mediated by other WHIRLY proteins. The loss-of-function of *WHIRLY1* affected expression of a small fraction of nuclear genes, which involved in various developmental processes and stress responses (Supplemental Table S2). Among differently expressed genes in the *why1-5* mutant compared to the wildtype, genes encoding enzymes of aliphatic GSL biosynthesis were significantly suppressed. However, knockout of *WHIRLY1* did not impact all aGSL-related genes, but mostly those encoding enzymes catalyzing the first steps of aGSL biosynthesis, such as *BCAT4,* the chloroplast-located enzymes *MAM1* and *IPMI1,* and the plastid-envelope-located transporter *BASS5* (Figure 4). This indicates that WHIRLY1 is a positive regulator of these genes. Interestingly, none of the previously described target genes of WHIRLY1, including *AtPR1* (Desveaux et al., 2004) and *AtWRKY53* (Miao et al., 2013) were differently expressed in seedlings of the *why1*-5 mutant.

Participation of WHIRLY1 in regulation of glucosinolate biosynthesis during seedling development

GSLs biosynthesis is regulated precisely at many levels and involves diverse regulators and pathways (Mitreiter & Gigolashvili, 2021). MYB28 is a central transcriptional regulator of aGSL biosynthesis, followed by MYB29 and MYB76 (Mérillon & Ramawat, 2017). Interestingly, similar genes were down-regulated due to the knock-out of *WHIRLY1* as well as in the *MYB28* knock-out line (Figure 4A). The influence of knock-out of *WHIRLY1* on GSL metabolism was only observed at the seedling stage, as there was no difference in GSL gene expression level in mature *why1-5* knock-out plant compared to the WT (Supplemental Figure S3). This indicates that during early development, when chloroplasts are developing, WHIRLY1 acts as an additional regulator of aGSL-related genes, while in mature plants, aGSL-related gene expression is regulated without WHIRLY1 and predominantly by MYB28.

It is well-known that chloroplast development requires intensive communication between developing plastids and the nucleus (Pfannschmidt & Munné-Bosch, 2013). Due to its dual localization in chloroplast and the nucleus, WHIRLY1 has been suggested as a communicator between these two organelles (Krupinska et al., 2020). It is possible that AtWHIRLY1 is involved in the biogenic signaling from the developing plastids to the nucleus. Due to their sulfur-rich chemical composition, GSLs are proposed to be, in addition to their defensive function, a sulfur storage (Mérillon & Ramawat, 2017; Mitreiter & Gigolashvili, 2021). Especially after germination, their turnover is important for supplying building blocks for other molecules, including many iron-sulfur clusters containing proteins in chloroplasts, some of which are components of the photosynthetic apparatus (Balk & Schaedler, 2014).

Knock-out of *WHIRLY1* did not affect the expression level of genes encoding MYBs indicating that WHIRLY1 is not an upstream regulator of the transcription of these genes (Figure 4B), but might regulate aGSL biosynthesis related genes independently from MYB28. One possible mechanism is that WHIRLY1 regulates these genes on an epigenetic level. There are few reports about the function of WHIRLY1/2/3 in epigenetic regulation. Barley plants deficient in WHIRLY1 due to RNAi-mediated knock-down of *HvWHIRLY1* showed changes in histone modification markers in some senescence-related genes (Janack et al., 2016b). All three WHIRLIES in *A. thaliana* were found to be connected to epigenetic reprogramming of meristem tissue, both in root and shoot (McCoy et al., 2021). Besides recently, WHIRLY1 was shown to interact with HDA15 to modify histone markers of some flowering-related genes (Huang et al., 2022).

It has been reported that WHIRLY1 might bind to melted promoter regions of nuclear target genes, whereby it has an affinity to certain binding motifs, such as the elicitor response element (ERE) including inverted repeat sequences (IR) and W-boxes (Desveaux et al., 2002, 2005).

Therefore, 1000-bp promoter regions of aGSL genes were analyzed by PLACE (https://www.dna.affrc.go.jp/PLACE/). The analysis revealed different motifs in the promoters of GSL-related genes. As expected, these promoters contain several MYB binding site motifs for transcriptional regulation by MYB28 and MYB29, and in additionally, many W-boxes and IR2 sequences that could serve as potential binding sites of WHIRLY1 are distributed throughout promoter sequences (Supplemental Figure S3). These findings are summarized in a working model of regulation of aGSL biosynthesis by AtWHIRLY1 (Figure 6).

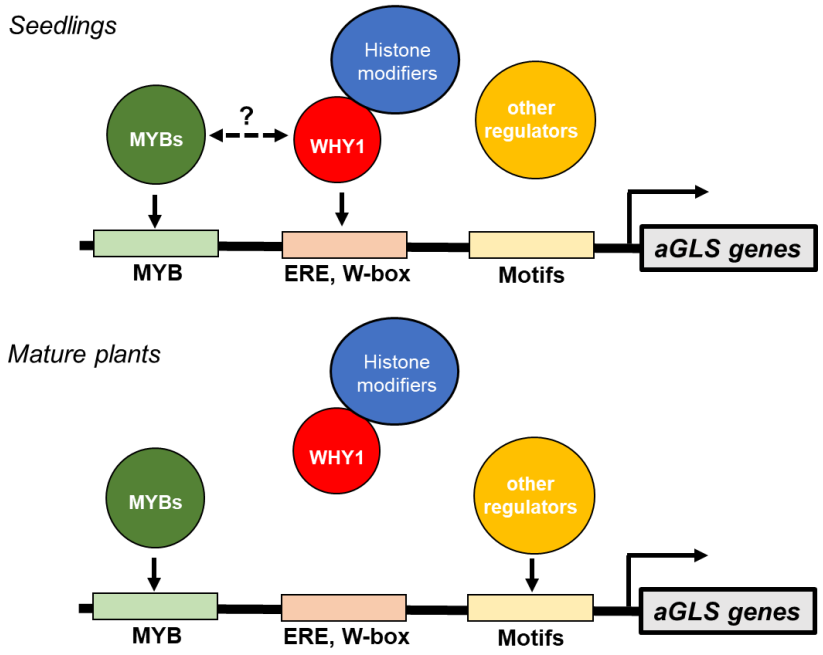

Figure 6. A proposed working model of AtWHIRLY1's function as an aliphatic glucosinolate (aGSL) regulator. In seedlings, WHIRLY1 binds to its binding motifs, for instances, W-boxes and ERE sequences, at the promoters of aGSL biosynthesis related genes to regulate their expressions. Whereby, it might be involved in epigenetic modification by interacting with histone modifiers. MYB transcription factors, such as MYB28, also interact with MYB motif on aGSL genes' promoters to drive their expression. The binding of both WHIRLY1 and MYBs are necessary to achieve an optimal level of aGSL gene expression in seedlings. However, in mature plant, WHIRLY1 no longer plays an important role in regulation of these gene. Perhaps, other developmental-related factors together with MYBs take over and modulate aGSL related genes.

Regulation of glucosinolate by WHIRLY1 - an evolutionary point of view

In the present work, the knock-out of *WHIRLY1* strictly affected genes in the first steps of Met-derived aGSL biosynthesis but did not influence the production of other types of GSL. From an evolutionary point of view, Met-derived aGSLs are novel GSLs that appeared rather recently after two whole genome duplication events, named At-α and At-β (discussed in (Edger et al.,

2018) (Supplemental Figure S5). Gene duplication at At-β, which occurred after the divergence of the Caricaceae family, was the origin of novel iGSL and Met-derived aGSL pathways (Barco & Clay, 2019). Whereas, the recent event At-α, which happened at the emergence of the Brassicaceae family, gave rise to many side-chain modification genes for both iGSLs and Met-derived aGSLs. Therefore, Met-derived aGSLs only are present in core families of Brassicales, i.e., those derived after the event At-β, such as Brassicaceae, Cleomaceae, etc. (Mérillon & Ramawat, 2017).

Papaya (*Carica papaya*), the only species of Caricaceae family with genomic data available in NCBI (till 08 August 2023), has only two WHIRLY proteins, in which the accession number XP_021897420.1 presents a WHIRLY1 protein and the other two sequences, XP_021912604.1 and XP_021912605.1, are two isoforms of WHIRLY2. Whereas, in the spider flower plant (*Tarenaya hassleriana*) belonging to the family Cleomaceae, which appeared after At-β, three WHIRLIES were reported in NCBI (accession numbers XP_010534870.1, XP_010551512.1, and XP_010527874.1). Hence, during evolution, the appearance of Met-derived aGSL biosynthesis coincided with the appearance of the third *WHIRLY* protein due to duplication of the *WHIRLY1* gene (Supplemental Figure S5).

Gene duplication is the primary drive for evolution as it provides material for natural selection. Duplicated genes could either maintain the ancestral function, become silent (pseudogene), or develop a new function. Therefore, the co-occurrence of the novel Met-derived aGSL biosynthesis and an additional *WHIRLY* locus in the genome suggests that the new WHIRLY protein might have obtained a novel function during evolution. In Brassicaceae, another chloroplast-targeted WHIRLY protein might have enabled retrograde control of GSL biosynthesis which is beneficial for chloroplast development. Synteny analysis showed that *AtWHIRLY1* and *AtWHIRLY3* share the same ancestral gene with monocot *WHIRLY1*, whereby AtWHIRLY3 seems to be more closely related to the monocot WHIRLY1 proteins than AtWHIRLY1 (Krupinska et al., 2022). Likely, after the duplication in the event At-β, AtWHIRLY3 evolved following its original ancestral gene function, while AtWHIRLY1 gained a novel function in the regulation of secondary metabolism. Its evolutionary co-appearance with aGSL biosynthesis-related genes and its dual-localization in chloroplasts and nucleus makes it an ideal candidate regulator of chloroplast-located steps in glucosinolate biosynthesis during seedling development.

Materials and methods

Plant materials and growth

Seeds of *A. thaliana* Columbia-0 and mutants were surface-sterilized in 70% (v/v) ethanol for 5 minutes and then rinsed several times with sterile water. Seeds were placed on Petri plates containing half-strength MS media supplemented with 1% (w/v) sucrose and 0.7% (w/v)

agar. After 3 days of stratification in darkness at low temperature, seeds germinated under a continuous light intensity of 15 µE/m^{2}s at 22 ± 2 °C and 60% relative humidity. Seedlings were investigated at 5 days after germination (DAG). Seedlings were grown vertically and then photographed from the side to determine morphological characteristics. First, a scale was set according to a ruler included in the picture using Fiji ImageJ version 1.53 (Schindelin et al., 2012). Hypocotyl and root were marked manually and their lengths were automatically measured.

Establishment and genotyping of CRISPR/Cas9-mediated knock-out mutants of AtWHIRLY1

Knock-out lines of *AtWHIRLY1* (*At1g14410*) were created by site-directed mutagenesis using CRISPR/Cas9 (Kumlehn et al., 2018; Li et al., 2021). Annealed oligos were integrated via *Bbs*I restriction enzyme sites in plasmid pSI57 yielding pGH510. The gRNA/Cas9 expression cassettes were introduced into p6i-d35S-TE9 (DNA-Cloning-Service, Hamburg, Germany) via *Sfi*I, generating plasmid pGH482. This construct was used for the *Agrobacterium*-mediated transformation of the ecotype Col-0. Primary transformants (T1) were screened on Murashige and Skoog (MS) media supplemented with 33.3 µg/ml hygromycin B. Resistant seedlings were checked for the presence of DNA encoding the Cas9 endonuclease by PCR with primers Cas9-F2/Cas9-R2. The next generations (T2 and T3) of *Cas9*-positive T1 plants were further analyzed. The mutated *WHIRLY1* locus was screened by PCR amplification and DNA sequencing. The mutated-*WHIRLY1*-containing homozygous lines were also checked for the absence of the Cas9 construct using specific primers.

For genotyping of the *WHIRLY1* knock-out mutants, a PCR-coupled cleavage amplified polymorphism site (CAPS) approach was used. Firstly, segments containing mutated site were amplified by PCR, whereby each reaction comprises 1 µl of diluted DNA (100 ng) as the template, 2 µl 10X *Taq* DNA Polymerase Buffer, 0.4 µl 25 mM MgCl2, 1.2 µl each primer including Why1-F1 and Why1-R1, and 0.1 µl 10 U/µl *Taq* DNA Polymerase (EURx, Poland). Then, 5 µl PCR products were digested with restriction enzyme in a 20 µl reaction with 2 µl 10X NEBuffer 3.1 and 1 µl 10 U/µl *Bsl*I (New England Biolab, USA). The reaction was performed at 37°C for 3 hours. The products obtained after digest were visualized by electrophoresis on an agarose gel.

Chlorophyll fluorescence measurements

To measure maximum photosystem II (PSII) efficiency, the FlourCam7 version 1.2.5.3 system (PSI, Czech Republic) was used. First, seedlings were dark-adapted for 15 min to determine the minimum chlorophyll fluorescence yield (F_O value). Then, plants were exposed to a saturating flash light, leading to the maximum fluorescence emission (F_M value). The

maximum PSII value was calculated as the ratio F_V/F_M, where F_V is the variable fluorescence ($F_V = F_M – F_O$).

Determination of pigments

To determine pigment contents, 20-30 mg of seedling shoots were collected and finely ground in liquid nitrogen. Then, plant materials were homogenized with 200 µl of Acetone/Tris buffer (80% acetone in 10 mM Tris-HCl pH7.8) by vigorously mixing. The mixture was incubated for 10 min in a cold ultrasonic bath before centrifuging at 13 000 rpm for 10 min at 4°C. Next, the supernatant was transferred into a new centrifuge tube. The absorption values of the supernatant were measured using a NanoPhotometer® NP80 (Implen GmbH, Germany) at the following wavelengths: 470, 537, 647, 663, and 750 nm. Pigment concentration was calculated using the equations of Sims & Gamon (2002).

RNA extraction and quantitative RT-PCR

For gene expression analysis, 5-day-old whole seedlings were ground in liquid nitrogen, and the total RNA was isolated using the RNeasy Plant Mini Kit (QIAGEN, Netherlands). 1 µg of total RNA was used to synthesize cDNA with the RevertAidTM H Minus First Strand cDNA Synthesis kit (Thermo Fisher Scientific, USA). The quantitative real-time PCR reaction was conducted using 1 µl cDNA (12.5 ng), 0.4 µl forward/reverse primer, 5 µl 2X SYBR green master mix (KAPA SYBR FAST Universal, KAPABIOSYSTEMS) and the CFX Connect Real-Time PCR Detection System (Bio-Rad, USA). The relative expression level was calculated using the method of Pfaffl (2001), in which three reference genes were used: *AtACTIN2, At5g46630,* and *AtPP2AA2*. The gene identifiers (IDs), gene names, gene-specific primers, and amplicon sizes are provided in Supplemental Table S3.

RNA sequencing

Total RNA was isolated using the RNeasy Plant Mini Kit (QIAGEN, Germany). Then, RNA quantity and quality were determined using the Agilent RNA 6000 Pico Kit with Bioanalyzer 2100 system (Agilent Technologies, USA). Samples with high purity and RNA Integrity Number (RIN) were sent to Novogene Co., Ltd (UK) for library preparation, RNA sequencing, and bioinformatics analysis according to their standard protocols. RNA-seq was done using an Illumina-based NovaSeq 6000 system with 150 bp paired-end read sequencing. Reads were checked with FastQC and cleaned up before mapping to the reference *A. thaliana* genome (TAIR10) by HISAT2 (Kim et al., 2019). The number of reads for each gene was counted using featureCounts of HISAT2. Then, the fpkm value (fragments per kilobase of transcript per million mapped reads) was calculated as the relative expression level of a gene in the sample, in which a gene is considered as expressed if the fpkm > 1. Differently expressed genes (DEGs) between the mutant and wildtype were determined using raw read counts by DESeq2 R

package. Functional annotation of DEGs was performed using the gene ontology enrichment tool (Aleksander et al., 2023; Ashburner et al., 2000).

Quantification of glucosinolates

3-5 mg of *A. thaliana* whole seedlings were ground in liquid nitrogen and subsequently mixed with 450 µl Methanol:Chloroform (2:1 volume ratio) extraction solution supplemented with 20 µg/ml sinigrin as an internal standard and 200 µl deionized water. Samples were incubated at room temperature for 60 min, followed by centrifugation at 10,000 x g for 20 min. The upper aqueous phase was used for GSL determination via Liquid Chromatography coupled with tandem Mass Spectrometry (LC-MS/MS). The separation of substrates was performed at 35°C with a Nucleoshell RP18 column (50 × 3mm, particle size 2.7 µm; Macherey-Nagel, Germany) using an Agilent 1290 High-Pressure Liquid Chromatography (HPLC) system. The analytes were detected by Electrospray Ionization MS/MS (ESI-MS/MS) by an API 3200 triple-quadrupole LC-MS/MS system with an ESI Turbo Ion Spray interface and operated in the negative ion mode (AB Sciex, Germany). GSL-specific signals were acquired via Multiple Reaction Monitoring (MRM), with a scan time of 15 msec. Peak areas were calculated automatically by the IntelliQuant algorithm of the Analyst 1.6 software (AB Sciex, Germany), with manual adjustment when necessary. GSL content in each sample was calculated and normalized to the internal standard and the fresh weight.

Statistical analysis

Physiological measurements, qRT-PCR, and GSL quantification were performed with three to four independent biological replicates. The mean of biological replicates was used to calculate the standard deviation. A two-tail paired Student's *t*-test was performed to estimate the statistically significant difference between lines using Microsoft Excel. Asterisks indicate statistical significance with * for p-value < 0.05, ** for p-value < 0.01, and *** for p-value <0.001.

Acknowledgement

We would like to thank Siska Herzklotz and Nancy Zimmermann for their excellent technical support. We also special thank Joerge Ziegler (IPB) for providing the optimized method for GSL quantification. This research was funded by the Deutsche Forschungsgemeinschaft (DFG, German Reasearch Foundaion) – 400681449/GRK2498.

Reference

Aleksander, S. A., Balhoff, J., Carbon, S., Cherry, J. M., Drabkin, H. J., Ebert, D., Feuermann, M., Gaudet, P., Harris, N. L., Hill, D. P., Lee, R., Mi, H., Moxon, S., Mungall, C. J., Muruganugan, A., Mushayahama, T., Sternberg, P. W., Thomas, P. D., Van Auken, K., … Westerfield, M. (2023). The Gene Ontology knowledgebase in 2023. *GENETICS*, *224*(1). https://doi.org/10.1093/genetics/iyad031

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T., Harris, M. A., Hill, D. P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J. C., Richardson, J. E., Ringwald, M., Rubin, G. M., & Sherlock, G. (2000). Gene Ontology: tool for the unification of biology. *Nature Genetics*, *25*(1), 25–29. https://doi.org/10.1038/75556

Augustine, R., & Bisht, N. C. (2017). Regulation of Glucosinolate Metabolism: From Model Plant Arabidopsis thaliana to Brassica Crops. In *Reference Series in Phytochemistry* (pp. 163–199). Springer Science and Business Media B.V. https://doi.org/10.1007/978-3-319-25462-3_3

Balk, J., & Schaedler, T. A. (2014). Iron Cofactor Assembly in Plants. *Annual Review of Plant Biology*, *65*(1), 125–153. https://doi.org/10.1146/annurev-arplant-050213-035759

Barco, B., & Clay, N. K. (2019). Evolution of Glucosinolate Diversity via Whole-Genome Duplications, Gene Rearrangements, and Substrate Promiscuity. *Annu. Rev. Plant Biol*, *70*, 585–604. https://doi.org/10.1146/annurev-arplant-050718

Cappadocia, L., Maréchal, A., Parent, J. S., Lepage, É., Sygusch, J., & Brisson, N. (2010). Crystal structures of DNA-whirly complexes and their role in Arabidopsis organelle genome repair. *Plant Cell*, *22*(6), 1849–1867. https://doi.org/10.1105/tpc.109.071399

Desveaux, D., Allard, J., Brisson, N., & Sygusch, J. (2002). A new family of plant transcription factors displays a novel ssDNA-binding surface. *Nature Structural Biology*, *9*(7), 512–517. https://doi.org/10.1038/nsb814

Desveaux, D., Després, C., Joyeux, A., Subramaniam, R., & Brisson, N. (2000). PBF-2 Is a Novel Single-Stranded DNA Binding Factor Implicated in PR-10a Gene Activation in Potato. In *The Plant Cell* (Vol. 12). www.plantcell.org

Desveaux, D., Maréchal, A., & Brisson, N. (2005). Whirly transcription factors: Defense gene regulation and beyond. In *Trends in Plant Science* (Vol. 10, Issue 2, pp. 95–102). Elsevier Ltd. https://doi.org/10.1016/j.tplants.2004.12.008

Desveaux, D., Subramaniam, R., Despré, C., Mess, J.-N., Vesque, C. L., Fobert, P. R., Dangl, J. L., & Brisson, N. (2004). A “Whirly” Transcription Factor Is Required for Salicylic Acid-Dependent Disease Resistance in Arabidopsis. In *Developmental Cell* (Vol. 6). http://www.developmentalcell.com/cgi/content/

Edger, P. P., Hall, J. C., Harkess, A., Tang, M., Coombs, J., Mohammadin, S., Schranz, M. E., Xiong, Z., Leebens-Mack, J., Meyers, B. C., Sytsma, K. J., Koch, M. A., Al-Shehbaz, I. A., & Pires, J. C. (2018). Brassicales phylogeny inferred from 72 plastid genes: A reanalysis of the phylogenetic localization of two paleopolyploid events and origin of novel chemical defenses. *American Journal of Botany*, *105*(3), 463–469. https://doi.org/10.1002/ajb2.1040

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. *Plant Direct*, *4*(5). 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*(4), 1800–1804. https://doi.org/10.1104/pp.108.122796

Huang, D., Lan, W., Ma, W., Huang, R., Lin, W., Li, M., Chen, C. Y., Wu, K., & Miao, Y. (2022). WHIRLY1 recruits the histone deacetylase HDA15 repressing leaf senescence and flowering in Arabidopsis. *Journal of Integrative Plant Biology*, *64*(7), 1411–1429. https://doi.org/10.1111/jipb.13272

HUMBECK, K., QUAST, S., & KRUPINSKA, K. (1996). Functional and molecular changes in the photosynthetic apparatus during senescence of flag leaves from field-grown barley plants. *Plant, Cell & Environment*, *19*(3), 337–344. https://doi.org/10.1111/j.1365-3040.1996.tb00256.x

Isemer, R., Krause, K., Grabe, N., Kitahata, N., Asami, T., & Krupinska, K. (2012). Plastid located WHIRLY1 enhances the responsiveness of Arabidopsis seedlings toward abscisic acid. *Frontiers in Plant Science*, *3*(DEC). https://doi.org/10.3389/fpls.2012.00283

Isemer, R., Mulisch, M., Schäfer, A., Kirchner, S., Koop, H. U., & Krupinska, K. (2012). Recombinant Whirly1 translocates from transplastomic chloroplasts to the nucleus. *FEBS Letters*, *586*(1), 85–88. https://doi.org/10.1016/j.febslet.2011.11.029

Janack, B., Sosoi, P., Krupinska, K., & Humbeck, K. (2016a). Knockdown of WHIRLY1 affects drought stress-induced leaf senescence and histone modifications of the senescence-associated gene HvS40. *Plants*, *5*(3). https://doi.org/10.3390/plants5030037

Janack, B., Sosoi, P., Krupinska, K., & Humbeck, K. (2016b). Knockdown of WHIRLY1 affects drought stress-induced leaf senescence and histone modifications of the senescence-associated gene HvS40. *Plants*, *5*(3). https://doi.org/10.3390/plants5030037

Kim, D., Paggi, J. M., Park, C., Bennett, C., & Salzberg, S. L. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nature Biotechnology*, *37*(8), 907–915. https://doi.org/10.1038/s41587-019-0201-4

Kretschmer, M., Croll, D., & Kronstad, J. W. (2017). Chloroplast-associated metabolic functions influence the susceptibility of maize to Ustilago maydis. *Molecular Plant Pathology*, *18*(9), 1210–1221. https://doi.org/10.1111/mpp.12485

Krupinska, K., Blanco, N. E., Oetke, S., & Zottini, M. (2020). Genome communication in plants mediated by organelle–nucleus-located proteins. In *Philosophical Transactions of the Royal Society B: Biological Sciences* (Vol. 375, Issue 1801). Royal Society Publishing. https://doi.org/10.1098/rstb.2019.0397

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*(5), 1337–1347. https://doi.org/10.1007/s00425-018-03085-z

Krupinska, K., Dähnhardt, D., Fischer-Kilbienski, I., Kucharewicz, W., Scharrenberg, C., Trösch, M., & Buck, F. (2014). Identification of WHIRLY1 as a Factor Binding to the Promoter of the Stress- and

Senescence-Associated Gene HvS40. In *Journal of Plant Growth Regulation* (Vol. 33, Issue 1, pp. 91–105). Springer New York LLC. https://doi.org/10.1007/s00344-013-9378-9

Krupinska, K., Desel, C., Frank, S., & Hensel, G. (2022). WHIRLIES Are Multifunctional DNA-Binding Proteins With Impact on Plant Development and Stress Resistance. In *Frontiers in Plant Science* (Vol. 13). Frontiers Media S.A. https://doi.org/10.3389/fpls.2022.880423

Kumlehn, J., Pietralla, J., Hensel, G., Pacher, M., & Puchta, H. (2018). The CRISPR/Cas revolution continues: From efficient gene editing for crop breeding to plant synthetic biology. *Journal of Integrative Plant Biology*, *60*(12), 1127–1153. https://doi.org/10.1111/jipb.12734

Li, M., Ruwe, H., Melzer, M., Junker, A., Hensel, G., Tschiersch, H., Schwenkert, S., Chamas, S., Schmitz-Linneweber, C., Börner, T., & Stein, N. (2021). The Arabidopsis AAC Proteins CIL and CIA2 Are Sub-functionalized Paralogs Involved in Chloroplast Development. *Frontiers in Plant Science*, *12*. https://doi.org/10.3389/fpls.2021.681375

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*(12). https://doi.org/10.3390/cells8121585

Liu, W., Yan, Y., Zeng, H., Li, X., Wei, Y., Liu, G., He, C., & Shi, H. (2018). Functional characterization of WHY–WRKY75 transcriptional module in plant response to cassava bacterial blight. *Tree Physiology*, *38*(10), 1502–1512. https://doi.org/10.1093/treephys/tpy053

Manh, M. B., Ost, C., Peiter, E., Hause, B., Krupinska, K., & Humbeck, K. (2023). WHIRLY1 Acts Upstream of ABA-Related Reprogramming of Drought-Induced Gene Expression in Barley and Affects Stress-Related Histone Modifications. *International Journal of Molecular Sciences*, *24*(7), 6326. https://doi.org/10.3390/ijms24076326

Maréchal, A., Parent, J. S., Véronneau-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 of the United States of America*, *106*(34), 14693–14698. https://doi.org/10.1073/pnas.0901710106

McCoy, R. M., Julian, R., Kumar, S. R. V., Ranjan, R., Varala, K., & Li, Y. (2021). A systems biology approach to identify essential epigenetic regulators for specific biological processes in plants. *Plants*, *10*(2), 1–16. https://doi.org/10.3390/plants10020364

Mérillon, J.-M., & Ramawat, · K G. (2017). *Glucosinolates* (J.-M. Mérillon & · K G Ramawat, Eds.; 1st ed.). Springer Cham. https://doi.org/https://doi.org/10.1007/978-3-319-25462-3

Miao, Y., Jiang, J., Ren, Y., & Zhao, Z. (2013). The single-stranded DNA-binding protein WHIRLY1 represses WRKY53 expression and delays leaf senescence in a developmental stage-dependent manner in arabidopsis. *Plant Physiology*, *163*(2), 746–756. https://doi.org/10.1104/pp.113.223412

Miao, Y., Laun, T., Zimmermann, P., & Zentgraf, U. (2004). Targets of the WRKY53 transcription factor and its role during leaf senescence in Arabidopsis. *Plant Molecular Biology*, *55*(6), 853–867. https://doi.org/10.1007/s11103-004-2142-6

Mitreiter, S., & Gigolashvili, T. (2021). Regulation of glucosinolate biosynthesis. In *Journal of Experimental Botany* (Vol. 72, Issue 1, pp. 70–91). Oxford University Press. https://doi.org/10.1093/jxb/eraa479

Monireh Saeid Nia, Susann Frank, Anke Schäfer, Christine Desel, Maria Mulisch, Ulrike Voigt, Daniela Nowara, Yudelsy Antonia Tandron Moya, Wolfgang Bilger, Nicolaus von Wiren, Götz Hensel, & Karin Krupinska. (2023). The balance between growth and resistance is shifted to the latter by over-accumulation of chloroplast-nucleus located WHIRLY1 in barley. *BioRxiv*. https://doi.org/10.1101/2023.02.22.529264

Pfaffl, M. W. (2001). A new mathematical model for relative quantification in real-time RT-PCR. In *Nucleic Acids Research* (Vol. 29, Issue 9).

Pfannschmidt, T., & Munné-Bosch, S. (2013). *Plastid Signaling During the Plant Life Cycle* (pp. 503–528). https://doi.org/10.1007/978-94-007-5724-0_22

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*(16), 5152–5165. https://doi.org/10.1093/nar/gkn492

Qiu, Z., Chen, D., Teng, L., Guan, P., Yu, G., Zhang, P., Song, J., Zeng, Q., & Zhu, L. (2022). OsWHY1 Interacts with OsTRX z and is Essential for Early Chloroplast Development in Rice. *Rice*, *15*(1). https://doi.org/10.1186/s12284-022-00596-y

Schaller, A. (2017). *Untersuchungen zur Rolle von WHIRLY-Proteinen in der Kommunikation zwischen Zellkern und Organellen*. Christian-Albrechts-University of Kiel.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.-Y., White, D. J., Hartenstein, V., Eliceiri, K., Tomancak, P., & Cardona, A. (2012). Fiji: an open-source platform for biological-image analysis. *Nature Methods*, *9*(7), 676–682. https://doi.org/10.1038/nmeth.2019

Sims, D. A., & Gamon, J. A. (2002). Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages. *Remote Sensing of Environment*, *81*(2–3), 337–354. https://doi.org/10.1016/S0034-4257(02)00010-X

Sønderby, I. E., Burow, M., Rowe, H. C., Kliebenstein, D. J., & Halkier, B. A. (2010). A complex interplay of three R2R3 MYB transcription factors determines the profile of aliphatic glucosinolates in Arabidopsis. *Plant Physiology*, *153*(1), 348–363. https://doi.org/10.1104/pp.109.149286

Taylor, R. E., West, C. E., & Foyer, C. H. (2022). WHIRLY protein functions in plants. In *Food and Energy Security*. John Wiley and Sons Inc. https://doi.org/10.1002/fes3.379

Yoo, H. H., Kwon, C., Lee, M. M., & Chung, I. K. (2007). Single-stranded DNA binding factor AtWHY1 modulates telomere length homeostasis in Arabidopsis. *Plant Journal*, *49*(3), 442–451. https://doi.org/10.1111/j.1365-313X.2006.02974.x

Yuri L. Negroni, Irene Doro, Alberto Tamborrino, Irene Luzzi, S. F., Götz Hense, Solmaz Khosravi, Laura Maretto, Piergiorgio Stevanato, Fiorella Lo Schiavo, Maria Concetta de Pinto, Karin Krupinska, &

Michela Zottini. (n.d.). The Arabidopsis mitochondrial nucleoid-associated protein WHIRLY2 is required for a proper response to salt stress. *Plant & Cell Physiology*, *under revision*.

Zhuang, K., Kong, F., Zhang, S., Meng, C., Yang, M., Liu, Z., Wang, Y., Ma, N., & Meng, Q. (2019). Whirly1 enhances tolerance to chilling stress in tomato via protection of photosystem II and regulation of starch degradation. *New Phytologist*, *221*(4), 1998–2012. https://doi.org/10.1111/nph.15532