Current Plant Biology xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Current Plant Biology

journal homepage: www.elsevier.com/locate/cpb

Co-expression and regulation of photorespiratory genes in *Arabidopsis thaliana*: A bioinformatic approach☆

Miriam Laxa[a,*], Steffanie Fromm[a,b]

[a] Institute of Botany, Leibniz University Hannover, Herrenhaeuser Strasse 2, 30419 Hannover, Germany
[b] Institute of Plant Genetics, Leibniz University Hannover, Herrenhaeuser Strasse 2, 30419 Hannover, Germany

ARTICLE INFO

Keywords:
Photorespiration
Co-expression
Gene regulation
Promoter
cis-Regulatory elements
Intron-mediated enhancement

ABSTRACT

Being a pathway tightly linked to photosynthesis, photorespiration is regulated by light, but also by both nutrients and metabolites on transcriptional level. However, only little is known about the signals and how they are integrated on promoter level to coordinate the whole pathway. Using a bioinformatic approach we analyzed the co-expression patterns of photorespiratory genes, the predicted *cis*-regulatory elements in their 5′ upstream regions and the existence of introns in their 5′UTRs. We found that there are groups of photorespiratory genes that are strongly co-expressed among each other. The analysis showed a high co-expression between photorespiration and ammonia re-fixation. However, a strong co-regulation between two genes, like *GDCH1* and *GDCH2*, did not necessarily mean that these genes share common *cis*-element in their 5′ upstream regions. TATA-box, MYB1AT, and MYB4 binding site motifs occurred in 16 out of 20 genes. Furthermore, photorespiratory genes are subjected to alternative splicing. We discuss the presence of *cis*-elements in the context of both stress responses and development. A genome wide analysis of *Arabidopsis* 5′UTRs revealed that 5′UTRs introns are overrepresented in photorespiratory genes. Promoter:*gusA* studies indicated that photorespiratory gene expression is also regulated by intron-mediated enhancement (IME). As already shown for glutamate: glyoxylate aminotransferase, transcript abundance of serine: glyoxylate aminotransferase was affected by IME on mRNA level. IME of gene expression of glycolate oxidase was shown to act on translational level.

1. Introduction

Photorespiration is an inevitable pathway in C_3-plants that detoxifies 2-phosphoglycolate formed by the oxygenation reaction of RuBisCO [1]. The pathway necessitates the cooperation of three different organelles: the chloroplast, the mitochondrion and the peroxisome. In the chloroplast, P-glycolate is first converted to glycolate *via* 2-phosphoglycolate phosphatase (*PGLP1*). In peroxisomes, glycolate is converted to glyoxylate by glycolate oxidase (*GOX1/2*). This reaction is oxygen-dependent and leads to the formation of hydrogen peroxide which is subsequently detoxified to water and oxygen by catalase (*CAT2*). Glyoxylate is a substrate for glutamate: glyoxylate amino transferase (*GGT1*). The glycine formed in this reaction is transported to mitochondria in which it is metabolized to serine by the joint action of the glycine decarboxylase complex (GDC) and serine hydroxymethyltransferase (*SHMT1*). This reaction results in the formation of serine, CO_2 and NH_3 and generates NADH. Transferred back to the peroxisome, serine is a substrate for serine: glyoxylate aminotransferase (*SGAT*). Reduction of the produced hydroxypyruvate to glycerate is catalyzed by the hydroxypyruvate reductase (*HPR1*) and consumes NADH. Finally, glycerate is transported back to the chloroplast where it is phosphorylated by glycerate kinase (GK) to phosphoglycerate.

Research on photorespiration begun with the observation of the so-called post-illumination burst (PIB), the release of CO_2 from photosynthetically active leaves upon darkening [2]. This burst in CO_2 evolution declines back to the level of mitochondrial (dark) respiration within a minute [2]. The magnitude of PIB is correlated with the light intensity, temperature and CO_2 concentration [3,4]. Tregunna and colleagues [5] were the first who observed a correlation between the initial rate of PIB in both C3 and C4 plants and photosynthesis. In addition, they introduced the term “photo-stimulated respiration” [5]. The discovery of photorespiration by Decker and others finally explained the observation that the rate of photosynthesis was inhibited by oxygen [6].

The elucidation of the pathway and the enzymes involved was initiated by Somerville and Ogren [7,8]. They postulated that mutants

☆ This article is part of a special issue entitled “Plant Genomics and Bioinformatics”.

* Corresponding author. Present address: Plant Biochemistry and Physiology, Bielefeld University, Universitätsstraße 25, 33615 Bielefeld, Germany.
E-mail address: miriam.laxa@uni-bielefeld.de (M. Laxa).

https://doi.org/10.1016/j.cpb.2018.09.001
Received 18 June 2018; Received in revised form 3 September 2018; Accepted 6 September 2018

Table 1
List primers and their respective sequences used in this study.

Primer	Sequence (5' to 3')	Purpose
*gusA*_mRNA_fw	GAAGCCGATGTCACGCCG	qPCR
*gusA*_mRNA_rv	TTGCCGTTTTCGTCGGTAATC	qPCR
*GAPDH*_mRNA_fw	TTGGTGACAACAGGTCAAGCA	qPCR
*GAPDH*_mRNA_rv	AAACTTGTCGCTCAATGCAATC	qPCR
*GOX1*_fw	CACCCCATCTTCTCAAATTTCAACAAC	Cloning
*GOX1*_rv	CTTCTGAAGTGTGTTTCTGTC	Cloning
*GOX1*_Δ5I_rv	CTTCTGAAGTGTGTTTACAGCTTTGATTGGAATGGAG	Cloning
*GOX2*_fw	CACCGCAAGCATCGATCAGATAGTATC	Cloning
*GOX2*_rv	CTTCCGAAGTGTGTTTCTGTC	Cloning
*GOX2*_Δ5I_rv	CTTCCGAAGTGTGTTTCTTGATGGGAATAGAGCTCTG	Cloning
*SGAT*_fw	CACCGGTTCGATCACCGTTCGGAA	Cloning
*SGAT*_rv	TTTTTCCTCTTTTCTTTTGGATCCTC	Cloning
*SGAT*_Δ5I_rv	TTTTTCCTCTTTTCTTTTGGATCCTCCTCTACACTCTGAATAAGGCGGAAGC	Cloning

with defects in the pathway of photorespiration will merely be viable under low photorespiratory conditions (high CO_2 concentrations) or low oxygen [8]. Mutants generated by an EMS treatment of *Arabidopsis thaliana* seeds were grown under high CO_2 and then shifted to low CO_2 concentrations. Plants exhibiting a visible phenotype were transferred back to high CO_2 and, thus, were rescued from the applied photorespiratory stress. The first enzyme involved in the pathway of photorespiration was identified as PGLP followed by SGAT, SHMT, GDC, and a chloroplast dicarboxylate transporter (DiT) [7,9–12]. Henceforward, the characterization of photorespiratory mutants was and is still under investigation. In 2007, Schwarte and Bauwe identified PGLP1 as the predominant isoform in photorespiration [13]. Dellero and colleagues [14] showed that the function of the two GOX isoforms in photorespiration is redundant. In addition, GOX greatly affected growth and glycolate levels, which were increased in mutants displaying 5% GOX activity in leaves [14]. Along with these examples, extensive research has been done on HPR1 [15–18], GGT1 [19,20], SGAT [21–23], GDC [24–26], SHMT1 [27,28] and glutamine synthetase 2 (GS2) [29]. However, knowledge in the field of transport remains elusive and is a topic of current research on photorespiration. The function of *Arabidopsis* plastidic 2-oxoglutarate/malate transporter (AtpOMT1; DiT1 according to [30]) and the general dicarboxylate transporter (AtpDCT1; DiT2.1 according to [30]) was shown to be important for photorespiration [31]. Interestingly, a transporter located in the endoplasmatic reticulum (ATP antiporter endoplasmic reticulum adenylate transporter 1, ER-ANT1) is essential for a functional photorespiratory pathway [32]. Loss of ER-ANT1 leads to a ROS-mediated inactivation of the GDC and, thus, to an inhibition of photorespiration. Eisenhut and colleagues [33] identified the *A BOUT DE SOUFFLE* (*BOU*) gene that encodes for a mitochondrial carrier protein with unknown substrate. In the same year, a chloroplastidic glycolate/glycerate transporter (PLGG1) was described whose function is required for photorespiration [34].

Alongside with the knowledge on the biochemical level, there are several studies that analyze the regulation of photorespiratory genes on promoter level [35–40]. Transcript profilings indicated that, with the exception of glycerate kinase (*GLYK*), photorespiratory genes are regulated by light, osmotic stress and low-nitrate treatment [30]. Timm and colleagues [41] identified serine as metabolic signal in the regulation of photorespiration since serine accumulation in the *hpr1* mutant led to perturbations in transcript levels of several photorespiratory genes. This study adds evidence that regulation of photorespiration occurs on the transcriptional level at least for *PGLP1*, the genes encoding the P- and T-protein of GDC, *SHMT1*, and *GLYK*. Thus, it is of great importance to study photorespiration at this level. Besides starting investigations of photorespiratory gene regulation on promoter level [39], we are curious whether the regulation is coordinated by *e.g.* common *cis*-elements in the 5′ upstream regions of photorespiratory genes. We therefore performed a bioinformatic approach in which we analyzed co-expression data, predicted *cis*-regulatory elements and gene regulation by intron-mediated enhancement.

2. Material and methods

2.1. Plant material and growth

Experiments were performed with *Arabidopsis* (*Arabidopsis thaliana*) ecotype Columbia-0 (Col-0). Col-0 was also used for stable transformations. Plants were grown on one-half-strength Murashige and Skoog (MS) medium including 1x vitamins and 0.7% plant agar (both Duchefa) in a Plate Percival (CU-41L5/D; CLF Plant Climatics) for 16 days under short-day conditions (115 µE, 8 h light/16 h darkness; 22/20 °C). For the selection of transformed plants, kanamycin (25 mg/ml) was added to the medium. After 16 days, plants were harvested for either qPCR analysis or GUS staining 1 h after illumination. Samples meant for qPCR analysis were immediately frozen in liquid nitrogen.

2.2. RNA isolation, cDNA synthesis and qPCR analysis

RNA isolation, cDNA synthesis and qPCR analysis were performed as described previously [39]. Primers used for quantification are listed in Table 1.

2.3. Cloning of promoter::gusA constructs

The 5′ upstream and coding sequences were obtained from Phytozome v9.1 [42]. *GOX1* (At3g14420), *GOX2* (At3g14415) and *SGAT* (At2g13360) 5′UTR intron deletions (Δ5I) were generated by a simple PCR using a primer that binds to the adjacent 5′ upstream region of the intron. In addition, the primer included the 3′ downstream region of the intron (basepairs upstream of the start codon). Primers used for cloning are listed in Table 1. Constructs *GGT1* and *GGT1*-Δ5I were already published [39]. Cloning was performed as described before [39].

2.4. Stable Arabidopsis transformation

The generation of stable *Arabidopsis* plants was performed as described [39].

2.5. GUS activity assay

GUS activity of 16-days-old plantlets was determined according to Jefferson et al. [43] using the fluorogenic substrate 4-methylumbelliferyl b-D-glucuronide.

2.6. GUS staining

Sixteen-days-old plants were stained for GUS activity by vacuum

Table 2
List of bioinformatic tools used to analyze the co-expression network of photorespiratory genes. The table gives information about the tool used including wegpage and reference, the type of data that have been extracted (co-expression PR, additional genes in network and network visualization), the parameters that were used and the results obtained. PCC: Pearsons correlation coefficient, SCC: Spearmans correlation coefficient, PR: photorespiratory genes, PPI: protein-protein interaction, Ath-m: *Arabidopsis thaliana* microarray data, Ath-r: *Arabidopsis thaliana* RNAseq data.

Tool	Analyzed parameter	Co-expression measure	Selected tools and parameters	Selected matrix (automatically - A or manual - M)	Results	Handling	Webpage	Reference
AraNet v2	Coexpression PR Additional genes in network Network visualization	Score[a] Score[a]	Find new members of a pathway', all PR genes, organism: Arabidopsis; further analysis: GeneSet analysis using 'Gene ontology'	Co-expression, structure, protein interaction data from 19 different data types (see publication) (A)	ROC curve, list and ranking of both query and additional genes that are linked to each other, visualized network, gene ontology according to four different enrichment analysis (Biological process, KEGG, MapMan and AraCyc terms)	datasets cannot be selected, simple network visualization (not interactive, only query genes are displayed)	http://www.inetbio.org/aranet	[44]
ATTED-II	Coexpression PR	PCC	'Edge Annotation'	Ath-m.c7.1 (A)	PCC and co-expression plot of two query genes displayed in 'detail'	many sub-tools to choose from, laborious because PCC values (Edge Annotation) need to be extracted manually ('detail')	http://atted.jp	[45]
	Additional genes in network Network visualization	MR	'CoEx Viewer', 'Network Drawer', all PR genes, select platform (Ath-m or Ath-r), display type: Cytoscape or Graphviz, co-expression options: add a few genes, PPI options: add a few genes	Ath-m, Ath-r (M), PPI (A)	List additional genes in network and visualization (Graphviz), force-directed network (Cytoscape)			
CORNET 3.0	Additional genes in network Network visualization	PCC > 0.95, $p \leq 0.05$	'Co-expression tool', 'predefined', all PR genes, methode: PCC > 0.95, $p \leq 0.05$, top 10 genes, 'correlation query gene(s) with neighbor(s)', output: Cytoscape, edge label: mean of correlation coefficient; PPI: pairwise interaction, in Cytoscape: force-directed layout	Co-expression: Microarray compendium 2 TAIR10 (111 exp - no bias) (M), PPI: TAIR, De Bodt (filteres - high stringency), AraNet (gene-gene association) - experimental data (M)	Network image *via* Cytoscape	difficult to compare to other tools, changing parameters ultimately leads to different network, additional genes are not ranked, AGI codes of additional genes had to be extracted manually from the network in Cytoscape	https://bioinformatics.psb.ugent.be/cornet/	[46]
CressExpress	Additional genes in network	PCC > 0.95	'Run the tool', all PR genes, 'select all microarrays from all experiments', threshold: 0.95, email details	Version 4.0 (8941 arrays - RMA processing) (A)	List of additional genes in network is emailed, genes are ranked based on the number of linked query genes	threshold can be chosen	http://www.cressexpress.org/	[47]
CSB.DB	Coexpression PR	PCC, SCC	'*A. thaliana* AthCoR', 'Multiple Gene Query (mGQ)', all PR genes, select matrix, select coefficient	nasc0271 (m0271): miscellaneous expt.; Ath1 chip; NASCArray; 9694 genes (M)	table with query genes ranked according to best pairwise PCC, can be copied and pasted into exel	four different matrices and three different coefficients can be selected	http://www.csbdb.de/csbdb/dbcor/cor.html	[48]
GeneMANIA	Coexpression PR Additional genes in network Network visualization	Score[b]	A) tool on webpage: select Arabidopsis, networks: Co-expression and physical interaction, network weighting: Assigned based on query genes, B) *via* Cytoscape: open Cytoscape, App manager: GeneMANIA, parameters as given in A)	various matrices (listed in each report along with the percentage of their contribution for the result) (A)	Webpage: Network image, report and images can be downloaded, list of additional genes ranked according to linkage to query genes; Cytoscape: Network image, node table	Webpage: gene annotation linked to ncbi webpage; Cytoscape: Node table can be exported, but rank is not given, query genes and resulting genes can be better distinguished from each other (color code) compared to the network visualized on the webpage	http://genemania.org/	[49]
PlaNet	Additional genes in network	PCC	A) 'NetworkComparer', 'user-sepecific analysis', all PR genes, select items to be compared; B) 'NetworkComparer', 'Standard analysis', At1g68010 (HPR1), 'Select one gene from each group that scores highly and significantly to the query'; C) 'EnsembleNet', 'Get set search', all PR genes	Affymetrix GeneChips (ATH1, Barley1, *Medicago*, Poplar, Rice, Soybean, Wheat) (A), B) Groups 19, 23, 36, 45, 46, 60, 66, 73, 75 and 81 (see also Suppl File 1) (M)	A) Network image; B) Network image of query genes in different species; C) Network image and 'gene-gene association table' (genes are ranked according to the number of associations with query genes)	A) not all genes are included in the analysis, network not interactive, a lot of genes added to the network; C) not all genes are included in the analysis	http://aranet.mpimp-golm.mpg.de/	[50]

[a] Sum of LLS (log likelihood scores) of network links to all other query genes.
[b] Reflects how often paths that start at a given gene node end up in one of the query nodes and how long and heavily weighted those paths are.

infiltration with GUS staining solution [GUS buffer containing 0.1 M Tris-HCl, pH 7.0, and 0.05 M NaCl, 100 mM hexacyanoferrate(III), 100 mM hexacyanoferrate(II), 50 mg/ml X-Gluc, 0.1% Triton X-100], followed by an incubation at 37 °C for at least 12 h. Chlorophyll was washed out with 70% ethanol, and plants were transferred to 10% glyercol. Images were taken using a binocular (Olympus SZ2-ILST) connected to a camera (Color View; Soft Imaging System) and visualized with the associated software program AnalySIS getIT Stereo.

2.7. Bioinformatics

2.7.1. Co-expression analysis

Co-expression analysis of photorespiratory genes and genes linked to photorespiration was performed with seven different bioinformatics tools: AraNet v2 (http://www.inetbio.org/aranet; 44), ATTED-II (http://atted.jp/, 45), CORNET 3.0 (https://bioinformatics.psb.ugent.be/cornet/; 46), CressExpress (http://www.cressexpress.org/; 47), CSB.DB (http://www.csbdb.de/csbdb/dbcor/cor.html; 48), http://genemania.org/; 49) and PlaNet (http://aranet.mpimp-golm.mpg.de/; 50) (Table 2). We analyzed genes previously listed by Foyer et al. [30] and Eisenhut et al. [33], respectively: 2-phosphoglycolate phosphatase 1 (*PGLP1*; At5g36700, ID: 833635), glycerate kinase (*GLYK*; At1g80380, ID: 844378), glutamine synthetase (*GS2*; At5g35630, ID: 833535), glutamate synthase (*GLU1*; At5g04140, ID: 830292), glycolate oxidase 1 (*GOX1*; At3g14420, ID: 820665), glycolate oxidase 2 (*GOX2*; At3g14415, ID: 820664), serine: glyoxylate aminotransferase (*SGAT*; At2g13360, ID: 815822), glutamate: glyoxylate aminotransferase (*GGAT*; At1g23310, ID: 838940), hydroxypyruvate reductase 1 (*HPR1*; At1g68010, ID: 843129), hydroxypyruvate reductase 2 (*HPR2*; At1g79870, ID: 843129), catalase 2 (*CAT2*; At4g35090, ID: 829661), glycine decarboxylase complex H-protein 1 (*GDCH1*; At2g35370, ID: 818104), glycine decarboxylase complex H-protein 2 (*GDCH2*; At1g32470, ID: 840141), glycine decarboxylase P-protein 1 (*GLDP1*; At4g33010, ID: 829438), glycine decarboxylase P-protein 2 (*GLDP2*; At2g26080, ID: 817149), glycine cleavage T-protein (*GDCT*; At1g11860, ID: 837733), mitochondrial lipoamide dehydrogenase 1 (*mLPD1*; At1g48030, ID: 841221), serine hydroxymethyltransferase 1 (*SHMT1*; At4g37930, ID: 829949), dicarboxylate transporter 1 (*DiT1*; At5g12860, ID: 831126), dicarboxylate transporter 2.1 (*DiT2.1*; At5g64290, ID: 836550), dicarboxylate transporter 2.1 (*DiT2.2*; At5g64280, ID: 836549). Selected tools and parameters are given in Table 2 and Supplemental Tables.

2.7.2. Prediction of regulatory cis-elements

Prediction of regulatory *cis*-elements in genes encoding photorespiratory enzymes was performed with the Athena tool [51]. Athena identified *cis*-elements in the 5′ upstream regions of the individual PR genes until the next upstream lying gene was reached, but sticked to an upper limit of 3000 bp.

2.7.3. Identification of Arabidopsis genes containing a 5′UTR intron

To identify *Arabidopsis* genes containing 5′UTR introns we downloaded the dataset “TAIR10_5_utr_20101028” from the TAIR database (https://www.arabidopsis.org). We analyzed the data in two different ways. i) Genes were sorted according to their AGI code. Afterwards, we removed all duplicate AGI codes, originating from different predicted splice forms, from the list. In the case that we found one splice form containing an intron while another did not, the gene was counted as “intron-containing”. Genes containing 5′UTR introns were assigned to their bin according to MapMan [52]. Over- and underrepresentation of introns in 5′UTR was calculated by comparing the expected and the real number of genes with 5′UTR intron relative to the total number of genes within the different bins. Significant difference of over- and underrepresentation relative to the average was tested by a statistical x^2-test. ii) 5′UTR intron containing genes were grouped into classes considering the number of introns present. Afterwards, duplicates within the individual classes were removed.

2.7.4. Confirmation of 5′UTR introns with RNA seq data

Predicted 5′UTR introns of genes within the bin “photorespiration” were confirmed by RNA seq data available in the EIN3 browser [53]. The EIN3 browser is a public accessible database of RNAseq data located on the SIGnAL: Salk Institute Genomic Analysis Laboratory Homepage (http://signal.salk.edu/).

2.7.5. Determination of IMEter scores

IMEter scores were determined with the IMEter v2.1 algorithm by Rose et al. [54].

3. Results and discussion

It is known that transcriptional co-expression of genes is one way to coordinate the expression of enzymes that either function in the same pathway [55] or directly interact with each other as shown for cyclophilin Cyp20-3 and peroxiredoxin 2-Cys Prx [56]. Furthermore, co-expression analysis is a suitable tool to find new players in pathways. For example, the *Arabidopsis* transporter A BOUT DE SOUFFLE is co-expressed with genes encoding enzymes in the photorespiratory pathway [33]. The metabolic pathway of photorespiration involves the three different organelles chloroplast, peroxisome and mitochondrion [1]. Thus, it appears likely that gene expression of the associated enzymes needs to be co-regulated. We therefore performed a co-expression analysis and listed the individual genes according to subcellular location of the corresponding proteins (Fig. 1).

A common measure for the strength of co-expression between two genes is the Pearson correlation coefficient (PCC). PCC describes the linear relationship between two quantitative dimensions, hence changes in gene expression. A downside of PCC is that it can easily give false-positive results. For instance, PCC can indicate a good correlation between two genes only based on an outlier [57]. A different measure is Spearman’s correlation coefficient (SCC). SCC describes a monotone relationship of two dimensions, no linear relationship is assumed. The CSB.DB enables a direct comparison between PCC and SCC of a given set of genes. Both, PCC and SCC can range from −1 (negative correlation) to +1 (positive correlation). The strength of co-expression is defined as no correlation ($0 \leq r \leq 0.2$), weak to moderate ($0.2 < r < 0.5$), distinct ($0.5 < r < 0.8$) and strong to perfect ($0.8 < r < 1.0$). We only concentrated on positive correlations in this study.

In a first approach, we analyzed the co-expression of photorespiratory genes among each other. For this, we used the tools AraNet v2, ATTED-II, CSB.DB and GeneMANIA (Table 2). Furthermore, the co-expression viewer in ATTED-II allowed the visualization of the correlation of expression patterns for each of the individual PR gene pairs. While AraNet v2 and GeneMANIA gave information about the linkage of each photorespiratory gene to the other genes in the pathway, ATTED-II and CSB.DB displayed PCC. With the exception of CSB.DB, the selected matrices (the PCC/link was calculated/determined from) were chosen automatically by the tools. In the case of CSB.DB, we chose the matrix ‘nasc0271’, thus a source of miscellaneous experiments and, therefore, best comparable to the other tools (Table 2). In general, most of the matrices based on microarray data. With RNAseq on the rise, many of the databases will include RNAseq data in the future, as it is already realized in ATTED-II. Fig. 1 summarizes links and PCC between PR genes based on the independent tools used. In the case of PCC, we chose a cut off 0.7 and, thus, a correlation coefficient indicating at least a very distinct coherence between two genes. For an easier overview and comparison of the four tools the genes were highlighted (Fig. 1).

In AraNet v2, the best linked genes were *GDCT*, *GLPD1*, *GDCH2* and *HPR1* as indicated by the highest scores (130.29, 104.51, 69.11 and 61.34, respectively) among all PR genes (Fig. 1). The given score reflects the sum of LLS (log likelihood scores) of network links to all other

AraNet v2

	At5g36700	At1g80380	At5g35630	At5g04140	At3g14420	At3g14415	At2g13360	At1g23310	At1g68010	At1g79870	At4g35090	At2g35370	At1g32470	At4g33010	At2g26080	At1g11860	At1g48030	At4g37930	At5g12860	At5g64290	At5g64280
GENE	PGLP1	GLYK	GS2	GLU1	GOX1	GOX2	SGAT	GGT1	HPR1	HPR2	CAT2	GDCH1	GDCH2	GLDP1	GLDP2	GDCT	mLPD1	SHM1	DiT1	DiT2.1	DiT2.2
PGLP1																					
GLYK																					
GS2																					
GLU1																					
GOX1																					
GOX2																					
SGAT																					
GGT1																					
HPR1																					
HPR2																					
CAT2																					
GDCH1																					
GDCH2																					
GLDP1																					
GLDP2																					
GDCT																					
mLPD1																					
SHM1																					
DiT1																					
DiT2.1																					
DiT2.2																					
linked/total	7/21	2/21	5/21	6/21	7/21	7/21	9/21	7/21	13/21	4/21	8/21	7/21	10/21	13/21	8/21	13/21	6/21	10/21	3/21	1/21	not linked
Score	25.42	7.41	17.79	18.63	34.64	38.40	58.3	39.55	61.34	13.75	19.42	58.73	69.11	104.51	71.13	130.29	29.09	55.22	11.77	6.92	

ATTED-II

	At5g36700	At1g80380	At5g35630	At5g04140	At3g14420	At3g14415	At2g13360	At1g23310	At1g68010	At1g79870	At4g35090	At2g35370	At1g32470	At4g33010	At2g26080	At1g11860	At1g48030	At4g37930	At5g12860	At5g64290	At5g64280
GENE	PGLP1	GLYK	GS2	GLU1	GOX1	GOX2	SGAT	GGT1	HPR1	HPR2	CAT2	GDCH1	GDCH2	GLDP1	GLDP2	GDCT	mLPD1	SHM1	DiT1	DiT2.1	DiT2.2
PGLP1																					
GLYK			0.244	0.343			0.609	0.337	0.299	0.005	0.374	0.215	0.243	0.187	0.160	0.378	0.296	0.099	0.243	0.219	0.246
GS2		0.244		0.789			0.558	0.773	0.746	0.479	0.214	0.824	0.819	0.739	0.585	0.710	0.686	0.804	0.661	0.626	0.211
GLU1		0.343	0.789				0.715	0.882	0.888	0.296	0.439	0.878	0.845	0.829	0.574	0.765	0.718	0.818	0.679	0.615	0.393
GOX1																					
GOX2																					
SGAT		0.609	0.558	0.715				0.762	0.761	0.200	0.516	0.592	0.551	0.582	0.360	0.556	0.450	0.524	0.399	0.370	0.450
GGT1		0.337	0.773	0.882			0.762		0.895	0.337	0.455	0.777	0.736	0.810	0.555	0.654	0.620	0.804	0.569	0.481	0.380
HPR1		0.299	0.746	0.888			0.761	0.895		0.331	0.415	0.857	0.842	0.793	0.539	0.701	0.618	0.808	0.546	0.518	0.397
HPR2		0.005	0.479	0.296			0.200	0.337	0.331		0.085	0.388	0.374	0.375	0.365	0.345	0.472	0.425	0.370	0.463	0.225
CAT2		0.374	0.214	0.439			0.516	0.455	0.415	0.085		0.234	0.198	0.409	0.427	0.332	0.355	0.250	0.290	0.228	0.252
GDCH1		0.215	0.824	0.878			0.592	0.777	0.857	0.388	0.234		0.935	0.763	0.514	0.755	0.714	0.829	0.649	0.638	0.320
GDCH2		0.243	0.819	0.845			0.551	0.736	0.842	0.374	0.198	0.935		0.748	0.544	0.813	0.691	0.822	0.692	0.658	0.281
GLDP1		0.187	0.739	0.829			0.582	0.810	0.793	0.375	0.409	0.763	0.748		0.745	0.728	0.732	0.862	0.632	0.595	0.282
GLDP2		0.160	0.585	0.574			0.360	0.555	0.539	0.365	0.427	0.514	0.544	0.745		0.616	0.642	0.588	0.538	0.564	0.237
GDCT		0.378	0.710	0.765			0.556	0.654	0.701	0.345	0.332	0.755	0.813	0.728	0.616		0.757	0.656	0.769	0.716	0.302
mLPD1		0.296	0.686	0.718			0.450	0.620	0.618	0.472	0.355	0.714	0.691	0.732	0.642	0.757		0.652	0.742	0.744	0.356
SHM1		0.099	0.804	0.818			0.524	0.804	0.808	0.425	0.250	0.829	0.822	0.862	0.588	0.656	0.652		0.596	0.525	0.203
DiT1		0.243	0.661	0.679			0.399	0.569	0.546	0.370	0.290	0.649	0.692	0.632	0.538	0.769	0.742	0.596		0.826	0.260
DiT2.1		0.219	0.626	0.615			0.370	0.481	0.518	0.463	0.228	0.638	0.658	0.595	0.564	0.716	0.744	0.525	0.826		0.313
DiT2.2		0.246	0.211	0.393			0.450	0.380	0.397	0.225	0.252	0.320	0.281	0.282	0.237	0.302	0.356	0.203	0.260	0.313	

CSB.DB

	At5g36700	At1g80380	At5g35630	At5g04140	At3g14420	At3g14415	At2g13360	At1g23310	At1g68010	At1g79870	At4g35090	At2g35370	At1g32470	At4g33010	At2g26080	At1g11860	At1g48030	At4g37930	At5g12860	At5g64290	At5g64280
GENE	PGLP1	GLYK	GS2	GLU1	GOX1	GOX2	SGAT	GGT1	HPR1	HPR2	CAT2	GDCH1	GDCH2	GLDP1	GLDP2	GDCT	mLPD1	SHM1	DiT1	DiT2.1	DiT2.2
PGLP1		0.5884	0.7435		0.8333		0.7995	0.9115	0.9328	-0.0906	0.4536			0.8496	0.4856	0.8534	0.4569	0.9426	0.7046	0.6644	0.5451
GLYK	0.5884		0.7077		0.7378		0.7672	0.7345	0.6917	-0.0285	0.6738			0.5839	0.4281	0.6653	0.0820	0.6769	0.5813	0.3931	0.4169
GS2	0.7435	0.7077			0.7859		0.7204	0.8436	0.7437	-0.0648	0.8260			0.6683	0.5349	0.8916	0.1343	0.8937	0.6917	0.4633	0.3971
GLU1																					
GOX1	0.8333	0.7378	0.7859				0.8629	0.8770	0.9005	-0.1695	0.6580			0.7197	0.4760	0.7733	0.1791	0.8464	0.6927	0.4627	0.6236
GOX2																					
SGAT	0.7995	0.7672	0.7204		0.8629			0.8903	0.8778	-0.1299	0.6110			0.7608	0.5583	0.7272	0.3111	0.8069	0.5541	0.5137	0.6574
GGT1	0.9115	0.7345	0.8436		0.8770		0.8903		0.9312	-0.0866	0.6330			0.8864	0.6331	0.8302	0.3629	0.9449	0.6931	0.5954	0.6485
HPR1	0.9328	0.6917	0.7437		0.9005		0.8778	0.9312		-0.1829	0.5540			0.7644	0.4236	0.7771	0.3178	0.8967	0.5999	0.5075	0.6001
HPR2	-0.0906	-0.0285	-0.0648		-0.1695		-0.1299	-0.0866	-0.1829		-0.0524			0.2061	0.2750	0.0901	0.5174	-0.0891	0.1900	0.3670	0.0508
CAT2	0.4536	0.6738	0.8260		0.6580		0.6110	0.6330	0.5540	-0.0524				0.4242	0.4717	0.6710	-0.0453	0.6293	0.5611	0.2931	0.2517
GDCH1																					
GDCH2																					
GLDP1	0.8496	0.5839	0.6683		0.7197		0.7608	0.8864	0.7644	0.2061	0.4242				0.7833	0.7671	0.6046	0.8536	0.7410	0.7526	0.6663
GLDP2	0.4856	0.4281	0.5349		0.4760		0.5583	0.6331	0.4236	0.2750	0.4717			0.7833		0.5932	0.4954	0.5898	0.6027	0.6084	0.6323
GDCT	0.8534	0.6653	0.8916		0.7733		0.7272	0.8302	0.7771	0.0901	0.6710			0.7671	0.5932		0.4273	0.9053	0.8236	0.7153	0.4813
mLPD1	0.4569	0.0820	0.1343		0.1791		0.3111	0.3629	0.3178	0.5174	-0.0453			0.6046	0.4954	0.4273		0.3328	0.4523	0.6989	0.4473
SHM1	0.9426	0.6769	0.8937		0.8464		0.8069	0.9449	0.8967	-0.0891	0.6293			0.8536	0.5898	0.9053	0.3328		0.7105	0.5919	0.5109
DiT1	0.7046	0.5813	0.6917		0.6927		0.5541	0.6931	0.5999	0.1900	0.5611			0.7410	0.6027	0.8236	0.4523	0.7105		0.8104	0.5096
DiT2.1	0.6644	0.3931	0.4633		0.4627		0.5137	0.5954	0.5075	0.3670	0.2931			0.7526	0.6084	0.7153	0.6989	0.5919	0.8104		0.4622
DiT2.2	0.5451	0.4169	0.3971		0.6236		0.6574	0.6485	0.6001	0.0508	0.2517			0.6663	0.6323	0.4813	0.4473	0.5109	0.5096	0.4622	

GeneMANIA

	At5g36700	At1g80380	At5g35630	At5g04140	At3g14420	At3g14415	At2g13360	At1g23310	At1g68010	At1g79870	At4g35090	At2g35370	At1g32470	At4g33010	At2g26080	At1g11860	At1g48030	At4g37930	At5g12860	At5g64290	At5g64280
GENE	PGLP1	GLYK	GS2	GLU1	GOX1	GOX2	SGAT	GGT1	HPR1	HPR2	CAT2	GDCH1	GDCH2	GLDP1	GLDP2	GDCT	mLPD1	SHM1	DiT1	DiT2.1	DiT2.2
PGLP1																					
GLYK																					
GS2																					
GLU1																					
GOX1																					
GOX2																					
SGAT																					
GGT1																					
HPR1																					
HPR2																					
CAT2																					
GDCH1																					
GDCH2																					
GLDP1																					
GLDP2																					
GDCT																					
mLPD1																					
SHM1																					
DiT1																					
DiT2.1																					
DiT2.2																					

Fig. 1. Co-expression of photorespiratory genes and genes linked to photorespiration. The table displays co-expression (Pearson's correlation coefficient) between the different genes encoding photorespiratory enzymes. All genes listed in Foyer et al. [30] were analyzed for co-expression with AraNet v2, ATTED-II, CSB.DB and GeneMANIA. Columns marked grey assign genes for which the tools displayed no information. Genes were grouped by organelle destination of the gene product with the exception of the transporters (black): green – chloroplast, red – peroxisome and orange – mitochondrion.

query genes [44]. The least linked genes were *GLYK* and the transporters. Both peroxisomal and mitochondrial PR genes individually appear to be tightly linked. Among the genes encoding mitochondrion located PR gene products, *GLDP1*, *GDCT* and *SHMT1* were the best linked genes to those encoding chloroplast and peroxisome located gene products. Fig. 1 shows that *HPR1* was most evenly linked to genes encoding PR proteins located in chloroplasts, mitochondria and peroxisomes, but there is no link to the transporters. In contrast to the pattern of links between PR genes observed in AraNet2, GeneMANIA did not allow any prediction of a tighter connection between genes encoding mitochondria located PR gene products for instance (Fig. 1). Information about the individual genes and their linkage to other PR genes was extracted from the interactive network graphic displayed on the GeneMANIA webpage. Hence, the links did not reflect any weighting of the individual links. The main conclusion drawn from the GeneMANIA analysis was that *DiT2.2* is not linked to most of the genes, thus, presumably is not part of the network. In addition, *CAT2* and *GLDP2* were the least connected among all PR genes.

Based on a cut-off of $r \geq 0.7$ for PCC in ATTED-II and CSB.DB, again there is evidence that genes encoding either mitochondria or peroxisome located PR gene products were co-expressed best (Fig. 1). Once more, *GLDP1*, *GDCT*, *SHMT1* and HPR1 were best co-expressed with all other PR genes. While genes encoding enzymes in the re-assimilation of ammonia were linked to only a few PR genes when using AraNet v2, they appeared very well co-expressed with the PR genes in ATTED-II and CSB.DB. Both ATTED-II and CSB.DB qualified *HPR2*, *CAT2*, *GLPD2* and the transporters to be weakly co-expressed with the other genes in photorespiration.

Because DCB.DB allowed a direct comparison of a linear (PCC) and non-linear (SCC) correlation, we also extracted co-expression data of the PR genes using SCC. Using an identical cut-off of $r \geq 0.7$ for both PCC and SCC, we only found minor differences in the co-expression patterns (Supplemental Fig. S1). In contrast to PCC, SCC found a better co-expression of the transporters with the other PR genes, but a generally weaker co-expression of GLYK with all other genes.

We also checked the correlation of expression patterns for each of the individual PR gene pairs with the 'CoEx Viewer' of ATTED-II. Supplemental Fig. S2 exemplarily shows the correlation of expression patterns for the best co-expressed genes *GDCH1* and *GDCH2* ($r = 0.935$) and the least co-expressed genes *GLYK* and *HPR2* ($r = 0.005$).

Disregarding data generated with GeneMANIA, the co-expression analysis revealed evidence that, firstly, *GLYK* is not strongly co-expressed with the other PR genes. Secondly, both *GS2* and *GLU1* showed an at least distinct co-expression to most of the genes listed (Fig. 1). This points out how tight nitrogen metabolism is connected to photorespiration [58]. In addition, this is in agreement with the finding that nitrate reduction is blocked as a consequence of inhibition of photorespiration [59]. This relationship between primary nitrogen assimilation and photorespiration was already observed in *Lotus japonicus* by Pérez-Delgado and colleagues [60] based on transcriptomics and co-expression analysis. Thirdly, *HPR2* and *GLDP2* showed the weakest co-expression and the lowest linkage to all other PR genes. However, both enzymes can compensate for the function of their isoform *HPR1* and *GLDP1* in photorespiration, respectively. Single knockout lines of *GLDP1* and *GLDP2* are indistinguishable from wild type plants under ambient CO_2 conditions, while the double mutant did not develop beyond the cotyledon stage even under non-photorespiratory conditions [26]. *HPR2* represents an alternative cytosolic pathway for the conversion of hydroxypyruvate to glycerate during photorespiration [17].

Forthly, the PR genes encoding enzymes located in the peroxisome (*SGAT*, *GGT1* and *HPR1*) showed strong co-expression among each other and a moderate co-expression with *CAT2* (Fig. 1). CAT2 detoxifies hydrogen peroxide produced by GOX during photorespiration, and *cat2* mutants develop a photorespiratory phenotype under ambient CO_2 concentrations [61,62]. The moderate co-expression of *CAT2* with *SGAT*, *GGT1* and *HPR1* can be explained by its general role in microbodies because CAT2 can be found in both glyoxysomes and peroxisomes [63]. The consensus of the co-expression analysis in Fig. 1 is that *HPR1* is predicted to be the best co-expressed gene with GDCH2, GLDP1 and GDCT (encoding GDC complex components) among the genes encoding PR enzymes located in the peroxisome (Fig. 1). Interestingly, *mLPD1* is only linked to genes encoding mitochondrial PR enzymes. Furthermore, co-expression of *mLPD1* is weak to moderate according to ATTED-II and CSB.DB (Fig. 1). This is in line with the observation that knockout mutants of *mLPD1* are vital, show an increased plant growth and an increased rate in both CO_2 assimilation and photorespiration [64]. In general, co-expression of GDC components makes sense because this co-expression ensures the stoichiometry between subunits of GDC. But, GDC gene expression is not regulated by over-expression of single components of the complex as observed for the T-protein [65]. This means, the plant cell does not sense the amount of GDC protein and coordinates gene expression hereupon. This fits to the picture that photorespiratory gene expression is regulated on transcriptional level [41]. The strong co-expression with *SHMT1* was also comprehensible. SHMT1 uses the methylene carbon of glycine generated by the GDC and an additional glycine to form serine. The GDC can only function when all four subunits are assembled [66].

Fifthly, an important point in photorespiration is the exchange of metabolites across membranes. The transporters listed in Fig. 1 are oxoglutarate/malate transporters and are predominantly distinctly co-expressed with genes encoding GDC enzymes [67]. According to Fig. 1 the best co-expression and linkage to the other PR genes was observed for *DiT1* followed by *DiT2.1*, while *DiT2.2* generally showed a moderate co-expression.

In a second approach, we visualized the co-expression and, if possible, the interaction network of PR genes/enzymes. Here, both the type of network presentation and nomenclature were as diverse as the underlying matrices of the tools (Table 2). With the exception of AraNet v2, the tools also displayed genes that were closely connected to either most of the PR genes (ATTED-II and GeneMANIA) or only a selection of them (CORNET 3.0 and PlaNet) (Fig. 2).

The network generated by AraNet v2 reflected the observations made in Fig. 1. The center of the network is made up of genes encoding the GDC components and SHMT1. *HPR* is tightly linked to the center. The second co-expressed unit includes PR genes with peroxisome located gene products as indicated in the upper part of the network. *GLU1* (GLUS in the network) appears better connected than *GS2*. *DiT1*, *DiT2.1* and *GLYK* are part of the network but show a lower degree of linkage. As indicated in Fig. 1, *DiT2.2* is not part of the network according to AraNet v2 (Fig. 2).

ATTED-II gave two different networks depending on the matrix used. The microarray-based network included 15 PR genes, while the RNAseq-based included 19 of the 21 genes (Fig. 2). A direct comparison of the two networks revealed similarities and differences. In both networks the two transporter *DiT1* and *DiT2.1* (*DCT* in ATTED-II) seemed to be linked to *mtLPD1*. While genes encoding mitochondria located PR gene products cluster better based on the Ath-m matrix, genes encoding peroxisome located PR gene products cluster better based on the Ath-r matrix. But using both matrices, genes encoding enzymes in ammonia refixation, *GS2* and *GLU1* (*GLUS* in ATTED-II), were linked to both GLDP isoforms. CSPD1, an oxidoductase, a transketolase and PDS2 were predicted to be part of the network based on either Ath-m or Ath-r (Fig. 2). Interestingly, the Ath-m based network displayed *BOU* as a member of the network of PR genes. The role of the BOU transporter in photorespiration has recently been elucidated experimentally [33].

CORNET 3.0 displayed a network containing only five of the 21 PR genes, namely *HPR*, *GGT1*, *GLU1*, *GDCH1* and *GDCH2*. Here, HPR and the two GDCH isoforms are linked best to the predicted members of the network (Fig. 2).

The GeneMANIA network resembled features of the networks described before, even though the network was built on both expression data and data on physical interaction. Like in AraNet v2, the two *GOX*

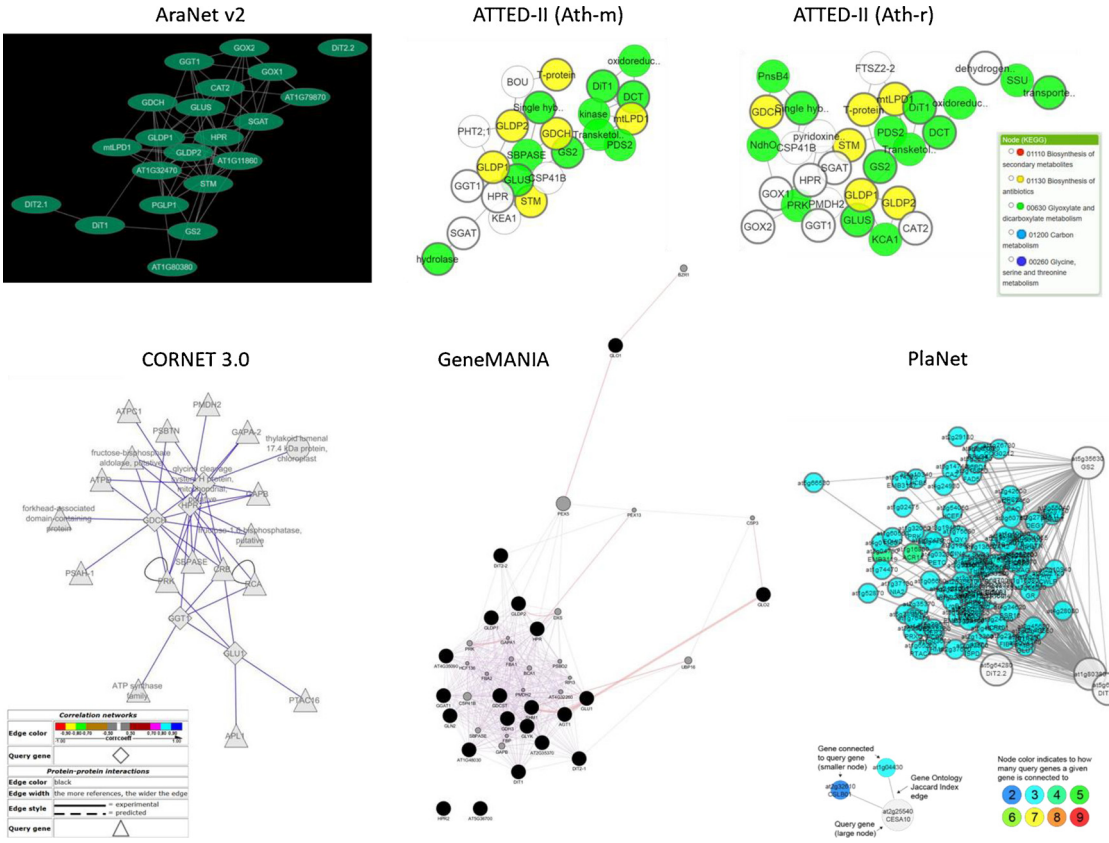

Fig. 2. Networks displayed with the different co-expression tools.

isoforms appeared loosely connected to the other PR genes (Fig. 2). The two transporters *DiT1* and *DiT2.1* as well as *mtLPD1* were located in the border area of the networks (AraNet v2, ATTED-II At-m and At-r and GeneMANIA). A distinct clustering of PR gene products located either in the mitochondria or peroxisome cannot be observed anymore using GeneMANIA. This might be the result of including data on physical interaction, even though the weighting of the data was 71.14%–28.86% in favor of the transcript data (Table 1). However, the GeneMANIA network nicely displayed both co-expression and interactions of the PR genes among each other and with the predicted network members.

PlaNet displayed the unclearest network with single members hardly to be separated from other members. Similar to CORNET 3.0 the network was built on the basis of only a few genes, namely *GLYK*, *GS2*, *PGLP1*, *DiT2.1* and *DiT2.2*. Thus, with the exception of *GS2* the PR gene network was based on members that are not the best linked and co-expressed genes within the network in general (Figs. 1 and 2).

Having a closer look on the individual network, we found that some genes like SBPase and CSP41B are displayed in different networks (Figs. 1 and 2). Despite using different matrices and algorithms, we therefore searched for genes enriched in multiple networks generated by the different tools. Table 3 summarizes predicted members of the PR network. With the exception of CORNET 3.0, the genes were ranked by the individual tools. Genes enriched in all data sets were marked grey (Table 3). In summary, five genes were found to be enriched, namely At5g09660 (*PMDH2*), At1g09340 (*CSP41B*), At3g55800 (*SBPase*), At1g42970 (*GAPB*) and At4g38970 (*FBA2*) (Table 4). *GDCT* appeared to be best co-expressed with these genes as indicated by strong correlation coefficients (Fig. 3).

PMDH2 encodes a peroxisomal NAD-malate dehydrogenase involved in ß-oxidation. Mutants laking both peroxisomal isoforms show a severe impairment in ß-oxidation, seedling establishment and a sucrose dependent growth phenotype [68]. However, the glyoxylate cycle functioned normally in the double mutant [68], even though PMDH

Table 3
List of genes linked to the network of PR genes. The genes were identified with the bioinformatics tools given. With the exception of CORNET 3.0, the rank of a gene displays their link to the PR genes. Genes that were enriched throughout the tools are marked grey. Ath-m: *Arabidopsis thaliana* microarray data, Ath-r: *Arabidopsis thaliana* RNAseq data.

Rank	AraNet v2	ATTED-II (Ath-m)	ATTED-II (Ath-r)	CORNET 3.0*	CressExpress	GeneMANIA	PlaNet
1	At5g09660	At3g55800	At5g09660	At2g21530	At1g09340	At5g56290	At3g04790
2	At3g13930	At1g01790	At4g38460	At5g53490	At3g55800	At1g09340	At2g42600
3	At1g54220	At5g46800	At1g18730	At5g19220	At3g54050	At4g24560	At2g42590
4	At2g35120	At1g15140	At5g10470	At4g32260	At1g32060	At1g75080	At1g56500
5	At3g17240	At3g11945	At3g52750	At4g09650	At1g42970	At4g15560	At1g16880
6	At1g42970	At3g26570	At1g74880	At4g38970	At1g12900	At1g42970	At5g54770
7	At2g47510	At3g60750	At1g15140	At3g46780	At2g35370	At1g32060	At5g09660
8	At1g09340	At1g56190	At3g11945	At1g12900	At5g09660	At3g55800	At4g39970
9	At1g03130	At1g09340	At3g60750	At5g09660	At3g55330	At3tg01500	At4g35260
10	At1g32080		At2g38230	At3g21055	At4g38970	At2g21330	At4g14680
11	At5g48230		At1g32060	At1g42970	At4g21280	At4g32260	At4g08900
12	At4g04640		At1g09340	At3g16140	At1g52230	At2g17870	At3g63160
13	At1g12900			At4g04640	At5g36700	At3g07560	At3g26650
14	At3g55800			At2g39730	At3g63140	At3g50820	At3g13450
15	At5g55070			At3g54050	At4g04640	At4g38970	At3g12780
16	At1g44575			At3g55800	At3g21055	At5g23120	At2g45630
17	At4g22890			At1g32060	At1g68010	At3g54050	At2g43750
18	At2g20420			At1g09340	At1g32470	At5g09660	At2g36885
19	At1g76150				At1g20020	At3g04790	At1g78050
20	At4g38970					At3g26650	At1g72610

* genes are not ranked

was considered to be a part of the photorespiratory pathway, providing NADH for HPR. The authors suggested that hydroxypyruvate can alternatively be reduced to glycerate in the cytosol as shown by Timm and colleagues [17]. Hence, there is ambiguity whether the enrichment of *PMDH2* within the predicted networks of PR genes is functionally relevant. Further experimental investigations are needed to clarify this observation in the future.

CSP41b (*chloroplast stem-loop binding protein 41 kDa*) is one of two *CSP41* isoforms (a, b) that share 52% sequence similarity on protein level. *CSP41b* belongs to the class of stroma proteins with the highest

Table 4
Genes found to be co-regulated or interacting with PR genes.

AGI code	Occurrence	Symbol	Gene	GO Biological Process according to TAIR
At5g09660	6	PMDH2	peroxisomal NAD-malate dehydrogenase 2	carbohydrate metabolic process, glyoxylate cycle, malate metabolic process, regulation of fatty acid beta-oxidation, regulation of photorespiration, response to cytokinin, tricarboxylic acid cycle
At1g09340	6	CSP41B	chloroplast RNA binding	cellular response to water deprivation, chloroplast organization, circadian rhythm, defense response to bacterium, monosaccharide metabolic process, plastid translation, polysaccharide catabolic process, positive regulation of transcription, DNA-templated, positive regulation of translation, rRNA processing, regulation of gene expression, response to cold, response to wounding
At3g55800	5	SBPase	sedoheptulose-bisphosphatase	carbohydrate biosynthetic process, carbohydrate metabolic process, defense response to bacterium, fructose 1,6-bisphosphate metabolic process, fructose 6-phosphate metabolic process, fructose metabolic process, gluconeogenesis, reductive pentose-phosphate cycle, starch biosynthetic process, sucrose biosynthetic process
At1g42970	4	GAPB	glyceraldehyde-3-phosphate dehydrogenase B subunit	glucose metabolic process, glycolytic process, oxidation-reduction process, reductive pentose-phosphate cycle, response to cadmium ion, response to cold, response to light stimulus, response to sucrose
At4g38970	4	FBA2	fructose-bisphosphate aldolase 2	gluconeogenesis, glycolytic process, pentose-phosphate shunt, response to abscisic acid, response to cadmium ion

level of abundance, while *CSP41a* groups into the second most abundant group [69]. CSP41a and CSP41b physically interact [70] and are members of higher molecular weight complexes that form in the dark and dissociate in light [71]. The model suggests CSP41 to bind non-translated mRNA and rRNA precursors to protect them from endonuclease activities in the dark. Like PR genes, CSP41b protein accumulates in mature leaves [72]. Loss of CSP41b led to a decrease in chlorophyll content in Arabidopsis and rice [73–75]. As described for *PMDH2*, there is no experimental evidence for a functionally relevant co-expression or interaction with PR genes/proteins. However, *CSP41b* was discovered as a target for *miR399f*, a micro RNA playing a crucial role in maintaining phosphate homeostasis in Arabidopsis [76]. Overexpression of *miR399f* Arabidopsis decreased *CSP41b* mRNA levels. *miR399f* overexpressing plants showed enhanced tolerance to salt stress and exogenous ABA supply, but were hypersensitive to drought [76]. Drought increases photorespiration. Whether the observed hypersensitivity to drought is caused by a disturbance of photorespiration which is linked to *CSP41b* expression, needs to be investigated in the future.

SPBase, GAPB and FBA2 are members of the Calvin cycle in chloroplasts. SBPase has a key function in the regulation of carbon flow through the Calvin cycle [77,78], and overexpression of SBPase in rice led to enhanced tolerance to both salt stress and high temperature [79,80]. In these plants, improved tolerance was associated with the content and activity of FBPase. In tomato, overexpression of SPBase enhanced tolerance to chilling stress [81]. In accordance with this, SBPase loss-of-function Arabidopsis mutants (*sbp*) suffered from severe growth retardation through inhibition of cell division and expansion [82]. ROS-mediated inactivation of SBPase following stress inhibited

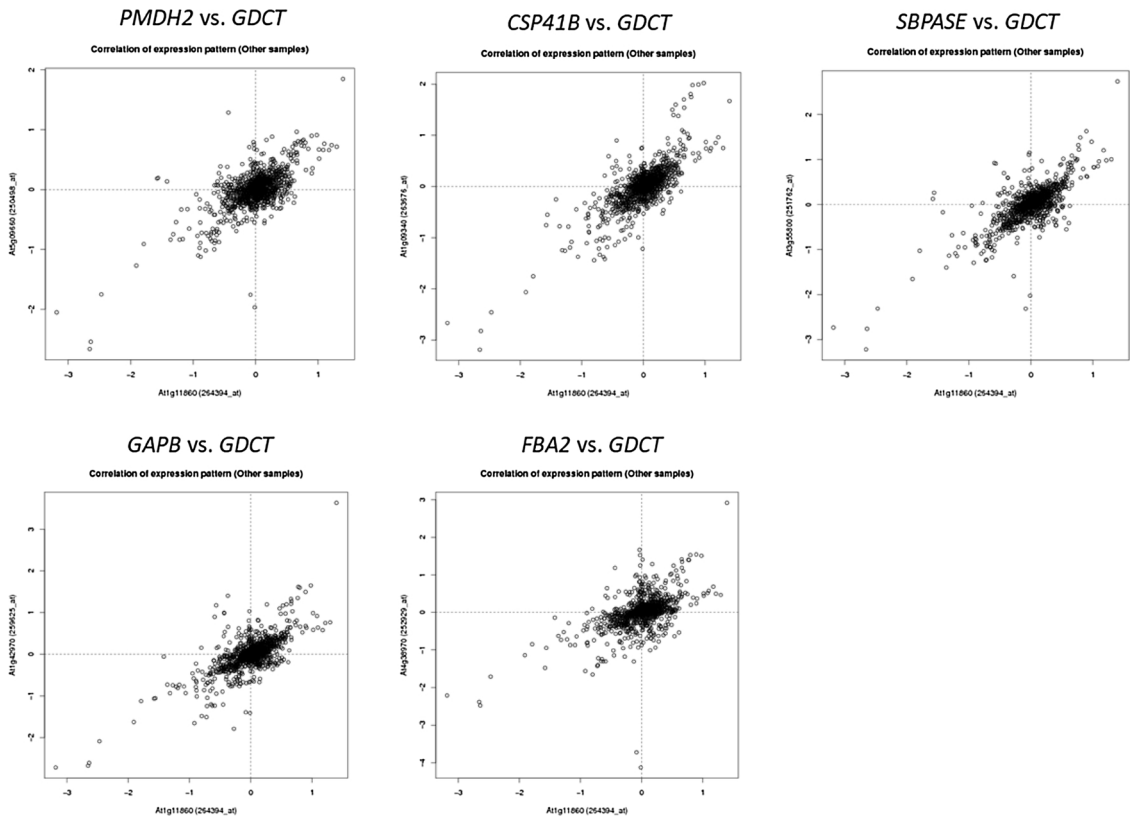

Fig. 3. Correlation of expression patterns of GDCT and the five most enriched genes according to AraNet v2, ATTED-II, CORNET 3.0, CressExpress, GeneMANIA and PlaNet.

the carbon assimilation efficiency. In this scenario, the author proposed that the inactivation of SBPase is an adaptation of plants to down-regulate the reductive pentose phosphate pathway under stress conditions [82]. Like SBPase, FBA2 is predicted to control carbon flux through the Calvin cycle and overexpression of FBA2 increased growth and photosynthesis [83]. Unfortunately, none of the studies analyzed either photorespiratory gene expression or the accumulation of photorespiratory metabolites in SBPase and FBA2 mutants. Photorespiration is tightly linked to the Calvin cycle *via* the oxygenase activity of RUBISCO [84]. Any increase in ribulose 1,5-bisphosphate regeneration will ultimately lead to an increase in CO_2 assimilation, but also in O_2 fixation under ambient CO_2 conditions. Thus, co-expression of PR genes and genes encoding regulatory enzymes in the regeneration step of the Calvin cycle makes perfectly sense. Reverse genetics revealed that disruption of the photorespiratory pathway leads to alteration in photosynthesis, sugar metabolism and sencescence [52,85–89]. There is strong experimental evidence that the primary product of the oxygenation reaction of RUBISCO, 2-phosphoglycolate (2-PG), feeds back on the Calvin cycle. Flügel and colleagues [90] used *pglp1* antisense lines to show that varying levels of 2-PG altered PSII efficieny, net CO_2 uptake and the CO_2 compensation point for example. Mechanistically, 2-PG regulates the Calvin cycle by inhibiting triose-phosphate isomerase (TPI) and FBPase [90]. Beside PGLP1 activity, GOX plays a role in regulating primary metabolism. Xu and colleagues [85] reported a linear relationship between GOX inhibition and the decrease in the rate of photosynthesis. This observation was strengthened by a recent publication by Dellero and colleagues [14] in which a decrease in glycolate oxidase activity altered both carbon allocation and leaf senescence. Beside this, GOX activity was shown to be involved in cell death and R-gene-mediated resistance [91]. Interestingly, Gilbert and Wolpert [92] identified six genes, that, when silenced, suppressed both the *LOV1* (*locus orchestrating victorin effects 1*) -mediated, victorin-induced and the *RPP8*-induced cell death in tobacco. According to sequence similarity, the genes were assigned to *GAPB*, *GOX1*, *GS2*, *GDCT*, *GLDP1* and *PHT*, the mitochondrial phosphate transporter 3, in Arabidopsis. However, the mechanism is still unclear. But based on their sensitivity to redox regulation, GDC (glutathionylation) [93] and GAPB (glutathionylation and oxidation) [94] are thought to play a role in ROS signaling. Based on this experimental data, the finding of GAPB within the network of PR genes in our bioinformatic approach fits well.

According to the co-expression analysis there are groups of genes that are co-regulated stronger than others among each other. Because gene expression is controlled by promoter elements, we analyzed the 5´ upstream regions of photorespiratory genes with Athena [51]. In particular, we were interested whether genes that showed a strong co-regulation contain identical *cis*-elements in their 5´ upstream sequences. Identical *cis*-elements in these promoters might offer a starting-point to analyze the regulation of co-expressed photorespiratory genes.

Fig. 4 gives an overview on putative *cis*-elements identified within the individual 5´ upstream regions. Comparing the number of putative *cis*-elements in the 5´ upstream regions of the individual gene, it becomes obvious that some genes like both *GDCH* isoforms contain only a low number of *cis*-elements, while others contain a large set of *cis*-elements as it is the case for *CAT2* (Fig. 4). The number of predicted *cis*-elements in the 5′ upstream region of *CAT2* can be directly linked to its function, the detoxification of hydrogen peroxide that is produced following several biotic and abiotic stresses [95]. Furthermore, the presence of multiple *cis*-elements in the 5´ upstream region, providing many sites to regulate gene expression, is accompanied with an open promoter structure of *CAT2* as indicated by a low number of nucleosomes that is not restricted to the transcription initiation start (TIS) [40]. There was no *cis*-element identified common to all twenty 5′ upstream regions tested, but a few *cis*-elements were enriched. The most abundant *cis*-elements were the TATA-box motif, the MYB1AT- and MYB4 binding site motif. Each motif was found in sixteen out of twenty 5′ upstream regions.

The TATA-box motif is one of two key regulatory *cis*-elements in the core promoter of plants. The TATA-box is located approximately 30 bp from TIS and a binding site for the TATA binding protein (TBP) [96,97]. TBP is a part of the general transcription factor TFIID which itself is a component of the RNA polymerase II preinitiation complex [98]. Eukaryotes do not necessarily need a TATA-box to initiate transcription. TATA-box-less promoters also exist. These promoters contain an Initiator (Inr) element instead. The Inr element is the second key element in core promoters, enriched in pyrimidine, and surrounding the TIS [99,100]. However, Inr elements can also be found in TATA-box containing promoters. There was no TATA-box predicted for *GLYK*, *HPR1*, *GDCH2* and *DiT1* (Fig. 4). Even though TATA-box is often seen as a sole binding site for the TBP protein, its presence is also associated with light responses as shown by Kiran and colleagues [101].

MYB1 und MYB4 belong to the family of R2R3-type MYBs [102]. R2R3-type MYBs are involved in the regulation of primary and secondary metabolism, cell fate and identity, developmental processes as well as responses to biotic and abiotic stress [103–105]. *Arabidopsis AtMYB1* is constitutively high expressed in all plant organs and following various treatments, while reverse Northern analyses revealed a weak expression of *AtMYB4* in most plant organs except siliques [103]. *AtMYB4* gene expression is regulated [103]. Jin and colleagues [106] found that *AtMYB4* negatively regulates cinnamate 4-hydroxylase (C4H, CYP73 A) expression. C4H catalyzes the conversion of cinnamic acid to *p*-coumaric acid in the general isoprenoid pathway [107,108]. AtMYB4 also negatively impacts gene expression of other steps in this pathway like the synthesis of sinapate esters [106]. In line with the function of sinapate esters, the protection of the plant towards UV-B irradiation, AtMYB4 transcript is downregulated following exposure to UV-B [106]. MYB1 TFs from other species were shown to regulate anthocyanin biosynthesis [109,110] and to be involved in light [111] and drought responses [112,113].

A regulation of photorespiratory genes by R2R3-type MYB factors is plausible because abiotic stress conditions like drought and high light increase photorespiration [1]. Furthermore, MYBs might be involved in the coordinated downregulation of photorespiratory genes following infection with virulent pathogens [114]. We hypothesize an antagonistic MYB-based regulation of photorespiratory gene expression. This idea is based on a publication by Schenke and colleagues [115]. Testing the influence of abiotic stress (UV-B), biotic stress (flagellin22) or a combination of both (UV-B and flagellin) on the expression of genes encoding enzymes in the flavonoid biosynthesis, they found that immune stress responses override abiotic stress responses. More precisely, the UV-B induced synthesis of protective flavonols [116] is suppressed when flagellin22 is present simultaneously [115]. Crosstalk among UV-B and pathogen stress responses has been observed before [117]. Schenke and colleagues [115] attributed the crosstalk to two antagonistically acting MYB factors AtMYB12 and AtMYB4. AtMYB12 is one of several R2R3-type MYB transcription factors (TF) regulating genes in the flavonol pathway and the main positive regulator in UV-B-induced responses [118]. On the other hand, as aforementioned, AtMYB4 negatively regulates C4H expression and, thus, suppresses flavonol expression [106]. Whether this hypothesis remains true and which TFs are involved in the regulation of photorespiratory genes, needs to be investigated in the future.

In accordance with the finding that virulent pathogens down-regulate photorespiratory gene expression [114], a W-box motif was found in eleven out of twenty genes (Fig. 4). W-box motifs ((T)TGAC(C/T)) are recognized and bound by WRKY transcription factors and found in promoters of genes that respond to wounding or pathogens [119].

Despite the strong co-expression of *GDCH1* and *GDCH2* (Fig. 1), they do not share common predicted *cis*-elements in their 5′ upstream region (Fig. 4). This observation points out that co-regulation is not necessarily based on common *cis*-regulatory elements. In contrast, *GS2* and *GLU1* share a set of fourteen different *cis*-elements.

TF motif	At5g36700 PGLP1	At1g80380 GLYK	At5g35630 GS2	At5g04140 GLU1	At3g14415 GOX1	At3g14420 GOX2	At2g13360 SGAT	At1g23310 GGT1	At1g68010 HPR1	At1g79870 HPR2	At4g35090 CAT2	At2g35370 GDCH1	At1g32470 GDCH2	At4g33010 GLDP1	At2g26080 GLDP2	At1g11860 GDCT	At4g37930 SHMT1	At5g12860 DiT1	At5g64290 DiT2.1	At5g64280 DiT2.2
ABRE binding site motif																	1			
ABRE-like binding site motif				1		1		2	1		4			2	1		1			
ACGTABREMOTIFA2OSEM				1		1		1	1		2						1			
AGATCONSENSUS							1													
AGL2ATCONSENSUS							1													
AP1 BS in AP3			1																	
ARF binding site motif		1	1	1				1					2		1				2	
ATHB2 binding site motif										1	2				1					
AtMYB2 BS in RD22						1	1												2	
AtMYC2 BS in RD22			2	1							3			1		1				
BoxII promoter motif			2	3	1	1	1				4	1				2			1	
CACGTGMOTIF								2			6			2						
CArG promoter motif	2		2				2										2			
CARGCW8GAT			8	6	2		4	8	2	2	8			4	4	2	4		4	
CCA1 binding site motif			1	1		3	1	1								1				
DRE core motif		1					1		1		1		1		1					
E2F binding site motif																			1	
EveningElement promoter motif							1				1									
GADOWNAT				1				1			1									
Gap-box Motif								1												
GAREAT	1		5	2		1	2	1	1		4				1	1			3	1
GCC-box promoter motif				1															1	
Hexamer promoter motif					1			1											1	
Ibox promoter motif	1		5	3		1	4				2				1		1		2	
L1-box promoter motif			1				1	1							2					
LEAFYATAG				1			1				1	1								
LTRE promoter motif		1																		
MYB binding site promoter	2	1									1			1	2	1	1		1	
MYB1 binding site motif											1									
MYB1AT	1	1	6	4	1	7	4	2	1		7		1		2	4	5	1	8	
MYB1LEPR		1			1			2			1			1			1		1	
MYB2 binding site motif																	1			
MYB2AT				1							2						1	1		
MYB3 binding site motif					1												2			
MYB4 binding site motif	3	2	1	3	2	1	5	4			5	1		1	7	3		2	4	2
MYCATERD1			1	1							3			1		1				
Nonamer promoter motif																		1		
PI promoter motif			1																	
RAV1-B binding site motif	1					1														
RY-repeat promoter motif				2										2			2			
SV40 core promoter motif		1	1				1	1	1		2					1		1	1	
TATA-box Motif	1		14	12	3	1	5	2		1	11	1		2	2	6	5		4	5
T-box promoter motif	1		4	1		2		2			2				2				1	
TELO-box promoter motif			1	1								1								
TGA1 binding site motif	1								1		1									
UPRMOTIFIAT	1							1	1		1									
UPRMOTIFIIAT											1									
W-box promoter motif			5	1		4	3	1			1				2	2	1	1	4	
Z-box promoter motif				1																

Fig. 4. Putative *cis*-elements predicted in the 5´ upstream regions of photorespiratory and photorespiration-linked genes. *Cis*-elements were ordered alphabetically, genes according to organelles except transporters (black). Green – chloroplast, red – peroxisome and orange – mitochondrion. The numbers indicate the count of a specific *cis*-element within a promoter. Data were obtained with the Athena [51].

Beside the already discussed MYB1AT and MYB4 binding sites as well as the W-box motif, the 5′ upstream regions contain predicted SV40 core promoter motives. This motif affects alternative splicing [120]. Thus, we searched the EIN3 browser [53] for alternatively spliced transcripts for selected genes. Except for *DiT2.1*, at least two different splice forms (SF) were indicated for *GS2* (3 SF), *GDCT* (3 SF) and *DiT1* (2 SF). Alternative splice site were mainly found in the 5′UTR and 3′UTR regions and were visible in the histograms representing the RNA seq reads. In the case of *DiT1*, ethylene led to the clear appearance of SF2 1 h after treatment. Comparing both isoforms in Phytozome v12 [42], it becomes obvious that the last 4 bp of the coding sequences (including the stop codon) in the first SF are part of an intron of 80 bp in length in the second SF. The result is a protein that is shortened by one amino acid, a tryptophan. More importantly, the 3′UTR is shortened. In *Arabidopsis*, the length of a 3′UTR is correlated with a specific gene function. For instance, the length of 3′UTRs of genes involved in response to salt and cadmium range between 1 and 500 bp, while 3′UTR of genes involved in the regulation of jasmonic acid signaling and systemic acquired resistance are typically 501–1000 bp long [121]. Thus, changing the length of a 3′UTR might change gene expression of a gene in favor of the environmental conditions. The length of 3′UTRs can be regulated by the mechanism of alternative polyadenylation (APA). It is estimated that 75% of all *Arabidopsis* genes are subjected to APA [121]. Prominent examples for APA are genes involved in the regulation of flowering [122,123]. Alternative splicing also includes the variation of 5′UTR length, alternative TIS usage, and intron retention [121].

PR genes encoding photorespiratory enzymes located in the peroxisome were shown to be co-expressed (Fig. 4). In addition to a MYB1AT binding site and a SV40 core promoter motif, *SGAT*, *GGT1* and *HPR1* share a CARGCW8GAT and a GAREAT in the 5′ upstream regions. CARGCW8GAT, a variant of the CArG motif, is predicted more frequently than the GAREAT motif (Fig. 4).

The presence of CARGCW8GAT motifs in promoters of peroxisomal photorespiratory genes is interesting, because this motif is bound by the MADS domain TF AGL15 (AGAMOUS-like 15) [124]. AGL15 accumulates in developing plant embryos [125]. It is known that peroxisomes have specific functions dependent on the cell type. Beevers [63] classified peroxisomes in three subtypes: glyoxysomes, leaf peroxisomes and unspecialized peroxisomes. Catalase is the only gene being present in all subtypes, while photorespiratory enzymes exist in leaf peroxisomes. Enzymes of fatty acid ß-oxidation are found in glyoxysomes. Glyoxysomes and leaf peroxisomes are interconvertible [126], this means that glyoxysomes are present in seedling until photosynthesis is established. Then, their set of enzymes is exchanged by enzymes functioning in photorespiration [127,128]. *Vice versa*, in the course of senescence leaf peroxisomes are transformed back into glyoxysomes [129]. Hence, upregulation of AGL15 during later stages of plant development might downregulate gene expression of *SGAT*, *GGT1* and *HPR1*. Downregulation of *HPR1* will in turn lead to the downregulation of other photorespiratory genes as observed in Timm et al. [41].

The GAREAT motif confers gibberellin (GA) responsiveness [130]. GA is a phytohormone that is essential for seed germination in *Arabidopsis* [131]. Whether the GA responsive element is important for inhibition of photorespiratory gene expression during seed germination, needs to be investigated.

When analyzing peroxisomal photorespiratory gene expression with the EIN3 browser [53], we once more observed alternative splicing in these genes. While *GGT1* and *HPR1* splice forms differed in 3′UTR sequences, *SGAT* splice forms varied in the 5′UTR, more precisely in the position and length of the 5′UTR intron. This was very interesting for us, because we previously showed that maximum transcript abundance of *GGT1* was controlled by the presence of its 5′UTR intron. We also identified 5′UTR introns in both GOX isoforms and SGAT and asked the question, whether 5′UTR introns are over-represented in photorespiratory genes. This prompted us to search the TAIR database for the presence of 5′UTR introns in the *Arabidopsis* genome.

As described in the Methods section we followed two different strategies. First, we evaluated the number of genes containing a 5′UTR intron regardless of the number of introns present in the 5′UTR. We downloaded 27101 5′UTR sequences from the TAIR database and removed duplicate AGI codes, originating from different predicted splice forms, from the list. The remaining 19737 genes were sorted into intron-containing or intron-less 5′UTR categories. Out of these genes, 21.4% contained a 5′UTR intron (Fig. 5a). 5′UTRs in *Arabidopsis* contain up to five introns but predominantly contain one intron (85.0% of all

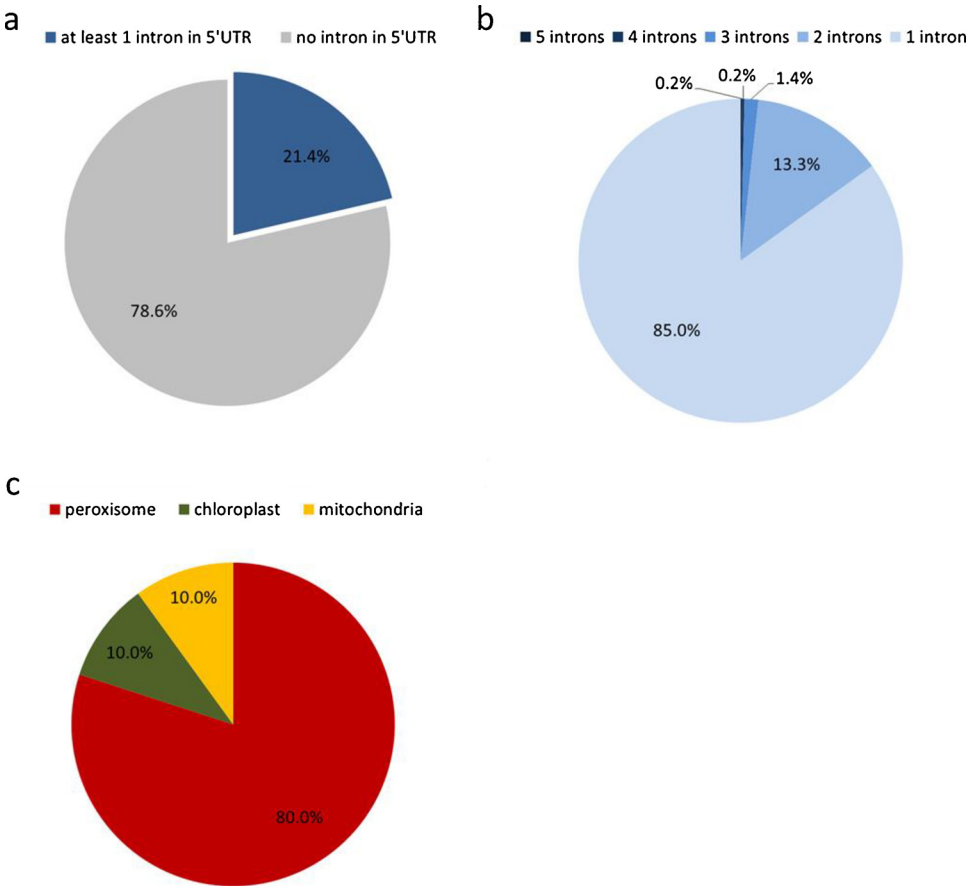

Fig. 5. 21.4% of all *Arabidopsis* genes contain at least one predicted intron in their 5′UTR. *Arabidopsis* 5′UTRs were analyzed for the presence or absence of 5′UTR introns. Results were displayed as (**a**) total amount of genes containing at least one 5′UTR intron, (**b**) partitioning of genes according to the predicted number of 5′UTR introns, and (**c**) 5′UTR intron containing photorespiratory genes (10 genes in total) sorted according to the final destination of the gene product.

intron containing genes) (Fig. 5b). 13.3% of the 5′UTRs contain two introns and only a small portion of 5′UTRs three (1.4%), four (0.2%) or five (0.2%) introns. In summary, we identified 4539 genes with 5′UTR intron. This number is higher than that calculated in Fig. 5a (4215 genes), because we allowed for multiple counts of 5′UTRs dependent on their different numbers of introns. Within one category multiple splice forms of one gene exist side by side. Here, the different splice forms either shared the same intron or the same amount of introns was alternatively spliced (data not shown). 1146 genes exist for which 5′UTRs with or without intron were predicted (data not shown).

To gain an overview of the genes function, the genes containing a 5′UTR intron (4215 genes) were assigned to their bins according to MapMan [52]. 4115 genes were successfully matched to the bins list. Table 5 summarizes the percentages of genes containing 5′UTR introns in each bin. The threshold for grouping a bin into the category 'over-represented' or 'underrepresented' was 21.4%, the average of 5′UTR intron containing genes within the whole *Arabidopsis* genome. A x^2-test confirmed whether the percentage of over- and underrepresented of 5′UTR intron containing genes was significant or not. According to Table 5 significant differences from the average (21.4%) were only observed for bins with an underrepresented number of 5′UTR intron containing genes. These bins included genes encoding enzymes involved in transport, photosynthesis, signaling, stress, development, hormone and secondary metabolism, and redox. Many of these processes are highly flexible in expression like signaling and stress responses following changing environmental conditions. In contrast, the category of overrepresented number of 5′UTR intron containing genes included major pathway like glycolysis, TCA cycle and amino acid metabolism, thus genes encoding enzymes functioning in C- and energy metabolism which are expressed constitutively in the lifecycle of a plant, here *Arabidopsis*. Even though these bins lack significance in overrepresentation, the results fit to the observation that 5′UTR introns often discriminate between vegetative and reproductive expression of genes within a gene family [132]. The gene family of *Arabidopsis* profilins contains five isoforms of which three (*PRF1, PRF2, and PRF3*) are expressed in vegetative tissues, while *PRF4* and *PRF5* are mainly expressed in reproductive tissues like pollen [132]. Jeong and colleagues found that the expression of *PRF1* and *PRF2*, representing the vegetative profilins, is solely mediated by the first intron. Furthermore, the *Arabidopsis* genes *ACT1* and *ACT2* represent vegetative and reproductive actin genes, respectively [133]. Intron deletion analysis showed that the expression of *ACT1* in pollen is strongly enhanced by the presence of its first intron. However, substituting the *ACT1* intron by the first intron of *ACT2* led to a strict repression of GUS activity in pollen [134].

Even though being a part of bin 1 "photosynthesis", the sub-bin photorespiration (1.02) showed an overrepresentation of 5′UTR intron containing genes (Table 5). This results caught our interest and we asked whether these 5′UTR introns present an additional level at which regulation of photorespiratory gene expression takes place. Eight out of ten photorespiratory genes with 5′UTR encode peroxisome located enzymes (*HAOX1*, *HAOX2*, *GOX1*, *GOX2*, *GOX3*, *GGT1*, *GGT2* and *SGAT*), while the gene products of *SHMT1* and *GLYK* are located in the mitochondrion and chloroplast, respectively (Table 6; Fig. 5c). To confirm the presence of 5′UTR introns in these genes, we analyzed their histograms, representing RNA seq reads, in the EIN3 browser [53]. The data confirmed the presence of spliced introns in all genes even though the expression of *HAOX1* and *HAOX2* was much lower compared to the other *GOX* genes. However, we could not determine the major splice form of *GLYK*. The first intron in the *GLYK* sequence is only part of the 5′UTR in splice form 3.

Introns were shown to enhance gene expression on different levels and at different strength [135]. A recent publication by Gallegos and Rose [136] demonstrated that introns even impact on transcription

Table 5
Introns in 5′UTRs of photorespiratory genes are overrepresented. Genes containing introns in their 5′UTR were assigned to their bins. The number of genes containing a 5′UTR intron was expressed as % of total genes. Significant over- and underrepresentation was verified by the x^2-test. The bold black line indicates the average percentage of all *Arabidopsis* genes containing a 5′UTR intron (21.4%; Fig. 5).

Bin	Bin number	Total number of genes in bin	Genes with 5'UTR intron	% of total genes	x^2-test
Polyamine metabolism	22	7	3	42.9	0.16629
Photorespiration	1.02	26	10	38.5	0.03390
Major CHO	2	103	29	28.2	0.09459
Glycolysis	4	61	16	26.2	0.35773
Minor CHO	3	123	31	25.2	0.30373
C1-metabolism	25	28	7	25.0	0.64231
TCA cycle	8	72	17	23.6	0.64734
Cell	31	616	143	23.2	0.27223
Amino acid metabolism	13	261	58	22.2	0.74602
RNA	27	3178	665	20.9	0.51391
N-metabolism	12	24	5	20.8	0.94603
Protein	29	3036	597	19.7	0.01969
Nucleotide metabolism	23	143	28	19.6	0.59574
OPPP	7	26	5	19.2	0.78739
Lipid metabolism	11	418	76	18.2	0.10865
mitochondrial ETC	9	109	19	17.4	0.31235
Gluconeogenesis	6	23	4	17.4	0.63924
Transport	34	928	149	16.1	0.00007
Tetrapyrrole synthesis	19	38	6	15.8	0.39907
Signalling	30	1120	172	15.4	0.00000
Photosynthesis	1	157	23	14.6	0.03918
Fermentation	5	14	2	14.3	0.51631
Cofactor & vitamine metabolism	18	38	5	13.2	0.21541
Not assigned	35	11994	1570	13.1	0.00000
Stress	20	923	111	12.0	0.00000
Development	33	353	39	11.0	0.00000
Hormone metabolism	17	600	66	11.0	0.00000
Metal handling	15	91	10	11.0	0.01545
Redox	21	187	19	10.2	0.00018
Secondary metabolism	16	437	43	9.8	0.00000
Cell wall	10	538	51	9.5	0.00000
S-assimilation	14	13	1	7.7	0.22817
Misc	26	1324	83	6.3	0.00000
DNA processes	28	3047	94	3.1	0.00000
Biodegradation of Xenobiotics	24	7	0	0.0	0.16742

Table 6
Confirmation of predicted 5′UTR introns of genes in bin 1.02 (photorespiration) by RNA seq data available in the EIN3 browser [53]. The table also includes information on the IMEter score of the individual introns as predicted by the IMEter algorithm v2.1 [54].

AGI	Gene	mRNA with spliced 5'UTR intron confirmed	IMEter score v2.1
At3g14130	*HAOX1*	yes	5.66
At3g14150	*HAOX2*	yes	5.28
At3g14415	*GOX1*	yes	3.53
At3g14420	*GOX2*	yes	4.21
At4g18360	*GOX3*	yes	10.01
At1g23310	*GGT1*	yes	12.62
At1g70580	*GGT2*	yes	17.30
At2g13360	*SGAT*	yes	8.36
At4g32520	*SHMT1*	yes	13.31
At1g80380	*GLYK*	yes, predicted for spliceform 3	6.07

initiation site (TIS) selection. We previously described that the 5′UTR intron of *GGT1* enhances maximum gene expression and influences the amount of RNA polymerase II bound to the TIS [39]. We also confirmed that the *GGT1* 5′UTR intron can be substituted by the 5′UTR intron of *GGT2* [39]. Thus, we hypothesized that intron-mediated enhancement plays a role in the expression of photorespiratory genes encoding peroxisome located enzymes. We focused on peroxisome located genes because they are all present in one organelle, are strongly co-regulated (Fig. 1) and, thus, represent a unity. Furthermore, with the exception of *CAT2*, the whole set of peroxisomal located enzymes contain an intron in their 5′UTRs (Table 6).

Introns that are able to mediate enhancement of gene expression are typically characterized by specific sequences like C/T-stretches [136–141] and the consensus motifs TTNGATYTG [54] and CGATT [142]. These motifs were found in a bioinformatic approach that compared the sequence composition of introns close to TIS to those located further downstream in gene sequences. Rose and colleagues [54] published the so-called IMEter algorithm which allows ranking introns for potential function in intron-mediated enhancement (IME). A positive score marks an intron to be most likely an enhancing intron [54].

Positive IMEter scores were calculated for all genes listed in Table 6. IMEter score and the enhancement of gene expression of an intron correlate in a linear way [54]. The higher the score, the higher the observed enhancement of gene expression [54]. For example, the first intron of *Arabidopsis* ubiquitin 10 (*UBQ10*) enhances mRNA expression by 12.5 fold and has an IMEter score of 97, while the non-enhancing intron 2 of *COR15a* has an IMEter score of -22 [54]. The IMEter scores

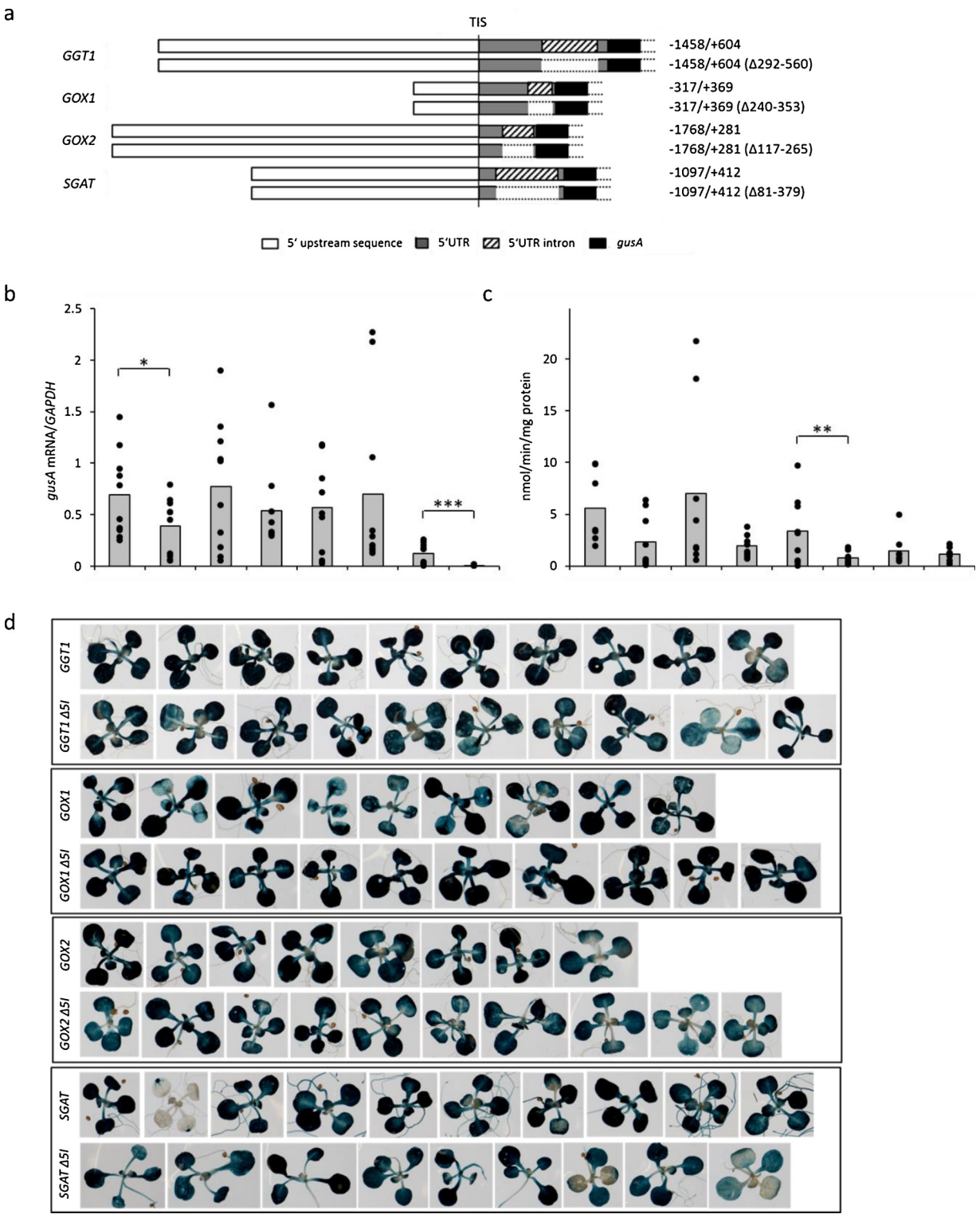

Fig. 6. Influence of the 5′UTR intron (5I) on the expression of genes encoding peroxisome located enzymes involved in photorespiration in vegetative tissue. *GGT1* and its intron-less version is a control in this experiment because the regulation of maximum transcript abundance by *GGT1* 5I was already shown [39]. (**a**) Schematic overview of the different constructs used in this experiment. The influence of 5I on both *gusA* transcript abundance (**b**) and GUS activity (**c**) was tested by qPCR and GUS activity measurements, respectively, according to Laxa et al. [39]. *Arabidopsis* plants were grown in short day conditions for 16 days before plants material was harvested 1 h after illumination. Individual transformation events of each construct are visualized by black (1 h L) circles, the mean by a grey bar. The mean represents the average of 9–10 (qPCR) and 8–9 (GUS activity assay) individual transformation events. Significance levels according to Student's *t*-test were as followed: *, $p \leq 0.1$; ***, $p \leq 0.01$. (**d**) GUS staining of individual transgenic lines.

for the 5′UTR introns of the photorespiratory genes ranged between 3.53 (*GOX1*) and 17.30 (*GGT2*) indicating that it the 5′UTR intron of *GGT2* enhances gene expression best. Based on the correlation between IMEter score and enhancement of gene expression [54] a 4- to 5-fold enhancement in mRNA levels can be assumed for *GOX1* and *GGT2*, respectively.

To test the hypothesis that gene expression of photorespiratory genes encoding peroxisome located enzymes are regulated by IME, 5′ upstream regions of *GOX1*, *GOX1* and *SGAT* were recombined into the binary GUS vector pSAG [38]. In parallel, constructs were generated in which the introns were deleted from the 5′UTRs (Fig. 6a). After stable transformation in *Arabidopsis* positive plants representing individual transformation events were selected on kanamycin containing medium. In the following experiments *GGT1*::*gusA* and *GGT1 Δ5I*::*gusA* lines were used as a control, because the 5′UTR intron of *GGT1* was already shown to enhance gene expression [39]. *GOX1*, *GOX2*, *GGT1* and *SGAT* were selected because these genes are high expressed in vegetative tissue according to the *Arabidopsis* eFP Browser 2.0 [143] (Fig. 6d).

Fig. 6 summarizes the influence of the 5′UTR introns of *GOX1*, *GOX2*, *GGT1* and *SGAT* on both mRNA level and protein level 1 h after illumination. The loss of the 5′UTR intron only caused significant changes in *gusA* mRNA levels for the control *GGT1* and for *SGAT*, which is also an aminotransferase (Fig. 6b). The 5′UTR intron of *GOX1* and *GOX2* did not affect *gusA* mRNA levels (Fig. 6b). As observed before the influence of the 5′UTR intron of *GGT1* was still visible on GUS activity level (Fig. 6c) [39]. However, the significant difference between lines expressing *gusA* under an intron-containing and intron-less SGAT 5′ upstream region was only seen by trend (Fig. 6c). The *GOX2* intron significantly affected the translational level of gene expression as it tends to be for *GOX1* (Fig. 6c). Changes on translational level could not be resolved *via* GUS staining due to the strong accumulation of GUS protein over time (Fig. 6d).

In summary, the data point out that the 5′UTR intron of the *GOX* genes enhances gene expression on translational level, while IME by the *SGAT* 5′UTR intron operates on mRNA level. Thus, expression of photorespiratory genes is not controlled by IME on the same level.

Screening the literature it becomes obvious that IME is more often reported on mRNA level than on translational level presumably, because scientists not always analyzed all transcriptional levels that can be targeted by IME [135]. Regulation on mRNA level by IME was also observed for the second major aminotransferase GGT1. Like *GGT1*, *SGAT* is expressed in green tissue [144]. *GGT1* transcript is strongly repressed in the dark and induced by light [145]. This repression was less distinct on GGT activity level. In contrast AlaAT1, another aminotransferase, showed a repression on activity level, while the *AlaAT1* transcript was not significantly changed. One could conclude that *GGT1* is regulated mainly on transcriptional level and *AlaAT1* predominantly on translational level. According to our data, *SGAT* is also regulated on transcriptional level by its 5′UTR intron (Fig. 6b). A publication by Taler and colleagues [146] also reported a regulation of SGAT on transcript level. They identified a SGAT isoform in wild melon that confers resistant to the pathogen *Pseudoperonospora cubensis*, an oomycete triggering downy mildew. The strength of resistance correlated with the amount of *SGAT* mRNA present in the resistant and hybrid (resistant x sensitive) lines [146]. Interestingly, we found that GOX2, and by trend GOX1, were regulated on translational level by their 5′UTR introns (Fig. 6b, c). This result is strengthened by an observation made by Barak et al. [147]. They showed that translation of a *gusA* transcript driven by a GOX 5′upstream region is enhanced approximately 30-fold over a CaMV 35S promoter-driven *gusA* transcript after a dark-to-light shift experiment in tobacco seedlings. The authors claimed *cis*-elements in the 5′UTR region to be responsible for this translational enhancement [147]. The second evidence for a translational regulation of GOX expression is based on a publication by Van Oosten and colleagues [148]. *GOX* mRNA abundance was not significantly changed when plants were exposed to high CO_2 concentrations, while GOX activity was decreased by this treatment [148,149]. Whether the 5′UTR intron plays an important role in the translational regulation needs to be investigated in future experiments.

4. Conclusion

Even though it is known that gene expression of photorespiratory genes is affected by light, nutrients, metabolites and pathogens, the underlying mechanism of integrating these signals on the promoters is highly complex. Photorespiratory genes do not necessarily share common *cis*-regulatory elements, but we found groups of genes with stronger correlation coefficients. Within these groups different *cis*-elements exist in a higher frequency compared to other groups. These groups will be of special interest in future experiments. We will work on the mechanistic basis of sub-regulatory units like the peroxisome located photorespiratory enzyme and the coordination of photorespiration with other pathways. In addition, IME will be of particular interest.

Conflict of interest

There is no conflict of interests.

Author contributions

M.L. designed, performed both experiments, analyzed data and wrote the article

S.F. performed experiments, analyzed data

Acknowledgements

The authors thank Prof. Dr. Hans-Peter Braun (Leibniz University of Hannover) for critically reading the manuscript and funding SF in the extended period of her PhD thesis. The publication of this article was funded by the Open Access Fund of the Leibniz Universität Hannover.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.cpb.2018.09.001.

References

[1] C. Peterhänsel, I. Horst, M. Niessen, C. Blume, R. Kebeish, S. Kürkcüoglu, F. Kreuzaler, Photorespiration, Arabidopsis Book 8 (2010) e0130.
[2] J.P. Decker, A rapid, postillumination deceleration of respiration in green leaves, Plant Physiol. 30 (1955) 82–84.
[3] J.P. Decker, Some effects of temperature and carbon dioxide concentration on photosynthesis of *Mimulus*, Plant Physiol. 34 (1959) 103–106.
[4] E.B. Tregunna, G. Krotkov, C.D. Nelson, Evolution of carbon dioxide by tobacco leaves during the dark period following illumination with the light of different intensities, Can. J. Bot. 39 (1961) 1045–1056.
[5] E.B. Tregunna, G. Krotkov, C.D. Nelson, Further evidence on the effects of light on respiration during photosynthesis, Can. J. Bot. 42 (1964) 989–997.
[6] J.S. Turner, E.G. Brittain, Oxygen as a factor in photosynthesis, Biol. Rev. Camb. Philos. Soc. 37 (1962) 130–170.
[7] C.R. Somerville, W.L. Ogren, A phosphoglycolate phosphatase-deficient mutant of *Arabidopsis*, Nature 280 (1979) 833–836.
[8] C. Somerville, W.L. Ogren, Genetic modification of photorespiration, Trends Biochem. Sci. 7 (1982) 171–174.
[9] C.R. Somerville, W.L. Ogren, Photorespiration mutants of *Arabidopsis thaliana* deficient in serine-glyoxylate aminotransferase activity, Proc. Natl. Acad. Sci. U. S. A. 77 (1980) 2684–2687.
[10] C.R. Somerville, W.L. Ogren, Photorespiration-deficient mutants of *Arabidopsis thaliana* lacking mitochondrial serine transhydroxymethylase activity, Plant Physiol. 67 (1981) 666–671.
[11] C.R. Somerville, W.L. Ogren, Mutants of the cruciferous plant *Arabidopsis thaliana* lacking glycine decarboxylase activity, Biochem. J. 202 (1982) 373–380.
[12] S.C. Somerville, W.L. Ogren, An *Arabidopsis thaliana* mutant defective in chloroplast dicarboxylate transport, Proc. Natl. Acad. Sci. U. S. A. 80 (1983) 1290–1294.
[13] S. Schwarte, H. Bauwe, Identification of the photorespiratory 2-phosphoglycolate phosphatase, PGLP1, in *Arabidopsis*, Plant Physiol. 144 (2007) 1580–1586.
[14] Y. Dellero, M. Jossier, N. Glab, C. Oury, G. Tcherkez, M. Hodges, Decreased

glycolate oxidase activity leads to altered carbon allocation and leaf senescence after a transfer from high CO_2 to ambient air in *Arabidopsis thaliana*, J. Exp. Bot. 67 (2016) 3149–3163.
[15] A.J. Murray, R.D. Blackwell, P.J. Lea, Metabolism of hydroxypyruvate in a mutant of barley lacking NADH-dependent hydroxypyruvate reductase, an important photorespiratory enzyme activity, Plant Physiol. 91 (1989) 395–400.
[16] C.V. Givan, L.A. Kleczkowski, The enzymic reduction of glyoxylate and hydroxypyruvate in leaves of higher plants, Plant Physiol. 100 (1992) 552–556.
[17] S. Timm, A. Nunes-Nesi, T. Pärnik, K. Morgenthal, S. Wienkoop, O. Keerberg, W. Weckwerth, L.A. Kleczkowski, A.R. Fernie, H. Bauwe, A cytosolic pathway for the conversion of hydroxypyruvate to glycerate during photorespiration in *Arabidopsis*, Plant Cell 20 (2008) 2848–2859.
[18] A.B. Cousins, B.J. Walker, I. Pracharoenwattana, S.M. Smith, M.R. Badger, Peroxisomal hydroxypyruvate reductase is not essential for photorespiration in *Arabidopsis* but its absence causes an increase in the stoichiometry of photorespiratory CO_2 release, Photosynth. Res. 108 (2011) 91–100.
[19] D. Igarashi, T. Miwa, M. Seki, M. Kobayashi, T. Kato, S. Tabata, K. Shinozaki, C. Ohsumi, Identification of photorespiratory glutamate: glyoxylate aminotransferase (GGAT) gene in *Arabidopsis*, Plant J. 33 (2003) 975–987.
[20] Y. Dellero, M. Lamothe-Sibold, M. Jossier, M. Hodges, *Arabidopsis thaliana ggt1* photorespiratory mutants maintain leaf carbon/nitrogen balance by reducing RuBisCO content and plant growth, Plant J. 83 (2015) 1005–1018.
[21] A.J. Murray, R.D. Blackwell, K.W. Joy, P.J. Lea, Photorespiratory N donors, aminotransferase specificity and photosynthesis in a mutant of barley deficient in serine: glyoxylate aminotransferase activity, Planta 172 (1987) 106–113.
[22] N.A. McHale, E.A. Havir, I. Zelitch, Photorespiratory toxicity in autotrophic cell cultures of a mutant of *Nicotiana sylvestris* lacking serine: glyoxylate aminotransferase activity, Planta 179 (1989) 67–72.
[23] K. Modde, S. Timm, A. Florian, K. Michl, A.R. Fernie, H. Bauwe, High serine:glyoxylate aminotransferase activity lowers leaf daytime serine levels, inducing the phosphoserine pathway in *Arabidopsis*, J. Exp. Bot. 68 (2017) 643–656.
[24] A.U. Igamberdiev, N.V. Bykova, P.J. Lea, P. Gardeström, The role of photorespiration in redox and energy balance of photosynthetic plant cells: a study with a barley mutant deficient in glycine decarboxylase, Physiol. Plant. 111 (2001) 427–438.
[25] N.V. Bykova, O. Keerberg, T. Pärnik, H. Bauwe, P. Gardeström, Interaction between photorespiration and respiration in transgenic potato plants with antisense reduction in glycine decarboxylase, Planta 222 (2005) 130–140.
[26] N. Engel, K. van den Daele, Ü. Kolukisaoglu, K. Morgenthal, W. Weckwerth, T. Pärnik, O. Keerberg, H. Bauwe, Deletion of glycine decarboxylase in *Arabidopsis* is lethal under non-photorespiratory conditions, Plant Physiol. 144 (2007) 1328–1335.
[27] K. Beckmann, C. Dzuibany, K. Biehler, H. Fock, R. Hell, A. Migge, T.W. Becker, Photosynthesis and fluorescence quenching, and the mRNA levels of plastidic glutamine synthetase or of mitochondrial serine hydroxymethyltransferase (SHMT) in the leaves of the wild-type and of the SHMT-deficient *stm* mutant of *Arabidopsis thaliana* in relation to the rate of photorespiration, Planta 202 (1997) 379–386.
[28] A. Kuhn, M.K. Engqvist, E.E. Jansen, A.P. Weber, C. Jakobs, V.G. Maurino, D-2-hydroxyglutarate metabolism is linked to photorespiration in the shm1-1 mutant, Plant Biol. 15 (2013) 776–784.
[29] M. Brestic, M. Zivcak, K. Olsovska, H.B. Shao, H.M. Kalaji, S.I. Allakhverdiev, Reduced glutamine synthetase activity plays a role in control of photosynthetic responses to high light in barley leaves, Plant Physiol. Biochem. 81 (2014) 74–83.
[30] C.H. Foyer, A.J. Bloom, G. Queval, G. Noctor, Photorespiratory metabolism: genes, mutants, energetics, and redox signaling, Annu. Rev. Plant Biol. 60 (2009) 455–484.
[31] M. Taniguchi, Y. Taniguchi, M. Kawasaki, S. Takeda, T. Kato, S. Sato, S. Tabata, H. Miyake, T. Sugiyama, Identifying and characterizing plastidic 2-oxoglutarate/malate and dicarboxylate transporters in *Arabidopsis thaliana*, Plant Cell Physiol. 43 (2002) 706–717.
[32] C. Hoffmann, B. Plocharski, I. Haferkamp, M. Leroch, R. Ewald, H. Bauwe, J. Riemer, J.M. Herrmann, H.E. Neuhaus, From endoplasmic reticulum to mitochondria: absence of the *Arabidopsis* ATP antiporter endoplasmic reticulum adenylate transporter 1 perturbs photorespiration, Plant Cell 25 (2013) 2647–2660.
[33] M. Eisenhut, S. Planchais, C. Cabassa, A. Guivarc'h, A.-M. Justin, L. Taconnat, J.-P. Renou, M. Linka, D. Gagneul, S. Timm, H. Bauwe, P. Carol, A.P.M. Weber, *Arabidopsis* A BOUT DE SOUFFLE is a putative mitochondrial transporter involved in photorespiratory metabolism and is required for meristem growth at ambient CO_2 levels, Plant J. 73 (2013) 836–849.
[34] T.R. Pick, A. Bräutigam, M.A. Schulz, T. Obata, A.R. Fernie, A.P.M. Weber, PLGG1, a plastidic glycolate glycerate transporter, is required for photorespiration and defines a unique class of metabolite transporters, Proc. Natl. Acad. Sci. U. S. A. 110 (2013) 3185–3190.
[35] P. Vauclare, D. Macherel, R. Douce, J. Bourguignon, The gene encoding T protein of the glycine decarboxylase complex involved in the mitochondrial step of the photorespiratory pathway in plants exhibits features of light-induced genes, Plant Mol. Biol. 37 (1998) 309–318.
[36] S. Engelmann, C. Wiludda, J. Burscheidt, U. Gowik, U. Schlue, M. Koczor, M. Streubel, R. Cossu, H. Bauwe, P. Westhoff, The gene for the P-subunit of glycine decarboxylase from the C4 species Flaveria trinervia: analysis of transcriptional control in transgenic *Flaveria bidentis* (C4) and Arabidopsis (C3), Plant Physiol. 146 (2008) 1773–1785.
[37] Y.Q. Hu, S. Liu, H.M. Yuan, J. Li, D.W. Yan, J.F. Zhang, Y.T. Lu, Functional comparison of catalase genes in the elimination of photorespiratory H_2O_2 using promoter- and 3'-untranslated region exchange experiments in the Arabidopsis *cat2* photorespiratory mutant, Plant Cell Environ. 33 (2010) 1656–1670.
[38] W. Adwy, M. Laxa, C. Peterhansel, A simple mechanism for the establishment of C_2-specific gene expression in *Brassicaceae*, Plant J. 84 (2015) 1231–1238.
[39] M. Laxa, K. Müller, N. Lange, L. Doering, J.T. Pruscha, C. Peterhänsel, The 5'UTR intron of *Arabidopsis* GGT1 aminotransferase enhances promoter activity by recruiting RNA polymerase II, Plant Physiol. 172 (2016) 313–327.
[40] M. Laxa, Regulatory cis-elements are located in accessible promoter regions of the *CAT2* promoter and affect activating histone modifications in *Arabidopsis thaliana*, Plant Mol. Biol. 93 (2017) 49–60.
[41] S. Timm, A. Florian, M. Wittmiß, K. Jahnke, M. Hagemann, A.R. Fernie, H. Bauwe, Serine acts as a metabolic signal for the transcriptional control of photorespiration-related genes in *Arabidopsis*, Plant Physiol. 162 (2013) 379–389.
[42] D.M. Goodstein, S. Shu, R. Howson, R. Neupane, R.D. Hayes, J. Fazo, T. Mitros, W. Dirks, U. Hellsten, N. Putnam, D.S. Rokhsar, Phytozome: a comparative platform for green plant genomics, Nucl. Acids Res. 40 (2012) D1178–D1186.
[43] R.A. Jefferson, T.A. Kavanagh, M.W. Bevan, GUS fusions: beta-glucuronidase as a sensitive and versatile gene fusion marker in higher plants, EMBO J. 6 (1987) 3901–3907.
[44] T. Lee, S. Yang, E. Kim, Y. Ko, S. Hwang, J. Shin, J.E. Shim, H. Shim, H. Kim, C. Kim, I. Lee, AraNet v2: an improved database of co-functional gene networks for the study of *Arabidopsis thaliana* and 27 other nonmodel plant species, Nucl. Acids Res. 43 (2015) D996–1002.
[45] T. Obayashi, K. Kinoshita, K. Nakai, M. Shibaoka, S. Hayashi, M. Saeki, D. Shibata, K. Saito, H. Ohta, ATTED-II: a database of co-expressed genes and *cis* elements for identifying co-regulated gene groups in *Arabidopsis*, Nucl. Acids Res. 35 (2007) D863–869.
[46] F. Van Bel, Coppens, Exploring plant co-expression and gene-gene interactions with CORNET 3.0, Methods Mol. Biol. 1533 (2017) 201–212.
[47] H. Wei, S. Persson, T. Mehta, V. Srinivasasainagendra, L. Chen, G.P. Page, C. Somerville, A. Loraine, Transcriptional coordination of the metabolic network in *Arabidopsis*, Plant Physiol. 142 (2006) 762–774.
[48] Steinhauser, B. Usadel, A. Luedemann, O. Thimm, J. Kopka, C.S.B. DB: a comprehensive systems-biology database, Bioinformatics 20 (2004) 3647–3651.
[49] D. Warde-Farley, S.L. Donaldson, O. Comes, K. Zuberi, R. Badrawi, P. Chao, M. Franz, C. Grouios, F. Kazi, C. Tannus Lopes, A. Maitland, S. Mostafavi, J. Montojo, Q. Shao, G. Wright, G.D. Bader, Q. Morris, The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function, Nucl. Acids Res. 38 (2010) W214–W220.
[50] M. Mutwil, S. Klie, T. Tohge, F.M. Giorgi, O. Wilkins, M.M. Campbell, A.R. Fernie, B. Usadel, Z. Nikoloski, S. Perssona, PlaNet: combined sequence and expression comparisons across plant networks derived from seven species, Plant Cell 23 (2011) 895–910.
[51] T.R. O'Connor, C. Dyreson, J.J. Wyrick, Athena: a resource for rapid visualization and systematic analysis of *Arabidopsis* promoter sequences, Bioinformatics 21 (2005) 4411–4413.
[52] B. Usadel, A. Nagel, O. Thimm, H. Redestig, O.E. Blaesing, N. Palacios-Rojas, J. Selbig, J. Hannemann, M. Conceição Piques, D. Steinhauser, W.-R. Scheible, Y. Gibon, R. Morcuende, D. Weicht, S. Meyer, M. Stitt, Extension of the visualization tool MapMan to allow statistical analysis of arrays, display of corresponding genes, and comparison with known responses, Plant Physiol. 138 (2005) 1195–1204.
[53] K.N. Chang, S. Zhong, M.T. Weirauch, G. Hon, M. Pelizzola, H. Li, S.S. Huang, R.J. Schmitz, M.A. Urich, D. Kuo, J.R. Nery, H. Qiao, A. Yang, A. Jamali, H. Chen, T. Ideker, B. Ren, Z. Bar-Joseph, T.R. Hughes, J.R. Ecker, Temporal transcriptional response to ethylene gas drives growth hormone cross-regulation in *Arabidopsis*, eLife 2 (2013) e00675.
[54] A.B. Rose, T. Elfersi, G. Parra, I. Korf, Promoter-proximal introns in *Arabidopsis thaliana* are enriched in dispersed signals that elevate gene expression, Plant Cell 20 (2008) 543–551.
[55] P. Schläpfer, P. Zhang, C. Wang, T. Kim, M. Banf, L. Chae, K. Dreher, A.K. Chavali, R. Nilo-Poyanco, T. Bernard, D. Kahn, S.Y. Rhee, Genome-wide prediction of metabolic enzymes, pathways, and gene clusters in plants, Plant Physiol. 173 (2017) 2041–2059.
[56] M. Liebthal, M. Strüve, X. Li, Y. Hertle, D. Maynard, T. Hellweg, A. Viehhauser, K.J. Dietz, Redox-dependent conformational dynamics of decameric 2-cysteine peroxiredoxin and its interaction with cyclophilin 20-3, Plant Cell Physiol. 57 (2016) 1415–1425.
[57] B. Usadel, T. Obayashi, M. Mutwil, F.M. Giorgi, G.W. Bassel, M. Tanimoto, A. Chow, D. Steinhauser, S. Persson, N.J. Provart, Co-expression tools for plant biology: opportunities for hypothesis generation and caveats, Plant Cell Environ. 32 (2009) 1633–1651.
[58] A.J. Bloom, Photorespiration and nitrate assimilation: a major intersection between plant carbon and nitrogen, Photosynth. Res. 123 (2015) 117–128.
[59] S. Rachmilevitch, A.B. Cousins, A.J. Bloom, Nitrate assimilation in plant shoots depends on photorespiration, Proc. Natl. Acad. Sci. U. S. A. 101 (2004) 11506–11510.
[60] C.M. Pérez-Delgado, M. García-Calderón, A.J. Márquez, M. Betti, Reassimilation of photorespiratory ammonium in *Lotus japonicus* plants deficient in plastidic glutamine synthetase, PLoS One 10 (2015) e0130438.
[61] G. Queval, E. Issakidis-Bourguet, F.A. Hoeberichts, M. Vandorpe, B. Gakière, H. Vanacker, M. Miginiac-Maslow, F. van Breusegem, G. Noctor, Conditional oxidative stress responses in the *Arabidopsis* photorespiratory mutant *cat2* demonstrate that redox state is a key modulator of day length-dependent gene expression, and define photoperiod as a crucial factor in the regulation of H_2O_2-induced cell death, Plant J. 52 (2007) 640–657.

[62] N.A. Anjum, P. Sharma, S.S. Gill, M. Hasanuzzaman, E.A. Khan, K. Kachhap, A.A. Mohamed, P. Thangavel, G.D. Devi, P. Vasudhevan, A. Sofo, N.A. Khan, A.N. Misra, A.S. Lukatkin, H.P. Singh, E. Pereira, N. Tuteja, Catalase and ascorbate peroxidase-representative H_2O_2-detoxifying heme enzymes in plants, Environ. Sci. Pollut. Res. 23 (2016) 19002–19029.

[63] H. Beevers, Microbodies in higher plants, Annu. Rev. Plant Physiol. 30 (1979) 159–193.

[64] S. Timm, M. Wittmiß, S. Gamlien, R. Ewald, A. Florian, M. Frank, M. Wirtz, R. Hell, A.R. Fernie, H. Bauwe, Mitochondrial dihydrolipoyl dehydrogenase activity shapes photosynthesis and photorespiration of *Arabidopsis thaliana*, Plant Cell 27 (2015) 1968–1984.

[65] S. Timm, J. Giese, N. Engel, M. Wittmiß, A. Florian, A.R. Fernie, H. Bauwe, T-protein is present in large excess over the other proteins of the glycine cleavage system in leaves of *Arabidopsis*, Planta 247 (2018) 41–51.

[66] R. Douce, J. Bourguignon, M. Neuburger, F. Rébeillé, The glycine decarboxylase system: a fascinating complex, Trends Plant Sci. 6 (2001) 167–176.

[67] A. Weber, U.I. Flügge, Interaction of cytosolic and plastidic nitrogen metabolism in plants, J. Exp. Bot. 53 (2002) 865–874.

[68] I. Pracharoenwattana, J.E. Cornah, S.M. Smith, Arabidopsis peroxisomal malate dehydrogenase functions in b-oxidation but not in the glyoxylate cycle, Plant J. 50 (2007) 381–390.

[69] B. Zybailov, H. Rutschow, G. Friso, A. Rudella, O. Emanuelsson, Q. Sun, K.J. van Wijk, Sorting signals, N-terminal modifications and abundance of the chloroplast proteome, PLoS One 3 (2008) e1994.

[70] U. Qi, C. Armbruster, E. Schmitz-Linneweber, A. Delannoy, T. Falcon de Longevialle, I. Rühle, P. Small, D. Jahns, Leister, *Arabidopsis* CSP41 proteins form multimeric complexes that bind and stabilize distinct plastid transcripts, J. Exp. Bot. 63 (2012) 1251–1270.

[71] D. Leister, Complex(iti)es of the ubiquitous RNA-binding CSP41 proteins, Front. Plant Sci. 5 (2014) 255.

[72] J. Fettke, A. Nunes-Nesi, A.R. Fernie, M. Steup, Identification of a novel heteroglycan-interacting protein, HIP1.3, from *Arabidopsis thaliana*, J. Plant Physiol. 168 (2011) 1415–1425.

[73] M.V. Beligni, S.P. Mayfield, *Arabidopsis thaliana* mutants reveal a role for CSP41a and CSP41b, two ribosome-associated endonucleases, in chloroplast ribosomal RNA metabolism, Plant Mol. Biol. 67 (2008) 389–401.

[74] T.J. Bollenbach, R.E. Sharwood, R. Gutierrez, S. Lerbs-Mache, D.B. Stern, The RNA-binding proteins CSP41a and CSP41b may regulate transcription and translation of chloroplast-encoded RNAs in *Arabidopsis*, Plant Mol. Biol. 69 (2009) 541–552.

[75] J. Mei, F. Li, X. Liu, G. Hu, Y. Fu, W. Liu, Newly identified CSP41b gene localized in chloroplasts affects leaf color in rice, Plant Sci. 256 (2017) 39–45.

[76] D. Baek, H.J. Chun, S. Kang, G. Shin, S.J. Park, H. Hong, C. Kim, D.H. Kim, S.Y. Lee, M.C. Kim, D.-J. Yun, A role for *Arabidopsis* miR399f in salt, drought, and ABA signaling, Mol. Cells 39 (2016) 111–118.

[77] E.P. Harrison, N.M. Willingham, J.C. Lloyd, C.A. Raines, Reduced sedoheptulose-1,7-bisphosphatase levels in transgenic tobacco lead to decreased photosynthetic capacity and altered carbohydrate accumulation, Planta 204 (1998) 27–36.

[78] C.A. Raines, E.P. Harrison, H. Olcer, J.C. Lloyd, Investigating the role of the thiol-regulated enzyme sedoheptulose-1, 7-bisphosphatase in the control of photosynthesis, Physiol. Plant. 110 (2000) 303–308.

[79] L.L. Feng, Y.J. Han, G. Liu, B.G. An, J. Yang, G.H. Yang, L. Yangsheng, Z. Yingguo, Overexpression of sedoheptulose-1,7-bisphosphatase enhances photosynthesis and growth under salt stress in transgenic rice plants, Funct. Plant Biol. 34 (2007) 822–834.

[80] L. Feng, K. Wang, Y. Li, Y. Tan, J. Kong, H. Li, Y. Li, Y. Zhu, Overexpression of SBPase enhances photosynthesis against high temperature stress in transgenic rice plants, Plant Cell Rep. 26 (2007) 1635–1646.

[81] F. Ding, M. Wang, S. Zhang, X. Ai, Changes in SBPase activity influence photosynthetic capacity, growth, and tolerance to chilling stress in transgenic tomato plants, Sci. Rep. 6 (2016) 32741.

[82] X.-L. Liu, H.-D. Yu, Y. Guan, J.-K. Li, F.-Q. Guo, Carbonylation and loss-of-function analyses of SBPase reveal its metabolic interface role in oxidative stress, carbon assimilation, and multiple aspects of growth and development in *Arabidopsis*, Mol. Plant 5 (2012) 1082–1099.

[83] K. Uematsu, N. Suzuki, T. Iwamae, M. Inui, H. Yukawa, Increased fructose 1,6-bisphosphate aldolase in plastids enhances growth and photosynthesis of tobacco plants, J. Exp. Bot. 63 (2012) 3001–3009.

[84] S. Kangasjärvi, J. Neukermans, S. Li, E.-M. Aro, G. Noctor, Photosynthesis, photorespiration, and light signalling in defence responses, J. Exp. Bot. 63 (2012) 1619–1636.

[85] Xu, J. Zhang, J. Zeng, L. Jiang, E. Liu, C. Peng, Z. He, X. Peng, Inducible antisense suppression of glycolate oxidase reveals its strong regulation over photosynthesis in rice, J. Exp. Bot. 60 (2009) 1799–1809.

[86] A. Timm, S. Florian, M. Arrivault, A.R. Stitt, H. Fernie, Bauwe, Glycine decarboxylase controls photosynthesis and plant growth, FEBS Lett. 586 (2012) 3692–3697.

[87] Y. Lu, Y. Lia, Q. Yangb, Z. Zhanga, Y. Chena, S. Zhangc, X.-X. Penga, Suppression of glycolate oxidase causes glyoxylate accumulation that inhibits photosynthesis through deactivating Rubisco in rice, Physiol. Plant. 150 (2014) 463–476.

[88] Y. Dellero, M. Lamothe-Sibold, M. Jossier, M. Hodges, *Arabidopsis thaliana ggt1* photorespiratory mutants maintain leaf carbon/nitrogen balance by reducing RuBisCO content and plant growth, Plant J. 83 (2015) 1005–1018.

[89] H.C. Lin, S. Karki, R.A. Coe, S. Bagha, R. Khoshravesh, C.P. Balahadia, J. Ver Sagun, R. Tapia, W.K. Israel, F. Montecillo, A. de Luna, F.R. Danila, A. Lazaro, C.M. Realubit, M.G. Acoba, T.L. Sage, S. von Caemmerer, R.T. Furbank, A.B. Cousins, J.M. Hibberd, W.P. Quick, S. Covshoff, Targeted knockdown of *GDCH* in rice leads to a photorespiratory-deficient phenotype useful as a building block for C4 rice, Plant Cell Physiol. 57 (2016) 919–932.

[90] F. Flügel, S. Timm, S. Arrivault, A. Florian, M. Stitt, A.R. Fernie, H. Bauwe, The photorespiratory metabolite 2-phosphoglycolate regulates photosynthesis and starch accumulation in *Arabidopsis*, Plant Cell 29 (2017) 2537–2551.

[91] C.M. Rojas, M. Senthil-Kumar, K. Wang, C.-M. Ryu, A. Kaundal, K.S. Mysore, Glycolate oxidase modulates reactive oxygen species-mediated signal transduction during nonhost resistance in *Nicotiana benthamiana* and *Arabidopsis*, Plant Cell 24 (2012) 336–352.

[92] B.M. Gilbert, T.J. Wolpert, Characterization of the LOV1-mediated, victorin-induced, cell-death response with virus-induced gene silencing, Mol. Plant Microbe Interact. 26 (2013) 903–917.

[93] M.C. Palmieri, C. Lindermayr, H. Bauwe, C. Steinhauser, J. Durner, Regulation of plant glycine decarboxylase by S-nitrosylation and glutathionylation, Plant Physiol. 152 (2010) 1514–1528.

[94] M. Zaffagnini, L. Michelet, C. Marchand, F. Sparla, P. Decottignies, P. Le Maréchal, M. Miginiac-Maslow, G. Noctor, P. Trostand, S.D. Lemaire, The thioredoxin-independent isoform of chloroplastic glyceraldehyde-3-phosphate dehydrogenase is selectively regulated by glutathionylation, FEBS J. 274 (2007) 212–226.

[95] A. Mhamdi, G. Queval, S. Chaouch, S. Vanderauwera, F. van Breusegem, G. Noctor, Catalase function in plants: a focus on *Arabidopsis* mutants as stress-mimic models, J. Exp. Bot. 61 (2010) 4197–4220.

[96] S.T. Smale, Core promoters: active contributors to combinatorial gene regulation, Genes Dev. 15 (2001) 2503–2508.

[97] M.C. Thomas, C.M. Chiang, The general transcription machinery and general cofactors, Crit. Rev. Biochem. Mol. Biol. 41 (2006) 105–178.

[98] T.I. Lee, R.A. Young, Transcription of eukaryotic protein-coding genes, Annu. Rev. Genet. 34 (2000) 77–137.

[99] J. Corden, B. Wasylyk, A. Buchwalder, P. Sassone-Corsi, C. Kedinger, P. Chambon, Promoter sequences of eukaryotic protein-coding genes, Science 209 (1980) 1406–1414.

[100] S.T. Smale, D. Baltimore, The "initiator" as a transcription control element, Cell. 57 (1989) 103–113.

[101] K. Kiran, S.A. Ansari, R. Srivastava, N. Lodhi, C.P. Chaturvedi, S.V. Sawant, R. Tuli, The TATA-box sequence in the basal promoter contributes to determining light-dependent gene expression in plants, Plant Physiol. 142 (2006) 364–376.

[102] R. Stracke, M. Werber, B. Weisshaar, The R2R3-MYB gene family in Arabidopsis thaliana, Curr. Opin. Plant Biol. 4 (2001) 447–456.

[103] H.D. Kranz, M. Denekamp, R. Greco, H. Jin, A. Leyva, R.C. Meissner, K. Petroni, A. Urzainqui, M. Bevan, C. Martin, S. Smeekens, C. Tonelli, J. Paz-Ares, B. Weisshaar, Towards functional characterisation of the members of the R2R3-MYB gene family from *Arabidopsis thaliana*, Plant J. 16 (1998) 263–276.

[104] S. Ambawat, P. Sharma, N.R. Yadav, R.C. Yadav, MYB transcription factor genes as regulators for plant responses: an overview, Physiol. Mol. Biol. Plants 19 (2013) 307–321.

[105] S. Roy, Function of MYB domain transcription factors in abiotic stress and epigenetic control of stress response in plant genome, Plant Signal. Behav. 11 (2016) e1117723.

[106] H. Jin, E. Cominelli, P. Bailey, A. Parr, F. Mehrtens, J. Jones, C. Tonelli, B. Weisshaar, C. Martin, Transcriptional repression by AtMYB4 controls production of UV-protecting sunscreens in *Arabidopsis*, EMBO J. 19 (2000) 6150–6161.

[107] D.W. Russell, E.E. Conn, The cinnamic acid 4-hydroxylase of pea seedlings, Arch. Biochem. Biophys. 122 (1967) 256–258.

[108] D.W. Russell, The metabolism of aromatic compounds in higer plants. X. Properties of the cinnamic acid 4-hydroxylase of pea seedlings and some aspects of its metabolic and developmental control, J. Biol. Chem. 246 (1971) 3870–3878.

[109] S.-S. Niu, C.-J. Xu, W.-S. Zhang, B. Zhang, X. Li, K. Lin-Wang, I.B. Ferguson, A.C. Allan, K.-S. Chen, Coordinated regulation of anthocyanin biosynthesis in Chinese bayberry (*Myrica rubra*) fruit by a R2R3 MYB transcription factor, Planta 231 (2010) 887–899.

[110] H. Du, B.R. Feng, S.S. Yang, Y.B. Huang, Y.X. Tang, The R2R3-MYB transcription factor gene family in maize, PLoS One 7 (2012) e37463.

[111] M. Feldbrugge, M. Sprenger, K. Hahlbrock, B. Weisshaar, PcMYB1 novel plant protein containing a DNA-binding domain with one Myb repeat, interacts in vivo with a light-regulatory promoter unit, Plant J. 11 (1997) 1079–1093.

[112] B.-J. Chen, Y. Wang, Y.L. Hu, Q. Wu, Z.P. Lin, Cloning and characterization of a drought inducible MYB gene from *Boea crassifolia*, Plant Sci. 168 (2005) 493–500.

[113] R.-K. Wang, Z.-H. Cao, Y.-J. Hao, Overexpression of a R2R3 MYB gene MdSIMYB1 increases tolerance to multiple stresses in transgenic tobacco and apples, Physiol. Plant. 150 (2014) 76–87.

[114] K. Sørhagen, M. Laxa, C. Peterhänsel, S. Reumann, The emerging role of photorespiration and non-photorespiratory peroxisomal metabolism in pathogen defence, Plant Biol. 15 (2013) 723–736.

[115] D. Schenke, C. Böttcher, D. Scheel, Crosstalk between abiotic ultraviolet-B stress and biotic (flg22) stress signalling in *Arabidopsis* prevents flavonol accumulation in favor of pathogen defence compound production, Plant Cell Environ. 34 (2011) 1849–1864.

[116] G.I. Jenkins, Signal transduction in responses to UV-B radiation, Annu. Rev. Plant Biol. 60 (2009) 407–431.

[117] E. Logemann, K. Hahlbrock, Crosstalk among stress responses in plants: pathogen defense overrides UV protection through an inversely regulated ACE/ACE type of light-responsive gene promoter unit, Proc. Natl. Acad. Sci. U. S. A. 99 (2002) 2428–2432.

[118] R. Stracke, H. Ishihara, G. Huep, A. Barsch, F. Mehrtens, K. Niehaus, B. Weisshaar, Differential regulation of closely related R2R3-MYB transcription factors controls

flavonol accumulation in different parts of the *Arabidopsis thaliana* seedling, Plant J. 50 (2007) 660–677.
[119] B. Ülker, I.E. Somssich, WRKY transcription factors: from DNA binding towards biological function, Curr. Opin. Plant Biol. 7 (2004) 491–498.
[120] E. Gendra, D.F. Colgan, B. Meany, M.M. Konarska, A sequence motif in the simian virus 40 (SV40) early core promoter affects alternative splicing of transcribed mRNA, J. Biol. Chem. 282 (2007) 11648–11657.
[121] A.K. Srivastava, Y. Lu, G. Zinta, Z. Lang, J.-K. Zhu, UTR-dependent control of gene expression in plants, Trends Plant Sci. 23 (2018) 248–259.
[122] C. Hornyik, C. Duc, K. Rataj, L.C. Terzi, G.G. Simpson, Alternative polyadenylation of antisense RNAs and flowering time control, Biochem. Soc. Trans. 38 (2010) 1077–1081.
[123] R. Macknight, M. Duroux, R. Laurie, P. Dijkwel, G. Simpson, C. Dean, Functional significance of the alternative transcript processing of the *Arabidopsis* floral promoter FCA, Plant Cell 14 (2002) 877–888.
[124] W. Tang, S.E. Perry, Binding site selection for the plant MADS domain protein AGL15: an in vitro and *in vivo* study, J. Biol. Chem. 278 (2003) 28154–28159.
[125] G.R. Heck, S.E. Perry, K.W. Nichols, D.E. Fernandez, AGL15, a MADS domain protein expressed in developing embryos, Plant Cell 7 (1995) 1271–1282.
[126] M. Hayashi, K. Toriyama, M. Kondo, A. Kato, S. Mano, L. de Bellis, Y. Hayashi-Ishimaru, K. Yamaguchi, H. Hayashi, M. Nishimura, Functional transformation of plant peroxisomes, Cell Biochem. Biophys. 32 (2000) 295–304.
[127] D.E. Titus, W.M. Becker, Investigation of the glyoxysome-peroxisome transition in germinating cucumber cotyledons using double-label immunoelectron microscopy, J. Cell Biol. 101 (1985) 1288–1299.
[128] M. Nishimura, J. Yamaguchi, H. Mori, T. Akazawa, S. Yokota, Immunocytochemical analysis shows that glyoxysomes are directly transformed to leaf peroxisomes during greening of pumpkin cotyledons, Plant Physiol. 81 (1986) 313–316.
[129] M. Nishimura, Y. Takeuchi, L. De Bellis, I. Hara-Nishimura, Leaf peroxisomes are directly transformed to glyoxysomes during senescence of pumpkin cotyledons, Protoplasma 175 (1993) 131–137.
[130] M. Ogawa, A. Hanada, Y. Yamauchi, A. Kuwahara, Y. Kamiya, S. Yamaguchi, Gibberellin biosynthesis and response during *Arabidopsis* seed germination, Plant Cell 15 (2003) 1591–1604.
[131] M. Koornneef, J.H. van der Veen, Induction and analysis of gibberellin sensitive mutants in *Arabidopsis thaliana* (L.) Heynh, Theor. Appl. Genet. 58 (1980) 257–263.
[132] Y.M. Jeong, J.H. Mun, I. Lee, J.C. Woo, C.B. Hong, S.G. Kim, Distinct roles of the first introns on the expression of *Arabidopsis* profilin gene family members, Plant Physiol. 140 (2006) 196–209.
[133] R.B. Meagher, E.C. McKinney, A.V. Vitale, The evolution of new structures: clues from plant cytoskeletal genes, Trends Genet. 15 (1999) 278–284.
[134] A. Vitale, R.J. Wu, Z. Cheng, R.B. Meagher, Multiple conserved 5′ elements are required for high-level pollen expression of the *Arabidopsis* reproductive actin ACT1, Plant Mol. Biol. 52 (2003) 1135–1151.
[135] M. Laxa, Intron-mediated enhancement: a tool for heterologous gene expression in plants? Front. Plant Sci. 7 (2017) 1977.
[136] J.E. Gallegos, A.B. Rose, Intron DNA sequences can be more important than the proximal promoter in determining the site of transcript initiation, Plant Cell 29 (2017) 843–853.
[137] S. Huang, Y.Q. An, J.M. McDowell, E.C. McKinney, R.B. Meagher, The Arabidopsis ACT11 actin gene is strongly expressed in tissues of the emerging inflorescence, pollen, and developing ovules, Plant Mol. Biol. 33 (1997) 125–139.
[138] A.B. Rose, J.A. Beliakoff, Intron-mediated enhancement of gene expression independent of unique intron sequences and splicing, Plant Physiol. 122 (2000) 535–542.
[139] M. Clancy, L.C. Hannah, Splicing of the maize *Sh1* first intron is essential for enhancement of gene expression, and a T-rich motif increases expression without affecting splicing, Plant Physiol. 130 (2000) 918–929.
[140] J.H. Mun, S.Y. Lee, H.J. Yu, Y.M. Jeong, M.Y. Shin, H. Kim, I. Lee, S.G. Kim, Petunia actin-depolymerizing factor is mainly accumulated in vascular tissue and its gene expression is enhanced by the first intron, Gene 292 (2002) 233–243.
[141] Y.M. Jeong, J.H. Mun, H. Kim, S.Y. Lee, S.G. Kim, An upstream region in the first intron of petunia actin-depolymerizing factor 1 affects tissue-specific expression in transgenic *Arabidopsis* (*Arabidopsis thaliana*), Plant J. 50 (2007) 230–239.
[142] G. Parra, K. Bradnam, A.B. Rose, I. Korf, Comparative and functional analysis of intron-mediated enhancement signals reveals conserved features among plants, Nucl. Acids Res. 39 (2011) 5328–5337.
[143] D. Winter, B. Vinegar, H. Nahal, R. Ammar, G.V. Wilson, N.J. Provart, An "Electronic Fluorescent Pictograph" browser for exploring and analyzing large-scale biological data sets, PLoS One 2 (2007) e718.
[144] A.H. Liepman, L.J. Olsen, Alanine aminotransferase homologs catalyze the glutamate:glyoxylate aminotransferase reaction in peroxisomes of *Arabidopsis*, Plant Physiol. 131 (2003) 215–227.
[145] M. Kendziorek, A. Paszkowski, B. Zagdańska, Differential regulation of alanine aminotransferase homologues by abiotic stresses in wheat (*Triticum aestivum* L.) seedlings, Plant Cell Rep. 31 (2012) 1105–1111.
[146] D. Taler, M. Galperin, I. Benjamin, Y. Cohen, D. Kenigsbuch, Genes that encode photorespiratory enzymes confer resistance against disease, Plant Cell 16 (2004) 172–184.
[147] S. Barak, A. Nejidat, Y. Heimer, M. Volokita, Transcriptional and posttranscriptional regulation of the glycolate oxidase gene in tobacco seedlings, Plant Mol. Biol. 45 (2001) 399–407.
[148] J.-J. van Oosten, D.W. Wilkins, R.T. Besford, Regulation of the expression of photosynthetic nuclear genes by CO_2 is mimicked by regulation by carbohydrates: a mechanism for the acclimation of photosynthesis to high CO_2? Plant Cell Environ. 17 (1994) 913–923.
[149] J.-J. van Oosten, D. Afif, P. Dizengremel, Long-term effects of a CO_2 enriched atmosphere on enzymes of the primary carbon metabolism of spruce trees, Plant Physiol. Biochem. 30 (1992) 541–547.