Phototropin links blue-light perception with starch accumulation in Chlamydomonas

Yizhong Yuan

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ GRENOBLE ALPES

Spécialité : Biologie Végétale

Arrêté ministériel : 25 mai 2016

Présentée par

Yizhong YUAN

Thèse dirigée par **Dimitris PETROUTSOS**
et co-encadrée par **Olivier BASTIEN**

préparée au sein du **Laboratoire LPCV - Laboratoire de Physiologie Cellulaire Végétale**
dans **l'École Doctorale Chimie et Sciences du Vivant**

Phototropine lie la perception de la lumière bleue et l'accumulation d'amidon chez Chlamydomonas

Phototropin links blue-light perception and starch accumulation in Chlamydomonas

Thèse soutenue publiquement le **29 juin 2022**,
devant le jury composé de :

Madame Angela Falciatore
DIRECTEUR DE RECHERCHE, CNRS - Sorbonne Université, UMR 7141, Examinatrice
Monsieur Christian Fankhauser
PROFESSEUR DES UNIVERSITES, University of Lausanne, Rapporteur
Monsieur Jean Alric
DIRECTEUR DE RECHERCHE, CNRS - Aix-Marseille University - CEA, UMR7265, Rapporteur
Madame Yonghua Li-Beisson
DIRECTEUR DE RECHERCHE, CEA - Aix-Marseille University, UMR 7265, Examinatrice
Madame Emilie Demarsy
MAITRE DE CONFERENCE, University of Geneva, Examinatrice
Monsieur Johannes Geiselmann
PROFESSEUR DES UNIVERSITES, University Grenoble Alpes, Président
Monsieur Arthur Grossman
PROFESSEUR DES UNIVERSITES, The Carnegie Institution for Science-Stanford University, Examinateur

Contents

Résumé

Les organismes photosynthétiques convertissent l'énergie lumineuse solaire en énergie chimique sous forme d'ATP et de NADPH pour alimenter la fixation du CO_2 en carbone organique par photosynthèse. Une partie du carbone organique se trouve sous forme d'amidon, l'une des principales formes de réserve d'hydrates de carbone chez les plantes et les algues. Outre son rôle évident de source d'énergie, la lumière joue également un rôle de signal informationnel, qui est perçu par les protéines photoréceptrices régulant un large éventail de fonctions cellulaires. Il a été démontré que la qualité de la lumière a un impact sur la répartition du carbone, la composition de la biomasse et le taux de photosynthèse chez différentes espèces de plantes et d'algues. Plus précisément, les algues vertes ont une teneur en amidon plus faible dans des conditions de lumière bleue que dans des conditions de lumière rouge ou blanche. Cependant, les mécanismes moléculaires de cette régulation restent mal connus.

Dans cette thèse, nous avons étudié le rôle du photorécepteur-phototropine (PHOT) de la lumière bleue dans le métabolisme de l'amidon en utilisant l'organisme modèle d'algues vertes *Chlamydomonas reinhardtii*. La délétion de PHOT conduit à une teneur en amidon plus élevée. Nous avons également découvert que l'isoforme GAP1 de la glycéraldéhyde 3-phosphate déshydrogénase (GAPDH) est un acteur essentiel de ce phénotype. Au total, nous avons trouvé trois différentes voies de signalisation en cascade reliant PHOT au métabolisme de l'amidon chez *Chlamydomonas reinhardtii*, résumées ici comme suit :

Voie 1 : La lumière bleue, via PHOT, réprime le facteur de transcription (TF) GAPR4 contenant le domaine bHLH, un activateur de GAP1. Le mutant PHOT knock-out *phot* et les lignées surexprimant GAP1 accumulent de grandes quantités d'amidon.

Voie 2 : La lumière bleue, via PHOT, active le TF RWP5 (appartenant à la famille des TFs RWP-RK), un répresseur de l'accumulation d'amidon indépendant de GAP1.

Voie 3 : La phosphorylation PHOT-dépendante d'une sérine/thréonine kinase que nous avons nommée Blue Starch Kinase 1 (BSK1 ; orthologue de la kinase HT1 d'Arabidopsis qui contrôle les mouvements stomatiques en réponse au CO_2) agit sur les niveaux de GAP1 et d'accumulation d'amidon indépendamment de GAPR4.

Notre travail a fourni une réponse complète à la façon dont la qualité de la lumière affecte la répartition du carbone chez l'algue verte modèle *Chlamydomonas reinhardtii*. Dans un contexte plus général, ce travail améliore la compréhension actuelle de la façon dont la signalisation lumineuse contrôle le métabolisme chez les microalgues. Il ajoute également un aspect de plus au rôle multiple de PHOT rapporté jusqu'à présent pour contrôler la gamétogenèse en milieu appauvri en azote, l'expression des gènes codant pour la biosynthèse de la chlorophylle et des caroténoïdes, la taille du tache oculaire et la photoprotection.

Abstract

Photosynthetic organisms convert solar light energy into chemical energy in the form of ATP and NADPH to fuel CO_2 fixation into organic carbon through photosynthesis. Part of the organic carbon is in the form of starch, one of the major carbohydrate reserve forms in plants and algae. Besides its evident role as energy source, light serves also a role of informational signal, that is perceived by photoreceptor proteins regulating a wide range of cellular functions. Light quality has been shown to impact carbon partitioning, biomass composition and photosynthesis rate in different plant and algae species. More specifically, green algae have a lower starch content under blue light conditions than red or white light conditions. However, the molecular mechanisms of this regulation are still unclear.

In this thesis, I investigated the role of blue light photoreceptor-phototropin (PHOT) in starch metabolism by using the green algae model organism *Chlamydomonas reinhardtii*. The deletion of PHOT leads to higher starch content. We further found that the GAP1 isoform of glyceraldehyde 3-phosphate dehydrogenase (GAPDH) is an essential player in this phenotype. Overall, we found three different signaling cascade pathways linking the PHOT with starch metabolism in *Chlamydomonas reinhardtii*, summarized here as follows:

Pathway 1: Blue light, via PHOT, represses the bHLH domain-containing transcription factor (TF) GAPR4, an activator of GAP1. Both the *PHOT* knock-out mutant *phot* and *GAP1* over-expressing lines accumulate high amounts of starch.

Pathway 2: Blue light, via PHOT, activates the RWP5 TF (belonging to the RWP-RK family of TFs), a repressor of starch accumulation in a *GAP1*-independent manner.

Pathway 3: PHOT-dependent phosphorylation of a serine/threonine kinase we provisionally name Blue Starch Kinase 1 (BSK1; orthologue of the Arabidopsis HT1 kinase that controls stomatal movements in response to CO_2) mediates GAP1 and starch accumulation levels in a GAPR4-independent manner.

Our work provided a comprehensive answer to how light quality affects the carbon partitioning in the model green alga *Chlamydomonas reinhardtii*. In a more general context, this work advances the current understanding of how light signaling controls

metabolism in microalgae. It also adds one more aspect to the multifaceted role of PHOT so far reported to control gametogenesis at low nitrogen, expression of genes encoding chlorophyll and carotenoid biosynthesis, the size of the eyespot and photoprotection.

Acknowledgments

First, I would like to thank Dr. Christian Fankhauser and Dr. Jean Alric for accepting to review my thesis. And also many thanks to my other jury members: Dr. Angela Falciatore, Dr. Arthur Grossman, Dr. Emilie Demarsy, Dr. Johannes Geiselmann and Dr. Yonghua Li-Beisson. Thank you all for taking the time to evaluate my work.

I want to express my deepest gratitude to my supervisor Dr. Dimitris Petroutsos. Without his help, I could not start and finish my Ph.D. study. During my Ph.D., he gave me the freedom to investigate and try different 'crazy' ideas. His patient guidance, daily care, and professional knowledge helped me tremendously to become an independent researcher.

I would also thank my co-supervisor Dr. Olivier Bastien for all the support in bioinformatics during my Ph.D. and Dr. Denis Falconet for the help in electron microscopy.

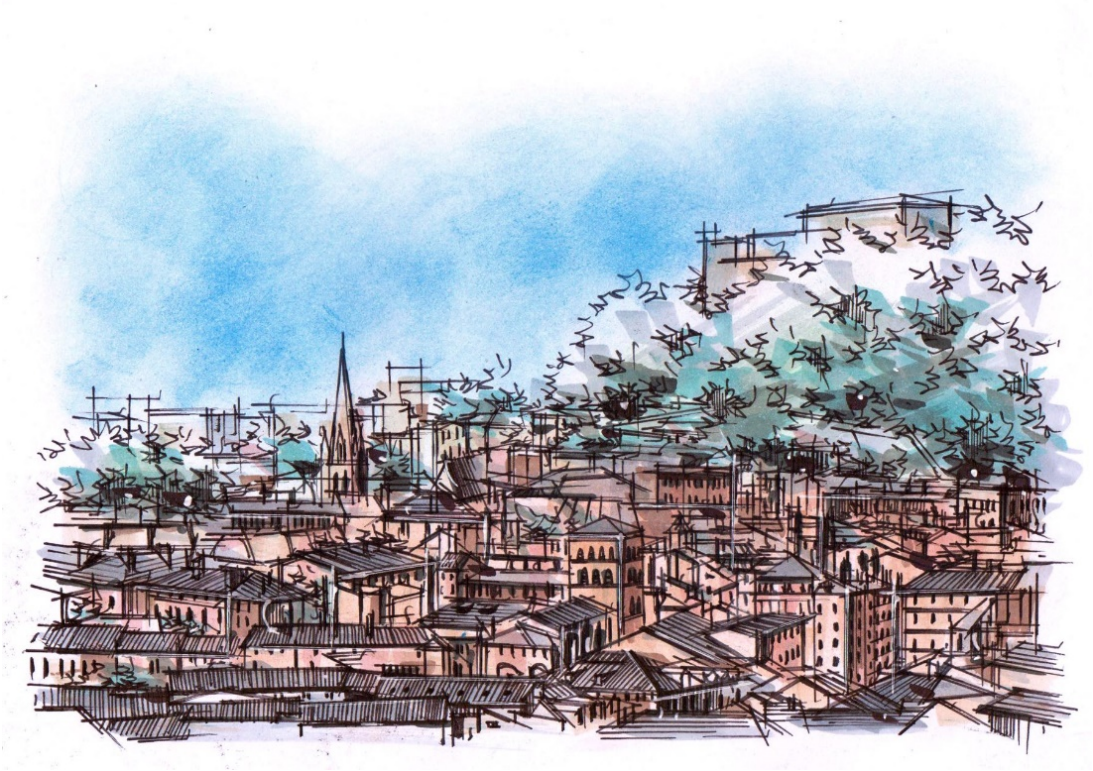
I also deeply appreciate all my CSI committee members: Dr. Olivier Vallon, Dr. Stephane Ravanel and Dr. Yonghua Li-Beisson (again). Their advice and help at the annual CSI meeting not only enriched the content of my thesis but also helped me to build my confidence in scientific research.

I would also like to thank all the present and previous lab members, especially to Anne-Flore Deton-Cabanillas (Big thanks for the correction of Résumé), Gaelle Villain, Georgios Kepesidis, Marcello De Mia, Dr. Xuelei Lai, Dr. Xue Zhao. Also, big thanks to all the collaborators, especially to Dr. Zoran Nikoloski, Marius Arend, Fabrizio Iacono, and Dr. Petra Redekop.

Million thanks to all my best best best friends in my daily life for their encouragement and accompany, especially after the loss of my family. It is a super tough period for me. Without their help and support, I would have already crashed.

Finally, I wish all the people I know stay in good health and live happily.

World peace!



(Marker painting by Yizhong-Grenoble, France, 15th December 2019)

Chapter 1

1.1 General Introduction

1.1.1 The model organism *Chlamydomonas reinhardtii*

Microalgae are one of the most diverse eukaryotic organisms occupying broad ecological niches. By converting atmospheric CO_2 to organic sugars, they provide much energy to many heterotrophs, including humans, and participate actively in the global carbon cycle, playing a crucial role in different life forms surviving on earth.

Chlamydomonas reinhardtii is a single-celled organism belongs to the *Chlamydomonas* genus which was named by C.G. Ehrenberg[1] in 1833. It is a green microalga with a large cup-shaped chloroplast, flagella, and eyespot. Because it can grow photoautotrophically or heterotrophically on acetate, mutations affecting genes for photosynthesis are not lethal but conditional, requiring acetate for growth[2]. This provides an important and convenient opportunity in the analysis of photosynthesis. As *Chlamydomonas reinhardtii* is vegetatively haploid, the effects of mutations are immediate without further crossing. *Chlamydomonas reinhardtii* has three genomes - nuclear[3], mitochondrial[4], and chloroplast genome[5], which are sequenced and subject to specific transformation. Many genetic tools have been developed for *Chlamydomonas reinhardtii*[6-11], such as zinc finger nuclease (ZFN)[12,13], transcription activator-like effector nuclease (TALEN)[14,15], and CRISPR/Cas systems[13,16]. Additionally, a genome-wide, indexed, and insertion mapped *Chlamydomonas reinhardtii* mutant library (CLiP) containing ca. 58,101 mutants affected in 14,695 genes is available in *Chlamydomonas* resource center[17,18].

Combining all the advantages, *Chlamydomonas reinhardtii has* become a popular and prominent model organism for different cellular processes research, such as photosynthesis[19-22], flagellar assembly[23], phototaxis[24-26], mitochondrial respiration[27], chloroplast biogenesis[28,29], metabolism[30-32], stress response[33-36], cell cycle[37-39], sexual life cycle[40-42].

1.1.2 Photosynthesis

Most life on earth depends on photosynthesis. Photosynthesis is carried out by

diverse plants, algae, and some bacteria[43]. It can convert light energy into chemical energy. Photosynthetic organisms can convert CO_2 and water into O_2 and carbohydrates, such that

$$6\ CO_2 + 6\ H_2O \rightarrow C_6H_{12}O_6 + 6\ O_2$$

It occurs in two integrated processes:1) the photochemical light reactions and 2 the CO_2 fixation reactions.

1.1.2.1 Photochemical light reactions

The photochemical light reactions use solar energy to extract electrons from water to produce reducing equivalent (NADPH). It requires several large multi-protein complexes, including PSI and PSII (light-harvesting photosystems), the cytochrome b_6f, and ATP synthase. Light reactions usually occur in thylakoid membranes, except for *Gloeobacter violaceus* - the only cyanobacterium without thylakoid membranes[44]. Photosynthesis in *G. violaceus* occurs in the cytoplasmic membrane[44].

When pigments in LHCs (light-harvesting complexes) absorb light, electrons become excited to a higher energy state than can be passed on to another molecule.

Electrons flow from water will go to the mobile electron carrier-plastoquinone (PQ) through PSII, and then go through cytochrome b6f, then to PSI by another mobile electron carrier-plastocyanin (PC), then via ferredoxin to reduce $NADP^+$ to NADPH. The linear electron transfer (LET) also contributes to a proton gradient to produce phosphorylating power (ATP) via the chloroplastic ATP synthase. The chemical energy produced in the form of ATP and NADPH is subsequently used to drive CO_2 assimilation.

1.1.2.2 Excess light damage avoidance

When chlorophylls absorb light, it will be excited from ground state (Chl) to singlet excited state (^{1}Chl*). In order to relax back to the ground state, it has three ways: 1) it can be emitted as fluorescence, 2) its excitation can be used to fuel photochemistry and 3) it can be dissipated by non-photochemical quenching (NPQ)[45]. The three ways can reduce the amount of fluorescence. The fate of captured light energy depends on temperature, nutrient availability, and many other factors. When captured light energy

exceeds the photosynthetic capacity, it may lead to photoinhibition and ROS (Reactive oxygen species) production, which can cause damage to photosystem components and even lead to cell death.

However, photosynthetic organisms have evolved diverse ways to cope with the absorption of excess light[46,47]. Plants can dynamically change the size of the antennae complex to decrease the light incidence. Higher plants can also move their chloroplasts called chloroplast positioning[44], while algae with flagella swim to move into optimal light conditions[48] a process called phototaxis.

Moreover, at the same time, plants can produce antioxidants to detoxify ROS, such as carotenoids, lutein, zeaxanthin, and violaxanthin[49,50]. Another photoprotective response is NPQ which encompasses three components: qE (Energy-dependent quenching), qT (State transition quenching), and qI (slow photoinhibitory quenching)[51]. The qE response is the dominant photoprotective response in *Chlamydomonas* and will be analyzed further in the next paragraph.

1.1.2.3 Energy-dependent quenching-qE

In the higher plants model organism *Arabidopsis thaliana*, qE activation needs de-epoxidized xanthophylls, a high pH gradient across the thylakoid membrane, and LHCs.

The de-epoxidized xanthophylls are produced by the xanthophyll cycle. The *npq1* mutant lacking the VDE (violaxanthin de-epoxidase) gene has reduced NPQ capacity than WT. And the *pgr1* and *pgr5* mutants, lacking *PROTON GRADIENT REGULATOR-LIKE 1* and *PROTON GRADIENT REGULATOR 5* respectively, which could not sustain a high pH gradient in the thylakoid lumen, have lower qE than WT. The high pH gradient also activates the xanthophyll cycle due to the thylakoid lumen located VDE enzyme. Moreover, a high pH gradient also protonates the LHC proteins. PSBS, one of the LHC proteins, plays a crucial role in qE in *Arabidopsis* thaliana. The *npq4* mutant lacking the PSBS gene shows a complete abolishment of qE phenotype[52] resulting in less seed set when placed in the natural environment[53].

In *Chlamydomonas reinhardtii,* qE components are similar to those in plants:

xanthophyll cycle, pH gradient, and specific proteins.

However, in *Chlamydomonas,* there is no VDE or ortholog of VDE in its genome[54]. The double mutant *npq1 lor1*, which lacks zeaxanthin and lutein, shows a severe lack of qE[55]. This indicated that lutein and zeaxanthin might play an essential role in *Chlamydomonas* quenching.

In *Chlamydomonas*, the critical photoprotective proteins of qE are LHCSR3 (product of the *LHCSR3.1* and *LHCSR3.2* genes)[56], LHCSR1[57] (product of the *LHCSR1* gene), and PSBS (product of the *PSBS1* and *PSBS2* genes)[58,59]. Those three proteins will accumulate in response to high light[55,57,59-62]. LHCSR3 plays a dominant role in *Chlamydomonas* qE. The *npq4* mutant lacking the *LHCSR3.*1 and *LHCSR3.2* gene showed a severe loss of qE and cannot survive under high light conditions[56]. When LHCSR3 is absent from the cell, LHCSR1 could also play a role in inducing qE[57,63,64]. While PSBS plays a crucial role in *Arabidopsis* qE, it seems less critical in *Chlamydomonas* qE[59] and its exact function needs further investigation[59,62]. In recent years, the expression of all photoprotective genes has been found to be controlled by photoreceptor proteins; the blue light receptor phototropin (PHOT) was initially found to control *LHCSR3.1* and *LHCSR3.2*[19], and it was shown later that PHOT also controls expression of *LHCSR1* and *PSBS1/2*[62,65]. UV-B radiation via the UVR8 mainly controls *LHCSR1* and *PSBS1/2* and to a lesser extent LHCSR3[62,66,67].

1.1.2.4 CO_2 fixation

The energy and reducing power from ATP and NADPH generated in photochemical light reactions will be transferred to CO_2 to produce high-energy, reduced sugars.

There are two types of photosynthesis to fix CO_2: C3 and C4 photosynthesis. In C3 photosynthesis, CO_2 reacts with RuBP (D-ribulose-1,5-bisphosphate) to produce two 3-carbon compounds 3-PGA (3-phosphoglyceric acid) via Calvin-Benson-Bassham cycle (CBB cycle). However, in C4 photosynthesis, CO_2 is first fixed to form a 4-carbon compound OAA (oxaloacetic acid), which is quickly converted to malic acid. The 4-carbon acid will be split into a 3-carbon compound and CO_2 that will enter the CBB cycle.

In *Chlamydomonas reinhardtii,* only C3 photosynthesis has been found, while C4 photosynthesis has been found in some marine algae[68-70]. No matter which type of photosynthesis, global outputs of photosynthesis are triose phosphate (i.e., glyceraldehyde 3-phosphate (G3P)), NADPH, and ATP, which constitute the primary metabolic intermediates in a photosynthetic cell.

The CBB cycle includes three stages. The first stage is the carboxylation step. In this step, CO_2 binds to RuBP and produce 3-PGA; The second stage is the reduction stage. Using the energy produced in light reactions, 3-PGA is phosphorylated and reduced to G3P and dihydroxyacetone phosphate. These triose phosphates exit the CBB cycle, becoming the precursors for starch synthesis, or are modified further by the CBB cycle and exit the CBB cycle as the precursors in the shikimate pathway, or stay in the CBB cycle to enter the third stage of the CBB cycle to regenerate the RuBP.

1.1.2.5 Glyceraldehyde-3-phosphate dehydrogenase (GAPDH)

In the second stage of the CBB cycle, the fate of G3P plays a crucial role in carbon portioning in the cell. GAPDH (Glyceraldehyde-3-phosphate dehydrogenase) is one of the critical enzymes to catalyze the reversible reaction of 1,3-bisphosphoglycerate (BGPA) to G3P. GAPDHs are highly conserved proteins in different organisms and play central roles in carbon metabolism.

Plants contain several isoforms of GAPDH.

In *Arabidopsis thaliana*, There are four isoforms of GAPDH: GAPA/B, GAPC, GAPCp, and the NADP-dependent non-phosphorylating cytosolic GAPDH (GAPN)[71]. However, GAPN belongs to the aldehyde dehydrogenase superfamily and has no close functional/structural relationship with other phosphorylating GAPDHs[72].

GAPA/B (encoded by *GAPa*1; AT3G26650, *GAPa2*; AT1G12900, and *GAPb*; AT1G42970) is the chloroplastic GAPDH playing a central role in the Calvin-Benson cycle of CO_2 assimilation[73,74]. GAPC (encoded by *GAPc1*; AT3G04120 and *GAPc2*; AT1G13440) is the cytosolic glycolytic GAPDH and plays a role in glycolysis. The GAPC1 knockout mutant showed delayed growth, seed and fruit development alteration phenotypes in *Arabidopsis*[75]. GAPCps (encoded by *GAPCp1;* AT1G79530

and *GAPCp2;* AT1G16300) are mainly localized in plastids[76,77], while GAPCp1 is also found in cell wall[78]. GAPCps deficiency affects amino acid and sugar metabolism and impairs plant development[79].

Chlamydomonas reinhardtii contains four distinct isoforms: three phosphorylating and one non-phosphorylating GAPDH. These include: GAP1 (Cre12.g485150), GAP2 (Cre07.g354200), GAP3 (Cre01.g010900) and the NADP-dependent non-phosphorylating cytosolic GAPDH (GAPN1, Cre12.g556600). However, as in *Arabidopsis*, GAPN1 has no close functional/structural relationship with other phosphorylating GAPDH in *Chlamydomonas*.

GAP1 is a chloroplastic GAPDH according to PredAlgo prediction[80] and as was later experimentally validated by a fluorescent protein-tagging approach[81]. Its transcription level is highly induced in nitrogen or sulfur depletion conditions[35] Although the exact function of GAP1 is still unknown, GAP1 showed interaction with FDX1, FDX2 and FDX5 (ferredoxin) and has been shown to be regulated post-translationally by FDX5[82,83].

GAP2 is the only cytosolic located GAPDH is *Chlamydomonas* which has been believed plays a role in the glycolysis. However, its exact function has not been experimentally confirmed.

GAP3 is located in the chloroplast in *Chlamydomonas* and functions like GAPA/B of *Arabidopsis*. GAP3 plays a central role in the CBB cycle[84], and its activity could be regulated by interaction with the highly unfolded and flexible protein CP12 and the phosphoribulokinase (PRK)[85-87].

1.1.3 Starch metabolism in *Chlamydomonas*

The output of the CBB cycle could be used for the alternative carbohydrates metabolism[88,89]. In higher plants and green algae, carbohydrates are formed in the chloroplast, while in red algae, glaucophyte, and dinoflagellate they are formed in the cytosol[90]. Furthermore, different species have different carbohydrate reserve forms. Green algae and higher plants synthesize amylose (long unbranched chain of D-glucose connected by α 1-4 linkages and very few α 1-6 linkages) and amylopectin (branched chain of D-glucose with both α 1-4 and 1-6 linkages) with ADP-glucose as

the donor[91], while red algae synthesize floridean starch (a hybrid of glycogen and starch) with UDP-glucose as the donor[92] and cyanobacteria synthesis glycogen (α 1-4 linked glucan)[93]. Phaeophyceae (brown algae) and Bacillariophyceae (diatom) synthesis glycogen in both β 1-3 and 1-6 linkages form[94].

1.1.3.1 Starch Synthesis in *Chlamydomonas*

Starch is the primary carbon sink and valuable storage polysaccharide in all the eukaryotic microalgae, although the localization and biosynthesis pathway vary among different phyla. Amylose comprises 10 to 30% of starch in a cell and is synthesized by granule-bound starch synthase (GBSSI)[95]. However, amylopectin comprises 70 to 90% of starch in a cell, and its synthesis requires more enzymatic processing than amylose[96,97].

In *Chlamydomonas reinhardtii*, starch accumulates in the chloroplast. Some closely packed starch granules are found surrounding the pyrenoid called starch sheath[98]. The starch sheath has been proved to be essential for pyrenoid function and CCM[99,100]. It prevents the leakage of CO_2 in low-CO_2 conditions from pyrenoid and minimizes oxygen exposure in the pyrenoid matrix[101].

Fructose 6-phosphate (F6P), the product of the reaction catalyzed by fructose 1,6-bisphosphatase (FBPase), is the branch point for metabolites leaving the CBB cycle and initiating starch synthesis[102]. Next, F6P is converted into glucose 6-phosphate (G6P) by phosphoglucose isomerase (PGI1). G6P is converted to glucose-1-phosphate (G1P) by phosphoglucomutase (PGM). The reaction of G1P with ATP leads to the formation of the ADP-glucose and pyrophosphate catalyzed by ADP-glucose pyrophosphorylase (AGPase). This step is one of the rate-limiting steps of the starch synthesis pathway (Fig. 1.1). After the production of ADP-glucose, starch is synthesized by three classes of enzymes: Starch synthases (SSS), starch debranching enzymes (DBE), and starch branching enzymes (BE).

1.1.3.2 Starch Degradation in *Chlamydomonas*

Starch is usually degraded in the dark to provide energy and substrates for different cellular activities. However, starch catabolism is not that clear in *Chlamydomonas.* There are two ways to degrade starch in green algae: the hydrolytic and the

phosphorolytic pathway[96,103-105].

Hydrolytic starch degradation generates maltose, mainly mediated by β-amylases (βAMY). This also leads to the generation of malto-oligosaccharides (MOS) as by-products which are mainly mediated by α-amylases (αAMY) and isoamylase (ISA). MOS could be modified again by βAMY or processed by disproportionating enzyme1 (DPE1).

Phosphorolytic starch degradation will first generate starch-(P)n by glucan-water dikinases (GWD) and phosphoglucan-water dikinases (PWD). Those phosphorylated glucans will be converted into maltose by βAMY. Maltose generated by starch degradation will be transported into the cytosol by Maltose Excess Protein1 (MEX1) and converted into heteroglycan by disproportionating enzyme2 (DPE2) and starch phosphorylase (PHO).

Table 1.1 showed the enzymes that are involved in the starch metabolism in *Chlamydomonas.*

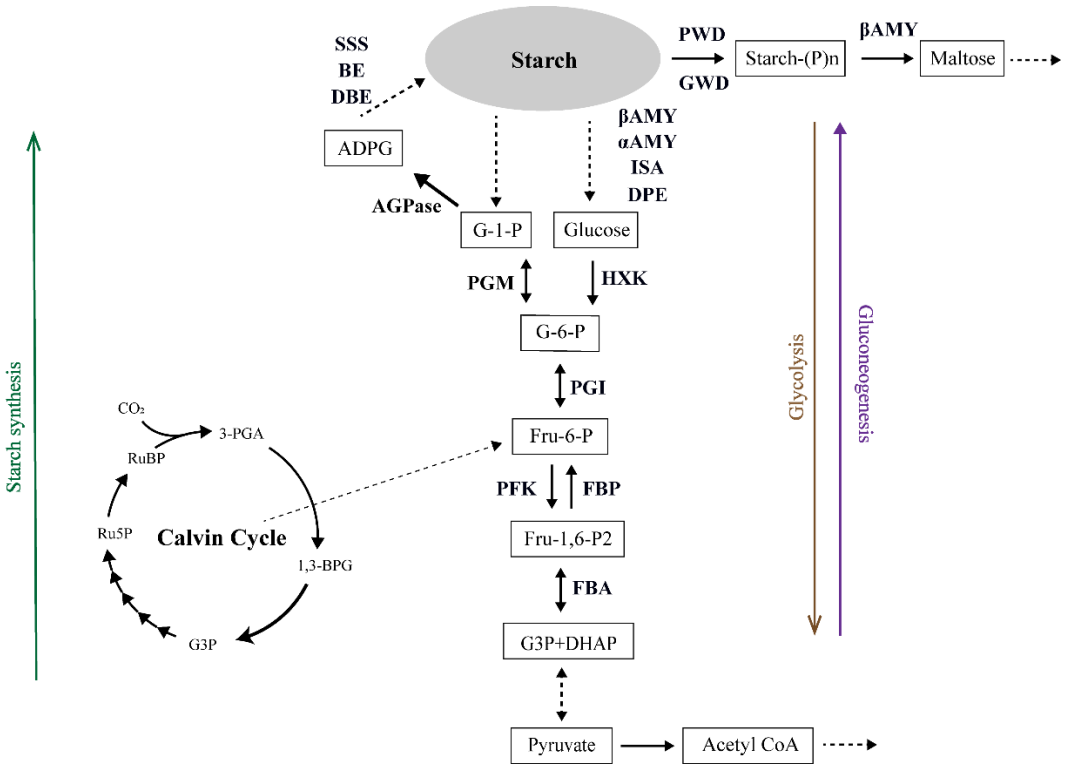


Figure 1.1. Scheme of starch metabolism in *Chlamydomonas reinhardtii.* (Figure modified from Koo et al.[106])

Table 1.1 Enzymes involved in starch metabolism in *Chlamydomonas*[104-109]

Enzyme	Synthesis (S) or Degradation (D)	Gene Name	Gene ID
ADP-glucose pyrophosphorylase	S	AGP2	Cre16.g683450
	S	STA1	Cre13.g567950
	S	STA6	Cre03.g188250
alpha glucosidase	S	AGL1	Cre03.g194700
alpha-amylase	S	AMYA3	Cre08.g384750
	S	AMYA1	Cre08.g385500
	S	AMYA2	Cre08.g362450
beta-amylase	S	AMYB1	Cre06.g307150
	S	AMYB2	Cre06.g270350
	S	AMYB3	Cre01.g044100
4-alpha-glucanotransferase	S	DPE1	Cre02.g095126
phosphoglucomutase	S	GPM1	Cre06.g278210
	S	GPM2	Cre01.g012600
glycosyl transferase	S	GTR11	Cre03.g198200
	S	GTR12	Cre04.g214650
	S	GTR12	Cre06.g302050
	S	GTR7	Cre12.g524000
hexokinase	S	HXK1	Cre02.g117500
Fructan fructosyltransferase	S	FFT4	Cre12.g488000
phosphoglucose isomerase	S	PGI1	Cre03.g175400
starch phosphorylase	S	PHOA	Cre07.g336950
	S	PHOB	Cre12.g552200
starch branching enzyme	S	SBE1	Cre06.g289850
	S	SBE2	Cre06.g270100
	S	SBE3	Cre10.g444700
NAD-dependent epimerase/dehydratase	S	SNE4	Cre06.g278143
soluble starch synthase	S	SSS1	Cre04.g215150
	S	SSS2	Cre03.g185250
	S	SSS3	Cre13.g579598
	S	SSS4	Cre16.g665800
	S	SSS6	Cre12.g521700
	S	STA3	Cre06.g282000
4-alpha-glucanotransferase	S	STA11	Cre03.g181500
alpha,alpha-trehalase	S		Cre03.g195600
UDP-glucose 6-dehydrogenase	S		Cre12.g532450
UDP-glucose dehydrogenase	S	UGD1	Cre07.g357200
	S	UGD2	Cre06.g278185
UDP-glucose pyrophosphorylase	S	UGP1	Cre04.g229700
fructose-1,6-bisphosphate aldolase	D	FBA1	Cre01.g006950
	D	FBA2	Cre02.g093450
	D	FBA3	Cre05.g234550
	D	FBA4	Cre02.g115650
fructose-1,6-bisphosphatase	D	FBP1	Cre12.g510650
dihydrolipoyl dehydrogenase	D	GCSL	Cre18.g749847
phosphoglucomutase	D	GPM1	Cre06.g278210
	D	GPM2	Cre01.g012600
hexokinase	D	HXK1	Cre02.g117500
phosphofructokinase family protein	D	PFK1	Cre06.g262900
	D	PFK2	Cre12.g553250
phosphofructokinase	D	PFK3	Cre11.g467557
	D	PFK3	Cre11.g467552

Enzyme	Synthesis (S) or Degradation (D)	Gene Name	Gene ID
enolase	D	PGH1	Cre12.g513200
phosphoglucose isomerase	D	PGI1	Cre03.g175400
phosphoglycerate kinase	D	PGK1	Cre11.g467770
phosphoglycerate mutase	D	PGM1	Cre06.g272050
	D	PGM2	Cre10.g460300
	D	PGM4	Cre05.g232550
	D	PGM5	Cre03.g166950
Alpha-glucan water dikinase	D	GWD1	Cre07.g319300
	D	GWD2	Cre07.g332300
Phosphoglucan water dikinase	D	PWD1	Cre17.g719900

1.1.4 Strategies to control or increase starch content in microalgae

The depletion of the nutrients, such as nitrogen (N), phosphorus (P), and sulfur (S), is the widely used way to increase the starch accumulation in microalgae[107,110]. N is essential for proteins, nucleic acids, and chlorophyll synthesis in green algae. The depletion or limitation of N leads to more than 50% (w/w, based on dry weight; DW) in starch-producing in green algae, although algae have slow growth rates and declined photosynthetic efficiency[109]. P is an essential component of biomembrane system synthesis and stabilization. P depletion leads to starch accumulation and lipid accumulation in microalgae[111]. S is widely distributed in polar lipids and coenzymes of green algae. Similar to N and P, S depletion leads to more starch than N or P depletion in *Chlorella*[110].

Except for nutrient limitation, other strategies such as temperature changing, specific inhibitors applying, light intensity, and light quality changing. A temperature higher than 30℃ will change the property of the starch fraction and accelerate the conversion of starch to sucrose in *Chlorella vulgaris.*[112]. By using the specific DNA replication inhibitor 5-fluorodeoxyuridine and protein synthesis inhibitor-cycloheximide, green algae cells have been shown to grow to large cell size and accumulate more starch[110,113].

Naturally, light is also a critical factor in starch accumulation. With the increase of light intensity, the starch content increase. Under high light intensities, the photosynthetic CO2 fixation capacity is enhanced, leading to the accumulation of more starch or lipids as shown for the green algae *Dunaliella* and *Chlorella*[108,114].

Besides light intensity, light quality could change the starch content in microalgae. In an industrial green algae-*Chlorella* sp. AE10, its total carbohydrate content (mainly starch) can go to 58.5% (w/w per DW) under red light conditions, while it is 45.17% (w/w per DW) under blue light conditions[115]. In a marine diatom- *Isochrysis* sp. CCAP 927/14, its total carbohydrate content (mainly starch) is 34% (w/w per DW) under white light conditions while it is only 22% (w/w per DW) under blue light conditions[116]. However, the mechanism of blue light decrease the starch content is still unclear. Some studies reported that blue light increased the starch degradation[117,118] in *Chlorella,* but this was not the case in *Phaeodactylum tricornutum*[119].

1.1.5 Starch metabolism regulator in *Chlamydomonas*

Few starch metabolism regulators have been found in *Chlamydomonas.* Li et al. reported that overexpression of *GNAT19*, a Gcn5-related N-acetyltransferase, leads to 72% and 118% higher biomass and starch content than its WT, respectively[120]. Torre-Romero et al. reported a cell cycle mutant- *crcdc5* (*Cell Division Cycle 5*)[38], can accumulate 1.3-1.4 times more starch than its parental strain. This might be due to the impairment in cell division, a process that needs much energy provided by starch degradation[105]. A mutant lacking *BIMODAL STARCH GRANULE 1* (*BSG1*) displayed an abnormal starch granule distribution and contained twice bigger starch granules than its WT[121]. Another two mutants lacking (*Starch Granules Abnormal 1*) *SAGA* [99] and (*Starch Granules Abnormal 2*) *SAGA2*[100] showed a more elongated and thinner starch sheath granules than its WT.

1.1.6 Light Perception

Light is not only the energy source for photosynthetic organisms but also a source of information. Light perceived by different photoreceptors can relay information to the plants about their environment such as light intensity[122], and length of the day[123], while recently, photoreceptors have also been found to play the role of temperature sensors[124,125]. Several classes of photoreceptors, including UV- and blue-light receptors and far-red receptors, are involved in the responses of cells to almost every different wavelength of light[61,126] (Fig. 1.2). These photoreceptors convert light, sensed by chromophore molecules, into biological signals controlling gene

expression[19], photo-orientation, developmental processes[40,127], entrainment of the circadian clock[128], and many other processes[66,129,130].

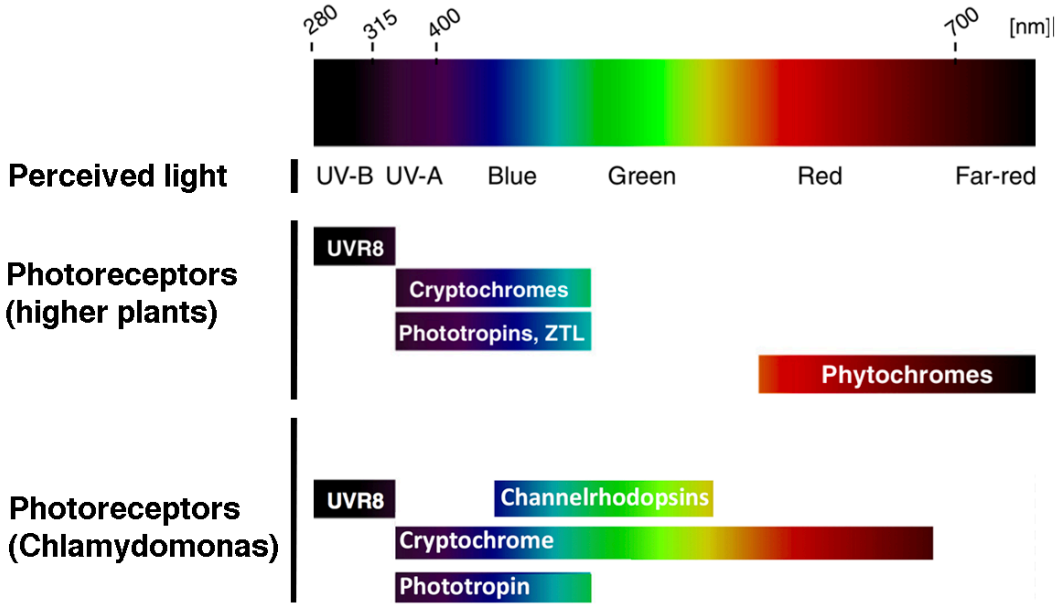


Figure 1.2. Solar spectrum and photoreceptors in higher plants and in *Chlamydomonas reinhardtii*. Figure modified from Heijde et al. [131]based on Petroutsos et al.[132]

1.1.7 Blue light photoreceptor - Phototropin

Phototropin is a blue light photoreceptor responsive to the UV-A and blue region of the electromagnetic spectrum, which was first identified in 1998 in *Arabidopsis*[133]. It is a plasma membrane-associated kinase activated by blue light (BL). Phototropin consists of three main domains, which are two tandemly arranged light, oxygen, or voltage sensing (LOV) domains (LOV1 and LOV2[134]) and a serine/threonine kinase domains[135] (Fig. 1.3). The LOV domains of phototropin function as light sensors and underwent a self-contained photocycle[134]. Although the function of phototropin LOV1 domain is still not very clear, the LOV2 domain represses the activity of its kinase domain[134,136]. The phototropin receptor is unphosphorylated and inactive in the dark or ground state. Absorption of light by the predominant light sensor LOV2 results in a disordering of the Jα-helix and activation of the C-terminal kinase domain, consequently leading to autophosphorylation of the photoreceptor and phosphorylation of protein substrates to start the following signal cascade[133,137-146].

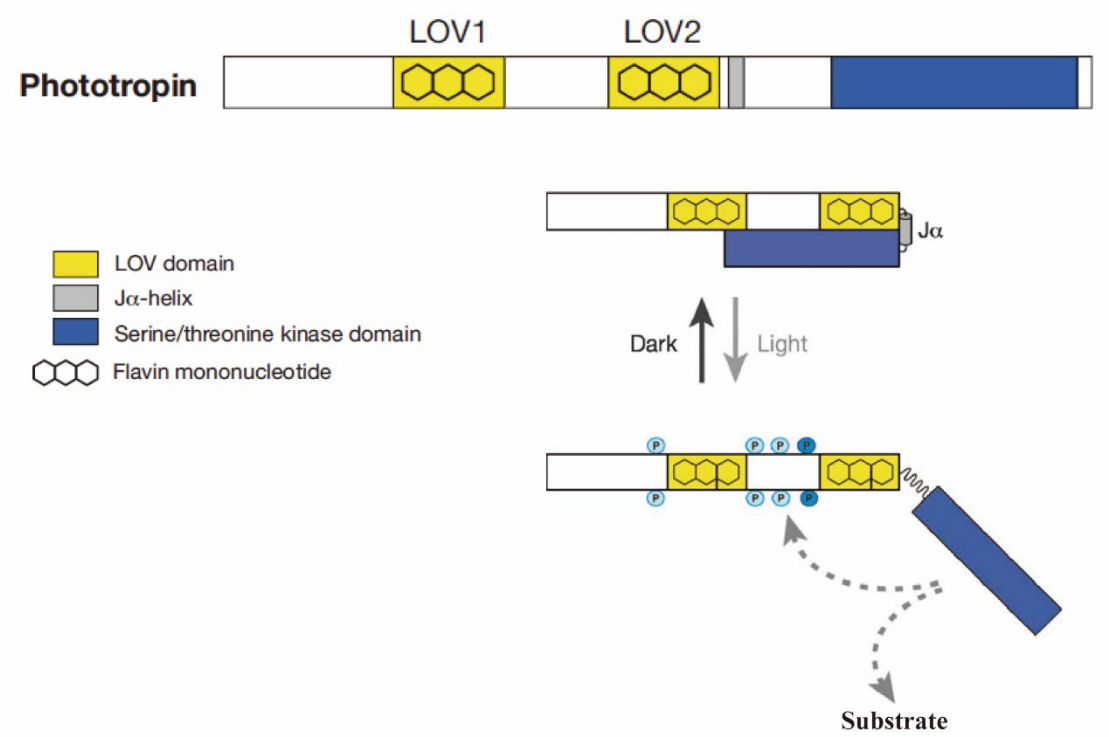


Figure 1.3. Schematic structure of phototropin and phototropin kinase regulation[137]

Phototropins are present in different organisms, including higher plants and algae. *Arabidopsis thaliana* contains two phototropins- PHOT1 and PHOT2. They are mainly located in the plasma membrane but can move into the cytosol[147-149]. PHOT2 localized weakly to the cytoplasm in the dark and exhibited a blue light-dependent association with the Golgi apparatus[149], while PHOT1 could release a fraction to the cytoplasm in a blue light-dependent manner[150]. Phototropins are activated by blue light and control many plant activities, such as temperature sensing[124], leaf flattening[151], stomatal opening[142,145,152-155], phototropism[156], mitochondrial and chloroplast positioning[156-161]. Chloroplast avoidance movement is only mediated by PHOT2[162,163]. Likewise, PHOT1 alone plays a role in mediating the rapid inhibition of hypocotyl growth[164] and promoting the destabilization of specific transcripts under high light intensities[165].

Unlike in *Arabidopsis thaliana*, there is only one single phototropin gene (PHOT, Cre03.g199000) in *Chlamydomonas reinhardtii*. It is located in the plasma membrane and flagella[166]. Its protein sequence is similar to *Arabidopsis thaliana* phototropins, which showed 38% and 35% identity with PHOT2 and PHOT1, respectively[167]. Moreover, phototropin from *Chlamydomonas reinhardtii* is also functional in

Arabidopsis thaliana[168]. The *Arabidopsis phot1 phot2* double mutant is almost totally deficient in phototropism, chloroplast positioning and blue light-induced stomatal opening[156,169]. All the above phenotypes could be rescued by expressing *Chlamydomonas* PHOT in the *phot1 phot2* double mutant of *Arabidopsis*[168].

In *Chlamydomonas reinhardtii*, PHOT has been reported to respond to low fluence blue light, mediating regulation of eyespot development[170], sexual life cycle[40,41], chlorophyll and carotenoid biosynthesis[130]. Eyespot size is significantly larger when PHOT is absent; this phenotype is fully rescued by the complementation with the missing gene. Interestingly, the phenotype can partially be rescued by complementation with a truncated PHOT lacking the kinase domain, indicating for the first time a kinase-independent function of PHOT[170]. Furthermore, channelrhodopsin-1 (ChR1) is also under the control of PHOT[170]. Blue light can induce the pregametes mature into gametes in the sexual life cycle of *Chlamydomonas.* The *PHOT* RNAi strains showed not only the impairment of gametes formation but also the impairment of dark-inactivated gametes reactivation and zygote germination[40,41]. PHOT is also involved in carotenoids and Chlorophyll-binding protein synthesis. The *PHOT* RNAi strains also showed a downregulation of genes encoding glutamate-1-semialdehyde aminotransferase (GSAT), Chlorophyll a-b binding protein (LHCBM6) and phytoene desaturase (PDS), which are critical for photosynthetic function[130]. Moreover, PHOT controls qE by mainly inducing the expression of LHCSR3, together with LHCSR1 and PSBS in excess light conditions[19,62,65].

1.1.8 Photoreceptors and carbon metabolism

Interesting links between photoperception and carbon metabolism have been found in higher plants.

In *Arabidopsis thaliana*, mutants devoid of phytochrome have a reduced CO_2 uptake but over-accumulate daytime sucrose and starch[171]. The stomatal opening and closure rely on the starch degradation of the guard cells. Hiyama et al. reported that blue light and CO_2 signals converge to regulate light-induced stomatal opening[154], while Horrer et al. reported that PHOT plays an essential role in stomatal opening via regulating starch degradation[155]. Both studies found that PHOT plays a vital role in the starch metabolism of guard cells in *Arabidopsis thaliana.* More proteins have been

shown to be involved in this PHOT-stomatal opening link, such as high leaf temperature1 (HT1)[172,173], convergence of blue light and CO_2 ½ (CBC1, CBC2)[154], blue light signaling 1 (BLUS1)[145,153], and CBL-interacting protein kinase 23 (CIPK23)[174].

In *Arabidopsis thaliana*, HT1 is a negative regulator in CO_2-dependent stomatal closure. In *ht1* mutants, the stomatal opening was impaired and constitutively showed high CO_2 responses[172,173,175]. CBCs stimulate stomatal opening by inhibiting the S-type anion channels (a major component of photosynthetically active radiation) in response to blue light and low concentrations of CO_2 resulting from photosynthesis. AtCBC1 and AtCBC2 function in the same pathway as AtHT1 for CO_2 signaling Both CBC1 and CBC2 interact with and are phosphorylated by HT1. Moreover, CBC1 is located in the PHOT-mediated signaling pathway. AtCBC1 physically interacted with PHOT1 and was directly phosphorylated by PHOT1 in vitro. AtCBC2 functions redundantly with AtCBC1 for stomatal opening[154].

PHOT can phosphorylate BLUS1 to activate its kinase activity and then activate the H^+-ATPase. Together with the decrease of intercellular CO_2 concentration, it can stimulate the stomatal opening[145,153]. CIPK23 stimulates stomatal opening in concert with BLUS1 but through an H^+-ATPase activation independent way. The *cipk23* mutant showed an impairment of blue light-dependent stomatal opening. CIPK23 can also interact with both PHOT1 and PHOT2, although it can not be phosphorylated by PHOT1[174].

However, in *Chlamydomonas reinhardtii,* the link between photoperception and carbon metabolism has not been explored so far.

1.2 Dissertation Organization

The overall aim of this dissertation was to understand the link between blue light perception and starch accumulation in *Chlamydomonas reinhardtii*. We first reported the link between photoreceptors and carbon metabolism in *Chlamydomonas reinhardtii*.

Chapter 2 describes the first and novel link between blue light photoreceptor-PHOT and starch metabolism in *Chlamydomonas.* PHOT inhibits starch accumulation in

Chlamydomonas. The PHOT knockout mutant could accumulate more starch than its WT in normal conditions. Moreover, the inhibition of starch accumulation requires the kinase activity of PHOT.

Chapter 3 describes the first signaling cascade pathway linking PHOT with starch accumulation in *Chlamydomonas.* Blue light, via PHOT, represses the bHLH domain-containing transcription factor (TF) GAPR4 (GAP1 Regulator 4), an activator of GAP1. Both GAP1 and GAPR4 overexpression lines could lead to higher starch content. The manuscript reporting the application of Gene Regulatory Network tools to predict the GAPR4 is also included in this chapter.

Chapter 4 describes the second signaling cascade pathway linking PHOT with starch accumulation in *Chlamydomonas*. Blue light, via PHOT, activates the RWP5 TF (belonging to the RWP-K family of TFs), a repressor of starch accumulation in a GAP1-independent manner. RWP5 is usually induced in -N conditions. Overexpression of RWP5 in the PHOT knockout mutant could partially rescue the high starch content phenotype of the PHOT knockout mutant.

Chapter 5 describes the third signaling cascade pathway linking PHOT with starch accumulation in *Chlamydomonas*. We reported the first PHOT-dependent dephosphorylated protein in *Chlamydomonas.* PHOT-dependent dephosphorylation of a serine/threonine kinase- BSK1 (Blue-light Starch Kinase 1) mediates GAP1 expression level and starch accumulation in a GAPR4-independent manner. Overexpression of dephosphorylated form of BSK1 could rescue the high starch content phenotype of the PHOT knockout mutant.

Chapter 2. Phototropin inhibits starch accumulation in *Chlamydomonas reinhardtii*

2.1 Abstract

In this chapter, I present the establishment of a novel link between blue light photoreceptor-phototropin (PHOT) and starch metabolism in *Chlamydomonas reinhardtii*. The experimental data showed that a PHOT knockout mutant (*phot*) over-accumulates starch. The complementation of the *phot* with the *PHOT* gene fully rescued this phenotype showing that PHOT is indeed responsible for regulating starch in *Chlamydomonas*.

2.2 Results

2.2.1 *PHOT*-dependent Genome-wide transcriptional profiling

In order to understand the role of *PHOT* in global gene expression control in *C. reinhardtii*, our lab has performed transcriptomic and proteomic analyses on *phot* and WT under low light (LL, 5 μmol/m^2/s) and high light (HL, 300 μmol/m^2/s) conditions. For the transcriptomics (RNA-seq), samples were collected 1h after the exposure to HL; for the proteomics, cells were exposed to HL for 4h before sampling. In LL conditions, there were 5020 down-regulated and 7380 upregulated genes in *phot*, while 6021 down-regulated and 6379 upregulated genes in HL conditions. Experiment design of collecting sample for transcriptomics and proteomics analysis were shown in Fig. 2.1.

Gene ontology analysis of transcriptomics and proteomics data indicated the involvement of PHOT in carbohydrate metabolism in *Chlamydomonas* (Fig. 2.1).

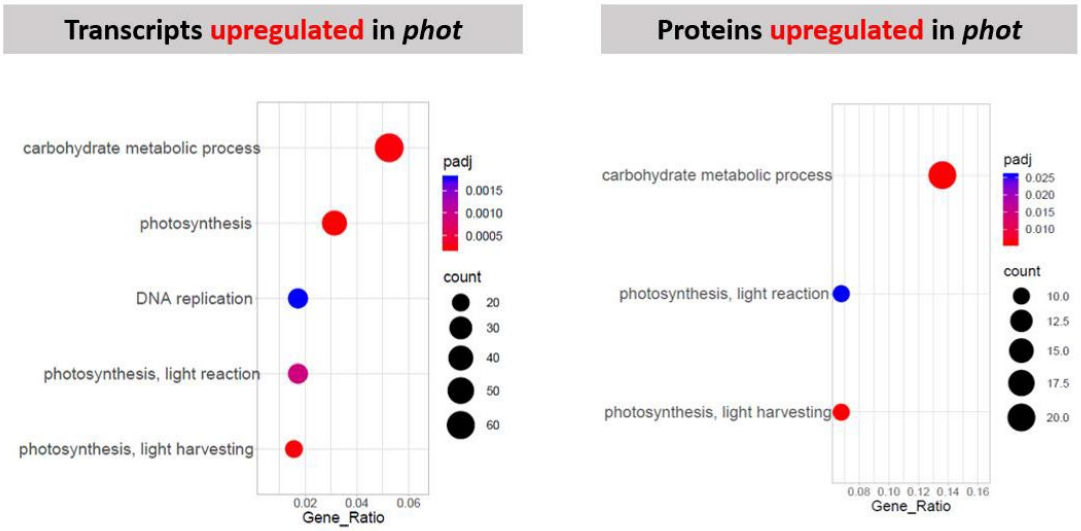


Figure 2.1. Experiment design of collecting sample for transcriptomics and proteomics analysis (Upper) and Gene Ontology (GO) Enrichment Analysis of DEGs between *phot* and WT (Lower). **Upper**: Cultures were grown in TAP medium until cell density to 2-3 million cells/ml. They were washed with fresh HSM medium and resuspended in HSM to a cell density of 1 million cells/ml. Cells were placed under low light conditions for 16-18 hours. After this period, Cells were sampled as LL samples and start the high light treatment. Samples were collected at 1h and 4h for transcriptomics and 37h, as indicated with green arrows. For each time point, 3 biological replicates were prepared for transcriptomics and proteomics analysis, respectively. Three biological replicates were prepared for omics analysis. **Lower**: The X-axis is the number of genes annotated with the given GO term; Point size gives the absolute number of genes; Color gives the statistical significance of the observed enrichment as multiple testing adjusted p-value, red means more significant.

Among those differentially expressed genes (DEGs), most genes involved in starch biosynthesis were upregulated in *phot* under both LL and HL conditions, such as *PGM1* (phosphoglucomutase 1), *STA1* (ADP-glucose pyrophosphorylase large subunit 1), *DPE1* (α-1,4-glucanotransferase), *SSS3* (Soluble starch synthase III), starch branching enzymes (*SBE1*, *SBE3*, *SBE4*) and starch debranching enzyme (*ISA2*, *ISA3*) (Table 2.1).

Table 2.1 Upregulated DEGs involved in *phot* under LL conditions (Calvin-Benson cycle**,** Starch metabolism)

Primary class	Secondary class	Locus ID	Gene name	*phot* LL /wt LL FC (log2)	*phot* HL/wt HL FC (log2)
Calvin-Benson		Cre12.g485150	*GAP1*	0.93	2.17

Primary class	Secondary class	Locus ID	Gene name	*phot* LL /wt LL FC (log2)	*phot* HL/wt HL FC (log2)
cycle					
		Cre01.g029300	*TPIC*	0.85	1.66
		Cre07.g354250	*PGK2*	0.83	1.55
		Cre02.g115650	*FBA4*	0.82	0.80
		Cre07.g354200	*GAP2*	0.34	1.34
		Cre02.g093450	*FBA2*	0.31	-0.45
		Cre12.g554800	*PRK1*	0.12	1.17
		Cre01.g006950	*FBA1*	0.09	1.23
Starch metabolism	synthesis	Cre16.g683450	*AGP2*	0.32	2.70
		Cre06.g272050	*GPM1*	0.60	0.43
		Cre13.g564000	*GPM3*	0.46	-0.52
		Cre17.g698850	*ISA2*	0.86	2.04
		Cre03.g207713	*ISA3*	0.39	0.79
		Cre01.g012600	*PGM1*	1.76	1.75
		Cre03.g199700	*PGM11*	1.55	0.61
		Cre12.g518950	*PGM13*	1.37	0.23
		Cre05.g236150	*PGM14*	0.64	-0.39
		Cre17.g711200	*PGM16*	0.30	-0.12
		Cre10.g431850	*PGM18*	1.27	0.54
		Cre10.g440250	*PGM18*	0.05	-0.92
		Cre12.g487150	*PGM2*	0.43	-0.09
		Cre03.g182300	*PGM7*	0.50	-0.25
		Cre14.g625150	*PGM8*	1.30	1.01
		Cre12.g527400	*PGM9*	0.34	0.93
		Cre06.g289850	*SBE1*	1.58	1.32
		Cre10.g444700	*SBE3*	0.22	0.44
		Cre08.g373450	*SBE4*	0.87	1.79
		Cre04.g215150	*SSS1*	1.39	2.03
		Cre13.g567950	*STA1*	0.15	1.31
		Cre06.g282000	*STA3*	0.58	1.30
		Cre12.g552200	*STA4*	0.39	1.51

Primary class	Secondary class	Locus ID	Gene name	*phot* LL /wt LL FC (log2)	*phot* HL/wt HL FC (log2)
	degradation	Cre08.g362450	*AMYA2*	1.12	2.28
		Cre03.g181500	*DPE1*	0.85	1.16
		Cre07.g319300	*GWD1*	0.05	1.69
		Cre17.g719900	*PWD1*	0.78	1.12

However, *STA6* (ADP-glucose pyrophosphorylase small subunit) was down-regulated in LL conditions and upregulated in HL conditions. In starch degradation, beta-amylase and alpha-amylase genes were down-regulated in *phot* under both LL and HL conditions, while *AMYA2* (α-amylase) was upregulated under both LL and HL conditions. However, *PWD1* (Phosphoglucan water dikinase) and *GWD1* (α-glucan water dikinase) were upregulated in *phot* under both LL and HL conditions (Table 2.2).

Table 2.2 Down-regulated DEGs involved in *phot* under LL conditions (Calvin-Benson cycle, Starch metabolism)

Primary class	Secondary class	Locus ID	Gene name	*phot* LL /wt LL FC (log2)	*phot* HL/wt HL FC (log2)
Calvin-Benson cycle		Cre01.g032650	*TAL1*	-2.14	1.00
		Cre05.g234550	*FBA3*	-1.16	0.66
		Cre03.g187450	*RPI1*	-0.98	-0.74
		Cre12.g511900	*RPE1*	-0.84	0.37
		Cre03.g185550	*SEBP1*	-0.83	1.66
		Cre01.g010900	*GAP3*	-0.58	0.69
		Cre02.g080200	*TRK1*	-0.41	1.38
		Cre12.g556600	*GAPN1*	-0.33	-1.02
		Cre12.g510650	*FBP1*	-0.12	0.49
		Cre02.g120100	*RBCS1*	-0.11	1.45
		Cre02.g116450	*RPE2*	-0.09	0.86
Starch metabolism	synthesis	Cre07.g331300	*AGP3*	-1.02	-0.12
		Cre05.g232550	*GPM4*	-0.33	-0.64

Primary class	Secondary class	Locus ID	Gene name	*phot* LL /wt LL FC (log2)	*phot* HL/wt HL FC (log2)
		Cre03.g155001	*ISA1*	-0.53	-1.45
		Cre03.g175400	*PGI1*	-0.68	0.55
		Cre02.g086500	*PGM10*	-0.22	-1.30
		Cre06.g293150	*PGM12*	-0.05	2.74
		Cre03.g165900	*PGM17*	-1.12	-1.05
		Cre16.g665800	*SS4*	-0.17	-0.76
		Cre03.g185250	*SSS2*	-3.33	-2.61
		Cre12.g521700	*SSS6*	-0.43	-1.00
		Cre17.g721500	*STA2*	-0.45	-0.68
		Cre03.g188250	*STA6*	-0.15	1.56
	degradation	Cre08.g385500	*AMYA1*	-1.51	-0.47
		Cre06.g307150	*AMYB1*	-1.14	-1.17
		Cre06.g270350	*AMYB2*	-1.61	-1.43
		Cre02.g095126	*DPE2*	-0.68	-1.29
		Cre07.g332300	*GWD2*	-0.29	0.16

Motivated by the omics data indicating a role of PHOT in carbohydrate metabolism and the involvement of PHOT in starch degradation in *Arabidopsis* guard cells[155] (Section 1.1.8), and starch is the main carbohydrate accumulating in *Chlamydomonas*, it is worth checking if there is a link between PHOT and starch in *Chlamydomonas*.

2.2.2 Generation of PHOT complemented strains

To address the link between PHOT and starch, we first checked the starch content of the WT (*CC125*) and *phot* (PHOT knockout mutant) in TAP medium under LL conditions. We found that *phot* could accumulate two times more starch than the wild type, although the two strains had identical photosynthetic efficiency under both LL (L1 in Fig. 2.2) and HL (L2 in Fig. 2.2) measured as YII (yield of photosystem II) (Fig. 2.2).

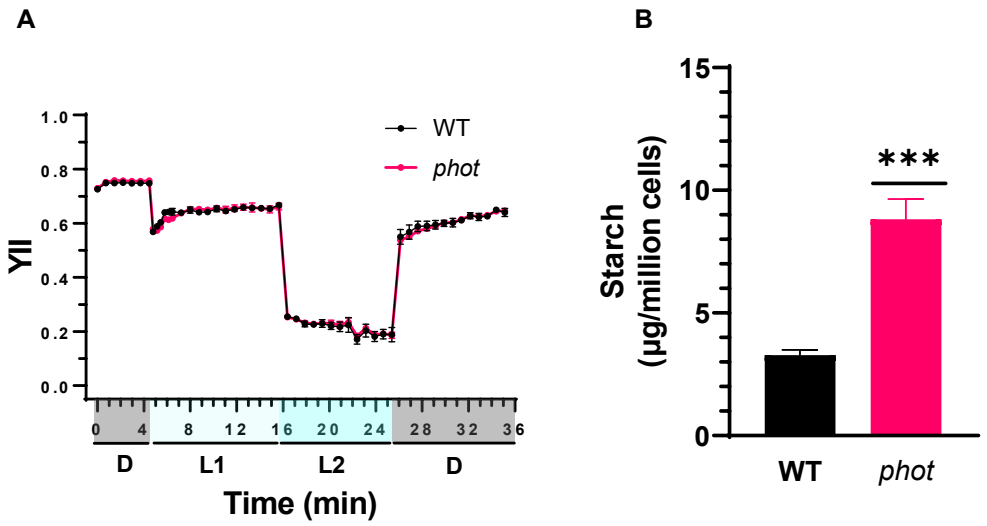


Figure 2.2. Y(II) values (A) and starch content (B) of WT and *phot* cells grown in TAP medium under LL (50 µmol/m²/s). A: Just prior to the onset of the measurements, cells were acclimated to darkness for 15 min. Chlorophyll fluorescence was recorded in the dark (labelled as "D"), at 21 (labelled as "L1") and 336 (labelled as "L2") µmol photons m^{-2} s^{-1} as indicated in the graphs. YII measurement setup details in 6.2.4. **B:** The p-values for the comparisons indicated in the graph are based on t-test (***, $P < 0.001$).

To ensure that the high starch content phenotype is linked to the PHOT deletion, the first step was to complement the *phot* with the *PHOT* gene.

The *phot* and *pp4* (Full PHOT complement strain) strains used before were generated in the *cw15* (cell wall-deficient mutant, flagella-less) genetic background[19]. As PHOT is located in the plasma membrane and in the flagella[166], and the *phot* strain we used for omics analyses was generated in the *CC125* genetic background[8], we decided to stick to the *CC125* background strains in the whole project.

Full-length *PHOT* CDS was amplified from a plasmid ble-phot preserved in our lab[19,170] and subcloned into phk330 plasmid through EcoRI/BamHI (Plasmid map is in section 6.3.1, the detail of amplification is in section 6.2.8) and then transformed into *phot* by electroporation. Cells were plated on Zeocin-containing TAP agar plates. As PHOT plays a key role in qE regulation[19], we decided to screen transformants based on qE phenotype. The Zeocin resistant transformants were picked and grew in the TAP medium under LL (15 µmol/m²/s). After two days of growth, the medium was changed into the fresh phototrophic medium-HSM and cells were exposed to HL (300 µmol/m²/s) for 4h to induce. One transformant completely rescued the low qE

phenotype after HL treatment (Fig. 2.3B). The western blot result showed it is a successful full PHOT complemented strain (Fig. 2.3A), and we named it *phot-C*[62].

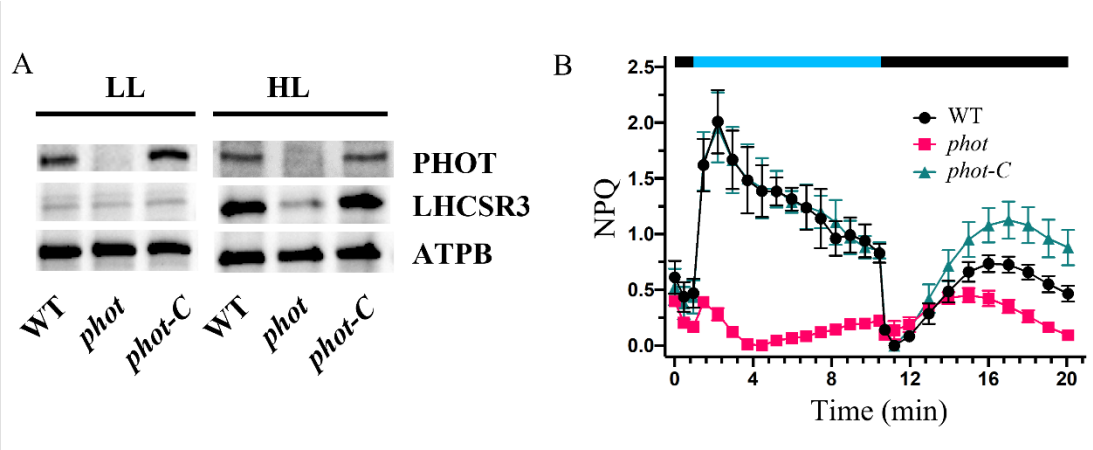


Figure 2.3. Phenotypes of new *phot* mutant and the *phot-C* complemented strain[62]. **A**: Immunoblot blot analyses of WT, *phot,* and *phot-C* acclimated to LL (15 µmol/m^2/s) and after exposure to HL (300 µmol/m^2/s) for 4 h; the proteins examined were PHOT, ATPB, and LHCSR3; ATPB served as a loading control. **B**: NPQ of WT, *phot,* and *phot-C* strains after exposure to HL for 4 h. Just prior to the onset of the measurements, cells were acclimated to darkness for 15 min. Chlorophyll fluorescence was recorded in the dark (black block), at 336 (blue block) µmol photons m^{-2} s^{-1} as indicated in the graphs. (NPQ measurement setup details in 6.2.4).

In addition, for other investigation goals and to get more PHOT complemented strains, I also tried to express full-length *PHOT* CDS with Venuse-Flag tag in *phot* by using pLM005[176] plasmid (Plasmid map is in Section 6.3.1). However, after screening more than 1000 colonies, we did not get any PHOT complemented strain. To solve this problem, we decided to use gDNA rather than CDS. Due to the long *PHOT* gDNA length (6540bp) and high GC content (68%), the *PHOT* gDNA was split into two fragments and amplified using different enzymes and primers G3-G6 (Amplification details are in Section 6.2.8). The two fragments of *PHOT* gDNA were assembled into pLM005 plasmid by Gibson assembly[177] (Plasmid map is in section 6.3.1). With this approach (use of genomic DNA), we got two PHOT complemented lines (*phot-C2* and *phot-C3*) which could rescue the low qE phenotype after screening only 100 colonies (Fig. 2.4). This also suggested that the introns play an important role in enhancing gene expression[178].

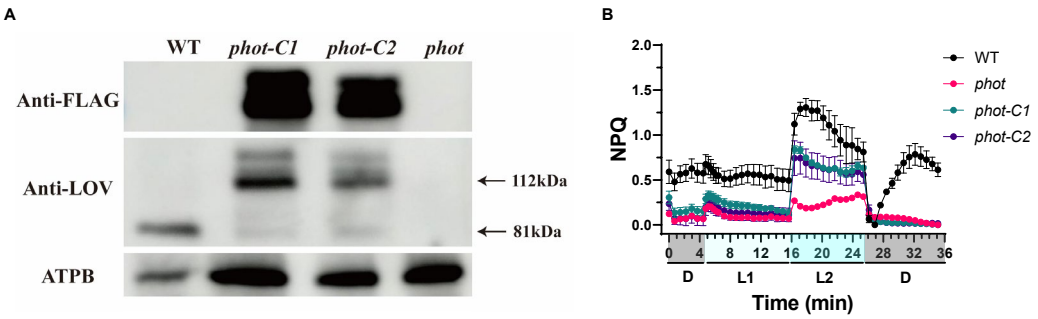


Figure 2.4. Phenotypes of two *phot-C* complemented strains. A: Immunoblot blot analyses of WT, *phot,* and *phot-C* after exposure to HL (300 µmol photons m^{-2} s^{-1}) for 4 h; the proteins examined were PHOT, ATPB, and FLAG; ATPB served as a loading control. **B**: NPQ values of WT, *phot* and *phot-C1/2* cells. Just prior to the onset of the measurements, cells were acclimated to darkness for 15 min. Chlorophyll fluorescence was recorded in the dark (labelled as "D"), at 21 (labelled as "L1") and 336 (labelled as "L2") µmol photons m^{-2} s^{-1} as indicated in the graphs. (NPQ measurement setup details in 6.2.4)

2.2.3 Total carbohydrate content is a reliable approximation of starch content in *Chlamydomonas*

In the starch content measurement method (Section 6.2.6; Fig. 2.2B), removing the ethanol after extraction is among the most time-consuming steps in determining starch in algae. Starch measurement will take half a day for 24 samples. As the main carbohydrate form in *Chlamydomonas* is starch and in order to check more samples in a shorter time, I investigated whether total carbohydrate content (TCC) will be a good proxy for starch content.

The phenol-sulfuric acid method[179,180] is a quick and reliable way to measure the total carbohydrate content, widely used in different microalgae species, such as *Scenedesumus*[181], *Chlorella*[182], *Porphyridium*[183], *Dunaliella salina*[184], *Synechocystis*[185], and *Chlamydomonas*[120].

I measured the starch content and total carbohydrate content of WT and *phot samples.* As Fig. 2.5 shows, the results based on total carbohydrate content and on starch content are in very good agreement, showing very similar differences (in starch or carbohydrates) between WT and *phot*. Because total carbohydrate content is only marginally higher than starch content, these results additionally confirm that starch is indeed the dominant carbohydrate accumulating in *Chlamydomonas.*

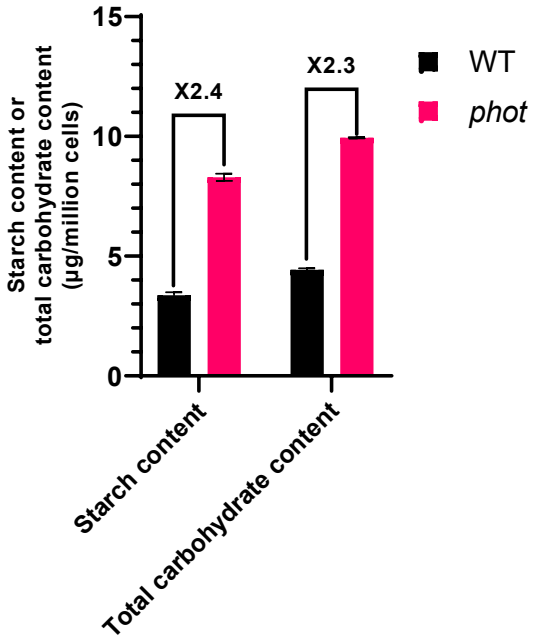


Figure 2.5. Starch and total carbohydrate content of WT and *phot* in TAP medium under LL conditions (50 µmol/m^2/s).

So the total carbohydrate content measurement is a reliable and convenient way to indicate the starch content change in *Chlamydomonas.* Unless otherwise mentioned, total carbohydrate content was used in all the following investigations.

2.2.4 PHOT inhibited starch accumulation in *Chlamydomonas*

Starch accumulated during the day and degraded during the night, so we checked the starch content of WT, *phot*, and *phot-C* strains grown phototrophically (HS medium; HSM) in the 12/12 light/dark cycle.

As Fig. 2.6 showed, *phot* accumulated around three times more starch than the WT, a phenotype that was fully reversed by the complementation with the *PHOT* gene (*phot-C* strain). Interestingly, the synthesis rate in the light and the degradation rate of starch in the dark are similar in these three strains.

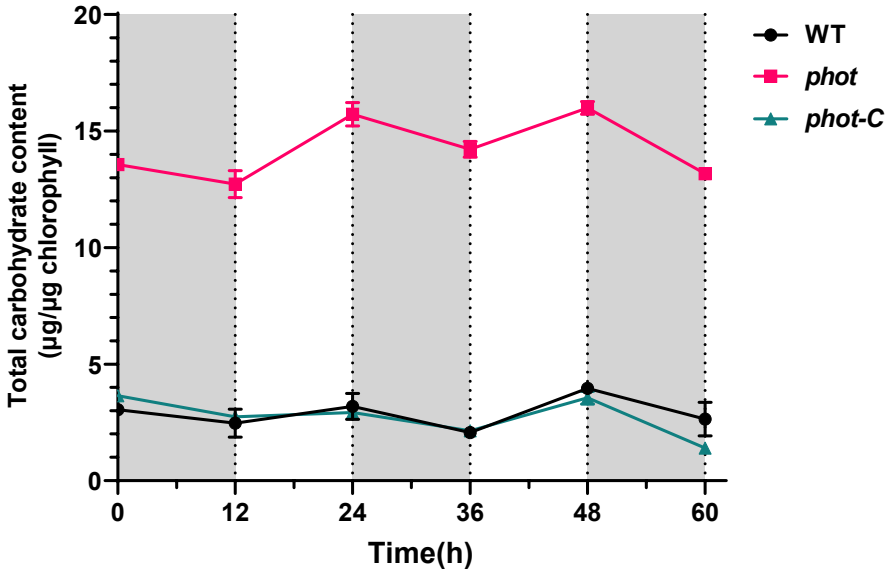


Figure 2.6. The total carbohydarte content of WT, *phot,* and *phot-C* in phototrophic HSM medium under 12/12 light/dark cycle.

In addition, we also checked the starch content of *phot-C2* and *phot-C3* under continuous LL conditions. They behave like *phot-C,* and their starch content was also similar to the WT (Fig. 2.7).

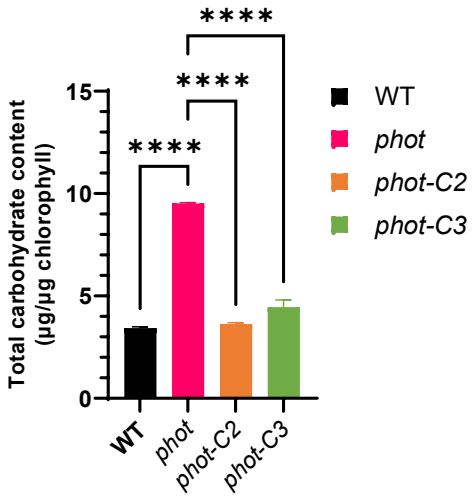


Figure 2.7. The total carbohydrate content of WT, *phot*, *phot*-C2, and *phot*-C3 in phototrophic HSM medium under continuous LL (50 µmol/m^2/s). The p-values for the comparisons indicated in the graph are based on ANOVA Dunnett's multiple comparisons test (**, P < 0.0001).**

Moreover, we used transmission electron microscopy (TEM) to get a high-resolution picture of the ultrastructure of cells. Many starch granules were located in the chloroplast of the cells (Fig. 2.8), and the starch sheath is much thicker in *phot* than in WT. The TEM pictures confirmed the high starch content phenotype in *phot.*

Combining all the results indicated that PHOT inhibits starch accumulation in *Chlamydomonas.*

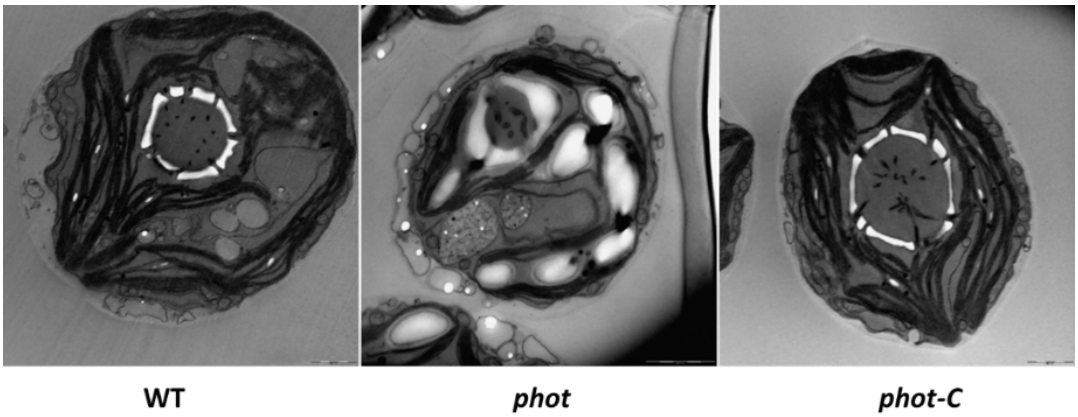


Figure 2.8. Transmission electron microscopy pictures of WT, *phot,* and *phot-C* under LL. Samples for TEM were collected at the end of the light phase of the 12/12 dark/light cycle. The white regions are starch granules.

2.2.5 PHOT inhibition of starch accumulation is kinase-activity-dependent

To explore whether the kinase activity of PHOT is involved in starch accumulation, we generated the *pkin* strain-complement with the kinase domain of PHOT (without the LOV domains) in *phot* by using plasmid ble-kin[170] (Section 6.2). The removal of the LOV domains has been shown to lead to a constitutively active kinase, as has been previously reported in plants[186] and in *Chlamydomonas*[19].

The *pkin* strain could rescue the qE phenotype of *phot* as described before[19] (Fig. 2.9A). Furthermore, the starch content of *pkin* went back to WT level (Fig. 2.9B).

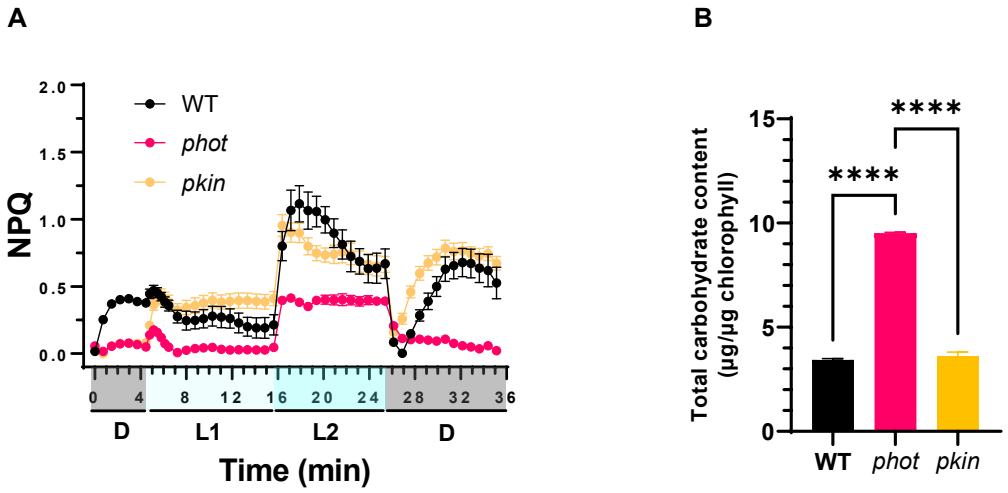


Figure 2.9. Phenotype of WT, *phot,* and *pkin*. A: NPQ values of WT, *phot,* and *pkin* strains after exposure to HL for 4 h. Just prior to the onset of the measurements, cells were acclimated to darkness for 15 min. Chlorophyll fluorescence was recorded in the dark (labelled as "D"), at 21 (labelled as "L1") and 336 (labelled as "L2") µmol photons m^{-2} s^{-1} as indicated in the graphs. (NPQ measurement setup details in 6.2.4). **B**: Total carbohydrate content of WT, *phot,* and *pkin* strains in phototrophic HSM medium under continuous LL (50 µmol/m^2/s) conditions. The p-values for the comparisons indicated in the graph are based on ANOVA Dunnett's multiple comparisons test (****, $P < 0.0001$).

Moreover, we used another set of *phot* mutants based on the *cw15* background[19,170] and checked their starch content. In accordance with my data (Fig. 2.7), the *cw15 phot* mutant accumulated around two times more starch than its WT. Furthermore, complementation with full-length PHOT (*pp4*) or only PHOT kinase domain (*pk1*) could rescue this phenotype. Conversely, complementation with LOV domains of PHOT (*pl1*) and the dead kinase form of PHOT (*pk3*) failed to rescue the starch phenotype (Fig. 2.10). However, due to the partial complementation of PHOT in *pp4*[19], the starch content of *pp4* stays a little higher than in WT.

Combined, the results of Fig.2.10 suggest that the inhibition of starch accumulation by PHOT is kinase-activity-dependent.

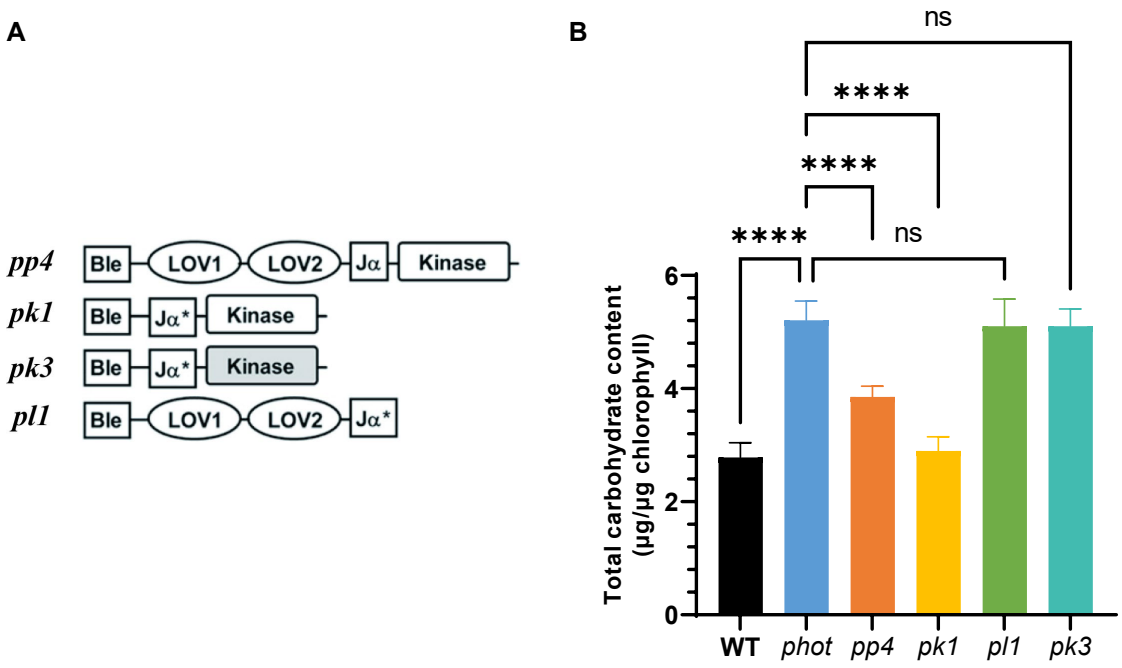


Figure 2.10. The total carbohydrate content (B) of WT, *phot*, *pp4*, *pk1*, *pl1*, *a*nd *pk*3 in phototrophic medium HSM under continuous LL (50 µmol/m^2/s) conditions. A: Schematic drawings of the ble::phot constructs for generating *pp4, pk1, pk3, pl1*[170]. The p-values for the comparisons indicated in the graph are based on ANOVA Dunnett's multiple comparisons test (****, P < 0.0001; ns: not significant).

2.3 Conclusion

In this chapter, we found that many genes involved in carbohydrate metabolism were upregulated in *phot* mutant by RNAseq analyses. As the major carbohydrate form in *Chlamydomonas* is starch, we want to know if there is a link between PHOT and starch in *Chlamydomonas*. First, we found the high starch content phenotype in *phot* mutant, and the full PHOT complement strains confirm that the deletion of PHOT leads to the high starch content phenotype. More importantly, by generating the different parts and dead kinase form of PHOT complemented strains, we confirmed that the enzymatically active kinase domain in PHOT is necessary for the PHOT-starch link.

In conclusion, these results provided the first evidence and a novel link between PHOT and starch metabolism in *Chlamydomonas*.

Chapter 3. Linking blue light perception via PHOT with GAP1 and starch

3.1 Abstract

Chapter 3 describes the first signaling cascade pathway linking PHOT with starch accumulation in *Chlamydomonas.*

In this Chapter, we found a chloroplast localized gene-*GAP1* (Cre12.g485150, glyceraldehyde 3-phosphate dehydrogenase) is controlled by PHOT and partially responsible for the high starch content phenotype of *phot*. Overexpression of *GAP1* leads to higher starch content. Moreover, using the Gene Regulatory Network (GRN) we recently constructed to elucidate the transcriptional responses to the availability of light and carbon source[63], a transcription factor (TF) regulating expression of GAP1 under the control of PHOT was identified. We named this TF *GAPR4* (GAPDH Regulator 4). Overexpression of *GAPR4* leads to higher starch content and *GAP1* expression level in all different light quality conditions.

The manuscript, which includes the GRN we used for predicting *GAP1* regulators, is entitled '*Widening the landscape of transcriptional regulation of algal photoprotection"* and is presented in section 3.2.8.

3.2 Results

3.2.1 PHOT represses GAP1

To get more insights into the starch over-accumulation phenotype of *phot*, we checked the expression level of genes related to starch metabolism and the Calvin cycle[187] (Fig. 3.1).

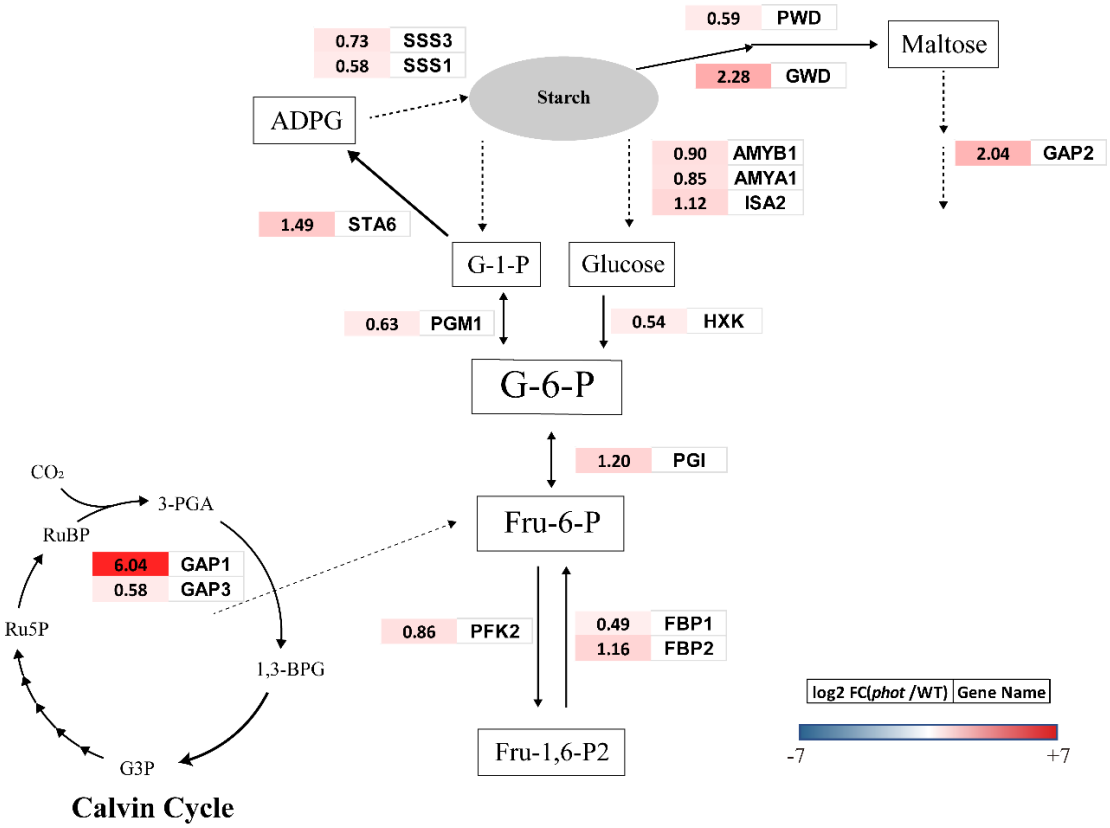


Figure 3.1. The transcription level of starch metabolism-related genes in WT and *phot*. Samples were collected in phototrophic HSM medium in the middle of the light (6h) under 12/12 dark/light cycle conditions. Fold changes between *phot* and WT were calculated as log2 FC (*phot*/WT). *GAP*, glyceraldehyde-3-phosphate dehydrogenase; *PFK*, phosphofructokinase; *FBP*, Fructose-1,6-bisphosphatase; *PGI*, phosphoglucose isomerase; *PGM*, phosphoglucomutase; *HXK*, Hexokinase; *STA6*, ADP-glucose pyrophosphorylase small subunit; *SSS*, starch synthase; *AMY*, α-amylase; *ISA*, isoamylase; *PWD*, phosphoglucan water dikinase; *GWD*, α-glucan water dikinase.

Interestingly, genes related to starch synthesis and degradation were both upregulated in *phot* compared with WT. This indicated that the *phot* mutant synthesizes more starch and consumes more starch. This result is in accord with the previous results in Fig. 2.6. However, one gene encoding one of the three isoforms of GAPDH in *Chlamydomonas*-GAP1 (Section 1.1.2.5) stands out among these genes.

The *GAP1* transcription level is around 100 times higher in *phot* than in WT (Fig. 3.2A), which is in accord with our RNA seq results (Table 2.1). Importantly, *GAP1* transcription levels in *phot-C* are similar to WT levels (Fig. 3.2A). Moreover, our proteomics data analysis showed that GAP1 protein level is 30 times and 150 times higher in *phot* than in WT under HL and LL conditions, respectively (Fig. 3.2B).

Therefore, phototropin inhibits starch accumulation and GAP1 in *Chlamydomonas*.

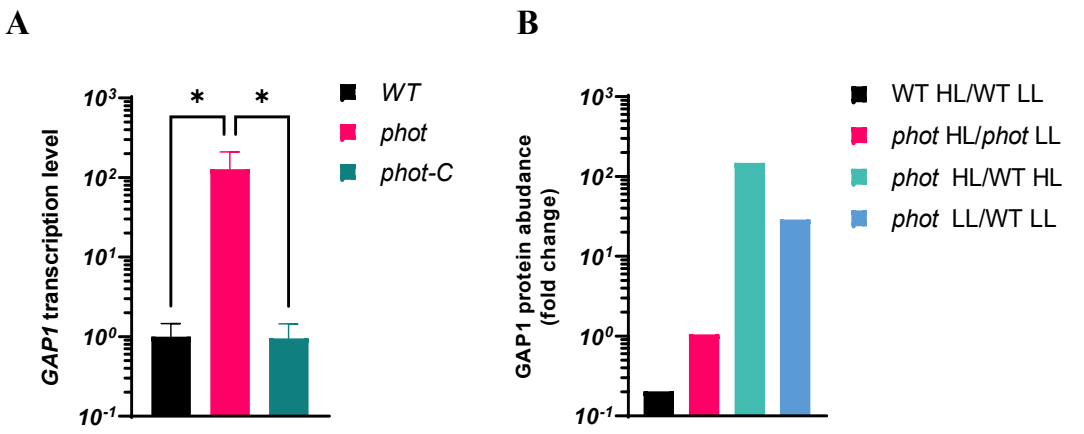


Figure 3.2. *GAP1* transcription level (A) in WT, *phot* and *phot-C* and protein abundance fold change (B) in WT and *phot*.. A: Samples were collected in the middle of the 12/12 dark/light cycle light phase in phototrophic medium HSM. Values relative to an endogenous control gene (GBLP) were normalized to wild-type samples (n = 3 biological samples, mean ± s.d.). The p-values for the comparisons indicated in the graph are based on ANOVA Dunnett's multiple comparisons test (*, P < 0.05)..; **B**: LL-5 µmol/m^2/s, HL-300 µmol/m^2/s.

Glyceraldehyde-3-phosphate dehydrogenases (GAPDHs) are highly conserved proteins in all organisms and play a central role in cellular carbon metabolism (Section 1.1.2.5). As GAP1 is chloroplast located[81] and probably involved in the Calvin cycle and glycolysis, it is worth checking if GAP1 plays a role in PHOT-dependent starch metabolism.

3.2.2 Blue light represses GAP1 in a PHOT-dependent way

As PHOT is a blue light photoreceptor, to test the role of GAP1 in the link between PHOT and starch, we decided to check the *GAP1* level and starch content under different light qualities (Red, Blue, White) in 12/12 light/dark cycle conditions.

In WT, in the middle of the light phase, *GAP1* expression levels were found to be around 100 times higher under red light conditions than under blue light-containing conditions (white light and blue light conditions). However, in *phot* mutant, *GAP1* levels do not vary amongst the different light quality conditions and remain as high as in WT under red light illumination. *GAP1* level in *phot-C* is similar to WT (Fig. 3.3). So we can conclude that blue light represses GAP1 in a phototropin-dependent way.

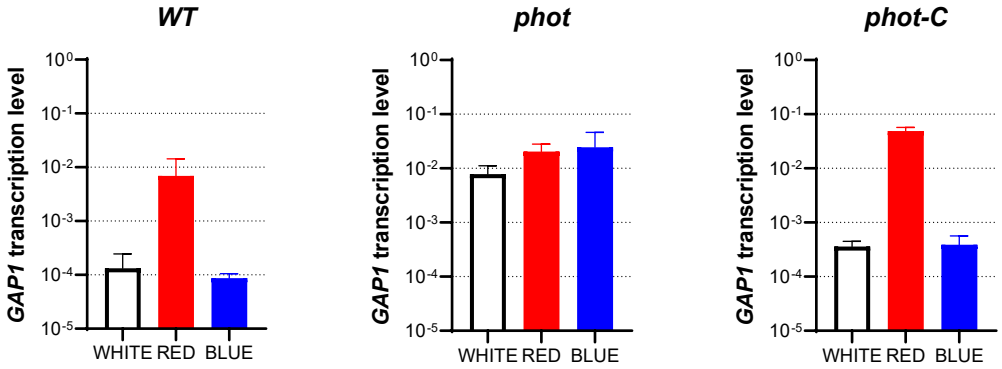


Figure 3.3. *GAP1* transcription level in WT, *phot* and *phot-C* under 12/12 light/dark cycle conditions. Samples were collected at the middle (6h) of the 12/12 dark/light cycle light phase. Values relative to GBLP were normalized to wild-type samples (n = 3 biological samples, mean ± s.d.).

We also checked the starch content of WT, *phot*, and *phot-C* under these conditions.

At the end of the light phase, starch accumulates higher when blue light is eliminated from the light spectrum (red light). This is true for all three strains, although starch levels are significantly higher in the *phot*. These results imply that blue light is a repressor of starch accumulation (Fig. 3.4).

However, in *phot* mutant, we still can see a further accumulation under red light, and this implies that perhaps PHOT is not the only mediator of this blue-light inhibition (Fig. 3.4). Some other blue-light photoreceptors or blue-light responsive proteins might also be involved.

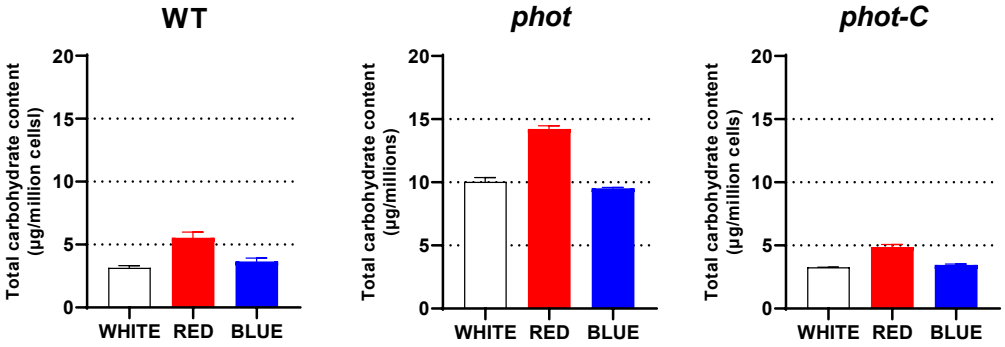


Figure 3.4. The total carbohydrate content in WT, *phot,* and *phot-C* under 12/12 light/dark cycle conditions. Samples were collected at the end (12h) of the 12/12 dark/light cycle light phase.

To summarize, blue light represses *GAP1* in a PHOT-dependent way and represses starch accumulation in a PHOT-dependent way and a PHOT-independent way.

3.2.3 Blue light repression of GAP1 is PHOT kinase-activity-dependent

To explore whether the kinase activity of PHOT is involved in GAP1 repression, we checked the *GAP1* transcription level in the *pkin* mutant. The *GAP1* transcription level also went back to the WT level in the *pkin* mutant (Fig. 3.5).

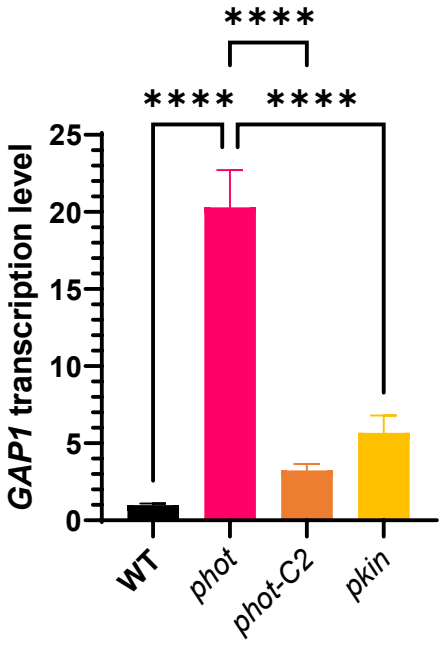


Figure 3.5. The *GAP1* transcription level of WT, *phot*, *phot-C2* and *pkin* in phototrophic HSM medium under continuous LL (50 µmol/m^2/s). Values relative to GBLP were normalized to wild-type samples (n = 3 biological samples, mean ± s.d.). The p-values for the comparisons indicated in the graph are based on ANOVA Dunnett's multiple comparisons test (****, P < 0.0001).

We got similar results using another set of *phot* mutants based on the *cw15* background. However, complementation with LOV domains of PHOT (*pl1*) and the dead kinase form of PHOT (*pk3*) failed to rescue the *GAP1* transcription level (Fig. 3.6).

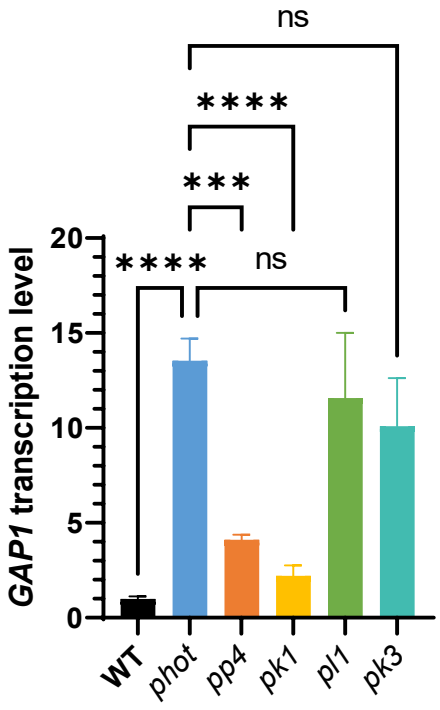


Figure 3.6. The *GAP1* transcription level of WT, *phot*, *pp4*, *pk1*, *pl1*, *a*nd *pk3* in phototrophic HSM medium under continuous LL (50 µmol/m²/s). Values relative to GBLP were normalized to wild-type samples (n = 3 biological samples, mean ± s.d.). The p-values for the comparisons indicated in the graph are based on ANOVA Dunnett's multiple comparisons test (****, P < 0.0001; ns, not significant).

Combining all the results, it suggested that PHOT represses GAP1 is kinase-activity-dependent.

3.2.4 GAP1 over-expression leads to increased starch contents

To further examine the link between *GAP1* and starch accumulation, we decided to generate *GAP1* overexpression lines.

Full-length *GAP1* CDS was amplified from WT cDNA and cloned into pLM005[176] and pRAM118 plasmid[188] (Amplification details in section 6.2.8, plasmid map in section 6.3). GAP1 CDS was under the control of the PSAD promoter. We got only one GAP1 overexpression line after screening more than 1,000 colonies. Therefore, I decided to change the strategy, and I used *GAP1* gDNA rather than CDS; as a result, I got two GAP1-OE expression lines after screening only 100 colonies. This situation is similar to PHOT complementation (Section 2.2.2). After changing from CDS to

gDNA, we screened fewer transformants but got more corrected mutants.

Overall, one *GAP1* CDS overexpression line (*gap1oe-C1)* and two *GAP1* gDNA overexpression lines (*gap1oe-C2* and *gap1oe-C3*) were generated (Fig. 3.7).

The *GAP1* transcription level in these OE lines was at least 12 times higher than in the WT. In these strains, overexpression of GAP1 increased the starch content by a factor of 1.1-1.4, while the disruption of PHOT led to an increase in *GAP1* transcription level and starch content by a factor of 90 and 3, respectively (Fig. 3.7).

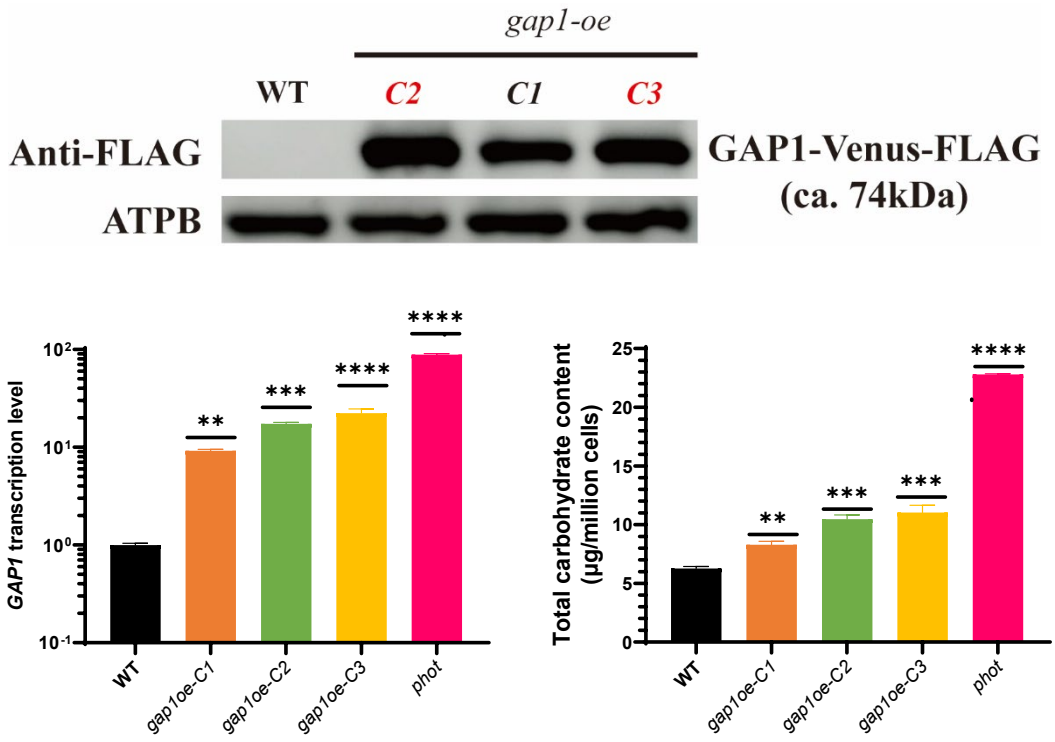


Figure 3.7. Immunoblot blot analyses of GAP1-OE lines (Top) and the *GAP1* transcription level and total carbohydrate content of WT, *phot,* and three *GAP1-OE* lines in phototrophic HSM medium (Bottom). Samples collected under continuous LL conditions (50 µmol/m^2/s). ATPB as a loading control. Values relative to GBLP were normalized to wild-type samples (n = 3 biological samples, mean ± s.d.). The p-values for the comparisons indicated in the graph are based on ANOVA Dunnett's multiple comparisons test (**, P < 0.01; ***, P<0.001; ****, P < 0.0001).

Combining all the results, it is clear that *GAP1* overexpression leads to increased starch accumulation.

3.2.5 GAP1 is not involved in PHOT-dependent LHCSR3 regulation

Another important *phot* phenotype is the qE phenotype[19], so we checked the qE and LHCSR3 levels in three GAP1-OE lines.

LHCSR3 levels did not change in none of the three *gap1oe* lines, and *gap1oe* showed the same qE value as WT (Fig. 3.8). This indicated that GAP1 is not responsible for the PHOT qE phenotype.

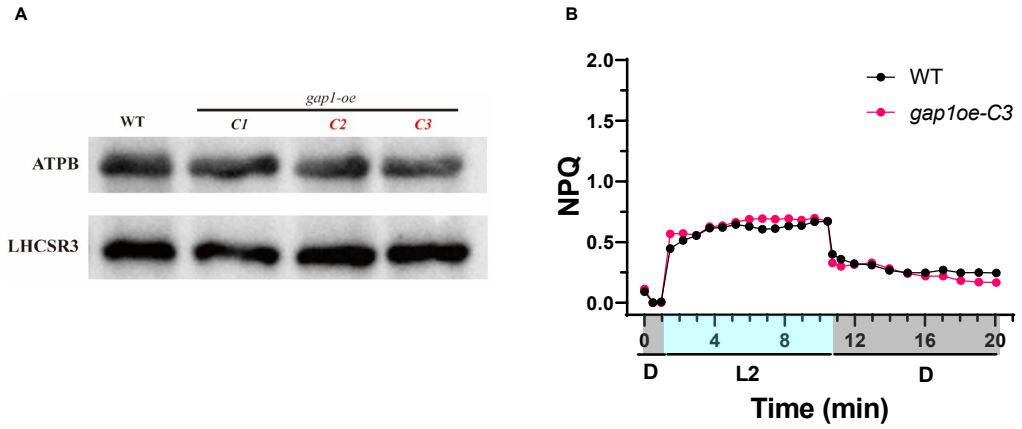


Figure 3.8. Phenotype of WT and *gap1oe lines*. A: Immune blot analyses of LHCSR3, ATPB as a loading control. **B**: NPQ values of for WT and *gap1oeC3* after exposure to HL for 4 h. Just prior to the onset of the measurements, cells were acclimated to darkness for 15 min. Chlorophyll fluorescence was recorded in the dark (labelled as "D"), at 336 (labelled as "L2") µmol photons m^{-2} s^{-1} as indicated in the graphs. (NPQ measurement setup details in 6.2.4).

3.2.6 Putative GAP1 regulator prediction

To get more insights into how PHOT regulates GAP1, by using our omics data and a mathematical modeling approach[63] (The manuscript[63] describing this tool is presented in section 3.2.8), we predicted few putative GAP1 regulators. These putative GAP1 regulators encode different family types of TFs, such as bHLH and RWP-RK family TFs.

We first combined data from two publicly available transcriptomics studies of synchronized chemostat WT cultures grown in a 12h/12h light dark scheme and sampled in 30 min to 2h intervals[189,190] with our RNAseq data generated from mixotrophically or autotrophically grown batch cultures of the WT and *phot* mutant acclimated to LL or exposed to HL; using these data we reconstructed the Gene

Regulatory Network 1 (**GRN1**). We also built the so-called PHOT-specific GRN (**GRN2**) using the RNAseq data from samples of *phot* and WT acclimated to LL and after 1h exposure to HL

In Table 3.1, we present the 19 predicted GAP1 regulators using GRN1 and 2). A smaller number means higher confidence in the prediction.

Table 3.1 Putative GAP1 regulators predicted results.

Locus ID	Gene Name	GRN1	GRN2	Check order (GAPRx)
Cre01.g000050	RWP14	7		
Cre01.g001600		1		*GAPR1*
Cre02.g079200		3		
Cre02.g118250	SWB1		6	*GAPR2*
Cre03.g189000			10	
Cre03.g197100			3	
Cre05.g240600		9		
Cre06.g262800		4		*GAPR3*
Cre06.g266850			4	
Cre06.g269100		10		
Cre07.g323000	TDP1	6		
Cre07.g349152		2	2	*GAPR4*
Cre07.g354500	HSF2	8		
Cre08.g383000			8	
Cre09.g397475			7	
Cre09.g399289			1	
Cre11.g480950			9	*GAPR5*
Cre13.g606750		5		*GAPR6*
Cre16.g672300			5	*GAPR7*

To validate the putative regulatory links, we took 7 TFs (marked in red in Table 3.1) and checked their transcription level in the WT and *phot* under RL and BL conditions by RT-qPCR (Fig. 3.9).

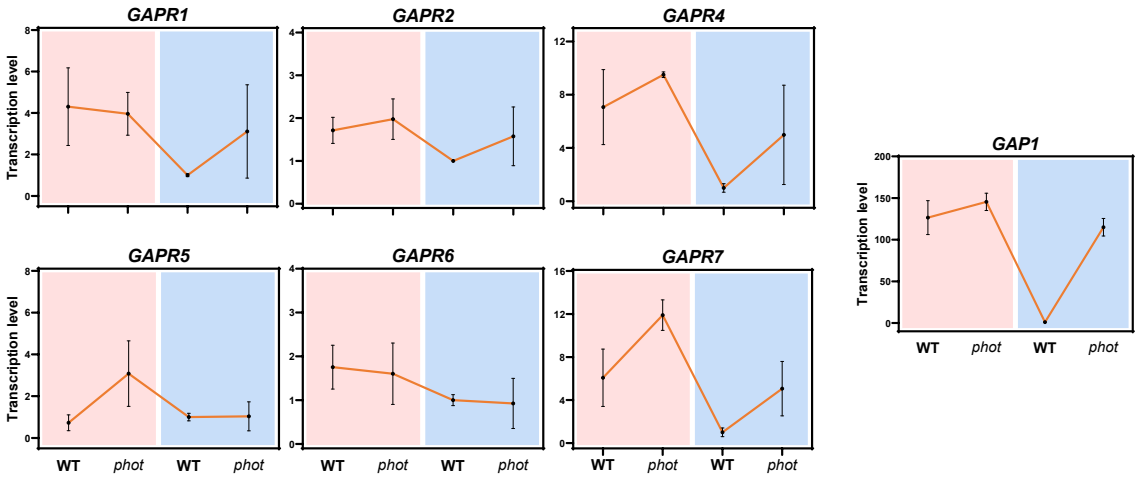


Figure 3.9. Putative GAP1 regulators transcription level in WT and *phot* under red (pink background) and blue light (blue background) conditions (50 µmol/m²/s). Values relative to GBLP were normalized to wild-type samples (n = 3 biological samples, mean ± s.d.).

GAPR1, *GAPR2*, *GAPR4,* and *GAPR7* transcription profiles were similar to GAP1 under red light and blue light conditions (Fig. 3.9) (The primers for GAPR3 did not work).

GAPR4 (Cre07.g349152) is the only gene in the top 10 gene list of two approaches predicted results (2nd position in both approaches, Table 3.1). Furthermore, its expression profile under LL and HL conditions is similar to GAP1 in our RNA-seq data (Fig. 3.10).

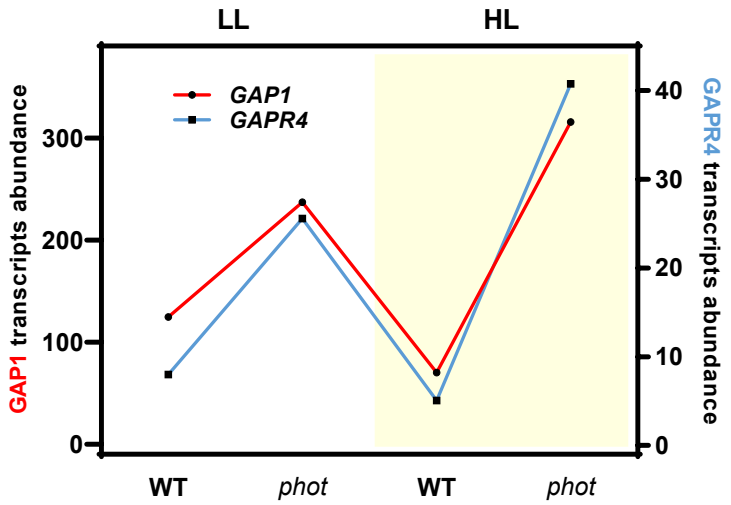


Figure 3.10. *GAPR4* (blue) and *GAP1* (red) transcripts abundance in RNA-seq dataset.

Moreover, *GAPR4* and *GAP1* transcription profiles under 12h/12h day/light cycle in published transcriptome datasets[189,190] were similar, e.g. see (Fig. 3.11).

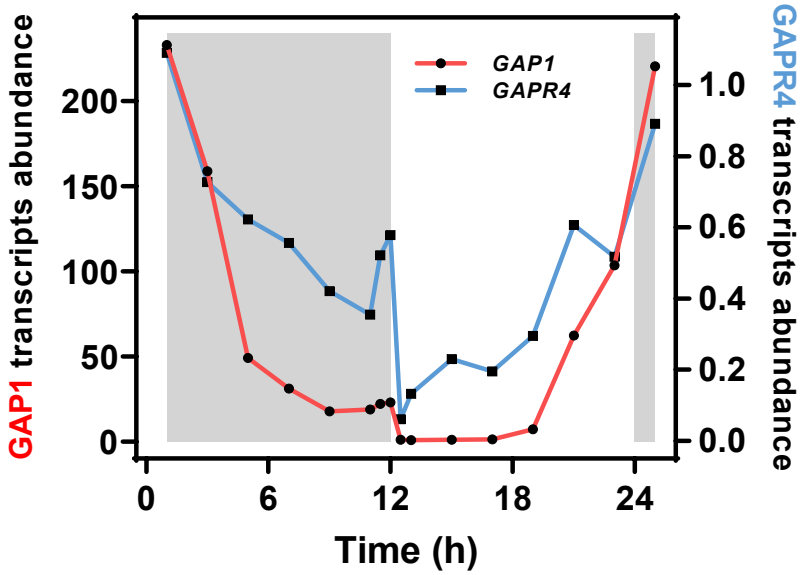


Figure 3.11. *GAPR4* and *GAP1* transcription profile from Strenkert et al. [190].

GAPR4 is a basic helix-loop-helix proteins domain-containing transcription factor (bHLHs). bHLHs are widely distributed in all three eukaryotic kingdoms and constitute one of the largest families of transcription factors[191-193]. bHLHs represent key regulatory components in transcriptional networks controlling many biological processes[193].

In animals, bHLHs have been involved in sensing environmental signals, regulating the cell cycle and circadian rhythms, and regulating diverse essential developmental processes, including neurogenesis, myogenesis, sex, cell lineage determination, proliferation, and differentiation[191,192].

In unicellular eukaryotes, such as yeast, bHLH proteins are involved in chromosome segregation, general transcriptional enhancement, and metabolism regulation[194].

So, we decided to take GAPR4 for further investigation.

3.2.7 GAPR4 - a novel GAP1 regulator under the control of PHOT

To probe the link between GAPR4 and PHOT, we checked *GAPR4* and *GAP1* transcription levels under different light qualities in the WT, *phot,* and *phot-C*. *GAPR4* expression profile is similar to *GAP1*(Fig. 3.12).

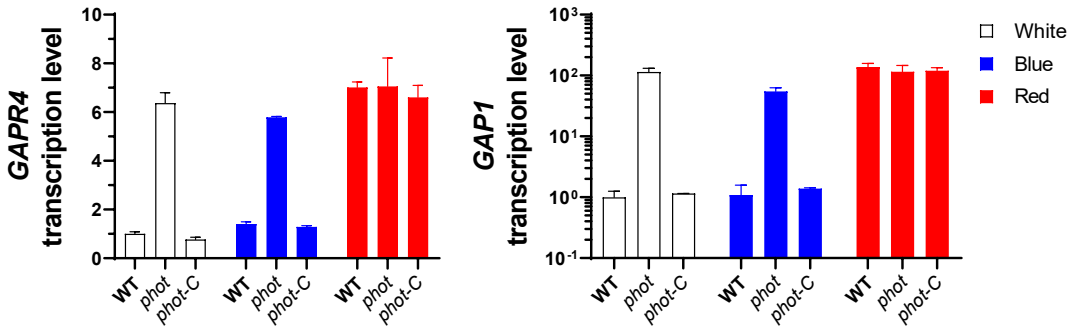


Figure 3.12. *GAPR4* (Left) and *GAP1*(Right) transcription levels in WT and *phot* under different light qualities conditions in phototrophic HSM medium in 12/12 light/dark cycle. Samples were collected at 6h after the start of the light phase. Values relative to GBLP were normalized to wild-type samples (n = 3 biological samples, mean ± s.d.).

In WT, the *GAPR4* transcription level is higher under red light conditions than blue light-containing conditions (white and blue light conditions). However, in *phot*, *GAPR4* levels do not vary amongst the different light quality conditions and remain as high as in WT under red light illumination. *GAPR4* level in *phot-C* is similar to WT (Fig. 3.12). Therefore, GAPR4 is under the control of PHOT.

To test whether GAPR4 is the regulator of GAP1, we generated one GAPR4

overexpression line (Amplification details in 6.2.8), which could increase the *GAP1* transcription level and leads to a bit higher starch than WT. (Fig. 3.13).

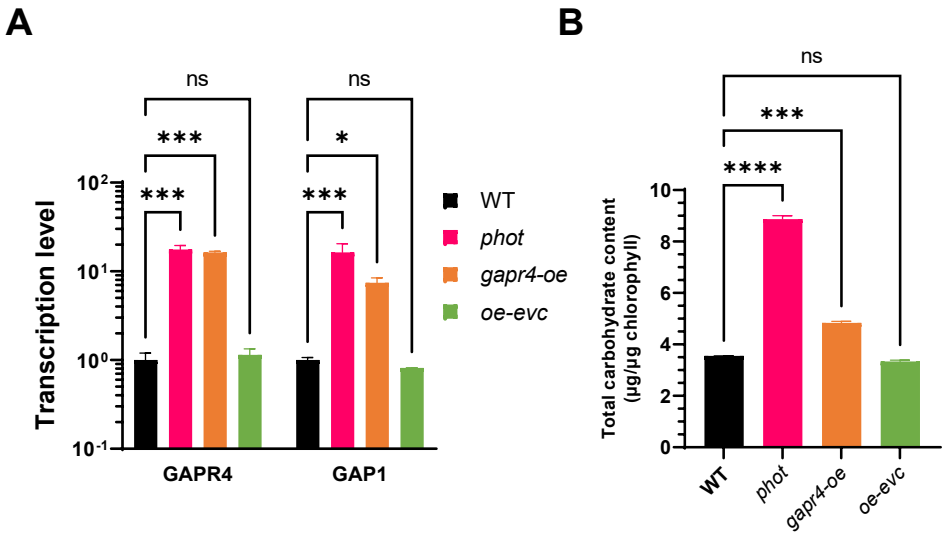


Figure 3.13. The *GAPR4* and *GAP1* transcription level (A) and total carbohydrate content (B) in WT, *phot*, *gapr4-oe* and *oe-evc* (empty vector overexpression control) in phototrophic HSM medium. Samples collected under continuous LL conditions (50 µmol/m^2/s). Values relative to GBLP were normalized to wild-type samples (n = 3 biological samples, mean ± s.d.). The p-values for the comparisons indicated in the graph are based on ANOVA Dunnett's multiple comparisons test (*, P<0.05; ***, P < 0.001; ****, P<0.0001; ns, not significant).

Moreover, the GAPR4 overexpression line leads to a higher *GAP1* transcription level in all different light quality conditions (Fig. 3.14).

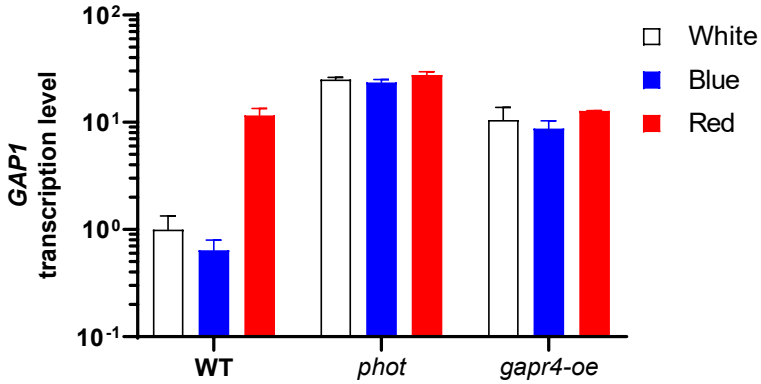


Figure 3.14. The *GAP1* transcription level in WT, *phot* and *gapr4-oe* under different light qualities conditions in phototrophic HSM medium in 12/12 light/dark cycle. (7h after the start of the light phase)

Combining all the results, we conclude that PHOT represses GAPR4 which acts as an activator of GAP1.

3.2.8 Gene regulator network analyses (GRN)

This section presents the manuscript describing the GRN tools used to predict GAP1 regulators. In this manuscript, I am the co-first author responsible for all the wet lab experiment work, including the mutant phenotyping, mutant generating, and experiment designing. This manuscript is available on bioRxiv[63] https://www.biorxiv.org/content/10.1101/2022.02.25.482034v3 and currently under review in *Nat Comms*.

In this manuscript,

1. We compiled a compendium of 158 samples from Chlamydomonas reinhardtii cultures exposed to different light and carbon supplies to infer its gene regulatory network by using an ensemble approach that integrates the results from five benchmarked methods for network inference. Further, we assessed the quality of the resulting light- and carbon-dependent gene regulatory network by employing curated, experimentally confirmed gene regulatory interactions.

2. We used the inferred gene regulatory network to rank candidate transcription factors for their involvement in the regulation of photoprotection and validated two of the identified six candidates in a series of experiments with respective mutants and complementation strains. Specifically, we showed that CCM regulator LCR1 also activates photoprotection and that QER7, a Squamosa Binding Protein, suppresses photoprotection- and CCM-gene expression. Because QER7 gene expression levels are suppressed by the blue light photoreceptor phototropin, our findings clearly suggest that blue light perception takes part in the regulation of CCM.

3. By combining our findings with the analysis of the inferred gene regulatory network, we demonstrated the existence of regulatory hubs that channel light- and CO2-mediated signals into a common response.

Widening the landscape of transcriptional regulation of algal photoprotection

Authors

Marius Arend[1,2†], Yizhong Yuan[3†], M. Águila Ruiz-Sola[3‡], Nooshin Omranian[1,2,4], Zoran Nikoloski[1,2,4*], Dimitris Petroutsos[3 *]

Affiliations

[1]Bioinformatics Group, Institute of Biochemistry and Biology, University of Potsdam, 14476 Potsdam, Germany

[2]Systems Biology and Mathematical Modeling Group, Max-Planck-Institute of Molecular Plant Physiology, 14476 Potsdam, Germany

[3]University of Grenoble Alpes, CNRS, CEA, INRAE, IRIG-LPCV, 38000 Grenoble, France

[4]Bioinformatics and Mathematical Modeling Department, Center of Plant Systems Biology and Biotechnology, 4000 Plovdiv, Bulgaria

[†] these authors contributed equally to this work

[‡] Current address: Instituto de Bioquímica Vegetal y Fotosíntesis, Universidad de Sevilla-CSIC, 41092, Sevilla, SPAIN

[*]**Corresponding Authors**: nikoloski@mpimp-golm.mpg.de (Z.N.), dimitris.petroutsos@cea.fr (D.P.)

Abstract

Availability of light and CO_2, substrates of microalgae photosynthesis, is frequently far from optimal. Microalgae activate photoprotection under strong light, to prevent oxidative damage, and the CO_2 Concentrating Mechanism (CCM) under low CO_2, to raise intracellular CO_2 levels. The two processes are interconnected; yet, the underlying transcriptional regulators remain largely unknown. By employing a large transcriptomics data compendium of *Chlamydomonas reinhardtii's* responses to different light and carbon supply we reconstructed a consensus genome-scale gene regulatory network from complementary inference approaches and used it to elucidate the transcriptional regulation of photoprotection. We showed that the CCM regulator LCR1 also controls photoprotection, and that QER7, a Squamosa Binding Protein, suppresses photoprotection- and CCM-gene expression under the control of the blue light photoreceptor Phototropin. Along with demonstrating the existence of regulatory hubs that channel light- and CO_2-mediated signals into a common response, our study provides a unique resource to dissect gene expression regulation in this microalga.

Introduction

Photosynthetic microalgae convert light into chemical energy in the form of ATP and NADPH to fuel the CO_2 fixation in the Calvin–Benson cycle[1]. They have evolved to cope with rapid fluctuations in light[2] and inorganic carbon[3] availability in their native habitats. When absorbed light exceeds the CO_2 assimilation capacity, the formation of harmful reactive oxygen species can lead to severe cell damage; this is prevented by the activation of photoprotective mechanisms, collectively called non-photochemical quenching (NPQ). NPQ encompasses several processes that are distinguished in terms of their timescales[2], among which the rapidly reversible energy-quenching (qE) is, under most circumstances, the predominant NPQ component[2,4]. The major molecular effector of qE in Chlamydomonas is the LIGHT HARVESTING COMPLEX STRESS RELATED protein LHCSR3, encoded by the *LHCSR3.1* and *LHCSR3.2* genes[5] that slightly differ only in their promoters; LHCSR1 can also contribute significantly to qE under conditions where LHCSR3 is not expressed[6,7]. PSBS, the key qE effector protein in higher plants[8] is encoded as two highly similar paralogues *PSBS1* and *PSBS2* in *Chlamydomonas*[9]. They are only transiently expressed in *Chlamydomonas* under high light (HL)[9,10] and accumulate under UV-B irradiation[11]; their precise contribution in *Chlamydomonas* photoprotective responses is still unresolved[12].

Intracellular levels of CO_2 are modulated by the availability of its gaseous and hydrated forms[3] in the culture media and the supply of acetate, that is partly metabolized into CO_2[7,13]. Under low CO_2, *Chlamydomonas* activates the CO_2-concentrating mechanism (CCM) to avoid substrate-limitation of photosynthesis by raising the CO_2 concentration at the site of RuBisCO, where CO_2 is assimilated[3]. The CCM mainly comprises of carbonic anhydrases (CAHs) and of inorganic carbon transporters. Almost all CCM-related genes are under the control of the nucleus-localized zinc-finger type nuclear factor CIA5 (aka CCM1)[14-16], including the Myb Transcription Factor LOW-CO_2-STRESS RESPONSE 1 (LCR1) that controls expression of genes coding for the periplasmic CAH1, the plasma membrane-localized bicarbonate transporter LOW CO_2-INDUCED 1 (LCI1), and the low-CO_2 responsive LCI6, whose role remains to be elucidated[17]. CIA5 is also a major qE regulator activating transcription of genes encoding LHCSR3 and PSBS, while repressing accumulation of LHCSR1 protein[7].

LHCSR3 expression relies on blue light perception by the photoreceptor phototropin (PHOT)[18], on calcium signaling, mediated by the calcium sensor CAS[19] and on active photosynthetic electron flow[18-20], likely via indirectly impacting CO_2 availability[7]. The critical importance of CO_2 in LHCSR3 expression is demonstrated by the fact that changes in CO_2 concentration can trigger LHCSR3 expression[21-23] even in the absence of light[7]. Accumulation of *LHCSR1* and *PSBS* mRNA is under control of the UV-B photoreceptor UVR8[11] and PHOT[24,25] and is photosynthesis-independent[20,25]. While LHCSR1 is CO_2/CIA5 independent at the transcript level[7,25], PSBS is responsive to CO_2 abundance and is under partial control of CIA5[7]. A Cullin (CUL4) dependent E3-ligase[24,26,27] has been demonstrated to post-translationally regulate the transcription factor (TF) complex of CONSTANS (CrCO)[26] and NF-Y isomers[27], which bind to DNA to regulate the transcription of *LHCSR1*, *LHCSR3*, and *PSBS*. The putative TF and diurnal timekeeper RHYTHM OF CHLOROPLAST 75 (ROC75) was shown to repress LHCSR3 under illumination with red light[28].

Here, we employed a large compendium of RNAseq data from *Chlamydomonas* to build a gene regulatory network (GRN) underlying light and carbon responses, and thus reveal the

transcriptional regulation of qE at the interface of these responses. The successful usage of RNAseq data to infer GRNs has been demonstrated in many studies[29-31], although the data pose some challenges that require careful consideration. All of the developed approaches to infer a GRN quantify the interdependence between the transcript levels of TF-coding genes and their putative targets; the resulting prediction model serves as a proxy for the regulatory strength that the product of the TF-coding gene exerts on its target(s). It is usually the case that the number of observations (samples) is considerably smaller than the number of TFs used as predictors, leading to collinearity of the transcript levels and associated computational instabilities; further, as an artifact of the computational techniques, some of the inferred regulations may be spurious[32-34]. To address these issues, here we took advantage of combining the outcome of multiple regularization techniques and post-processing to increase the robustness of identified interactions[33-35]. In contrast to our approach, the existing predicted GRNs of *Chlamydomonas* either focused on nitrogen starvation[36] or used a broad RNAseq data compendium, not tailored to inferring regulatory interactions underlying responses to particular cues[37]. Moreover, these GRNs were not obtained by combining the outcomes from multiple inference approaches, shown to increase accuracy of predictions[30], and their quality was not gauged against existing knowledge of gene regulatory interactions.

We used an RNAseq data compendium of 158 samples (**Supplementary Table 1**) from *Chlamydomonas* cultures exposed to different light and carbon supply as input to seven benchmarked GRN inference approaches that employ complementary inference strategies[29,30]. We assessed the performance of each approach based on a set of curated TF-target gene interactions with experimental evidence from *Chlamydomonas.* Based on this assessment, we integrated the outcome of the five best performing approaches into a unique resource, a consensus network of *Chlamydomonas* light- and carbon-dependent transcriptional regulation. We used the consensus network to reveal regulators of qE genes and demonstrated the quality of predictions by validating two of the six tested candidates. We show here that LCR1 regulates not only CCM, as previously reported[17] but also qE by activating the expression of LHCSR3, and demonstrate that qE-REGULATOR 7 (QER7), belonging to the SQUAMOSA-PROMOTER BINDING PROTEIN-LIKE gene family, is a repressor of qE and CCM-gene expression. Our work consolidates the extensive co-regulation of CCM and photoprotection[7] based on the untargeted assessment of the obtained genome-scale GRN.

Results

Computationally inferred GRN recovered known regulatory interactions underlying qE and CCM in *Chlamydomonas*

We first aimed to employ published and here generated RNAseq data sets capturing the transcriptional responses of *Chlamydomonas* to light and acetate availability to infer the underlying GRN. To this end, we obtained data from two publicly available transcriptomics studies of synchronized chemostat wild-type (WT) cultures grown in a 12h/12h light dark scheme and sampled in 30 min to 2h intervals[38,39]. We combined these with our RNAseq data generated from mixotrophically or autotrophically grown batch cultures of the WT and *phot* mutant acclimated to low light (LL) or exposed to HL (**Methods, Supplementary Table 1**). These data sets capture the expected expression patterns of the key genes involved on CCM and qE (**Fig. 1a**) in response to changes in acetate availability and light intensity.

Specifically, we found strong up-regulation of these genes in the light[7,25], and a marked inhibition of *LHCSR3.1/2* and CCM genes by acetate as previously described[7,40].

We employed these data together with a list of 407 transcription factors from protein homology studies[41,42] **(Methods, Supplementary Table 2)**, as input to seven GRN inference approaches to robustly predict TF-target interactions, as shown in benchmark studies[30]. Since there exists no study that experimentally probes TF binding to DNA on a genome-scale, we first curated a list of known, experimentally validated regulatory interactions underlying CCM and qE in *Chlamydomonas*[22,26-28], to assess the quality of GRNs inferred by the different approaches (**Fig. 1b**). As a negative control, we considered the lack of effect of the SINGLET OXYGEN RESISTANT 1 (SOR1) TF on *PSBS1* transcript levels in diurnal culture[43]. The comparative analysis of the predicted and known interactions demonstrated that two of the applied approaches for GRN inference (i.e. ARACNE[44] and global silencing[45]) are unable to recover any literature interactions when using a network density threshold of 10% of all possible TF-TF and TF-target interactions. This poor performance is due either to over-trimming or issues with the validity of the underlying assumptions, as seen in other case studies[46]. We therefore considered only the remaining five approaches, namely: Graphical Gaussian Models (GGM), Context Likelihood of Relatedness (CLR), Elastic Net regression, Gene Network Inference with Ensemble of Trees (GENIE3), and Network Deconvolution to infer a consensus GRN by using the Borda count election method[30,47] (**Methods**). We then ranked the inferred interactions within each approach and quantified the variability of ranks for the known TF-target gene interactions. We found that the average standard deviation of the ranks of the TF-target gene interactions within an approach is larger than the average standard deviation for the rank of a TF-target gene interaction across the five approaches (**Extended Data Fig. 1**, **Fig. 1b**). This observation suggested that the properties of a given TF-target gene interaction have a stronger influence on its assigned rank than the inference approach used. More specifically, we noted that the regulation of LHCSR genes by the two NF-Y paralogues and the induction of LCR1 by CIA5 are not recovered by any of the used approaches; this is in line with reports showing that CIA5 is constitutively expressed and regulated post-translationally[16]-not reflected in the transcriptomics data. Further, NF-Y factors that rely on complex formation with CrCO to regulate their targets[27], may also act via unresolved posttranslational mechanisms. Importantly, the regulatory interactions of the CCM effector genes *LCI1* and *CAH1* by *LCR1* are assigned very high ranks (top 1%) by the approaches considered in the consensus GRN (except for GGM); moreover, the experimentally falsified interaction of *SOR1* and *PSBS1* transcripts[43] is correctly discarded by all approaches (**Fig. 1b**). In addition, we observed that CLR and GENIE3 demonstrated the best performance with respect to the set of known interactions. For instance, they identified the regulation of *LHCSR3.1* by CrCO[26,27] and of *PSBS1* by NF-YB[27] (**Fig. 1b**). Generalization of this ranking beyond the known interactions underlying qE and CCM processes is challenging, due to the lack of genome-scale gold standard, and we therefore opted to combine the results of the five approaches, that showed comparable performance, in the consensus GRN **(Methods, Supplementary Table 3)** to increase robustness of the predictions. Our analyses of the overlap between the consensus and individual GRNs and the enrichment of TF-TF interactions demonstrated the robustness of the inferred interactions **(Supplementary Note 1, Extended Data Fig. 2)**.

Consensus GRN pinpoints LCR1 as a regulator of qE-related genes

Using the consensus GRN, we inferred direct regulators of *LHCSR* and *PSBS* genes and ranked them according to the score resulting from the Borda method (**Methods**)[30,47]. Mutants were available for four of the top ten of TFs with strongest cumulative regulatory effect on qE-related genes (**Fig. 2a, Supplementary Table 4):** Two knock-out mutants of previously uncharacterized genes were ordered from the CliP library[48], which we termed *qE-regulators 4* and *6* (*qer4*, *qer6*; see **Extended Data Fig. 3** for the genotyping of these mutants). Additionally, we obtained an over-expressor line of the N-acetyltransferase LCI8[49] and the knock-out strain of the known CCM regulator LCR1[17]. We tested for a regulatory effect by switching LL-acclimated mutant strains and their respective WT background to HL for 1h and quantified transcript levels of qE-related genes. For *qer4*, *qer6*, and *lci8-oe* we could not observe an effect on the transcript levels of investigated genes after HL exposure (**Extended Data Fig. 4**). Thus, qer4 and qer6 are considered false positive predictions of the GRN, despite the fact that *qer4* accumulated 1.5 times more *LHCSR3.1* under LL than the WT. A review of the closest orthologs of LCI8 together with the experimental data indicate that it is likely involved in arginine synthesis[49] and wrongly included as histone acetylase in the list of TFs. Interestingly, LCR1, the highest ranking among the tested regulators showed significantly decreased expression of LHCSR3 at both the gene (3 times lower, **Fig. 2b**) and protein level (4 times lower, **Fig. 2c,d**) compared to the WT; as a result, *lcr1* developed very low NPQ and qE **(Fig. 2e)**. Complementation of *lcr1* with the knocked-out gene (strain *lcr1-C*) restored LHCSR3 gene and protein expression as well as the qE phenotype (**Fig. 2b-e**). Interestingly, the *lcr1* mutant over-accumulated LHCSR1 and PSBS both at the transcript and at the protein level (**Fig. 2b-d**); Complementation of *lcr1* with the knocked-out gene (strain *lcr1-C*) restored LHCSR3 gene and protein expression as well as the qE phenotype (**Fig. 2b-e**). Because pre-acclimation conditions impact qE gene expression[25] we conducted independent experiments in which cells were acclimated to darkness before exposure to HL and we obtained very similar results. Our data demonstrated that *lcr1* showed significantly lower expression of LHCSR3 and higher expression of LHCSR1/PSBS at both the gene (**Extended Data Fig. 5a**) and protein level (**Extended Data Fig. 5b,c**), and had lower NPQ and qE (**Extended Data Fig. 5d**) than the WT, although the higher expression levels of *LHCSR1* gene were not rescued by the complementation with the missing *LCR1* gene (**Extended Data Fig. 5a**). Altogether, our data show that LCR1 is a regulator of qE by activating *LHCSR3.1* transcription and repressing LHCSR1 and PSBS accumulation.

Further, we revisited the role of LCR1 in regulating CCM genes[17] by analyzing expression of selected CCM genes in WT, *lcr1* and *lcr1-C* cells shifted from LL or darkness to HL, conditions favoring CCM gene expression[7]. We first confirmed that under our experimental conditions *lcr1* could not fully induce *LCI1* (**Extended Data Fig. 6a, b**) in accordance to the report of the discovery of LCR1[17]. Our analyses further showed a statistically significant impairment of *lcr1* in inducing genes encoding the Ci transporters LOW-CO_2-INDUCIBLE PROTEIN A (LCIA), HIGH-LIGHT ACTIVATED 3 (HLA3), CHLOROPLAST CARRIER PROTEIN 1 (CCP1) and BESTROPHINE-LIKE PROTEIN 1 (BST1) as well as the carbonic anhydrase CAH4, when shifted from LL or dark to HL (**Extended Data Fig. 6a, b**), indicating that the role of LCR1 in low-CO_2 gene expression extends beyond the regulation of gene expression of *CAH1*, *LCI1* and *LCI6*[17].

PHOT-specific GRN reveals a novel repressor of qE and CCM

The light-dependent induction of LHCSR3 is predominantly mediated by the blue light photoreceptor PHOT[18]. To analyze the PHOT-dependent transcriptional regulators, we first identified the interactions shared by the consensus network and a GRN inferred only by

GENIE3, due to its good performance, using the RNAseq data from samples of *phot* and WT acclimated to LL and after 1h exposure to HL (**Methods**). This so-called PHOT-specific GRN (**Methods**, **Supplementary Table 5**) comprises regulatory interactions underlying the transcriptomic changes observed in the *phot* mutant while burrowing the statistical power of the RNAseq compendium.

When investigating the top 10 regulators of qE genes in this PHOT GRN the interactions where weighted based on the importance score from GENIE3 (**Supplementary Table 6, Fig. 3a**). Among these interactions we recovered the previously discussed false positive LCI8 gene (see above, **Fig. 3a**) (**Extended Data Fig. 4**). In addition, two previously reported regulators of qE, ROC75[28] and CrCO[26,27], are also included in this list of PHOT-dependent regulators. These observations are in line with an existing hypothesis[24] suggesting that a CUL4-dependent E3-ligase targeting CrCO,[26] acts downstream of PHOT. ROC75 has been previously reported to act independently of the PHOT signal based on qPCR studies of the mutant grown synchronously under different light spectra[28]. In our RNAseq data, gathered under continuous white light, we observed a significant difference in expression levels of ROC75 between WT and *phot* (log2 fold-change = 1.03, adj. p-value = $1.80*10^{-7}$).

The fact that several regulators showed larger regulatory strength than CrCO in the PHOT GRN indicates the existence of yet unreported regulators of qE effector genes in the PHOT signaling pathway. This is in line with existing results[26], showing that the knock-out of CrCO is insufficient to fully abolish light-dependent activation of LHCSR3. Following this reasoning we obtained *qer1* and *qer7*, the available regulator candidates mutants, from the CLiP library[48] , (for genotyping see **Extended Data Fig. 3**). Our results show higher mRNA and protein levels of LHCSR1 in the *qer1* mutant (**Extended Data Fig. 7**); however, this could not be rescued by ectopic expression of the *QER1* gene in the *qer1* mutant background (**Extended Data Fig. 7**). We found significant upregulation of *LHCSR3.1* gene expression in the *qer7* mutant (1.7 times, **Extended Data Fig. 8a**) also reflected in higher NPQ (**Extended Data Fig. 8b)** and qE levels (**Extended Data Fig. 8c**) which we followed up in more detail. To this end, we ectopically expressed the WT *QER7* gene in the *qer7* mutant and generated the complemented strain *qer7-C* that expressed *QER7* to levels similar to those WT (**Extended Data Fig. 3c**). As a result, the *qer7-C* strain showed reduced *LHCSR3* gene expression, NPQ and qE levels as compared with the *qer7* mutant (**Extended Data Fig. 8a-c**). *LHCSR1* and *PSBS* seemed to be unaffected in the qer7 in these LL to HL transition experiments (**Extended Data Fig. 8a**). As with LCR1, we also performed dark to HL experiments to further characterize the photoprotective responses of *qer7*; under these conditions, *qer7* accumulated significantly more *LHCSR1* (1.7 times) and *PSBS1* (2.2 times) while *LHCSR3* remained unaffected (**Extended Data Fig. 8d**). As in the LL to HL experiments (**Extended Data Fig. 8b, c**) *qer7* showed more NPQ and qE (**Extended Data Fig. 8e, f**). Complementation of *qer7* with the missing *QER7* gene restored all phenotypes (*LHCSR1*, *PSBS*, NPQ, qE; **Extended Data Fig. 8d-f**). These data validate the prediction of QER7 as regulator of qE gene expression (**Fig. 3a**) and indicate that QER7 regulates different subsets of qE genes depending on the pre-acclimation conditions; *LHCSR3* when preacclimated under LL, *LHCSR1* and *PSBS* when pre-acclimation occurs in darkness. Motivated by these findings and given the fact that most of the Chlamydomonas transcriptome, undergoes diurnal changes according to biological function[39] we decided to address the role of QER7 in regulating qE genes under light/dark cycles. We synchronized WT, *qer7* and *qer7-C* cells in 12h L/12h D cycle and exposed them to HL right after the end of the dark phase. Our results revealed

that under these conditions QER7 functions as a repressor of all qE-related genes; the *qer7* mutant expresses significantly higher LHCSR3, LHCSR1 and PSBS both at the gene (**Fig. 3b**) and protein (**Fig. 3c, d**) level, and exhibits higher NPQ and qE (**Fig. 3e**), with all phenotypes rescued in the *qer7-C* complemented line. Previous protein homology studies identified QER7 as Squamosa Binding Protein[50] or bZIP TF[51], and here, we provide the first functional annotation of QER7 as a novel qE regulator.

QER7 co-regulates qE-related and CCM genes

The recent findings that the regulatory role of CIA5[7] and LCR1 (**Fig. 2 and Extended Data Fig. 5**) extends beyond CCM to also control qE-related gene expression, prompted us to also inspect the expression levels of CCM genes in synchronized *qer7* cells (**Fig. 4a**). Indeed, for four of these transcripts (*CAH4*, *LCIA*, *CCP1*, *HLA3*) we observed a significant upregulation in *qer7* after HL exposure that was reversed by complementation with the *QER7* gene, indicating that QER7 suppresses expression of CCM genes; the suppression role of QER7 on CCM genes was only observable under HL, conditions that favor CCM gene expression[7] and not under LL (**Fig. 4a**). Interestingly, and in line with QER7 acting as suppressor of CCM gene expression (**Fig. 4a**), we found LCR1 and QER7 among the top 10 regulators of the CCM genes probed by qPCR in the inferred GRNs (**Extended Data Fig. 9**).

After observing the inhibition of CCM gene transcription by QER7, we investigated the signaling pathway upstream of QER7. To this end, we quantified *QER7* gene expression in synchronized *phot* cultures. As expected, we observe a strong effect of the *PHOT* knock-out showing that PHOT suppresses *QER7* expression (**Fig. 4b**) implying that blue-light perception via PHOT is involved in the regulation of CCM related gene expression. Interestingly, *phot* expressed *LCR1* to levels similar to those in WT (**Fig. 4c**). Thus, while sharing part of their target genes, the two TFs mediate different signals. These data demonstrate that the PHOT-specific GRN successfully pinpoints TFs that depend on PHOT activity while retaining the statistical power to recover genuine regulatory interactions.

The two qE regulators that we validated in this study also regulate CCM genes. Therefore, we next investigated to what extent the observed coregulation pattern applies to the global, known transcriptional regulation of low CO_2 and light stress responsive genes. To this end, we took advantage of the size of the presented genome-scale GRNs and compiled a list of genes putatively involved in photoprotection (**Supplementary Table 7**) or the CCM (**Supplementary Table 8**); we then extracted the 10 TFs exhibiting the strongest regulatory strength on the genes in the compiled lists. We found six (empirical p-value<0.001, **Methods**) and four (empirical p-value < 0.01) of the top 10 regulators to be shared between these two responses in the consensus **(Fig. 5a, b**) and the PHOT-specific GRN, respectively (**Figure 5c, d**). The significant, large number of shared regulators is a strong indication that co-regulation of photoprotective and carbon assimilatory processes is a principal feature of Chlamydomonas' transcriptional regulatory program.

Discussion

The molecular actors and structure of the transcriptional regulatory mechanisms that shape Chlamydomonas' response to differential light and carbon availability are largely unknown, although they are paramount to survival of Chlamydomonas and offer valuable targets for biological engineering. Here we set out to elucidate the GRN underlying the response to light and carbon availability by combining the results from five complementary inference

approaches and data from 158 RNAseq samples of cultures responding to these cues. The obtained GRN enabled us to identify novel regulators controlling qE. Experimentally testing available mutant strains for six of these candidates we were able to validate two TFs: QER7, suppressing qE- and CCM-gene expression, and LCR1, activating LHCSR3 and suppressing LHCSR1 and PSBS expression. The two TFs belong to independent signalling pathways: LCR1 has been known to be controlled by CIA5 while we found that QER7 is controlled by Phototropin; this last finding adds a new layer of regulation of CCM by blue light sensing. We complemented the finding of the involvement of LCR1 in the regulation of qE related genes with an analysis of the genome-scale co-regulation of CCM and photoprotective genes. Indeed, we observed significant number (six out of the top 10) of regulators targeting both processes in the inferred consensus GRN. Together with previous studies, our results provide evidence for extensive signaling crosstalk between light and carbon sensing and name a set of TFs that likely integrate these disparate signals into a common transcriptional response. The GRN gives an unbiased representation of the genome-scale regulators acting in light and carbon sensing. Thus, it represents a powerful resource for future dissection of the transcriptional regulation of responses of *Chlamydomonas* to light and carbon availability.

Methods

Transcriptome analysis

We assembled a compendium of RNAseq data (**Supplementary Table 1**) that capture regulation of light-dependent processes by combining in-house produced RNAseq measurements with publicly available data from two studies of densely sampled diurnal cultures of Chlamydomonas[38,39]. For the samples in the acetate time-resolved experiment, adapter sequences were specifically trimmed from raw reads using BBduk[52] (ktrim=r k=30 mink=12 minlen=50). Raw reads of the diurnal transcriptome study from Strenkert et al.[39] were obtained from NCBI GEO database (GSE112394). Reads were aligned to the Chlamydomonas reference transcriptome[53] available from JGI Phytozome (Assembly version 5) using RNA STAR aligner. The BAM files obtained from these measurements were analyzed using HTSeq-count[54] (standed=reverse) to create raw read count files. The raw read counts from Zones et al.[38] were obtained as .tsv from NCBI GEO (GSE71469). The final data set consists of 158 samples from 62 experimental conditions or time points (**Supplementary Table 1**). Genes with less than 1 count per million in at least 9 measurements where discarded and the remainder were voom[53] transformed and normalized using library normalization factors based on the TMM[55] approach as implemented in the R Bioconductor package edgeR[56].

Transcription factor set from comparative genomics

To reduce the set of parameters in our network model, we compiled transcription factor (TF) annotations for the Chlamydomonas genome based on proteome homology studies. We obtained the proteomes and protein IDs of predicted Chlamydomonas TFs from Pérez-Rodríguez et al.[41] Since these predictions were built based on the older Chlamydomonas assembly, we first used the conversion table provided by Phytozome to convert JGI4 to Crev5.6 IDs. For the TFs that could not be recovered by this approach we used the Phytozome BLAST tool to align these sequences against the Crev5.6 proteome (BLASTP, E threshold: -1, comparison matrix: BLOSUM62, word length: 11, number of reported

alignments: 5). The reported hits were filtered for sequence identity > 97% and gaps ≤ 1 . If sequences mapped multiple times to the same Crev5.6 gene ID, only the hit alignment closest to the N-terminus of the query sequence was kept. The hit was only accepted if the alignment started at least six residues from the N-terminus of the hit sequence. For Crev5.6 loci that had multiple JGI4 TF queries assigned to them the best hit was selected manually. This set was then extended by the TFs found in the study of Jin et al.[42] and the regulators in the manually curated set of CCM and qE regulatory interactions (**Supplementary Tables 7 and 8**). Using this procedure, we compiled a list of 407 Chlamydomonas TFs (**Supplementary Table 2**) to be considered as regulators in the inferred networks.

Gene regulatory network inference

The CLR and ARACNE approach were based on all replicate measurements; for all other inference methods the median from each condition was used as input. All input matrices where standardized gene-wise. If not explicitly stated in the respective paragraph the implementations of all GRN inference approaches were applied with their default settings.

Graphical Gaussian Models

The network inferred from a Graphical Gaussian model of gene regulation was obtained using the implementation of the partial correlation estimate from Schäfer et al.[34] as implemented in the R GeneNet package. All interactions between TFs and another gene/TF with non-zero partial correlations were included as network edges.

GENIE3

The random forest-based network from GENIE3 was generated using the R Bioconductor implementation provided by the authors[57]. We used only expression levels of TFs as predictors.

Elastic net regression

A linear regression based network was obtained using the elastic net algorithm[35]. A model was fit for each gene using the expression levels of all TFs as predictors. The two hyperparameters $\lambda 2$ (quadratic penalty) and s (fraction of L1 norm coefficients) were tuned for each gene model using 6-fold cross validation. The 2D parameter space scanned was $\lambda 2$={0,0.001,0.01,0.05,0.1,0.5,1,1.5,2,10,100} and s={0.1,0.2,0.3,0.4,0.5,0.6,0.7,0.8,0.9}. The R2 value for each model was calculated as

$$R^2 = 1 - \frac{\sum(y - \hat{y})}{var(y)} \quad (1)$$

with y marking the vector of observed expression values and $\hat{y}$ the model predictions. Models with a negative $R2$ value were discarded as regularization artifacts. The results of the remaining models were assembled into a network in which interactions were ranked by regression coefficients β normalized by the maximum absolute coefficient.

$$\beta_n = \frac{\beta}{\max|\beta|} \quad (2)$$

CLR and ARACNE

The implementation of mutual information (MI)-based network inference approaches from the R package minet[58] was used. Pairwise MI was estimated based on the Spearman correlation as proposed by Olsen et al.[59]. Two networks were constructed based on these MI estimates. Using the CLR approach[60] non-significant interactions were removed based on the z-scores calculated from the marginal distributions of MI values for each gene pair. Alternatively, the ARACNE algorithm[44] was used to prune the network based on the data processing inequality. For both networks only interactions originating from a TF were taken into consideration and edges were ranked according to the assigned MI value.

Deconvolution and Silencing

For the two networks based on decomposition of the interaction matrix *G* the Pearson correlation matrix obtained from gene expression values was used as input.

The deconvolution approach introduced by Feizi et al.[61] was implemented as previously described[46]. The eigenvalue scaling factor β was initialized as β=0.9 and iteratively reduced in increments of 0.05 until the largest eigenvector of the direct interaction matrix generated by deconvolution was smaller than 1. Edges were ranked according to the deconvoluted interaction matrix.

The Silencing approach as described by Barzel et al.[45] was implemented in R. The proposed approximation of the direct interaction matrix *S* in which spurious interactions are silenced relies on the invers of the observed correlation matrix *G*. In our implementation we used the Moore-Penrose pseudoinverse in case *G* was close to singular. In the resulting network edges were ranked according to the approximated silenced interaction matrix.

Consensus network construction

To improve network quality[30], we built a consensus network integrating the GRN models inferred by the different approaches introduced above. To this end, we used the Borda count election method[47] whereby the rank *r* of an interaction *I* in the consensus network built on the predictions from *k* approaches is given by arithmetic mean of the ranks in the individual networks

$$r_{consens}(I) = \frac{\sum_{i=1}^{k} r_i(I)}{k}. \quad (3)$$

Following the reasoning of Feizi et al. [61] in this integration only the top 10% all possible edges in the GRN (625815) were considered from each individual ranking. For an edge that was not assigned a rank by some approaches, the missing ranks were set to 10% of all possible edges plus one.

Using this integration method, we assembled a consensus network based on all approaches to compare predictions from all GRN inference approaches (**Extended Data Fig. 1**). Due to this comparison and their inability to recover known interactions (**Fig. 1b**) the rankings derived from ARACNE and Silencing were only considered in **Extended Data Fig. 1** and excluded from the final consensus network used for all other analyses. As with the individual networks returned by the different approaches the consensus network **(Supplemental Table 3)** was trimmed to the top 10% of all possible edges according to the integrated ranks. The weight of edges in the final network was set as $r_{consens}^{-1}$.

PHOT-specific network

To investigate the PHOT-specific regulatory interactions genes that are differentially expressed between phot mutant and wt under low and high light were inferred. To this end, transcript counts of genes with more than 1 count per million in at least four replicates from these conditions (**Supplementary Table 1**) were tested for differential expression using the R packages limma[62], DeSeq2[63], and edgeR[56]. Only genes deemed significant by all three tools after Benjamini-Hochberg correction for a false discovery rate of 0.05 were considered differentially expressed with respect to PHOT mutation

In the next step, we focused on the normalized and scaled expression levels from these differentially expressed genes and the previously mentioned conditions, to infer a PHOT-specific GRN using GENIE3. To improve robustness of this network, which was obtained from a comparably sparse data set, we only considered the edges in the intersection with the final consensus network. Again, for both networks only the top 10% of possible edges were taken into account. Therefore, the obtained PHOT-network represents a subnetwork of the final consensus in which edges are weighted by the "PHOT specific" GENIE3 importance measure **(Supplementary Table 5)**.

Identification of major regulators

We compiled a manually curated list of possible target genes known to be involved in the processes of qE (*LHCSR1, LHCSR3.1/2, PSBS1/2*), photoprotection (**Supplementary Table 7**), and CCM (**Supplementary Table 8**). Based on the assumption that major regulators act on several genes important for a biological process, the regulatory strength of a candidate regulator (for the given process) was determined by the sum of edge weights w_{ij}between this regulator and the k genes in the respective target gene set

$$C(TF) = \sum_{j=1}^{k} w_{TFj}. \quad (4)$$

Empirical p-value calculation using Monte-Carlo simulation

The one-sided p-value for the overlap between the regulators of CCM and photoprotective genes was approximated by sampling the overlaps of random gene sets. To this end, we compiled two gene sets with the same cardinality as the curated CCM and photoprotective genes. The genes in these sets where randomly sampled without replacement from all targets in the respective networks. The 10 strongest regulators of these two gene sets where then obtained as previously described and the overlap was calculated as our sample statistic. This process was repeated 10,000 times and an empirical p-value was calculated from the number of iterations, r, where the overlap was higher or equal to the observed value, and the total number of iterations, n [64]:

$$p = \frac{r+1}{n+1}. \quad (5)$$

Strains and conditions

C. reinhardtii strains were grown under 15 µmol photons m^{-2} s^{-1}) in Tris-acetate-phosphate (TAP) media[65] at 23 °C in Erlenmeyer flasks shaken at 125 rpm. For all experiments cells were transferred to Sueoka's high salt medium (HSM)[66] at 1 million cells mL^{-1} and exposed to light intensities as described in the text and figure legends. For the investigation of the impact of acetate on the genome-wide transcriptome, HSM was supplemented with 20 mM sodium

acetate. *C. reinhardtii* strain CC-125 mt+ was used as WT. The *phot* (depleted from *PHOT1*; gene ID: Cre03.g199000)*, was* previously generated[67] and recently characterized[25]. For synchronized cultures, the cells were grown in HSM for at least 5 days under a 12h light/12h dark cycle (light intensity was set at 15 µmol photons m^{-2} s^{-1}; temperature was 18 °C in the dark and 23 °C in the light). All CLiP mutant strains used in this study and their parental strain (CC-4533) were obtained from the CLiP library (REF); *qer1* (LMJ.RY0402.072278), *qer4* (LMJ.RY0402.202963), *qer6* (LMJ.RY0402.162350), *qer7* (LMJ.RY0402.118995). The *lcr1* (strain C44)*, lcr1-C* (strain C44-B7) and its parental strain Q30P3 as described in[17] were a kind gift from Hideya Fukuzawa. Before performing phenotyping experiments, we first confirmed that *lcr1* shows no expression of *LCR1* and that this is rescued in the *lcr1-C* strain (**Extended Data Fig. 10a**). The *lci8* overexpressing line was purchased from the Chlamydomonas Resource center; strain CSI_FC1G01, expressing pLM005-Cre02.g144800-Venus-3xFLAG in the CC-4533 background. Overexpression of LCI8-FLAG was verified by immunoblotting against FLAG (**Extended Data Fig. 10b**).

To complement *qer1*, a 1152 bp genomic DNA fragment from *Chlamydomonas* CC-4533 was amplified by PCR using KOD hot start DNA polymerase (Merck) and primers P11 and P12 (**Supplementary Table 9**). To complement *qer7*, a 5755 bp fragment DNA fragment from *Chlamydomonas* CC-4533 was amplified by PCR with Platinum superfii DNA Polymerase (Thermo Fisher Scientific) and primers P13 and P14 (**Supplementary Table 9**). The PCR products were gel purified and cloned into pRAM118[68] by Gibson assembly[69] for expression under control of the *PSAD* promoter. Junctions and insertion were sequenced and constructs were linearized by EcoRV before transformation. Eleven ng/kb of linearized plasmid[70] mixed with 400 µL of 1.0 $x10^7$ cells mL^{-1} were electroporated in a volume of 120 mL in a 2-mm-gap electro cuvette using a NEPA21 square-pulse electroporator (NEPAGENE, Japan). The electroporation parameters were set as follows: Poring Pulse (300V; 8 ms length; 50 ms interval; one pulse; 40% decay rate; + Polarity), Transfer Pluse (20V; 50 ms length; 50 ms interval; five pulses; 40% decay rate; +/- Polarity). Transformants were selected onto solid agar plates containing 20 µg/ml hygromycin and screened for fluorescence by using a Tecan fluorescence microplate reader (Tecan Group Ltd., Switzerland). Parameters used were as follows: YFP (excitation 515/12 nm and emission 550/12 nm) and chlorophyll (excitation 440/9 nm and 680/20 nm). Transformants showing high YFP/chlorophyll value were further analyzed by real time qPCR.

Unless otherwise stated, LL conditions corresponded to 15 µmol photons m^{-2} s^{-1} while HL conditions corresponded to 300 µmol photons m^{-2} s^{-1} of white light (Neptune L.E.D., France; see[7] for light spectrum). All experiments were repeated at least three times to verify their reproducibility.

DNA Isolation and genotyping of CLiP mutants.

Total genomic DNA from CLiP mutants and corresponding wild-type strain CC-4533 was extracted according to the protocol suggested by CLiP website (https://www.chlamylibrary.org/). One µl of the extracted DNA was used as a template for the PCR assays, using Phire Plant Direct PCR polymerase (Thermo Fisher Scientific). To confirm the CIB1 insertion site in the CLiP mutants, gene-specific primers were used that anneal upstream and downstream of the predicted insertion site of the cassette (primer pairs P3-P4, P7-P8, P9-P10 and P5-P6 for *qer6*, *qer1*, *qer7* and *qer4* respectively*;* **Supplementary Table 9**). While all these primers worked in DNA extracted from WT, they

did not work in the DNA extracted from the mutants, with the exception of *qer4* (**Extended Data Fig. 3**), therefore primers specific for the 5' and 3' ends of the CIB1 Cassette were additionally used. All the primers used for genotyping were shown in **Supplementary Table 9**. We further confirmed the disruption of the genes of interest by quantifying their mRNA accumulation (**Extended Data Fig. 3**).

mRNA quantification

Total RNA was extracted using the RNeasy Mini Kit (Qiagen) and treated with the RNase-Free DNase Set (Qiagen). 1 µg total RNA was reverse transcribed with oligo dT using Sensifast cDNA Synthesis kit (Meridian Bioscience, USA). qPCR reactions were performed and quantitated in a Bio-Rad CFX96 system using SsoAdvanced Universal SYBR Green Supermix (BioRad). The primers (0.3 µM) used for qPCR are listed in **Supplementary Table 10**. A gene encoding G protein subunit-like protein (GBLP)[71] was used as the endogenous control , and relative expression values relative to *GBLP* were calculated from three biological replicates, each of which contained three technical replicates. All primers using for qPCR (**Supplementary Table 10**) were confirmed as having at least 90% amplification efficiency. In order to conform mRNA accumulation data to the distributional assumptions of ANOVA, i.e. the residuals should be normally distributed and variances should be equal among groups, Two-Way Analysis of Variance were computed with log-transformed data Y=logX where X is mRNA accumulation[72].

Immunoblotting

Protein samples of whole cell extracts (0.5 µg chlorophyll or 10 µg protein) were loaded on 4-20% SDS-PAGE gels (Mini-PROTEAN TGX Precast Protein Gels, Bio-Rad) and blotted onto nitrocellulose membranes. Antisera against LHCSR1 (AS14 2819), LHCSR3 (AS14 2766), ATPB (AS05 085), CAH4/5 (AS11 1737) were from Agrisera (Vännäs, Sweden); antiserum against PSBS was from ShineGene Molecular Biotech (Shanghai, China) targeting the peptides described in Ref.[9]. ATPB was used as a loading control. An anti-rabbit horseradish peroxidase–conjugated antiserum was used for detection. The blots were developed with ECL detection reagent, and images of the blots were obtained using a CCD imager (ChemiDoc MP System, Bio-Rad). For the densitometric quantification, data were normalized with ATPB.

Fluorescence-based measurements

Fluorescence-based photosynthetic parameters were measured with a a pulse modulated amplitude fluorimeter (MAXI-IMAGING-PAM, HeinzWaltz GmbH, Germany). Prior to the onset of the measurements, cells were acclimated to darkness for 15 min. Chlorophyll fluorescence was recorded during 10 min under 570 µmol m^{-2} s^{-1} of actinic blue light followed by finishing with 10 min of measurements of fluorescence relaxation in the dark. A saturating pulse (200 msec) of blue light (6000 µmol photons m^{-2} sec^{-1}) was applied for determination of Fm (the maximal fluorescence yield in dark-adapted state) or Fm' (maximal fluorescence in any light-adapted state). NPQ was calculated as (Fm − Fm')/Fm' based on[73]; qE was estimated as the fraction of NPQ that is rapidly inducible in the light and reversible in the dark.

References

1. Eberhard, S., Finazzi, G. & Wollman, F. A. The dynamics of photosynthesis. *Annu. Rev. Genet.* **42,** 463–515 (2008).

2. Erickson, E., Wakao, S. & Niyogi, K. K. Light stress and photoprotection in Chlamydomonas reinhardtii. *Plant J.* **82,** 449–465 (2015).

3. Moroney, J. V. & Ynalvez, R. A. Proposed carbon dioxide concentrating mechanism in Chlamydomonas reinhardtii. *Eukaryotic Cell* **6,** 1251–1259 (2007).

4. Niyogi, K. K. PHOTOPROTECTION REVISITED: Genetic and Molecular Approaches. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **50,** 333–359 (1999).

5. Peers, G. *et al.* An ancient light-harvesting protein is critical for the regulation of algal photosynthesis. *Nature* **462,** 518–521 (2009).

6. Dinc, E. *et al.* LHCSR1 induces a fast and reversible pH-dependent fluorescence quenching in LHCII in Chlamydomonas reinhardtii cells. *Proc. Natl. Acad. Sci. U.S.A.* **113,** 7673–7678 (2016).

7. Ruiz-Sola, M. *et al.* Photoprotection is regulated by light-independent CO2 availability. *bioRxiv* (2021). https://doi.org/10.1101/2021.10.23.465040

8. Li, X. P. *et al.* A pigment-binding protein essential for regulation of photosynthetic light harvesting. *Nature* **403,** 391–395 (2000).

9. Correa-Galvis, V. *et al.* Photosystem II Subunit PsbS Is Involved in the Induction of LHCSR Protein-dependent Energy Dissipation in Chlamydomonas reinhardtii. *J. Biol. Chem.* **291,** 17478–17487 (2016).

10. Tibiletti, T., Auroy, P., Peltier, G. & Caffarri, S. Chlamydomonas reinhardtii PsbS Protein Is Functional and Accumulates Rapidly and Transiently under High Light. *Plant Physiol.* **171,** 2717–2730 (2016).

11. Allorent, G. *et al.* UV-B photoreceptor-mediated protection of the photosynthetic machinery in Chlamydomonas reinhardtii. *Proc. Natl. Acad. Sci. U.S.A.* **113,** 14864–14869 (2016).

12. Redekop, P. *et al.* PsbS contributes to photoprotection in Chlamydomonas reinhardtii independently of energy dissipation. *BIOCHIMICA ET BIOPHYSICA ACTA-BIOENERGETICS* **1861,** 148183 (2020).

13. Hanawa, Y., Watanabe, M., Karatsu, Y., Fukuzawa, H. & Shiraiwa, Y. Induction of a high-CO2-inducible, periplasmic protein, H43, and its application as a high-CO2-responsive marker for study of the high-CO2-sensing mechanism in Chlamydomonas reinhardtii. *Plant Cell Physiol.* **48,** 299–309 (2007).

14. Moroney, J. V. *et al.* Isolation and Characterization of a Mutant of Chlamydomonas reinhardtii Deficient in the CO2 Concentrating Mechanism. *Plant Physiol.* **89,** 897–903 (1989).

15. Fukuzawa, H. *et al.* Ccm1, a regulatory gene controlling the induction of a carbon-concentrating mechanism in Chlamydomonas reinhardtii by sensing CO2 availability. *Proc. Natl. Acad. Sci. U.S.A.* **98,** 5347–5352 (2001).

16. Xiang, Y., Zhang, J. & Weeks, D. P. The Cia5 gene controls formation of the carbon concentrating mechanism in Chlamydomonas reinhardtii. *Proc. Natl. Acad. Sci. U.S.A.* **98,** 5341–5346 (2001).

17. Yoshioka, S. *et al.* The novel Myb transcription factor LCR1 regulates the CO2-responsive gene Cah1, encoding a periplasmic carbonic anhydrase in Chlamydomonas reinhardtii. *Plant Cell* **16,** 1466–1477 (2004).

18. Petroutsos, D. *et al.* A blue-light photoreceptor mediates the feedback regulation of photosynthesis. *Nature* **537,** 563–566 (2016).

19. Petroutsos, D. *et al.* The chloroplast calcium sensor CAS is required for photoacclimation in *Chlamydomonas reinhardtii*. *Plant Cell* **23,** 2950–2963 (2011).

20. Maruyama, S., Tokutsu, R. & Minagawa, J. Transcriptional regulation of the stress-responsive light harvesting complex genes in *Chlamydomonas reinhardtii*. *Plant Cell Physiol.* **55,** 1304–1310 (2014).

21. Miura, K. *et al.* Expression profiling-based identification of CO2-responsive genes regulated by CCM1 controlling a carbon-concentrating mechanism in Chlamydomonas reinhardtii. *Plant Physiol.* **135,** 1595–1607 (2004).

22. Fang, W. *et al.* Transcriptome-wide changes in Chlamydomonas reinhardtii gene expression regulated by carbon dioxide and the CO2-concentrating mechanism regulator CIA5/CCM1. *Plant Cell* **24,** 1876–1893 (2012).

23. Brueggeman, A. J. *et al.* Activation of the carbon concentrating mechanism by CO2 deprivation coincides with massive transcriptional restructuring in Chlamydomonas reinhardtii. *Plant Cell* **24,** 1860–1875 (2012).

24. Aihara, Y., Fujimura-Kamada, K., Yamasaki, T. & Minagawa, J. Algal photoprotection is regulated by the E3 ligase CUL4-DDB1DET1. *Nature Plants* **5,** 34–40 (2019).

25. Redekop, P. *et al.* Transcriptional regulation of photoprotection in dark-to-light transition- more than just a matter of excess light energy. *BioRxiv* (2021) https://doi.org/10.1101/2021.10.23.463292 (accepted for publication in *Science Advances*)

26. Gabilly, S. T. *et al.* Regulation of photoprotection gene expression in Chlamydomonas by a putative E3 ubiquitin ligase complex and a homolog of CONSTANS. *Proc. Natl. Acad. Sci. U.S.A.* **116,** 17556–17562 (2019).

27. Tokutsu, R., Fujimura-Kamada, K., Matsuo, T., Yamasaki, T. & Minagawa, J. The CONSTANS flowering complex controls the protective response of photosynthesis in the green alga Chlamydomonas. *Nat Comms* **10,** 655–10 (2019).

28. Kamrani, Y. Y., Matsuo, T., Mittag, M. & Minagawa, J. ROC75 is an Attenuator for the Circadian Clock that Controls LHCSR3 Expression. *Plant Cell Physiol.* **59,** 2602–2607 (2018).

29. Omranian, N., Eloundou-Mbebi, J. M. O., Mueller-Roeber, B. & Nikoloski, Z. Gene regulatory network inference using fused LASSO on multiple data sets. *Sci Rep* **6,** 1–14 (2016).

30. Marbach, D. *et al.* Wisdom of crowds for robust gene network inference. *Nat. Methods* **9,** 796–804 (2012).

31. Fang, L. *et al.* GRNdb: Decoding the gene regulatory networks in diverse human and mouse conditions. *Nucl. Acids Res.* **49,** D97–D103 (2021).

32. Omranian, N. & Nikoloski, Z. in *Methods in Molecular Biology* **1629,** 283–295 (2017).

33. Huynh-Thu, V. A. & Sanguinetti, G. in *Methods in Molecular Biology* **1883,** 1–23 (2019).

34. Schäfer, J. & Strimmer, K. A shrinkage approach to large-scale covariance matrix estimation and implications for functional genomics. *Stat Appl Genet Mol Biol* **4,** Article32 (2005).

35. Zou, H. & Hastie, T. Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **67,** 301–320 (2005).

36. Gargouri, M. *et al.* Identification of regulatory network hubs that control lipid metabolism in Chlamydomonas reinhardtii. *J. Exp. Bot.* **66,** 4551–4566 (2015).

37. Lämmermann, N. *et al.* Ubiquitin ligase component LRS1 and transcription factor CrHy5 act as a light switch for photoprotection in <em>Chlamydomonas</em&gt. *bioRxiv* 2020.02.10.942334 (2020).

38. Zones, J. M., Blaby, I. K., Merchant, S. S. & Umen, J. G. High-Resolution Profiling of a Synchronized Diurnal Transcriptome from Chlamydomonas reinhardtii Reveals Continuous Cell and Metabolic Differentiation. *Plant Cell* **27,** 2743–2769 (2015).

39. Strenkert, D. *et al.* Multiomics resolution of molecular events during a day in the life of Chlamydomonas. *Proc. Natl. Acad. Sci. U.S.A.* **116,** 2374–2383 (2019).

40. Fett, J. P. & Coleman, J. R. Regulation of Periplasmic Carbonic Anhydrase Expression in Chlamydomonas reinhardtii by Acetate and pH. *Plant Physiol.* **106,** 103–108 (1994).

41. Pérez-Rodríguez, P. *et al.* PlnTFDB: Updated content and new features of the plant transcription factor database. *Nucl. Acids Res.* **38,** (2009).

42. Jin, J. *et al.* PlantTFDB 4.0: Toward a central hub for transcription factors and regulatory interactions in plants. *Nucl. Acids Res.* **45,** D1040–D1045 (2017).

43. Fischer, B. B. *et al.* SINGLET OXYGEN RESISTANT 1 links reactive electrophile signaling to singlet oxygen acclimation in Chlamydomonas reinhardtii. *Proc. Natl. Acad. Sci. U.S.A.* **109,** E1302–11 (2012).

44. Margolin, A. A. *et al.* ARACNE: An algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context. *BMC Bioinformatics* **7,** 1–15 (2006).

45. Barzel, B. & Barabási, A.-L. Network link prediction by global silencing of indirect correlations. *Nature Biotechnology* **31,** 720–725 (2013).

46. Omranian, N., Eloundou-Mbebi, J. M. O., Mueller-Roeber, B. & Nikoloski, Z. Gene regulatory network inference using fused LASSO on multiple data sets. *Sci Rep* **6,** 20533 (2016).

47. de Borda, J.-C. Mémoire sur les élections au scrutin. *Histoire de lAcademie Royale des Sciences* (1781).

48. Li, X. *et al.* A genome-wide algal mutant library and functional screen identifies genes required for eukaryotic photosynthesis. *Nat. Genet.* **51,** 627–635 (2019).

49. UniProt Consortium. UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Res.* **49,** D480–D489 (2021).

50. Mi, H. *et al.* PANTHER version 16: A revised family classification, tree-based classification tool, enhancer regions and extensive API. *Nucl. Acids Res.* **49,** D394–D403 (2021).

51. Thiriet-Rupert, S. *et al.* Transcription factors in microalgae: genome-wide prediction and comparative analysis. *BMC Genomics* **17,** 282 (2016).

52. Bushnell, B. BBMap: A Fast, Accurate, Splice-Aware Aligner. *Lawrence Berkeley National Laboratory. LBNL Report LBNL-E* (2014).

53. Law, C. W., Chen, Y., Shi, W. & Smyth, G. K. Voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biology* **15,** 1–17 (2014).

54. Anders, S., Pyl, P. T. & Huber, W. HTSeq-A Python framework to work with high-throughput sequencing data. *Bioinformatics* **31,** 166–169 (2015).

55. Robinson, M. D. & Oshlack, A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biology* **11,** 1–9 (2010).

56. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26,** 139–140 (2009).

57. Huynh-Thu, V. A., Irrthum, A., Wehenkel, L. & Geurts, P. Inferring regulatory networks from expression data using tree-based methods. *PLoS ONE* **5,** e12776 (2010).

58. Meyer, P. E., Lafitte, F. & Bontempi, G. Minet: A r/bioconductor package for inferring large transcriptional networks using mutual information. *BMC Bioinformatics* **9,** 1–10 (2008).

59. Olsen, C., Meyer, P. E. & Bontempi, G. On the impact of entropy estimation on transcriptional regulatory network inference based on mutual information. *Eurasip Journal on Bioinformatics and Systems Biology* **2009,** (2009).

60. Faith, J. J. *et al.* Large-scale mapping and validation of Escherichia coli transcriptional regulation from a compendium of expression profiles. *PLoS Biol.* **5,** 0054–0066 (2007).

61. Feizi, S., Marbach, D., Médard, M. & Kellis, M. Network deconvolution as a general method to distinguish direct dependencies in networks. *Nature Biotechnology* **31,** 726–733 (2013).

62. Ritchie, M. E. *et al.* Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucl. Acids Res.* **43,** e47 (2015).

63. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15,** 1–21 (2014).

64. North, B. V., Curtis, D. & Sham, P. C. A note on the calculation of empirical P values from Monte Carlo procedures [1]. *American Journal of Human Genetics* **71,** 439–441 (2002).

65. Gorman, D. S. & Levine, R. P. Cytochrome f and plastocyanin: their sequence in the photosynthetic electron transport chain of *Chlamydomonas reinhardii*. *Proc. Natl. Acad. Sci. U.S.A.* **54,** 1665–1669 (1965).

66. Sueoka, N. Mitotic replication of deoxyribonucleic acid in *Chlamydomonas reinhardii*. *Proc. Natl. Acad. Sci. U.S.A.* **46,** 83–91 (1960).

67. Greiner, A. *et al.* Targeting of Photoreceptor Genes in Chlamydomonas reinhardtii via Zinc-Finger Nucleases and CRISPR/Cas9. *Plant Cell* **29,** 2498–2518 (2017).

68. Kaye, Y. *et al.* The mitochondrial alternative oxidase from Chlamydomonas reinhardtii enables survival in high light. *J. Biol. Chem.* **294,** 1380–1395 (2019).

69. Gibson, D. G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6,** 343–345 (2009).

70. Mackinder, L. C. M. *et al.* A repeat protein links Rubisco to form the eukaryotic carbon-concentrating organelle. *Proc. Natl. Acad. Sci. U.S.A.* **113,** 5958–5963 (2016).

71. Schloss, J. A. A Chlamydomonas gene encodes a G protein beta subunit-like polypeptide. *Molec. Gen. Genet.* **221,** 443–452 (1990).

72. Schlesselman, J. J. Data Transformation in Two-Way Analysis of Variance. *Journal of the American Statistical Association* (2012). doi:10.1080/01621459.1973.10482435

73. Genty, B., Briantais, J.-M. & Baker, N. R. The relationship between the quantum yield of photosynthetic electron transport and quenching of chlorophyll fluorescence. *Biochimica et Biophysica Acta (BBA)* **990,** 87–92 (1989).

Acknowledgments

We are grateful to Prof. Hideya Fukuzawa for sending us *lcr1*, *lcr1-C* and their respective WT strains and to Prof. Peter Jahns for the antibody against PSBS.

Funding:

The authors would like to thank the following agencies for funding: The Human Frontiers Science Program through the funding of the project RGP0046/2018 (DP, ZN); the French National Research Agency in the framework of the Young Investigators program ANR-18-CE20-0006 through the funding of the project MetaboLight (DP); the French National Research Agency in the framework of the Investissements d'Avenir program ANR-15-IDEX-02, through the funding of the "Origin of Life" project of the Univ. Grenoble-Alpes (DP, YY); the French National Research Agency through the funding of the Grenoble Alliance for Integrated Structural & Cell Biology GRAL project ANR-17-EURE-0003 (DP, MAR-S), the Prestige Marie-Curie co-financing grant PRESTIGE-2017-1-0028 (MAR-S); the International Max Planck Research School 'Primary Metabolism and Plant Growth' at the Max Planck Institute of Molecular Plant Physiology (MA, ZN).

Competing interests: Authors declare that they have no competing interests.

Figures

Fig. 1: Characterization of the consensus GRN inferred by employing a compendium of RNA-seq data from diverse light and culture conditions. **a.** Expression levels of representative CCM and qE related genes are plotted over all samples used for network inference (z-scaled log values are depicted). The column annotation gives information on the culture conditions. The values of the rows of the heatmap correspond to the z-scores of representative CCM and qE genes plotted over all samples (columns) used in the network inference. **b.** The heatmap rows correspond to experimentally validated or falsified (neg) gene regulatory interactions involved in qE and CCM, curated from literature. The heatmap indicates ranking of these interactions by different approaches and the consensus network. Edges are considered highly ranked (depicted in blue) if they are above the 10% network density threshold. Edges ranked below this threshold are depicted in red. Edges that were not included in the given network are marked in grey. ARACNE and Silencing columns were only plotted for comparison and were not used in building the consensus GRN (see Methods).

Fig. 2: Consensus GRN for light and acetate responses pinpoints LCR1 as regulator of qE-related genes. a. Dot plot of the relative regulatory strength of the top 10 regulators of qE-related genes in the consensus GRN (see Methods). TFs are marked in green if qE transcript levels were affected in the respective knock-out strain and this effect was reversed by complementation with the missing gene. TFs for which no effect was observed are marked in red. TFs for which no mutant lines were available are plotted in grey. **b.** WT, *lcr1* and *lcr1-C* cells were acclimated for 16h in LL (15 µmol photons m^{-2} s^{-1}). After sampling for the LL conditions, light intensity was increased to 300 µmol photons m^{-2} s^{-1} (HL); samples were taken 1 h (RNA) or 4 h (protein and photosynthetic measurements) after exposure to HL. Shown are relative expression levels of qE-related genes at the indicated conditions normalized to WT LL (*n* = 3 biological samples, mean ± sd). **c**. Immunoblot analyses of LHCSR1, LHCSR3, PSBS and ATPB (loading control) of one of the three biological replicates, under the indicated conditions. **d.** Quantification of immunoblot data of all replicates in panel **c** after normalization to ATPB. Shown are the HL treated samples; WT protein levels were set as 1. **e.** NPQ and calculated qE (as an inset) 4h after exposure to HL (*n* = 3 biological samples, mean ± s.d). The p-values for the comparisons are based on ANOVA Dunnett's multiple comparisons test and as indicated in the graphs (*, $P < 0.005$, **, $P < 0.01$, ***, $P < 0.001$, ****, $P < 0.0001$). Statistical analyses for panel **b** and **d** were applied on log10-transformed values.

Fig. 3: A PHOT-specific GRN pinpoints QER7 as a suppressor of the expression of qE-related genes. a. Dot plot of the relative regulatory strength of the top 10 regulators of qE-related genes in the PHOT-specific GRN (see **Methods**). TFs are marked in green if qE transcript levels were affected in the respective knock-out strain and this effect was reversed by complementation with the knocked-out gene, in yellow, if the effect was not reversed by complementation, and in red, if no mutant effect was observed in the mutant. TFs for which no mutant lines were available are plotted in grey. **b**. WT, *qer7* and *qer7-C* cells were synchronized under 12h light (15 µmol m^{-2} s^{-1})/12h dark cycles. After sampling for the dark conditions (end of the dark phase), cells were exposed to 300 µmol photons m^{-2} s^{-1} (HL); samples were taken 1 h (RNA) or 4 h (protein and photosynthetic measurements) after

exposure to HL. Shown are relative expression levels of qE-related genes at the indicated conditions normalized to WT LL (*n* = 3 biological samples, mean ± sd). **c**. Immunoblot analyses of LHCSR1, LHCSR3 and ATPB (loading control) of one of the three biological replicate set of samples, under the indicated conditions. **d.** Quantification of immunoblot data of all replicates in panel **c** after normalization to ATPB. WT protein levels at HL were set as 1. **e.** NPQ and qE, measured 4h after exposure to HL (*n* = 3 biological samples, mean ± sd). The p-values for the comparisons are based on ANOVA Dunnett's multiple comparisons test and are indicated in the graphs (*, P < 0.005, **, P < 0.01, ***, P < 0.001, ****, P < 0.0001). Statistical analyses for panel **b** and **d** were applied on log10- transformed values.

Fig:4: QER7 suppresses transcription of CCM genes and depends on PHOT. WT, *qer7* and *qer7-C* cells were synchronized under 12h light (15 µmol m^{-2} s^{-1})/12h dark cycles. After sampling at the end of the dark phase, cells were exposed to 300 µmol photons m^{-2} s^{-1} (HL) and samples were taken 1 h after HL exposure. **a.** Relative expression levels of CCM genes at the indicated conditions normalized to WT LL (*n* = 3 biological samples, mean ± sd). **b and c.** Relative expression levels of *QER7* (**b**) and *LCR1* (**c**) in synchronized WT, *phot* and *phot-C* (phot cells complemented with the *PHOT* gene) cells. The p-values for the comparisons are based on ANOVA Dunnett's multiple comparisons test on log10- transformed values and are indicated in the graphs (*, P < 0.005, **, P < 0.01, ***, P < 0.001, ****, P < 0.0001).

Fig. 5: Consensus and PHOT-specific GRNs indicate extensive co-regulation of CCM and photoprotective genes. Top: Venn diagram depicting the overlap of the top 10 predicted TFs of the curate CCM and photoprotective (PhPr) genes based on **a.** the consensus or **c.** PHOT-specific GRN. Bottom: Network representation of the top 10 TF sets of **b.** the consensus network or **d.** the phot GRN (center nodes, same color code as in panel **a** and the target genes used for prediction (left and right columns of nodes, photoprotection genes are shown in green, qE-related genes in blue, and CCM genes in red. Larger TF nodes indicate higher regulatory strength on target genes. The plotted edge width is proportional to the strength of the regulatory interaction.

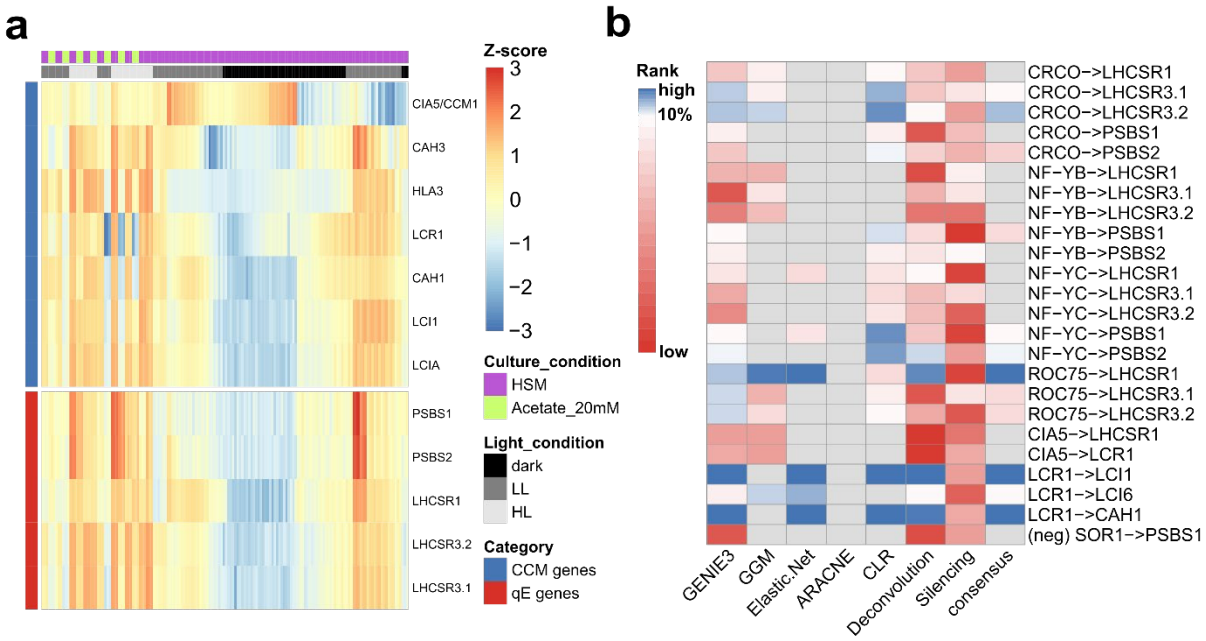


Fig. 1: Characterization of the consensus GRN inferred by employing a compendium of RNA-seq data from diverse light and culture conditions. **a.** Expression levels of representative CCM and qE related genes are plotted over all samples used for network inference (z-scaled log values are depicted). The column annotation gives information on the culture conditions. The values of the rows of the heatmap correspond to the z-scores of representative CCM and qE genes plotted over all samples (columns) used in the network inference. **b.** The heatmap rows correspond to experimentally validated or falsified (neg) gene regulatory interactions involved in qE and CCM, curated from literature. The heatmap indicates ranking of these interactions by different approaches and the consensus network. Edges are considered highly ranked (depicted in blue) if they are above the 10% network density threshold. Edges ranked below this threshold are depicted in red. Edges that were not included in the given network are marked in grey. ARACNE and Silencing columns were only plotted for comparison and were not used in building the consensus GRN (see Methods).

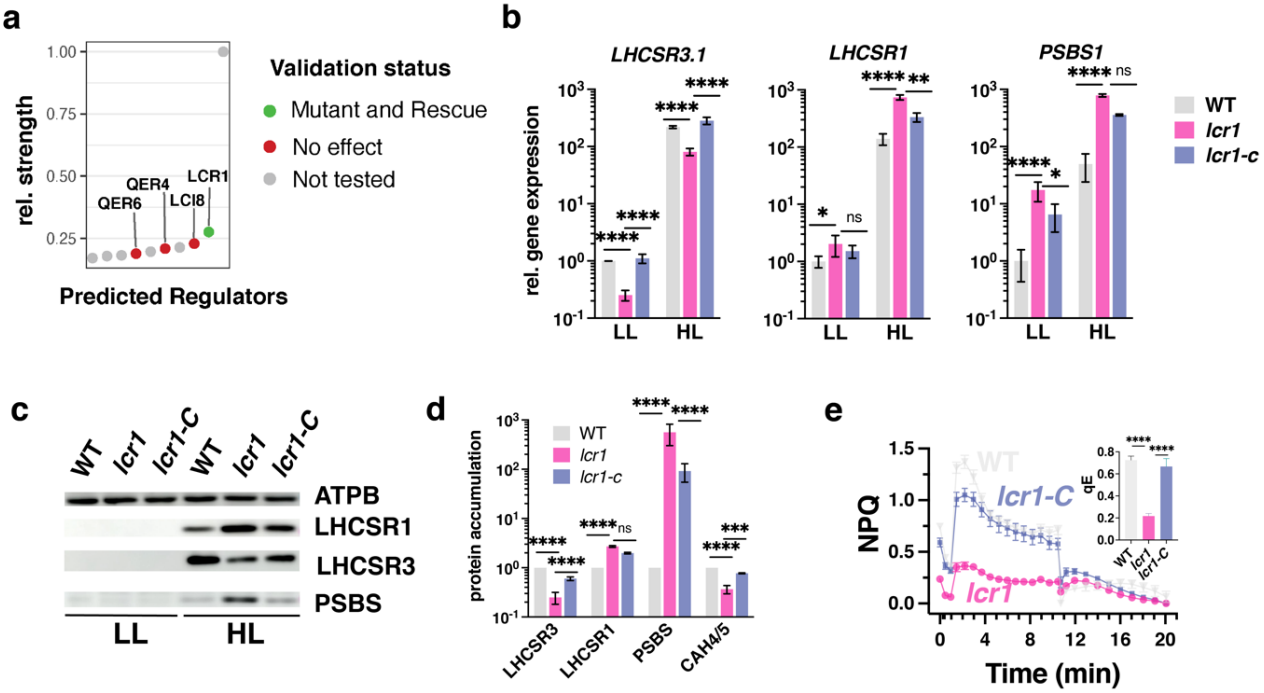


Fig. 2: Consensus GRN for light and acetate responses pinpoints LCR1 as regulator of qE-related genes. a. Dot plot of the relative regulatory strength of the top 10 regulators of qE-related genes in the consensus GRN (see Methods). TFs are marked in green if qE transcript levels were affected in the respective knock-out strain and this effect was reversed by complementation with the missing gene. TFs for which no effect was observed are marked in red. TFs for which no mutant lines were available are plotted in grey. **b.** WT, *lcr1* and *lcr1-C* cells were acclimated for 16h in LL (15 µmol photons m^{-2} s^{-1}). After sampling for the LL conditions, light intensity was increased to 300 µmol photons m^{-2} s^{-1} (HL); samples were taken 1 h (RNA) or 4 h (protein and photosynthetic measurements) after exposure to HL. Shown are relative expression levels of qE-related genes at the indicated conditions normalized to WT LL (*n* = 3 biological samples, mean ± sd). **c**. Immunoblot analyses of LHCSR1, LHCSR3, PSBS and ATPB (loading control) of one of the three biological replicates, under the indicated conditions. **d.** Quantification of immunoblot data of all replicates in panel **c** after normalization to ATPB. Shown are the HL treated samples; WT protein levels were set as 1. **e.** NPQ and calculated qE (as an inset) 4h after exposure to HL (*n* = 3 biological samples, mean ± s.d). The p-values for the comparisons are based on ANOVA Dunnett's multiple comparisons test and as indicated in the graphs (*, P < 0.005, **, P < 0.01, ***, P < 0.001, ****, P < 0.0001). Statistical analyses for panel **b** and **d** were applied on log10-transformed values.

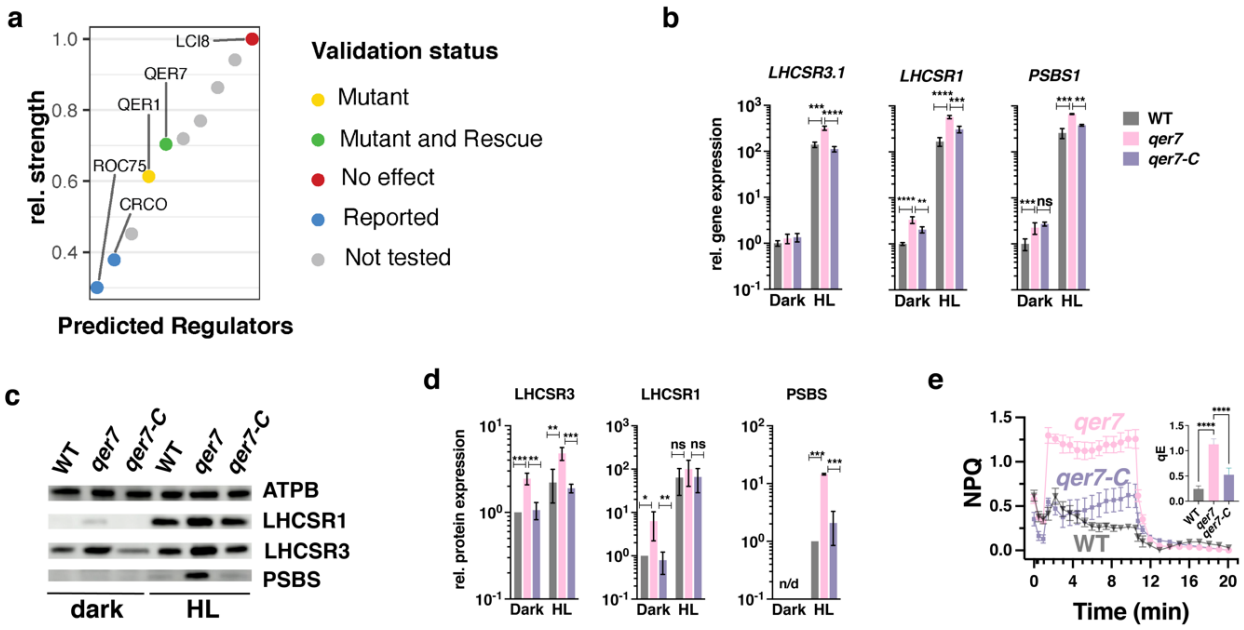


Fig. 3: A PHOT-specific GRN pinpoints QER7 as a suppressor of the expression of qE-related genes. a. Dot plot of the relative regulatory strength of the top 10 regulators of qE-related genes in the PHOT-specific GRN (see **Methods**). TFs are marked in green if qE transcript levels were affected in the respective knock-out strain and this effect was reversed by complementation with the knocked-out gene, in yellow, if the effect was not reversed by complementation, and in red, if no mutant effect was observed in the mutant. TFs for which no mutant lines were available are plotted in grey. **b**. WT, *qer7* and *qer7-C* cells were synchronized under 12h light (15 µmol m^{-2} s^{-1})/12h dark cycles. After sampling for the dark conditions (end of the dark phase), cells were exposed to 300 µmol photons m^{-2} s^{-1} (HL); samples were taken 1 h (RNA) or 4 h (protein and photosynthetic measurements) after exposure to HL. Shown are relative expression levels of qE-related genes at the indicated conditions normalized to WT LL (*n* = 3 biological samples, mean ± sd). **c**. Immunoblot analyses of LHCSR1, LHCSR3 and ATPB (loading control) of one of the three biological replicate set of samples, under the indicated conditions. **d.** Quantification of immunoblot data of all replicates in panel **c** after normalization to ATPB. WT protein levels at HL were set as 1. **e.** NPQ and qE, measured 4h after exposure to HL (*n* = 3 biological samples, mean ± sd). The p-values for the comparisons are based on ANOVA Dunnett's multiple comparisons test and are indicated in the graphs (*, P < 0.005, **, P < 0.01, ***, P < 0.001, ****, P < 0.0001). Statistical analyses for panel **b** and **d** were applied on log10- transformed values.

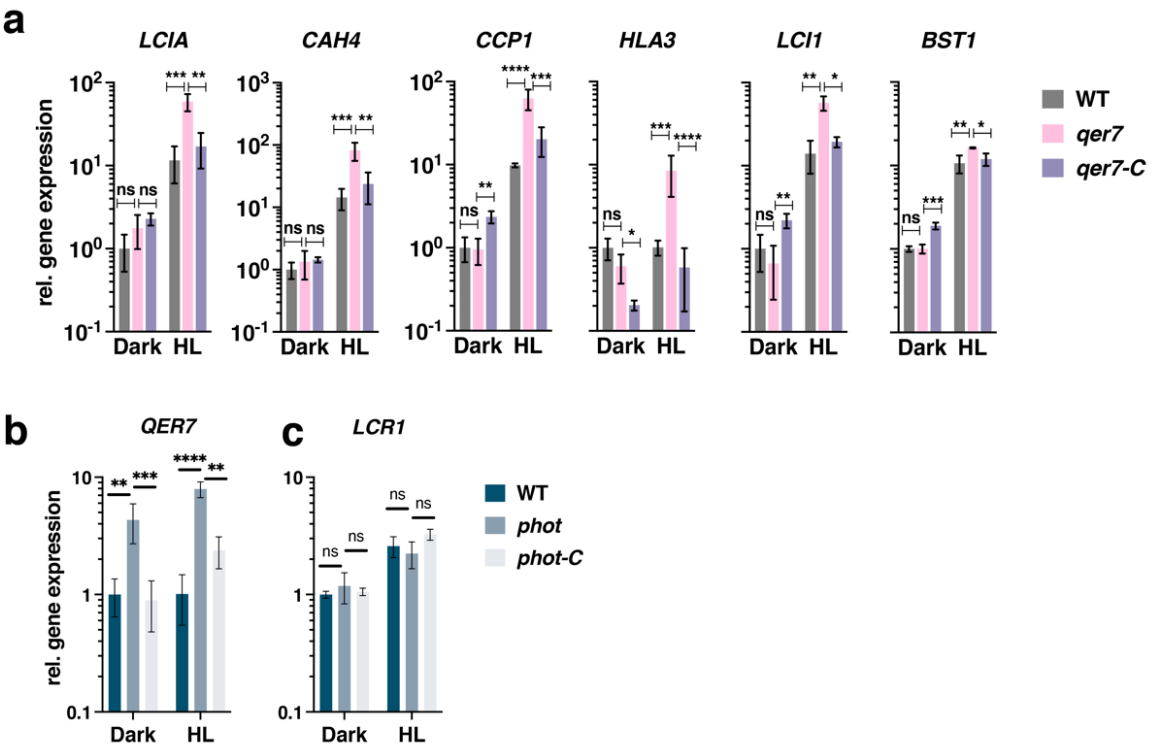


Fig:4: QER7 suppresses transcription of CCM genes and depends on PHOT. WT, *qer7* and *qer7-C* cells were synchronized under 12h light (15 µmol m^{-2} s^{-1})/12h dark cycles. After sampling at the end of the dark phase, cells were exposed to 300 µmol photons m^{-2} s^{-1} (HL) and samples were taken 1 h after HL exposure. **a.** Relative expression levels of CCM genes at the indicated conditions normalized to WT LL (*n* = 3 biological samples, mean ± sd). **b and c.** Relative expression levels of *QER7* (**b**) and *LCR1* (**c**) in synchronized WT, *phot* and *phot-C* (phot cells complemented with the *PHOT* gene) cells. The p-values for the comparisons are based on ANOVA Dunnett's multiple comparisons test on log10- transformed values and are indicated in the graphs (*, P < 0.005, **, P < 0.01, ***, P < 0.001, ****, P < 0.0001).

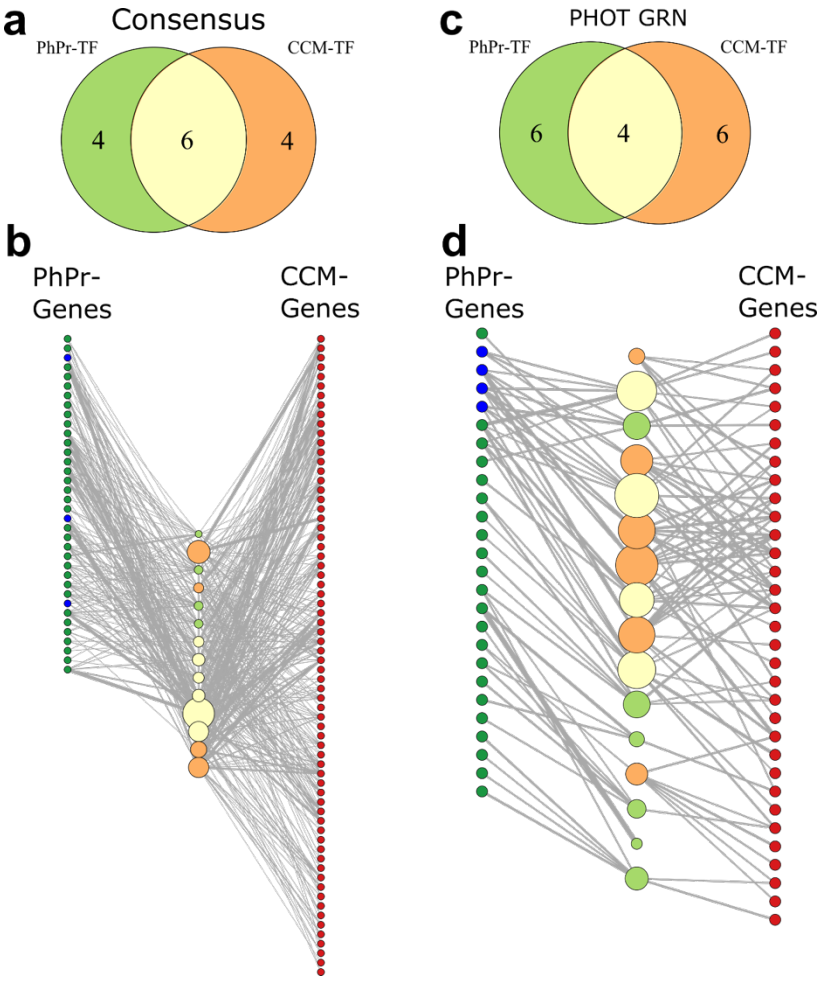


Fig. 5: Consensus and PHOT-specific GRNs indicate extensive co-regulation of CCM and photoprotective genes. Top: Venn diagram depicting the overlap of the top 10 predicted TFs of the curate CCM and photoprotective (PhPr) genes based on **a.** the consensus or **c.** PHOT-specific GRN. Bottom: Network representation of the top 10 TF sets of **b.** the consensus network or **d.** the phot GRN (center nodes, same color code as in panel **a** and the target genes used for prediction (left and right columns of nodes, photoprotection genes are shown in green, qE-related genes in blue, and CCM genes in red. Larger TF nodes indicate higher regulatory strength on target genes. The plotted edge width is proportional to the strength of the regulatory interaction.

Supplementary Materials for

Widening the landscape of transcriptional regulation of algal photoprotection

Authors

Marius Arend[1,2†], Yizhong Yuan[3†], M. Águila Ruiz-Sola[3‡], Nooshin Omranian[1,2,4], Zoran Nikoloski[1,2,4*], Dimitris Petroutsos[3 *]

[†] these authors contributed equally to this work

[*]**Corresponding Authors**: nikoloski@mpimp-golm.mpg.de (Z.N.), dimitris.petroutsos@cea.fr (D.P.)

This PDF file includes:

Supplementary Note

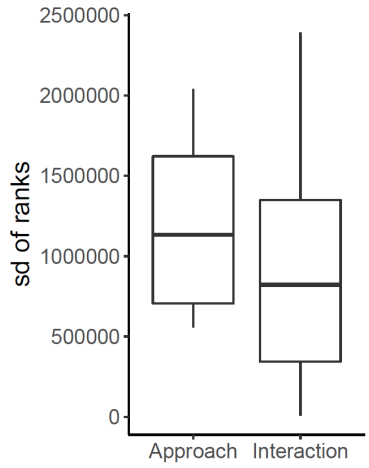
Extended Data Figs. 1 to 10

Other Supplementary Materials for this manuscript include the following:

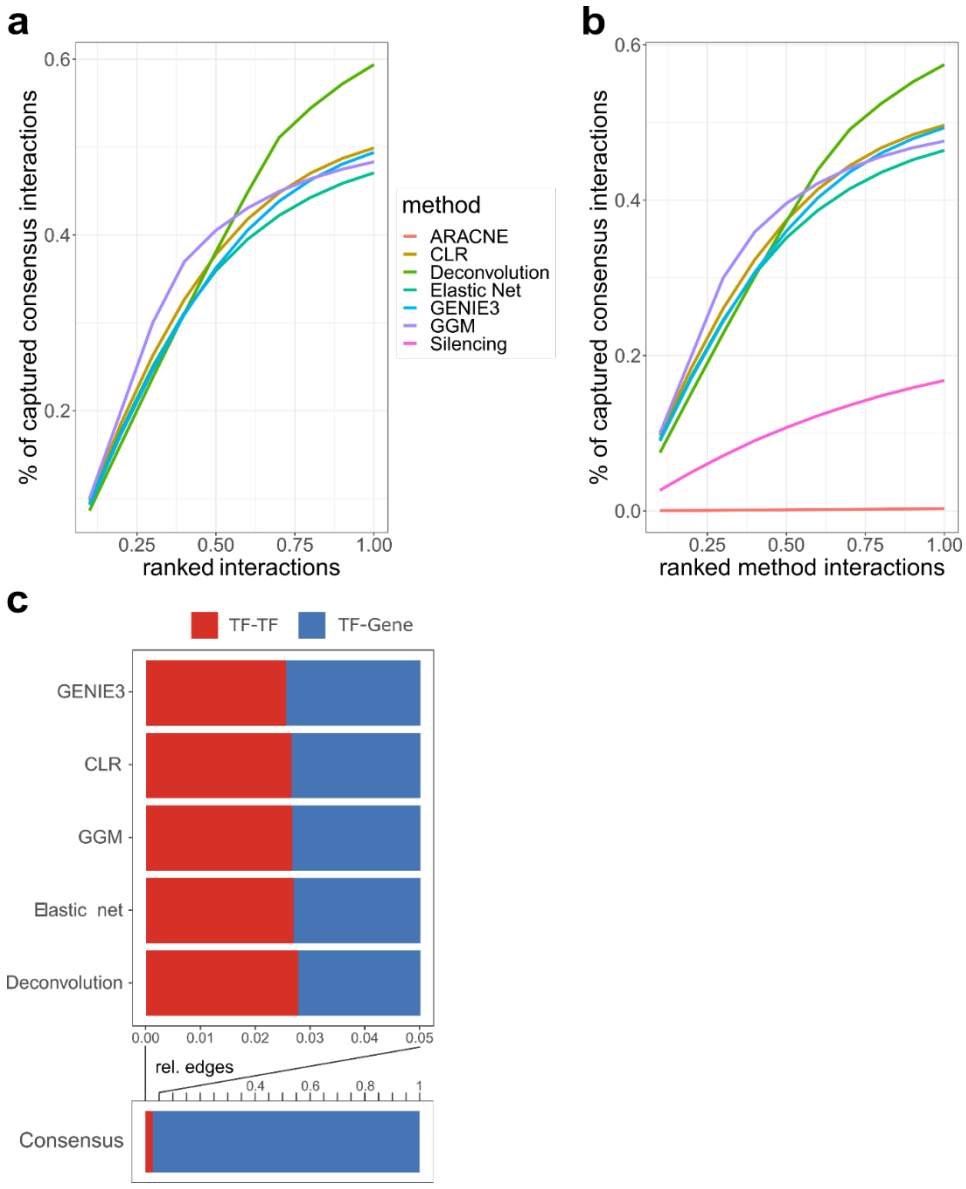
Tables S1 to S10 (to be downloaded as excel file)

Supplementary Note

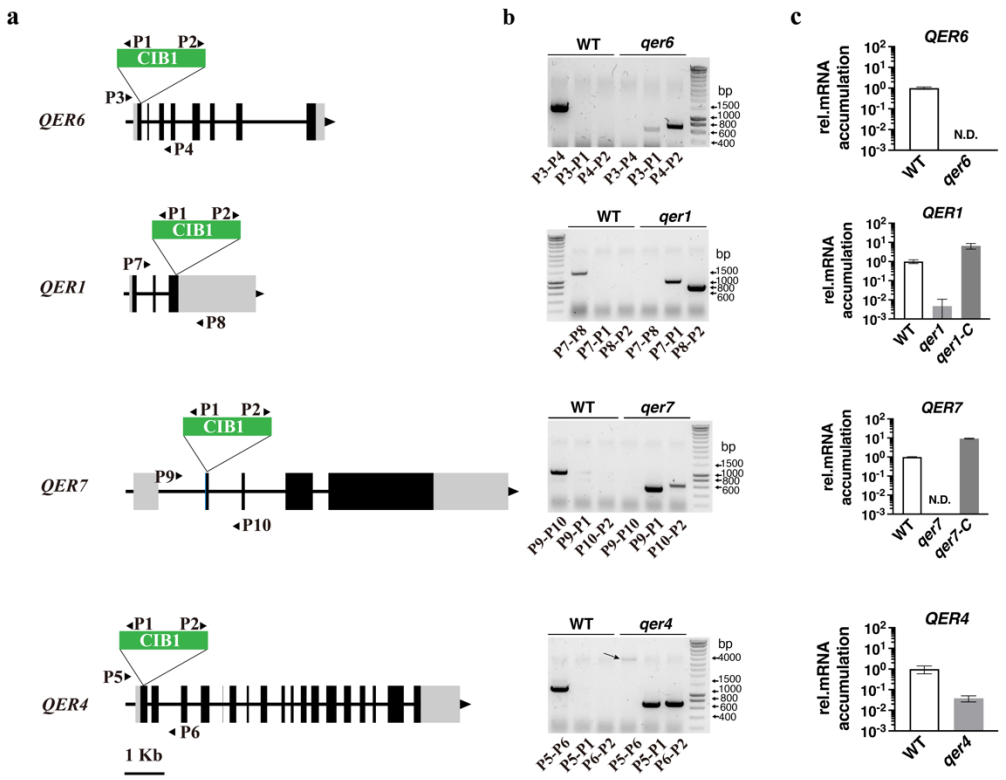
Simulation studies[30] indicated that the top 10% of the edges in the consensus network are enriched in positive interactions if the underlying approaches perform better than guessing. Thus, a higher overlap with the consensus networks is expected to result in improved predictions. We found that the interactions from the network deconvolution approach exhibited the largest overlap with the consensus network, while GGM assigns the highest rank to consensus interactions (**Extended Data Fig. 2a**). Repeating this analysis by including the two worse performing approaches (i.e. ARACNE and Silencing) resulted in no qualitative changes in the overlaps (**Extended Data Fig. 2b**) and only 7.47% of difference to the original consensus network, demonstrating the robustness of the inferred interactions. Considering the 10% network density threshold, we also inspected the proportion of TF-TF interactions in the different networks and their consensus; we found that it ranges from 2.55% in the GENIE3 network to 2.77% in the network deconvolution approach (**Extended Data Fig. 2c**). Since the fully connected network containing all possible TF-TF and TF-target interactions has a relative TF-TF content of 2.52%, these findings suggest an enrichment of TF-TF interactions in GRNs inferred by all five approaches considered in the consensus.



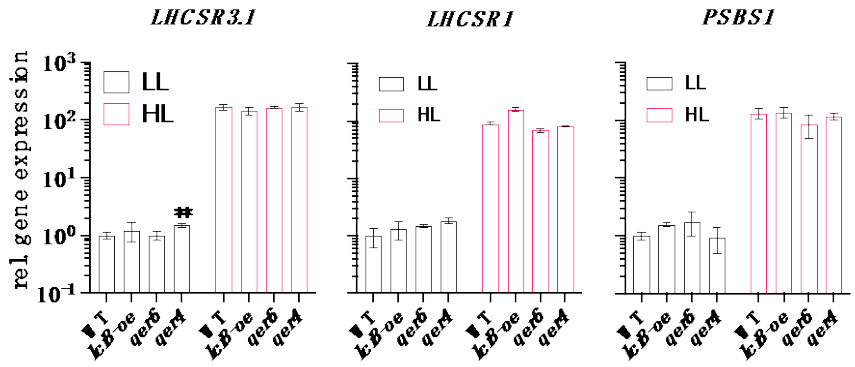
Extended Data Fig. 1: Variation in ranks of known interactions across approaches and interactions from different inference approaches. Boxplots of the standard deviation of ranks calculated for the curated interactions plotted in **Fig. 1b** over all ranks assigned by one approach or over all ranks assigned to one interaction.



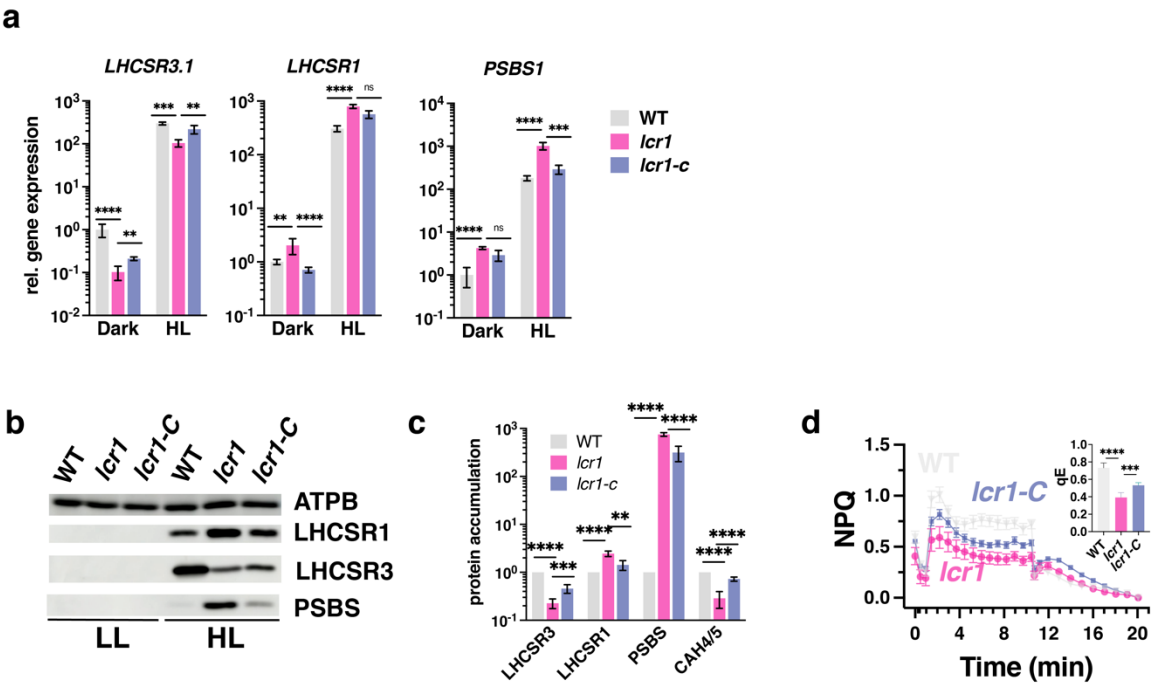
Extended Data Fig. 2: Overlap of consensus network with individual inference methods and proportion of TF-TF interactions in the obtained GRNs. a. The graph shows the overlap between the edges of the consensus network and the ranked interactions of the individual approaches normalized to the total number of edges in the consensus network used for regulator prediction. **b.** Same as in a but plotting the overlap with a consensus of all used approaches (including ARACNE and Silencing). **c**. Bar plots provide the proportion of TF-TF interactions in comparison to TF-target gene interactions contained in the GRNs inferred by the different approaches and the consensus GRN resulting from their integration. All depicted analyses considered only the interactions within in the 10% network density threshold (see **Methods**).



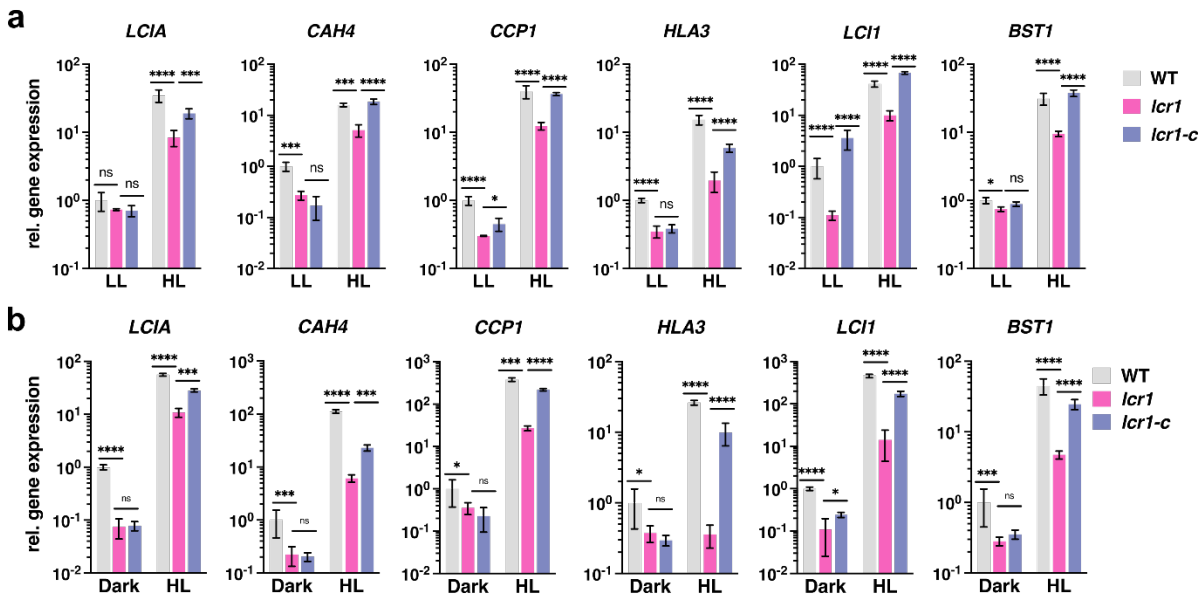
Extended Data Fig. 3: Genotyping of CLiP mutants affected in the predicted regulators of qE-related genes. a. Insertion map of the CIB1 cassette in the different genes. Exons are shown in black, introns as interconnecting lines, 5'UTR and 3'UTR in light gray, and primers in arrows. The insertion site of the CIB1 cassette is indicated by the triangle; **b**. PCR-validation of the insertion site in the different CLiP mutants using genomic DNA. To confirm the CIB1 insertion site, gene-specific primers were used that anneal upstream and downstream of the predicted insertion site of the cassette (primer pairs P3-P4, P7-P8, P9-P10 and P5-P6 for *qer6*, *qer1*, *qer7* and *qer4* respectively*;* **Supplementary Table 9**). Pairs of primers used are indicated at the bottom of the agarose gels used to separate the PCR products. Note that the PCR product of P5-P6 is indicated by an arrow. **c.** Relative expression levels of predicted qE regulator genes in the different CLiP mutants after exposure to HL (300 µmol photons m^{-2} s^{-1}) for 1h.



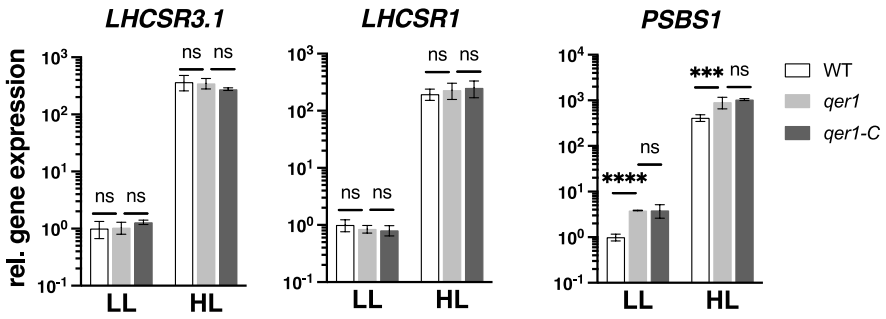
Extended Data Fig. 4: qE gene and protein expression in CLiP mutants bearing mutations in predicted qE regulator genes. WT, *lci8-oe, qer6* and *qer4* cells were acclimated for 16h in LL (15 µmol photons m^{-2} s^{-1}). After sampling for the LL conditions, light intensity was increased to 300 µmol photons m^{-2} s^{-1} (HL); samples were taken 1 h after exposure to HL. Shown are relative expression levels of of *LHCSR3.1*, *LHCSR1* and *PSBS1* at the indicated conditions normalized to WT LL (*n* = 3 biological samples, mean ± sd). The p-values for the comparisons between the mutants and the WT are based on ANOVA Dunnett's multiple comparisons test on log10- transformed values and are indicated in the graphs (**, P < 0.01).



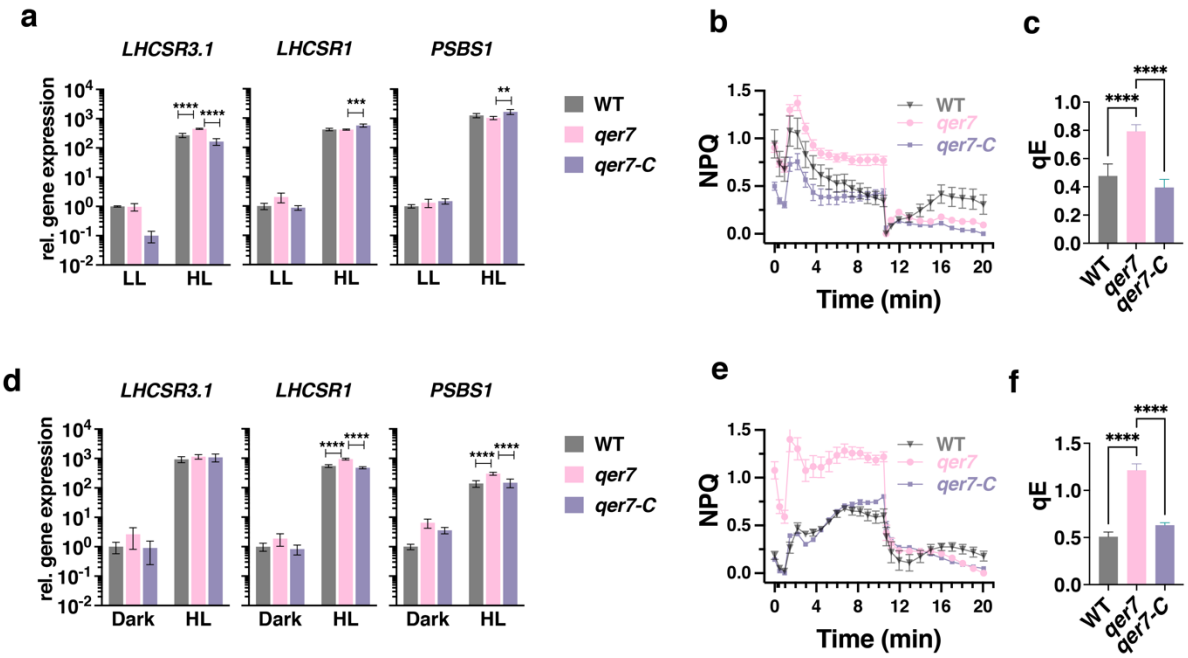
Extended Data Fig. 5: LCR1 is required for proper expression of qE-related genes during transitions from dark-to-light. WT, *lcr1* and *lcr1-C* cells were acclimated for 16h in darkness. After sampling for the dark conditions, light intensity was increased to 300 µmol photons m^{-2} s^{-1} (HL); samples were taken 1 h (RNA) or 4 h (protein and photosynthetic measurements) after exposure to HL. **a**. Relative expression levels of qE genes at the indicated conditions normalized to WT LL (*n* = 3 biological samples, mean ± sd). **b**. Immunoblot analyses of LHCSR1, LHCSR3, PSBS and ATPB (loading control) of one out of the three biological replicate samples, under the indicated conditions. **c.** Summary graph of immunoblots of all replicate samples of Extended Data Fig. 3b after normalization to ATPB. Shown are the HL treated samples; WT protein levels were set as 1. **d.** NPQ and calculated qE, 4h after exposure to HL (*n* = 3 biological samples, mean ± s.d). The p-values for the comparisons are based on ANOVA Dunnett's multiple comparisons test and are indicated in the graphs (*, P < 0.005, **, P < 0.01, ***, P < 0.001, ****, P < 0.0001). Statistical analyses for panel **a** and **c** were applied on log10- transformed values.



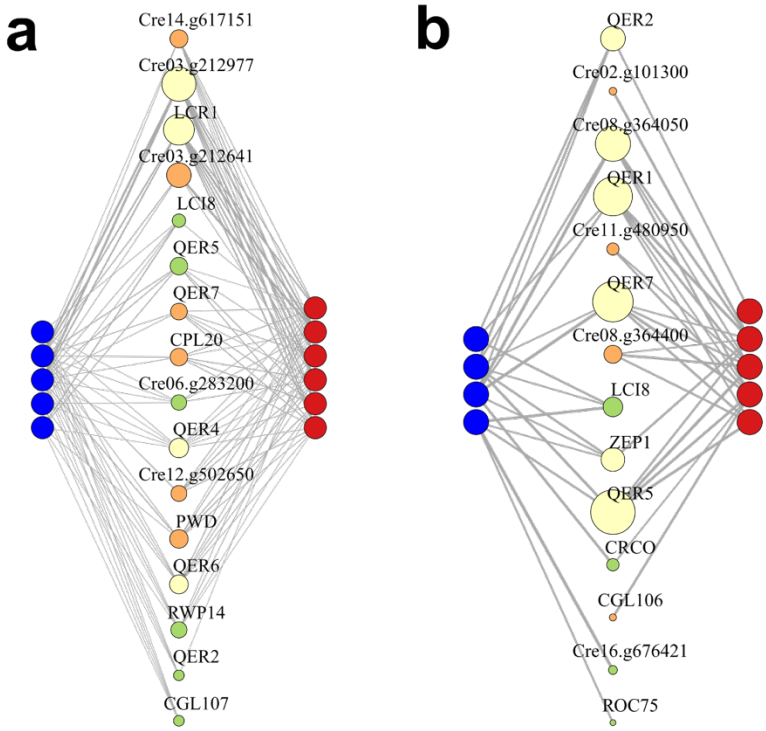
Extended Data Fig. 6: LCR1 activates the transcription of CCM genes in LL and dark-acclimated cells. a. WT, *lcr1* and *lcr1-C* cells were acclimated for 16h in LL (15 µmol photons m^{-2} s^{-1}). After sampling under LL conditions, light intensity was increased to 300 µmol photons m^{-2} s^{-1} (HL) and samples were taken 1 h after exposure to HL. Relative expression of CCM genes at the indicated conditions normalized to WT under LL. **b**. Relative expression of CCM genes acclimated in 16h of darkness (*n* = 3 biological samples, mean ± sd). The p-values for the comparisons are based on ANOVA Dunnett's multiple comparisons test on log10-transformed values and are indicated in the graphs (*, P < 0.005, **, P < 0.01, ***, P < 0.001, ****, P < 0.0001).



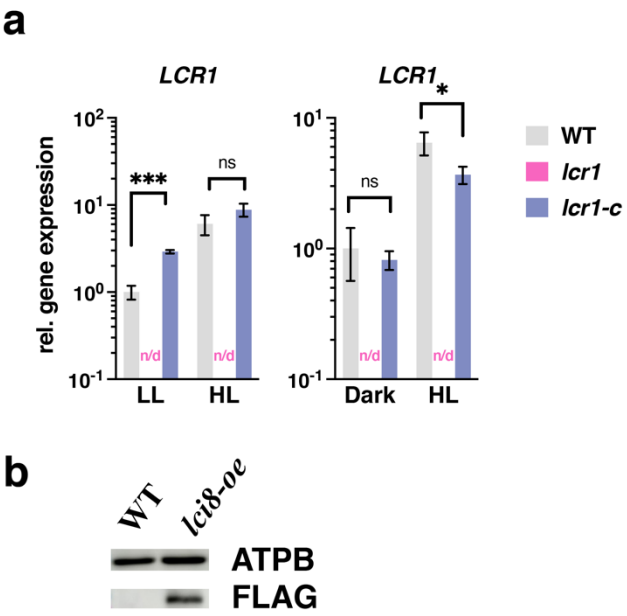
Extended Data Fig.7: Complementation of *qer1* with *QER1* (*qer1-C* strain) fails to rescue the *LHCSR1* and *PSBS* phenotypes. WT, *qer1* and *qer1-C* cells were acclimated for 16h in LL (15 µmol photons m^{-2} s^{-1}). After sampling for the dark conditions, light intensity was increased to 300 µmol photons m^{-2} s^{-1} (HL); RNA samples were taken 1 h after exposure to HL. Shown are relative expression levels of qE genes at the indicated conditions normalized to WT LL (*n* = 3 biological samples, mean ± sd). The p-values for the comparisons are based on ANOVA Dunnett's multiple comparisons test on log10- transformed values and are indicated in the graphs (***, P < 0.001, ****, P < 0.0001).



Extended Data Fig.8: Relative expression of qE-related genes in asynchronous *qer7* cells. WT, *qer7* and *qer7-C* cells were acclimated for 16h in LL (15 µmol photons m^{-2} s^{-1}; **a-c**) or darkness (**d-f**). After sampling for the LL or dark conditions, light intensity was increased to 300 µmol photons m^{-2} s^{-1} (HL) and samples were taken 1 h (RNA) or 4 h (photosynthetic measurements) after exposure to HL. **a, d**. Relative expression of qE and CCM genes at the indicated conditions normalized to WT LL (**a**) or dark (**d**) respectively (*n* = 3 biological samples, mean ± sd). **b, e.** NPQ and **c, f.** calculated qE, 4h after exposure to HL (*n* = 3 biological samples, mean ± s.d. The p-values for the comparisons are based on ANOVA Dunnett's multiple comparisons test and are indicated in the graphs (**, P < 0.01, ***, P < 0.001, ****, P < 0.0001). Statistical analyses for panel **a** and **d** were applied on log_{10}-transformed values.



Extended Data Fig. 9: Reconstructed GRNs recapitulate experimentally found regulators of qE and CCM. Network representation of the top 10 regulators of CCM genes quantified via qPCR (**Fig. 4),** and qE genes as predicted from **a**. the concensus or **b**. PHOT-specific GRN. Regulators are represented by centeral nodes (same color code as in **Fig. 5**), target qE genes present in the respective network are shown in blue (*LHCSR1*, *LHCSR3.1*, *LHCSR3.2*, *PSBS1*, *PSBS2*) and target CCM genes are shown in red (*LCI1*, *LCIA*, *CAH4*, *CCP1*, *HLA3*, *BST1*). Larger TF nodes indicate higher regulatory strength on target genes. The plotted edge width is proportional to the strength of the regulatory interaction.



Extended Data Fig. 10: LCR1 mutation and LCI8 overexpression confirmation. **a.** Relative expression of *LCR1* in WT, *lcr1* and *lcr1-C* strains. Cells were acclimated for 16h in LL or darkness. After sampling for the LL or dark conditions, light intensity was increased to 300 µmol photons m^{-2} s^{-1} (HL), and samples for RNA purification were taken 1 h after exposure to HL (*n* = 3 biological samples, mean ± sd). *LCR1* gene expression was non-detectable (indicated as n/d in the graphs) in the *lcr1* mutant. **b.** Immunoblot analysis of LCI8-FLAG fused protein in WT and lci8-oe cells grown in TAP, LL. The p-values for the comparisons are based on unpaired t-tests on log10- transformed values and are indicated in the graphs (*, P < 0.005, ***, P < 0.001).

Supplementary Table Captions

Supplementary Table 1: RNAseq data sets used in this study; HSM: Sueoka's high salt medium (1); HSM – new trace = modified Sueoka's high salt medium with trace elements as described in (2); Rep= biological replicates

(1) Sueoka, N. Mitotic replication of deoxyribonucleic acid in Chlamydomonas reinhardii. Proc. Natl. Acad. Sci. U.S.A. 46, 83–91 (1960).

(2) Kropat, J.et al. A Revised Mineral Nutrient Supplement Increases Biomass and Growth Rate in Chlamydomonas Reinhardtii. Plant J. 2011, 66 (5), 770.

Supplementary Table 2: List of Transcription factors used in this study; Family info was adapted from Pérez-Rodríguez et al. (2010) and Jin et al. (2017)

Supplementary Table 3: Edge list representation of the consensus network with mean ranks resulting from Borda count election method (Borda (1781), Marbach et al. (2011)) as edge attributes.

Supplementary Table 4: Top 10 predicted regulators of qE genes in the consensus GRN. Regulation strength was calculatedd based on the edge rank in the consensus network.

Supplementary Table 5: Edge list representation of the PHOT GRN with importance score from GENIE 3 (Huyn-Thu et al., 2010) as edge attributes.

Supplementary Table 6: Top 10 predicted regulators of qE genes in the PHOT GRN. Regulation strength was calculated based on the edge weight assigned by GENIE3.

Supplementary Table 7: List of genes putatively involved in photoprotection used for regulator prediction

Supplementary Table 8: List of genes putatively involved in CCM used for regulator prediction.

Supplementary Table 9: PCR primers for CLiP mutant validation and complementation in this study

Supplementary Table 10: RT-qPCR primers for all genes in this study

3.3 Conclusion

By taking advantage of proteomics and RNA-seq data, we found that GAP1 encoding the chloroplast form GAPDH is overexpressed in *phot* mutant. As phototropin is a blue light photoreceptor and the *GAP1* transcription level reaches its peak at dusk, we checked the *GAP1* expression level under white, blue, and red light conditions. We found that GAP1 is repressed by blue light in a PHOT-dependent way.

GAP1 is a chloroplast form of GAPDH and is probably involved in the Calvin cycle, so we generated GAP1 overexpression lines to check the link between GAP1 and starch. All GAP1 overexpression lines could accumulate at least 1.4 times more starch than WT while not influencing LHCSR3 expression level. These results support the link between PHOT, GAP1, and starch.

To get more insights into how PHOT regulates GAP1, by using our omics data and gene regulatory network (GRN), we found a novel TF-GAPR4 that regulates GAP1. Overexpression of GAPR4 could upregulate the *GAP1* expression level.

Section 3.2.8 presents the manuscript describing the GRN tool used to predict the GAP1 regulators. The GRN developed in this manuscript is first based on qE regulator regulation. In this study, we used it to predict GAP1 regulators and experimentally proved the reliability of the prediction. It suggested that this GRN tool can be a powerful and reliable tool to predict putative regulators of genes that may be regulated by light and/or CO_2 availability.

In this chapter, we further confirmed the link between PHOT and starch. We proposed a novel pathway involved in starch accumulation: Blue light represses GAPR4 which activates GAP1 in a PHOT-dependent way. PHOT knockout mutant, GAP1 and GAPR4 overexpression lines could accumulate higher starch content than WT.

Chapter 4. A member of the RWP-RK family of transcription factors links PHOT and starch accumulation

4.1 Abstract

Chapter 4 describes the second signaling cascade pathway linking PHOT with starch accumulation in *Chlamydomonas.*

Blue light could activate RWP5 (a TF belonging to the RWP-RK family of TFs) via PHOT. RWP5, which is usually induced in nitrogen depletion or limitation conditions in *Chlamydomonas*[35,195], is downregulated in *phot* mutant. Moreover, its deletion leads to high starch content in the cell. Overexpressing *RWP5* in *phot* could partially rescue the high starch content phenotype. These results suggest that RWP5 is involved in PHOT-dependent starch metabolism, and PHOT might also play a role in nitrogen metabolism in *Chlamydomonas*.

4.2 Results

4.2.1 PHOT-dependent Genome-wide transcriptional profiling - Nitrogen metabolism

Nitrogen depletion is the most popular and widely used method to increase starch or lipid accumulation in microalgae. Taking advantage of our RNAseq data, we found that many genes involved in nitrogen metabolism[195] are differentially expressed genes (DEGs) between WT and *phot* (Fig. 4.1, Table 4.1 & 4.2).

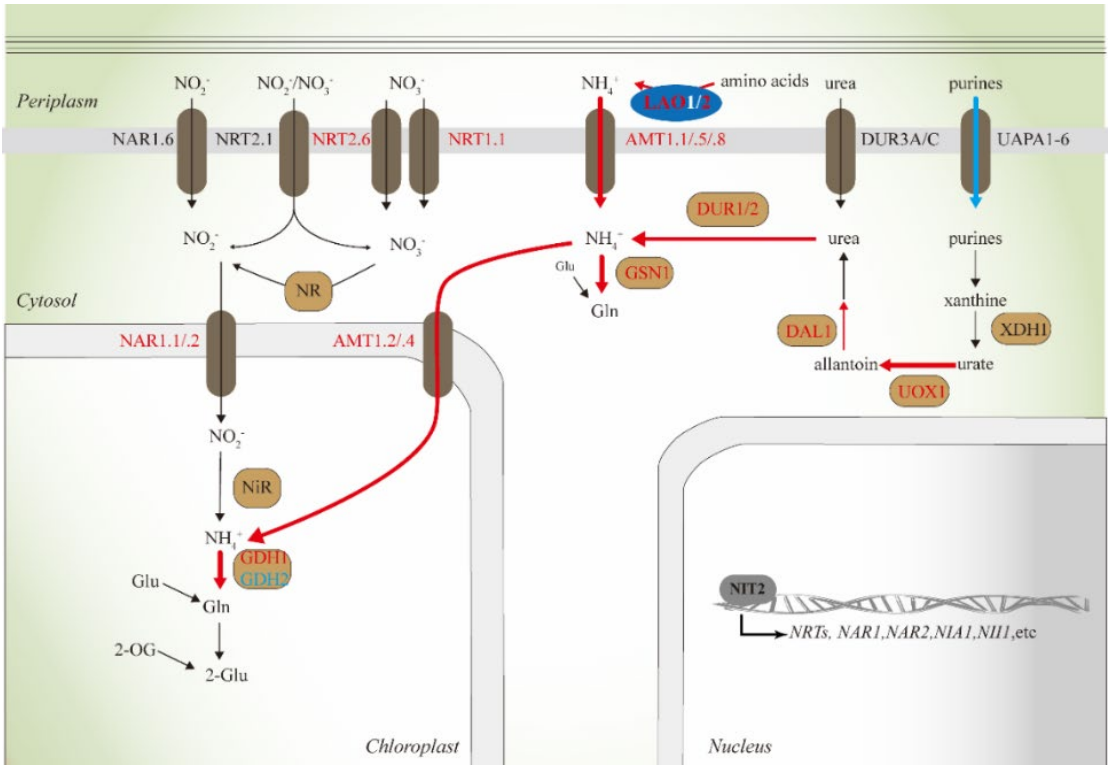


Figure 4.1. Enzymes related to nitrogen metabolism in microalgae. The upregulated or downregulated enzymes in *phot* are marked in red and blue, respectively (RNAseq and Proteomics data).

Among those differentially expressed genes, we found that several nitrogen metabolism-related genes were upregulated in *phot* under both LL and HL conditions, such as ammonium transporter (*AMT4*, *AMT6*, *AMT7*), nitrogen import (*AOC3*, *AOC7*), NO_3^-/NO_2^- transporter (*NRT1.1*, *NRT2.6*), ammonium fixation (*GSN1*, *GDH1*) (Table 4.1). Usually, these genes are upregulated in -N conditions, although *AMT6* and *AMT7* are downregulated[195].

Table 4.1 Upregulated DEGs involved in *phot* under LL conditions (Nitrogen metabolism)

Primary class	Secondary class	Locus ID	Gene name	*phot* LL /wt LL FC (log2)	*phot* HL/wt HL FC (log2)
Nitrogen metabolism	Import	Cre01.g023650	*AOC7*	1.58	1.52
		Cre02.g108550	*AOC3*	1.36	1.95
		Cre06.g298750	*AOT4*	0.88	0.36
		Cre02.g112000	*AOT7*	0.82	1.46

Primary class	Secondary class	Locus ID	Gene name	*phot* LL /wt LL FC (log2)	*phot* HL/wt HL FC (log2)
		Cre01.g041050	*AOC6*	0.74	-0.71
		Cre07.g329050	*AOC5*	0.65	1.75
		Cre02.g115300	*AOT6*	0.57	0.03
		Cre02.g145700	*AOT1*	0.21	0.72
		Cre03.g168550	*AOT3*	0.12	0.43
		Cre06.g264450	*AOT5*	0.04	0.78
	Known regulator	Cre16.g673250	*NRR1*	1.04	-1.04
		Cre07.g357350	*GLB1 (NII)*	0.46	0.25
	NH4 fixation	Cre13.g592200	*GSN1*	1.74	3.56
		Cre09.g388800	*GDH1*	0.42	2.29
		Cre03.g207250	*GLN4*	0.35	0.47
		Cre12.g530600	*GLN3*	0.11	0.73
	NH4 transporter	Cre13.g569850	*AMT4*	2.21	1.94
		Cre02.g111050	*AMT7*	0.95	0.76
		Cre06.g284100	*RHP1*	0.71	2.63
		Cre07.g355650	*AMT6*	0.32	0.77
	NO3 NO2 transporter	Cre04.g224700	*NRT1.1*	0.87	1.52
		Cre02.g110800	*NRT2.6*	0.25	0.51
	Periplasmic deamination of amino acids	Cre12.g551350	*LAO2*	0.99	1.83
		Cre05.g246550	*LAO3*	0.25	0.96

NAR1.5 for NO_3^-/NO_2^- transportation and GDH2 for ammonium fixation, downregulated in -N conditions[195], are also downregulated in *phot* under both LL and HL conditions (Table 4.2).

However, a known nitrogen regulator-NRR1 (Nitrogen Response Regulator 1) was up-regulated in *phot* in LL conditions while down-regulated in HL conditions (Table 4.1).

Table 4.2 Down-regulated DEGs involved in *phot* under LL conditions (Nitrogen metabolism)

Primary class	Secondary class	Locus ID	Gene name	*phot* LL /wt LL FC (log2)	*phot* HL/wt HL FC (log2)
Nitrogen metabolism	NH4 fixation	Cre12.g514050	*GSF1*	-0.18	-1.07
		Cre02.g113200	*GLN1*	-0.22	1.09
		Cre12.g530650	*GLN2*	-1.04	-0.90
		Cre05.g232150	*GDH2*	-3.65	-1.43
	NO metabolism	Cre07.g315400	*NOA2*	-0.06	0.46
		Cre09.g400550	*NOA1*	-0.19	-0.45
	NO3 NO2 transporter	Cre12.g541250	*NAR1.5*	-1.69	0.18

The fact that many DEGs between WT and *phot* are upregulated or downregulated in -N conditions, and that GAP1 is also highly induced in -N conditions[35], may suggest that PHOT might be involved in nitrogen metabolism.

4.2.2 Identification of putative downstream effectors of PHOT controlling starch metabolism

To answer the above question, we took advantage of the RNAseq dataset (WT vs *phot*; LL vs HL) to pinpoint putative regulators of starch metabolism and *GAP1* expression.

First, I chose the DEGs with $\log_2$FC value > 1 between *phot* and WT under LL and HL conditions from the RNAseq dataset; this resulted in a list of 2201 genes. After removing the genes that showed the opposite direction of regulation under LL and HL conditions, I got a list of 1957 genes.

For all these genes, I used python to do the web crawling according to gene name or gene number, checked related publications, and known function of each gene (Source code in Section 6.2.12). For unknown genes in *C. reinhardtii*, I used the homologous genes in *Arabidopsis thaliana* of these genes and then repeated the checking step. After that, I divided all the genes into different groups depending on the selection criterion. For putative GAP1 regulator selection, I selected GAP1 co-expressed genes (from *Phytozome* v13) and genes with a similar expression profile as

GAP1 in the published RNA-seq dataset[189]. For starch metabolism, I divided genes into two groups: TFs and non-TFs. For the TFs group, I ranked them according to their FC value. As mentioned in Section 4.2.1, we hypothesize that PHOT might be linked to nitrogen metabolism. So I chose the top 100 up-regulated genes and top 100-down-regulated genes according to the published datasets of *Chlamydomonas* cells exposed to nitrogen limitation[109,195]. Within those 200 genes, I found that the top four nitrogen sparing responsive TFs families are RWP-RK, MYB-like DNA-binding protein, Squamosa promoter-binding-link protein, and Acetyltransferase. I check particularly TFs belonging to these four families in our dataset. For Non-TFs, I checked them one by one.

After the selection, I got a list with 24 genes, and we ordered the existing seven CDS mutants lacking the genes listed in Table 4.3, which might act as regulators of GAP1 or starch accumulation.

Table 4.3. List of predicted putative starch or GAP1 regulators

Gene ID	Name in this study	Annotation	*phot* LL/WT LL FC (log2)	*phot* HL/WT HL FC (log2)
Cre06.g285600	*RWP5*	RWP-RK domain-containing protein	-2.61	-4.37
Cre11.g475550	*UMM10*	SAM-dependent methyltransferases superfamily protein	-0.68	-2.95
Cre02.g085150	*SBP*	Squamosa promoter binding protein-like 1	2.24	2.26
Cre04.g218050	*RWP8*	RWP-RK domain-containing protein	1.55	2.11
Cre01.g034350	*MYB*	MYB domain-containing protein	2.05	1.66
Cre03.g207800	*ADH5*	Alcohol dehydrogenase	-2.21	-3.84
Cre09.g398400	*TRP5*	Transient receptor potential ion channel protein	-7.2	-5.1

Before performing any experiment with these mutants, we first verified the presence of the transforming cassette in different strains by colony PCR (*myb and rwp5* validation will be shown in 4.2.4, validation of other mutants is not shown). All the mutants shown in Fig. 4.2 have been confirmed as being real CDS mutants.

Among these mutants, *myb* and *rwp5* showed significantly higher starch content than WT at t=0d (Fig. 4.2A) and t=4d in N-deplete conditions (Fig. 4.2B). Therefore,

we decided to choose *myb* and *rwp5* for further investigation.

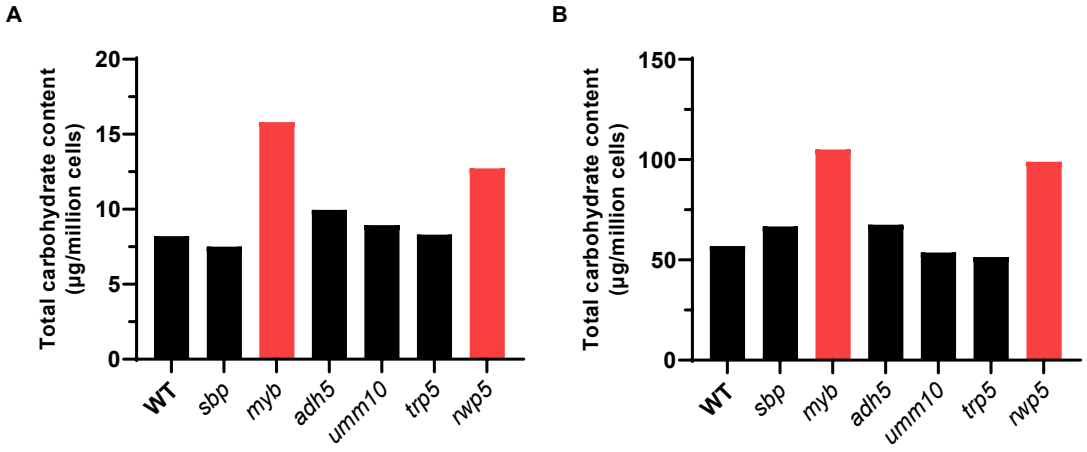


Figure 4.2. The total carbohydrate content of different CLiP mutants. (**Left**: day 0; **Right**: day 4 in -N conditions)

4.2.3 RWP5 and MYB are under the control of PHOT

From our RNA-seq dataset, the *MYB* transcription level was up-regulated in *phot* in both LL and HL conditions, while the *RWP5* transcription level was down-regulated in *phot* (Table 4.3). Thus, combining the results (Fig. 4.2) and our RNA-seq dataset, we hypothesized that RWP5 and MYB are under the control of PHOT and involved in starch metabolism.

To check if MYB and RWP5 are under the control of PHOT, we checked the transcription levels of *MYB* and *RWP5* in WT, *phot,* and *phot-C* under the same conditions in the RNA-seq experiment was performed. The results confirmed the RNA-seq data; the addition of the *phot-C* strain in the qPCR experiment confirmed that both MYB and RWP5 are under the control of PHOT (Fig. 4.3).

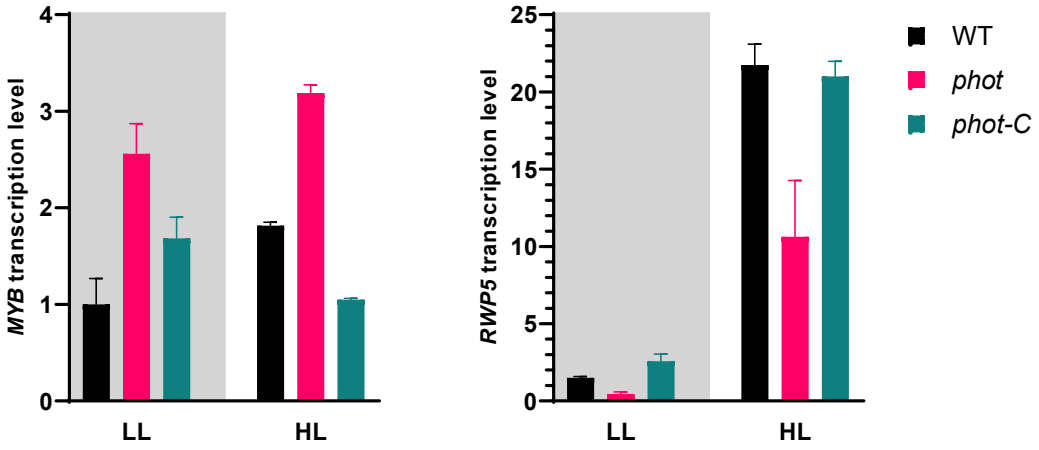


Figure 4.3. The transcription level of *MYB* (Left) and *RWP5* (Right) in WT, *phot,* and *phot-C*. LL (5 µmol/m2/s) and HL (300 µmol/m2/s) conditions. Values relative to GBLP were normalized to wild-type samples under low light intensity (n = 3 biological samples, mean ± s.d.).

4.2.4 The *rwp5* and *myb* are starch over accumulators.

To further investigate the role of RWP5 and MYB in starch metabolism in *C. reinhardtii*, we took CDS insertion mutants of *RWP5* and *MYB* from CLiP library[17,18] and validated the insertion site.

MYB CDS mutant, LMJ.RY0402.112219 (hereafter *myb*) has one insertion of the paromomycin resistance cassette (CIB1) in the middle of exon 6 of the *MYB* gene (Fig. 4.4); *RWP5* CDS mutant, LMJ.RY0402.184466 (hereafter *rwp5*) also holds the same CIB1 cassette in the middle of exon 5 of the *RWP5* gene (Fig. 4.5). The site of CIB1 cassette integration in the CLiP mutants had been validated by the insertion site in these two mutants by the LEAP-Seq method[17]. Using primers annealing 1kb away from the flanking sequence on either side of the insert, no fragment was observed in genomic DNA samples isolated from wild type *cc-4533* and mutant strain. It is presumably due to the presence of complex or large insertions, which are typically found in CLiP mutants[196]. So, we decided to change the amplification region by trying other primers.

For *MYB*, using primers designed to amplify the 5' and 3' junction sites of the CIB1 cassette, specific 767- and 944-bp fragments could be detected in the *myb* mutant (Fig. 4.4). The sequencing analysis of the PCR amplicons confirmed the insertion site (Fig. 6.5). The *MYB* transcription level of WT and *myb* has also been

checked by RT-qPCR in -N and +N conditions. However, when we tried to check the *MYB* transcription level in *myb* using the primers designed to amplify the upstream region of the CIB1 cassette insertion site, we still could see the induction of *MYB* in *myb*. Then we decided to use primers designed to amplify the downstream region of the CIB1 cassette insertion site and low *MYB* transcription level in *myb* mutant in both -N and +N conditions (Fig. 4.4). The results showed that *myb* is more likely a knockdown mutant than a knockout mutant.

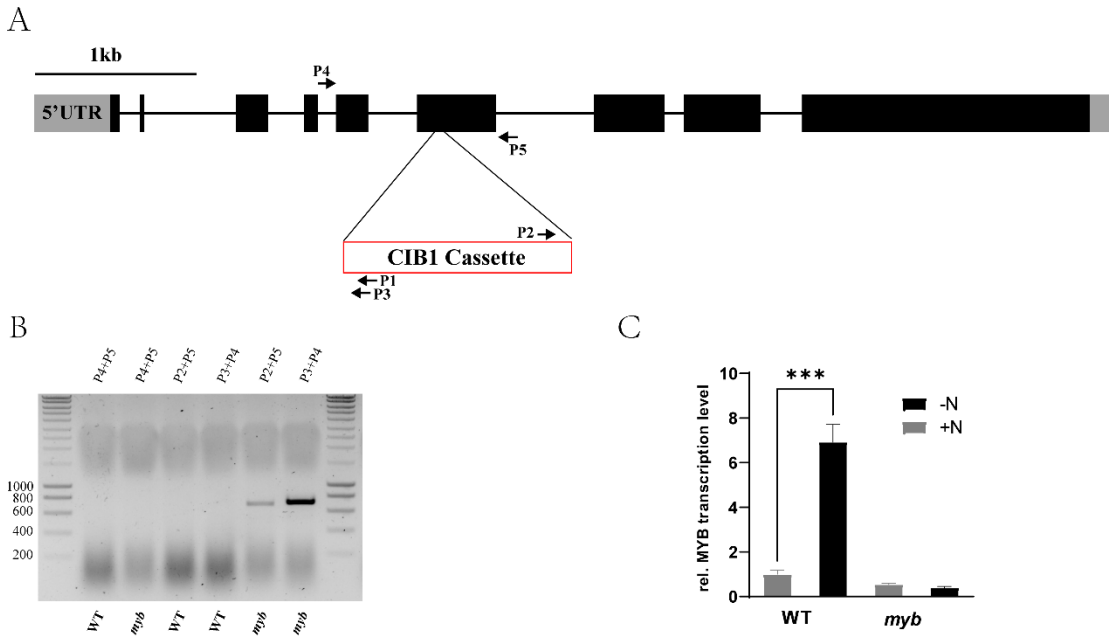


Figure 4.4. Characterization of a CLiP mutant-*myb* (LMJ.RY0402.112219) affected in the MYB gene. A. Insertion map of the CIB1 cassette in MYB gene. Exons were shown in black, introns in line, 5'UTR and 3'UTR in light gray, and primers in arrows. The triangle indicates the insertion site of the CIB1 cassette; **B**. Validation of the insertion site in *myb*. Primers used are at the top. The presented PCR results are based on different single colonies, and all colonies showed identical results; **C**. Transcript levels of *MYB*, as determined by RT-qPCR in culture under nitrogen replete or depleted conditions using primers binding the 3'UTR region of *MYB*. Values relative to GBLP were normalized to wild-type samples (n = 3 biological samples, mean ± s.d.).

For *RWP5*, using primers designed to amplify the 5' and 3' junction sites of the CIB1 cassette, specific 585- and 785-bp fragments could be detected in the *rwp5* mutant (Fig. 4.5). The sequencing analysis of the PCR amplicons confirmed the predicted insertion site of the CIB1 cassette (Fig. 6.6). Moreover, no *RWP5* transcription has been detected by RT-qPCR in neither -N or +N conditions. All the

results showed that *rwp5* is a knockout mutant.

All the primers used for colony PCR are shown in Table 6.2.

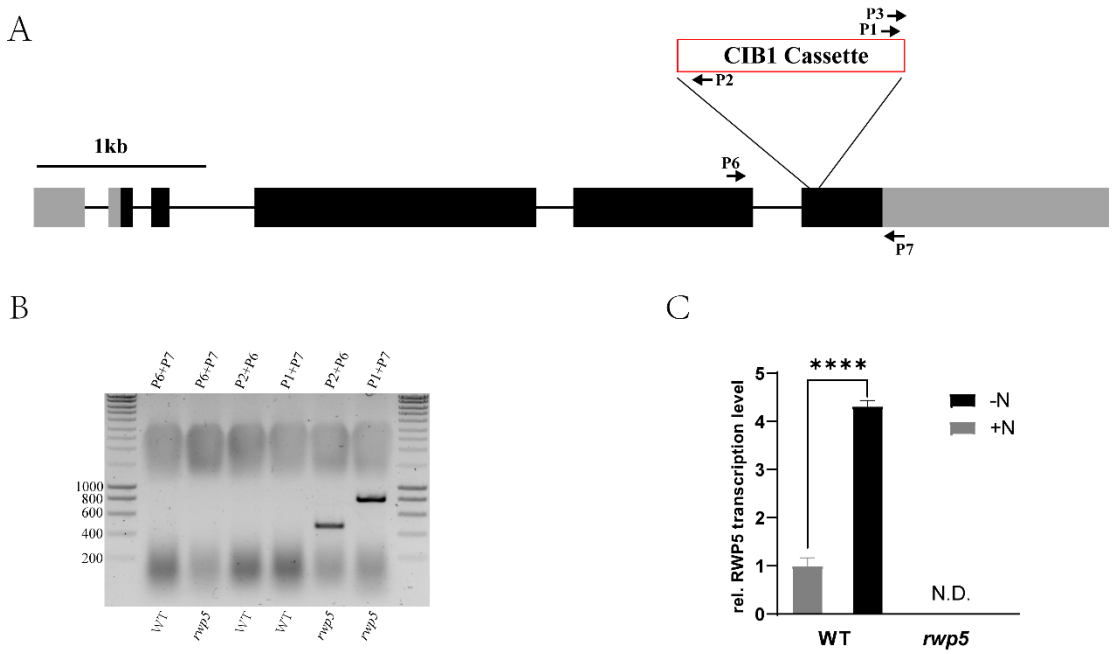


Figure 4.5. Characterization of a CLiP mutant-*rwp5* (LMJ.RY0402.184466) affected in the RWP5 gene. A. Insertion map of the CIB1 cassette in the *RWP5* gene. Exons were shown in black, introns as interconnecting lines, 5'UTR and 3'UTR in light gray, and primers in arrows. The triangle indicates the insertion site of the CIB1 cassette; **B**. Validation of the insertion site in *rwp5*. The primers used are at the top. The presented PCR results are based on different single colonies, and all colonies showed identical results; **C**. Transcript levels of *RWP5*, as determined by RT-qPCR in culture under nitrogen replete or deplete conditions using primers binding the 3'UTR region of *RWP5*. Values relative to GBLP were normalized to wild-type samples (n = 3 biological samples, mean ± s.d.).

After confirming they are real CDS mutants, we checked their starch content in HSM medium under light/dark cycle conditions at different time points. The *rwp5* and *myb* mutants could accumulate more starch than WT (Fig. 4.6).

In the 12/12 light-dark cycle, the rapid decrease of starch content at the start of the experiment might be due to some macro or mineral nutrition deficiency after six days of synchronization in HSM medium. Starch content is increased during the day and decreased during the night. However, *myb* did not show this fluctuation like WT and *rwp5* (Fig. 4.6). The total carbohydrate content experiences a rapid decline after 24h and a more quick increase rate from 60h. It might imply that *myb* has difficulty in cell

division or disruption in the cell cycle. Thus, it might be worth checking if *MYB* plays a role in cell division in the future.

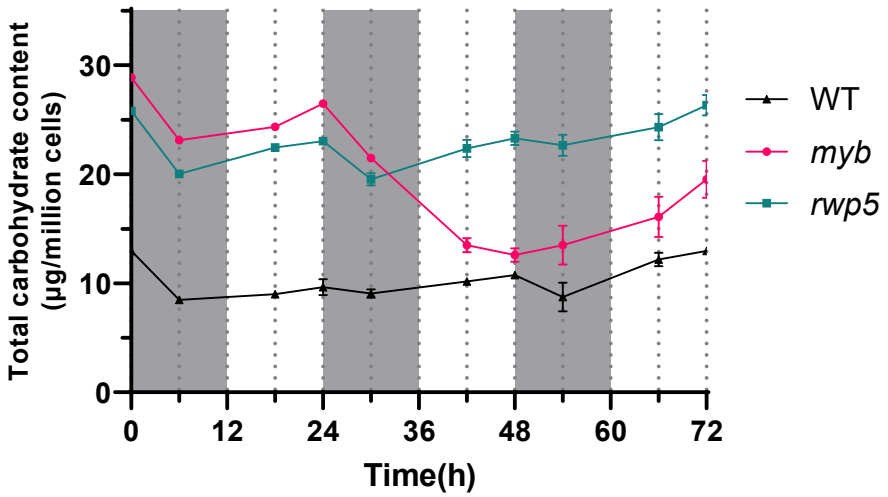


Figure 4.6. Total carbohydrate content in WT, *rwp5,* and *myb*. Cells were grown under the 12/12 light-dark cycle in phototrophic HSM medium.

4.2.5 RWP5 is a starch metabolism regulator

To confirm it is the deletion of RWP5 and MYB leads to the high starch phenotype, it is necessary to generate the complementation lines.

For *MYB*, another lab also found that the starch content in *myb* was significantly higher than in its WT[197]. However, instead of getting the complement line, they tested two sets of progenies derived from *myb* and another WT for co-segregation between antibiotic resistance and high starch content; there was no co-segregation[197]. Therefore, the high starch content of *myb* might be due to some non-detected gene disruptions.

For *RWP5*, this gene is 5236 bp with 68% GC content. We first tried to synthesize it. However, due to the gene structure complexity, high GC content, and repeated regions, direct synthesis of the gene could not be achieved.

It is impossible to get RWP5 gDNA by PCR cloning in one step with long primers for the Gibson assembly, so we split the RWP5 gene into three fragments reassembled

following PCR amplification. To get each fragment, we tried different enzymes, buffers, programs, chemicals, and programs. However, the cloning efficiency dropped off rapidly due to the multiplicative effect. Fragment 1 from 1bp to 1710bp and fragment 2 from 1691bp to 3411bp were amplified using Platinum and Phusion enzymes. Fragment 3 from 3391bp to 5233bp was amplified using Platinum and KOD enzymes. (Amplification details in section 6.2.8 and Table 4.4, primers in Table 6.2). Cloning RWP5 gDNA was sequence-verified and assembled into pRAM118 plasmid[198] and pLM005 plasmid[176] by Gibson assembly.

Table 4.4. RWP5 gDNA amplification strategy

Name	Template	Enzyme	add	Tm value	GET?
RWP5-F1	gDNA	Platinum		60	Y
	gDNA	KOD		55	N
	gDNA	Phusion+HF buffer		66	N
	gDNA	Phusion+GC buffer		66	N
	gDNA	Platinum	5%DMSO	60	N
	gDNA	KOD	5%DMSO	55	N
	gDNA	Phusion+HF buffer	5%DMSO	66	N
	gDNA	Phusion+GC buffer	5%DMSO	66	N
RWP5-F2	gDNA	Platinum		60	N
	gDNA	KOD		55	N
	gDNA	Phusion+HF buffer		66	N
	gDNA	Phusion+GC buffer		66	N
	gDNA	Platinum	5%DMSO	60	Y
	gDNA	KOD	5%DMSO	55	N
	gDNA	Phusion+HF buffer	5%DMSO	66	N
	gDNA	Phusion+GC buffer	5%DMSO	66	N
	gDNA	Platinum	2.5MBetaine	60	N
	gDNA	KOD	2.5MBetaine	55	N
	gDNA	Phusion+HF buffer	2.5MBetaine	66	Y
	gDNA	Phusion+GC buffer	2.5MBetaine	66	Y
RWP5-F3	gDNA	Platinum		60	N
	gDNA	KOD		55	N
	gDNA	Phusion+HF buffer		66	N
	gDNA	Phusion+GC buffer		66	N
	gDNA	Platinum	5%DMSO	60	Y
	gDNA	KOD	5%DMSO	55	Y
	gDNA	Phusion+HF buffer	5%DMSO	66	N
	gDNA	Phusion+GC buffer	5%DMSO	66	N
RWP5-F1+F2	F1+F2	Platinum		60	N
	F1+F2	KOD		55	N
	F1+F2	Phusion+HF buffer		66	N
	F1+F2	Phusion+GC buffer		66	N
	F1+F2	Platinum	5%DMSO	60	N
	F1+F2	KOD	5%DMSO	55	Y
	F1+F2	Phusion+HF buffer	5%DMSO	66	N
	F1+F2	Phusion+GC buffer	5%DMSO	66	N
	F1+F2	Platinum	2.5MBetaine	60	N
	F1+F2	KOD	2.5MBetaine	55	N

Name	Template	Enzyme	add	Tm value	GET?
	F1+F2	Phusion+HF buffer	2.5MBetaine	66	**Y**
	F1+F2	Phusion+GC buffer	2.5MBetaine	66	**N**
RWP5-F1+F2+F3	F1+F2+F3	Platinum		60	**N**
	F1+F2+F3	KOD		55	**N**
	F1+F2+F3	Phusion+HF buffer		66	**N**
	F1+F2+F3	Phusion+GC buffer		66	**N**
	F1+F2+F3	Platinum	5%DMSO	60	**N**
	F1+F2+F3	KOD	5%DMSO	55	**N**
	F1+F2+F3	Phusion+HF buffer	5%DMSO	66	**N**
	F1+F2+F3	Phusion+GC buffer	5%DMSO	66	**N**
	F1+F2+F3	Platinum	2.5MBetaine	60	**N**
	F1+F2+F3	KOD	2.5MBetaine	55	**N**
	F1+F2+F3	Phusion+HF buffer	2.5MBetaine	66	**N**
	F1+F2+F3	Phusion+GC buffer	2.5MBetaine	66	**N**
RWP5-F2+F3	F2+F3	Platinum		60	**N**
	F2+F3	KOD		55	**N**
	F2+F3	Phusion+HF buffer		66	**N**
	F2+F3	Phusion+GC buffer		66	**N**
	F2+F3	Platinum	5%DMSO	60	**N**
	F2+F3	KOD	5%DMSO	55	**Y**
	F2+F3	Phusion+HF buffer	5%DMSO	66	**N**
	F2+F3	Phusion+GC buffer	5%DMSO	66	**N**
RWP5-F1+F2,F2+F3	F1+F2,F2+F3	Platinum		60	**N**
	F1+F2,F2+F3	KOD		55	**N**
	F1+F2,F2+F3	Phusion+HF buffer		66	**N**
	F1+F2,F2+F3	Phusion+GC buffer		66	**N**
	F1+F2,F2+F3	Platinum	5%DMSO	60	**Y**
	F1+F2,F2+F3	KOD	5%DMSO	55	**N**
	F1+F2,F2+F3	Phusion+HF buffer	5%DMSO	66	**N**
	F1+F2,F2+F3	Phusion+GC buffer	5%DMSO	66	**N**
	F1+F2,F2+F3	Platinum	2.5MBetaine	60	**Y**
	F1+F2,F2+F3	KOD	2.5MBetaine	55	**N**
	F1+F2,F2+F3	Phusion+HF buffer	2.5MBetaine	66	**N**
	F1+F2,F2+F3	Phusion+GC buffer	2.5MBetaine	66	**N**
gib-RWP5-F1+F2,F2+F3	F1+F2,F2+F3	Platinum		60	**N**
	F1+F2,F2+F3	KOD		55	**N**
	F1+F2,F2+F3	Phusion+HF buffer		66	**N**
	F1+F2,F2+F3	Phusion+GC buffer		66	**N**
	F1+F2,F2+F3	Platinum	5%DMSO	60	**N**
	F1+F2,F2+F3	KOD	5%DMSO	55	**N**
	F1+F2,F2+F3	Phusion+HF buffer	5%DMSO	66	**N**
	F1+F2,F2+F3	Phusion+GC buffer	5%DMSO	66	**N**
	F1+F2,F2+F3	Platinum	2.5MBetaine	60	**N**
	F1+F2,F2+F3	KOD	2.5MBetaine	55	**N**
	F1+F2,F2+F3	Phusion+HF buffer	2.5MBetaine	66	**N**
	F1+F2,F2+F3	Phusion+GC buffer	2.5MBetaine	66	**N**

*: **Y** means I could get the correct size band. **N** means I did not get the correct size band.

After electroporation, we got one complementation line of *rwp5* strain (*rwp5-C*) after screening around 500 transformants. Due to the two mismatched base pairs in the flag region of the pRAM118 plasmid, we can not detect the protein by using the anti-FLAG antibody. So we confirmed the complementation by RT-qPCR (Fig. 4.7).

Moreover, the starch content of *rwp5-C* went back to the WT level (Fig. 4.7).

Combining all the results, RWP5 is under the control of PHOT, and its deletion leads to high starch content in *Chlamydomonas reinhardtii*.

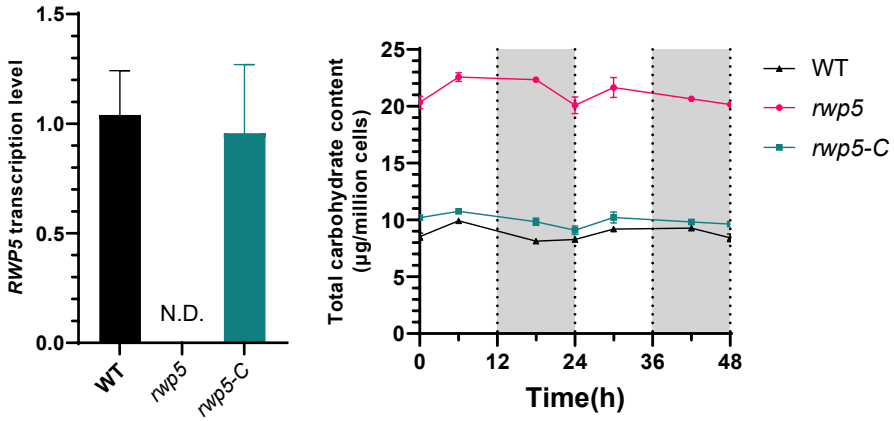


Figure 4.7. The *RWP5* transcription level (left) and total carbohydrate content (right) in WT, *rwp5,* and *rwp5-C*. Left: Samples collected after 1h high light treatment in phototrophic HSM medium. **Right**: Cells were grown under the 12/12 light-dark cycle in phototrophic HSM medium. Values relative to GBLP were normalized to wild-type samples (n = 3 biological samples, mean ± s.d.). The p-values for the comparisons indicated in the graph are based on ANOVA Dunnett's multiple comparisons test (*, P<0.05; ***, P < 0.0001).

4.2.6 RWP5 is involved in the PHOT-dependent starch metabolism

As RWP5 is downregulated in *phot* under high light conditions (Fig. 4.3 & 4.7) and cannot be induced in -N conditions (Fig. 4.5), it is worth checking if RWP5 is involved in the PHOT-dependent starch metabolism.

To address this question, we overexpressed RWP5 in *phot* and got two mutants (*phot::rwp5-C1* and *phot::rwp5-C2*) after screening around 1000 transformants. After confirming the RWP5 expression level by RT-qPCR (Fig. 4.8), we checked the starch content of these two mutants.

Surprisingly, the high starch phenotype of *phot* in these two mutants was partially rescued (Fig. 4.9). However, the *GAP1* transcription levels could not be rescued by the overexpression of RWP5 in *phot*; they remained high as in *phot*. These data indicate that RWP5 is involved in PHOT-dependent starch metabolism, and that PHOT-RWP5-starch and PHOT-GAP1-starch are two independent pathways.

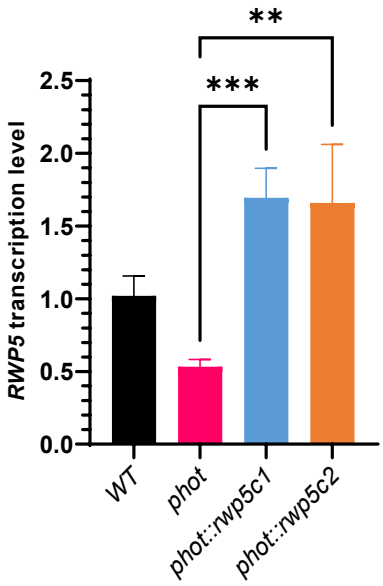


Figure 4.8. The *RWP5* transcription level in WT, *rwp5*, and two *phot::rwp5-C* lines. Samples were collected after 1h HL treatment in phototrophic HSM medium. Values relative to GBLP were normalized to wild-type samples (n = 3 biological samples, mean ± s.d.). The p-values for the comparisons indicated in the graph are based on ANOVA Dunnett's multiple comparisons test (**, P<0.01; ***, P < 0.001).

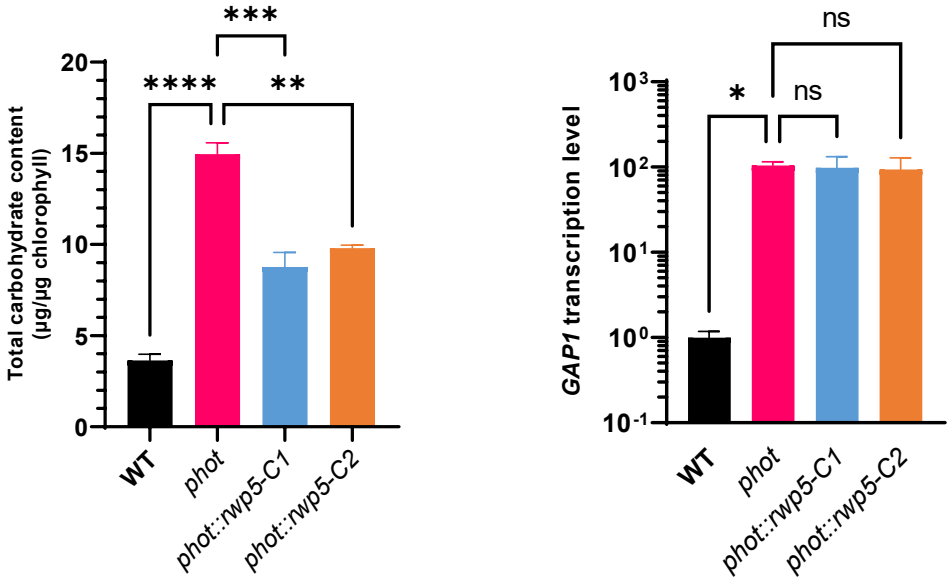


Figure 4.9. The total carbohydrate content (Left) and *GAP1* transcription level (Right) of WT, *rwp5*, and two *phot::rwp5-C* lines. Samples were collected in phototrophic HSM medium under LL conditions. Values relative to GBLP were normalized to wild-type samples (n = 3 biological samples, mean ± s.d.). The p-values for the comparisons indicated in the graph are based on ANOVA Dunnett's multiple comparisons test (*, P<0.05; **, P<0.01; ***, P<0.001; ****, P < 0.0001; ns, not significant).

4.2.7 Blue light-dependent activation of RWP5 requires full-length PHOT

To explore whether the kinase activity of PHOT is involved in RWP5 activation, I checked the *RWP5* transcription level in the *pkin* mutant. The *RWP5* transcription level in *pkin* mutant is the same as in *phot* mutant. (Fig. 4.10).

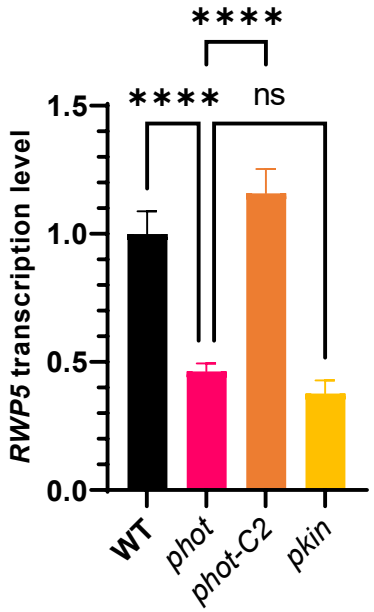


Figure 4.10. The *RWP5* transcription level of WT, *phot*, *phot-C2, a*nd *pkin* in phototrophic HSM medium after 1h HL treatment. Values relative to GBLP were normalized to wild-type samples (n = 3 biological samples, mean ± s.d.). The p-values for the comparisons indicated in the graph are based on ANOVA Dunnett's multiple comparisons test (****, P < 0.0001; ns, not significant).

Moreover, we got similar results by using another set of *phot* mutants based on the *cw15* background. Interestingly, complementation with LOV domains of PHOT (*pl1*) and the dead kinase form of PHOT (*pk3*) also failed to rescue the *RWP5* transcription level (Fig. 4.11).

These results indicated that the activation of *RWP5* by blue light requires the presence of the full length PHOT and that for this activation the kinase activity is not playing an important role..

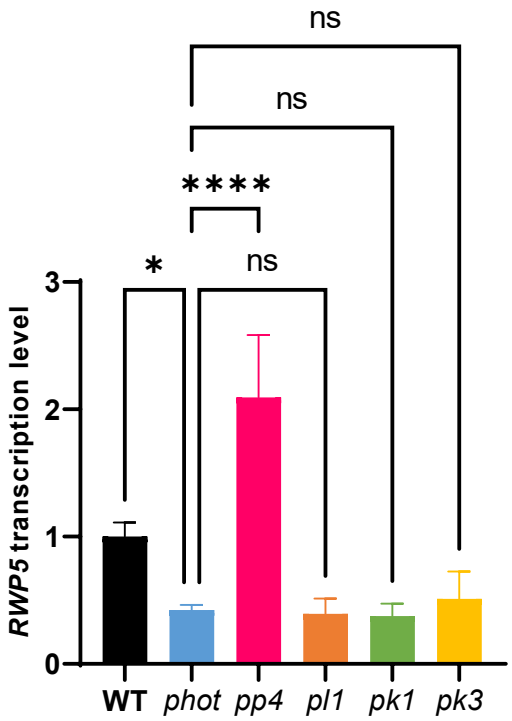


Figure 4.11. The *RWP5* transcription level of WT, *phot*, *pp4*, *pk1*, *pl1*, *a*nd *pk3* in phototrophic HSM medium after 1h HL treatment. Values relative to GBLP were normalized to wild-type samples (n = 3 biological samples, mean ± s.d.). The p-values for the comparisons indicated in the graph are based on ANOVA Dunnett's multiple comparisons test (*, P<0.05; ****, P < 0.0001; ns, not significant).

4.2.8 RWP5 is not involved in PHOT-dependent photoprotection genes regulation

As RWP5 is under the control of PHOT, we also tested if RWP5 plays a role in PHOT-dependent qE regulation (Fig. 4.12).

As Fig. 4.12 showed, the deletion of RWP5 did not affect the expression of LHCSR3, LHCSR1, PSBS, and qE.

Although MYB is not involved in the PHOT-starch link, we also test if MYB is involved in photoprotection gene regulation. Surprisingly, we found that MYB could accumulate more LHCSR1 and PSBS protein than WT, although *myb* did not show higher qE (Fig. 4.12). MYB deserves a more profound investigation as a potential PSBS and LHCSR1 regulator.

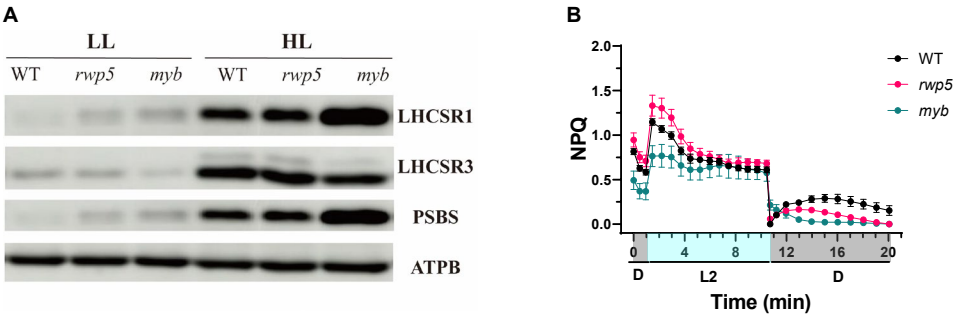


Figure 4.12. Phenotype of WT, *rwp5,* and *myb.* A: Immune blot analyses of LHCSR3, LHCSR1, PSBS. ATPB as a loading control. **B**: NPQ values of for WT, *rwp5* and *myb* after exposure to HL for 4 h. Just prior to the onset of the measurements, cells were acclimated to darkness for 15 min. Chlorophyll fluorescence was recorded in the dark (labelled as "D"), at 336 (labelled as "L2") µmol photons m^{-2} s^{-1} as indicated in the graphs. (NPQ measurement setup details in 6.2.4).

4.3 Conclusion

In this chapter, we found a novel starch metabolism regulator-RWP5 under the control of PHOT.

RWP5 is highly induced in -N conditions[195]. Moreover, RWP5 is downregulated in *phot* mutant in HL conditions and can not be induced in *phot* mutant in -N conditions. Deletion of RWP5 leads to higher starch content than its WT in normal conditions. By overexpressing RWP5 in *phot*, the high starch content phenotype of *phot* could be partially rescued. This suggested that RWP5 is not only a starch metabolism regulator but also a linker between blue light perception and nitrogen metabolism. Furthermore, RWP5 is not involved in the pathway PHOT-GAPR4-GAP1-STARCH presented in chapter 3. Additionally, combining all the results, we also propose that PHOT might play a role in nitrogen metabolism by activating RWP5.

Chapter 5. PHOT-dependent phosphorylation of the BSK1 impacts starch accumulation

5.1 Abstract

This Chapter established a third novel link between PHOT and starch metabolism in *Chlamydomonas reinhardtii*. We identified the first PHOT-dependent phosphorylation protein - BSK1 (Blue-light Starch Kinase 1) in *Chlamydomonas* by taking advantage of phosphoproteomics data. Blue light could de-phosphorylate BSK1 in a PHOT-dependent manner, and this dephosphorylation process plays an essential role in PHOT-dependent starch metabolism. Overexpression of the de-phosphorylated form of BSK1 in *phot* could rescue its high starch content phenotype. Moreover, GAP1 is also under the control of BSK1.

Overall, this work describes for the first time a PHOT-dependent phosphorylation protein in *Chlamydomonas* and investigates its role in starch metabolism.

5.2 Results

5.2.1 PHOT-dependent phosphoproteome analyses

In order to explore the early PHOT-dependent phosphorylation events, our team performed a phosphoproteomics analysis of WT and *phot* after an overnight acclimation to darkness and subsequent exposure to blue light (100 μE) for 5 min.

Based on the analysis, 6273 unique phosphopeptides from 2618 phosphoproteins were identified. From the trend analysis that has been performed across the conditions, 12 clusters (from A to L) with notable trends have emerged.

Among different clusters, cluster H contains the phosphopeptides, which increase in abundance only in blue-light exposed WT samples. Cluster I contains phosphopeptides that decrease in abundance after blue-light exposure in the WT but not in the *phot*.

In Cluster H and I, we found several phosphoproteins with interesting biological significance, which may mediate some of the PHOT-initiated signaling cascades.

Among them, in cluster I, one of the phosphoproteins that attracted our attention is the gene product of Cre16.g659400 which encodes a Ser/Thr protein kinase of around 61.33 kDa in size (Fig. 5.1). We named this protein BSK1 (Blue-light Starch Kinase 1). The reason will be explained in the following sections.

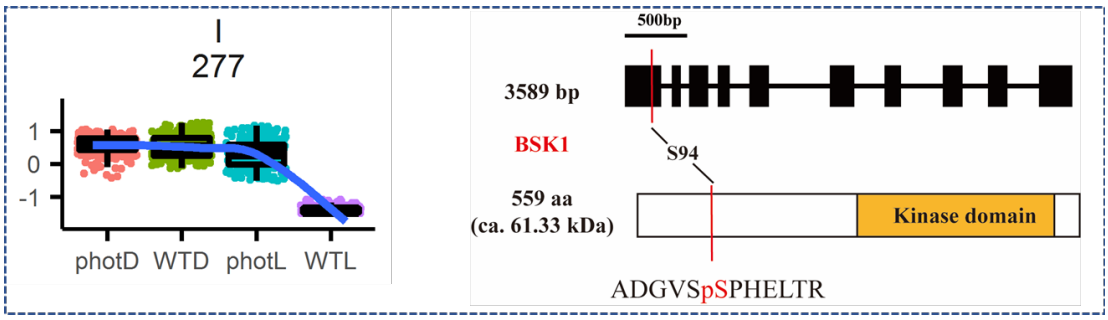


Figure 5.1. Schematic structures of BSK1. Left: Cluster I from the trend analysis results of phosphoproteomics data. photD, *phot* mutant in the dark; WTD, WT after 5 min blue light exposure; photL, *phot* mutant after 5 min blue light exposure; WTL, WT in the dark. One-way ANOVA with BH correction was performed for 6273 unique phosphopeptides. Y axis is the unsupervised hierarchical clustering was performed on phosphopeptide abundances normalized to z-scores. **Right**: Red lines indicate the phosphorylated residues in BSK1. Black boxes indicate the exons; Orange boxes indicate the kinase domains.

The detected phosphorylated sequence of BSK1 in the dataset is ADGVSS^{P}HELTR (Ser94) which does not belong to the kinase domain (Fig. 5.1). The phosphorylated BSK1 was less abundant after 5 min exposure to blue light from dark in WT, while it remains at almost the same levels in *phot* (Fig. 5.2). Furthermore, in another phosphoproteomics dataset of our team (WT and *phot*, in continuous low light conditions (5 µE) and after exposure to high light (300 µE) for 20 min), we found the same phosphorylated site and dephosphorylated trend of BSK1 (Fig. 5.2).

Therefore, BSK1 is an interesting candidate worth investigating whether it is involved in any PHOT-mediated pathways.

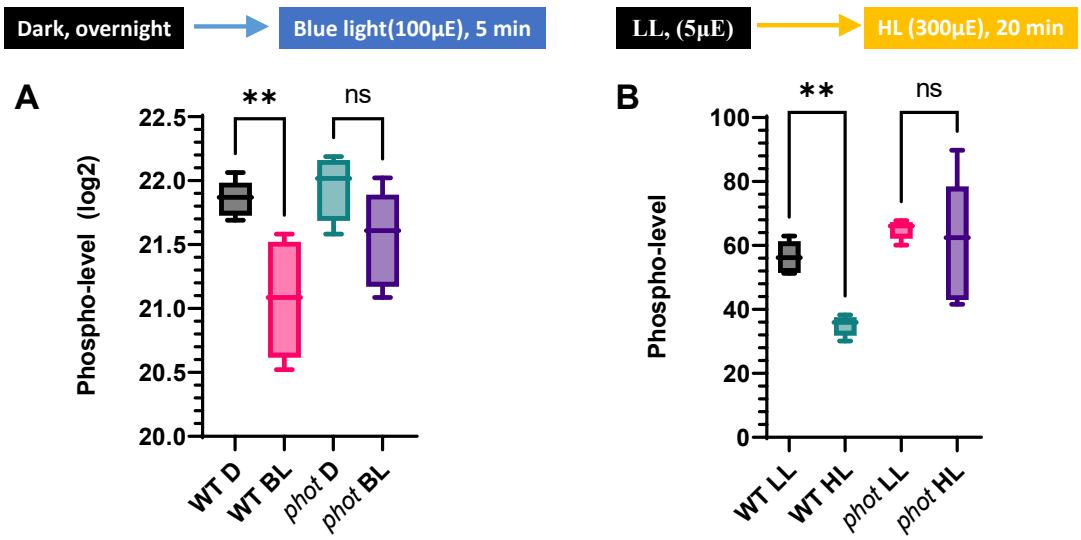


Figure 5.2. The phosphorylation level of BSK1 in two phosphoproteomics data. A: D, dark; BL, blue light; **B:** LL, low light; HL, high light. Y axis is the Changes in the phosphorylation levels of BSK1 proteins.

5.2.2 Phylogenetic analysis of BSK1

Before investigating the role of BSK1 in *Chlamydomonas,* we first took the protein sequence of BSK1 and blast it in *Arabidopsis thaliana* to find its orthologous proteins. From the first blast results, we found three BSK1 orthologous proteins in *Arabidopsis*- HT1, CBC1 and CBC2. Then we did a phylogenetic analysis of BSK1 with 131 proteins in different organisms. (Details in 6.2.14, Fig. 5.3).

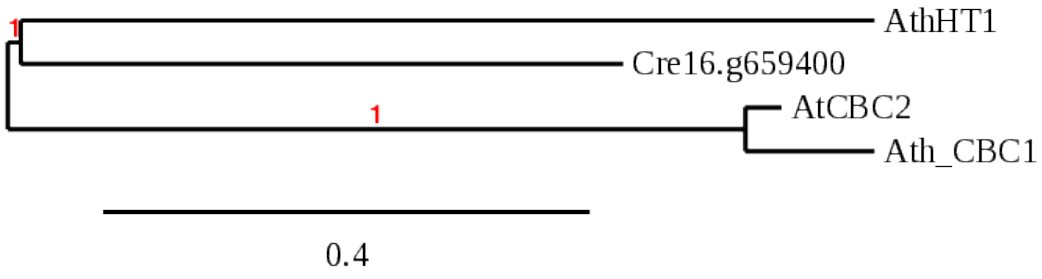


Figure 5.3. Phylogenetic tree of BSK1 by comparing amino acids. (Full tree is in Fig. 6.7)

The results showed that BSK1 is evolutionarily closer to AtHT1 than to AtCBCs.

HT1 is a negative regulator in CO_2-dependent stomatal closure In *Arabidopsis*

thaliana . In *ht1* mutants, the stomatal opening was impaired. AtCBC1 and AtCBC2 function in the same pathway as AtHT1[154] .(Section 1.1.8, Fig. 5.4).

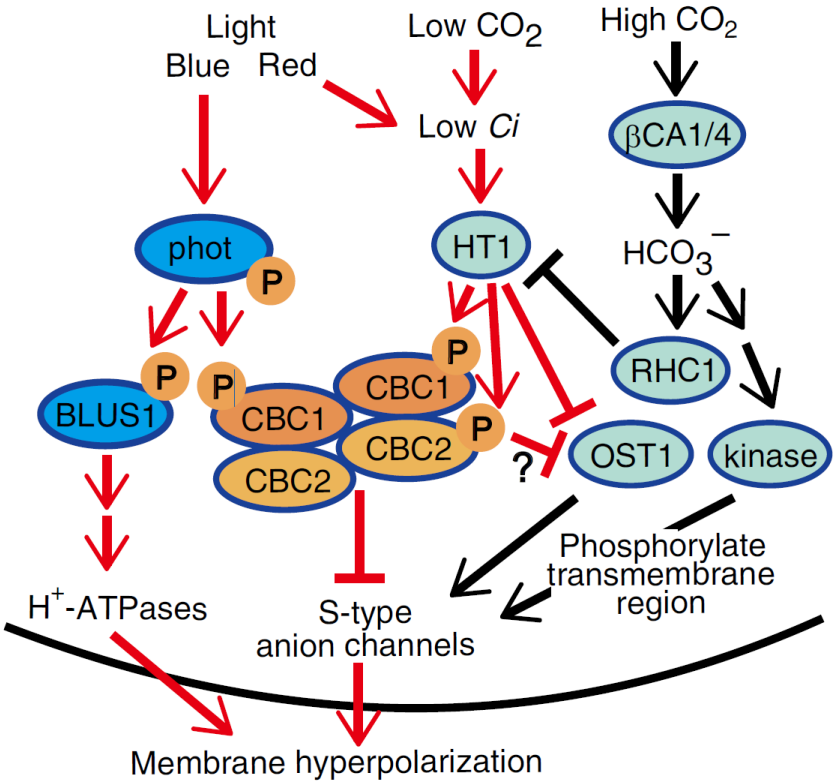


Figure 5.4. AtHT1 and AtCBCs function in Blue-light and CO2 signaling in *Arabidopsis thaliana*[154].

5.2.3 BSK1 is involved in the PHOT-dependent starch metabolism

To understand the function of BSK1 in *Chlamydomonas,* we first checked the *BSK1* transcription level in WT, *phot*, *phot-C,* and *pkin* mutants (Fig. 5.5).

The *BSK1* transcription level showed no difference in all the cell lines. In addition, we also checked the *BSK1* transcription level in *cw15* background mutants and got similar results (Fig. 5.5). These results indicated that the *BSK1* transcription level is unrelated to the PHOT-dependent regulation. Moreover, our proteomics data showed no change in WT and *phot* under LL and HL conditions in terms of BSK1 protein abundance (Data not shown).

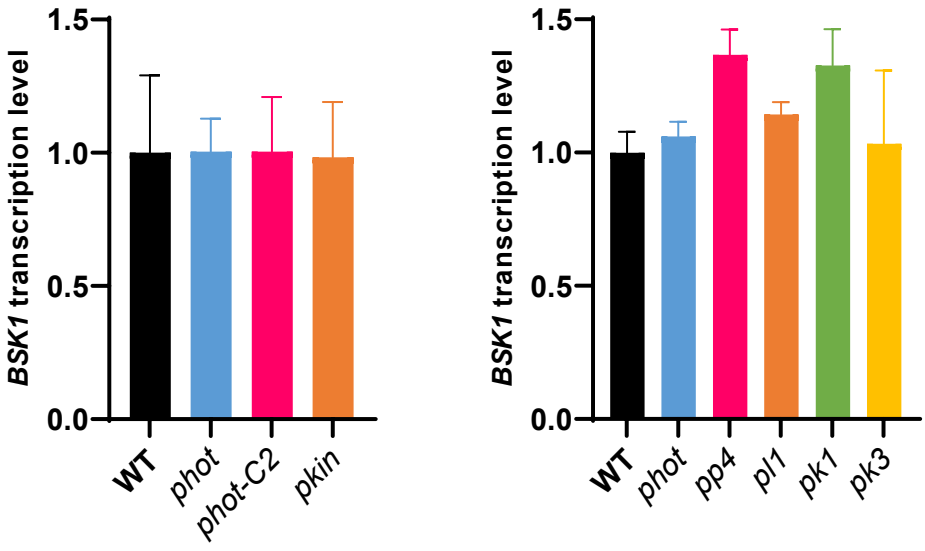


Figure 5.5. *BSK1* transcription level in different cell lines. Samples were collected in phototrophic HSM medium under LL conditions. Values relative to GBLP were normalized to wild-type samples (n = 3 biological samples, mean ± s.d.).

Then we checked the influence of the different phosphorylated forms of BSK1. As mentioned above, BSK1 is phosphorylated at Serine 94 in *Chlamydomonas*. We decided to generate phospho-specific BSK1 overexpression strains in *phot* backgrounds.

BSK1, BSK1^{S94A,} and BSK1^{S94D} represent the endogenous, phospho-deficient, and phospho-mimetic variants. BSK1 genomic sequence was amplified from genomic DNA using primers G19 and G20, and point mutation was introduced using primers G21, G22, G23, and G24 (Amplification details in section 6.2.8, primers in Table 6.4). The cloned gene was sequence-verified and assembled into pLM005 plasmid by Gibson assembly.

After transformation, we got six *phot::bsk1^{S94A}-FLAG* mutants with different protein expression levels. To determine the biological consequences of the expression of *BSK1^{S94A}* in *phot*, functional analyses were performed aiming at PHOT-starch signaling.

First, we checked the starch content of WT, *phot,* and six *phot::bsk1^{S94A}* mutants (Fig. 5.6). As the expression of BSK1^{S94A}-Venus-FLAG protein increased, the starch

content of six *phot::bsk1*^{S94A} mutants gradually decreased (Fig. 5.6). Due to the absence of PHOT, the main BSK1 form in *phot* is the phosphorylated form. The ectopic BSK1^{S94A} expression increases the proportion of the de-phosphorylated form of BSK1 in *phot* and decreases starch content.

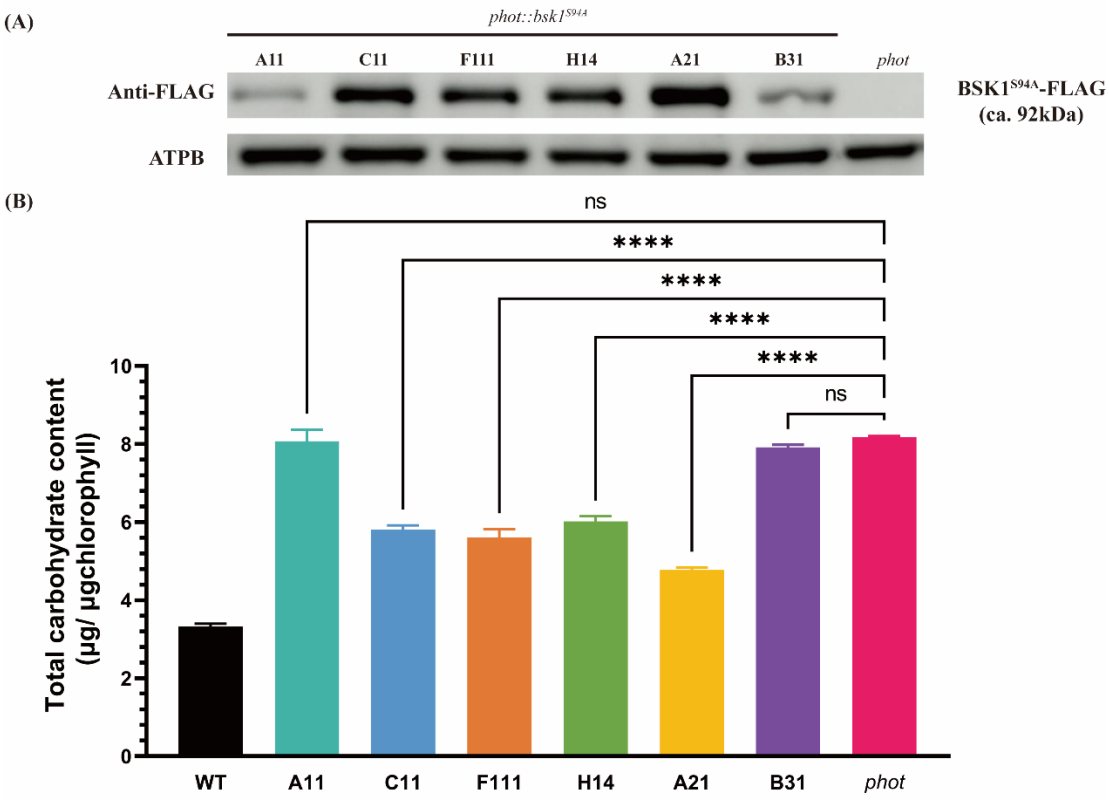


Figure 5.6. Immunoblot analysis (A) and total carbohydrate content (B) of *phot* and six *phot::bsk1*^{S94A} mutants under continuous LL conditions in phototrophic HSM medium. ATPB as a loading control. The p-values for the comparisons indicated in the graph are based on ANOVA Dunnett's multiple comparisons test. (****, P < 0.0001; ns, not significant).

To further confirm the role of BSK1 in PHOT-dependent starch metabolism, we generated more BSK1 overexpression lines in WT and *phot* background: *bsk1-oe* (an endogenous form of BSK1 overexpression line in WT), *phot::bsk1* (an endogenous form of BSK1 overexpression line in *phot*), *phot:: bsk1*^{S94D} (a phospho-mimic form of BSK1 overexpression line in *phot*). We checked the starch content of these mutants together with WT and *phot::bsk1*^{S94A} (*phot::bsk1*^{S94A}*-A21*).

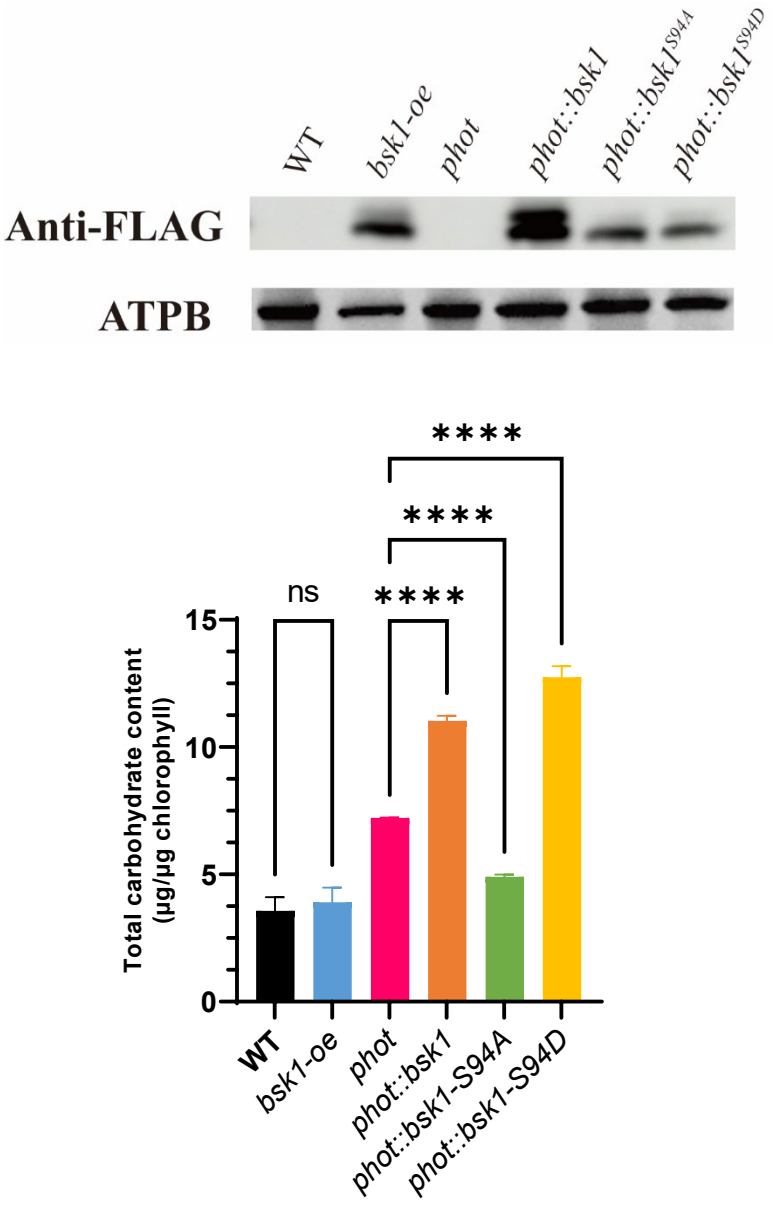


Figure 5.7. Immunoblot analysis (Top) and total carbohydrate content (Bottom) of WT, *bsk1-oe*, *phot*, *phot::bsk1*, *phot::bsk1*^{S94A}, *phot:: bsk1*^{S94D} under continuous LL conditions in phototrophic HSM medium. ATB as a loading control. The p-values for the comparisons indicated in the graph are based on ANOVA Dunnett's multiple comparisons test (****, P < 0.0001; ns, not significant).

Interestingly, *phot::bsk1*^{S94D} and *phot::bsk1* both could accumulate even 1.5 times more starch than *phot*. However, *bsk1-oe* showed a similar starch level with WT. This might be due to the presence of PHOT (Fig. 5.7). In *phot* mutant, BSK1 mainly stays in the phosphorylated form under light conditions due to the absence of PHOT. In WT, because of the presence of PHOT, BSK1 could be normally de-phosphorylated by PHOT, so it will not change the proportion of different phosphorylated forms of

BSK1 in WT under light conditions.

Combining all the results, BSK1 is indeed involved in starch metabolism in *phot,* and its phosphorylation plays a crucial role in this regulation.

5.2.4 GAP1 is under the control of BSK1

To check if BSK1 is involved in PHOT-GAPR4-GAP1-Starch and PHOT-RWP5-starch link, we checked *GAP1, GAPR4,* and *RWP5* transcription levels in WT, *phot,* and different forms of BSK1 overexpression lines.

Surprisingly, the *GAP1* expression level went down due to the increasing BSK1^{S94A} expression level and increased due to the increasing BSK1 and BSK1^{S94D} expression levels (Fig. 5.8). More important, the *GAP1* transcription level is positively correlated with the starch content (Fig. 5.7 & 5.8).

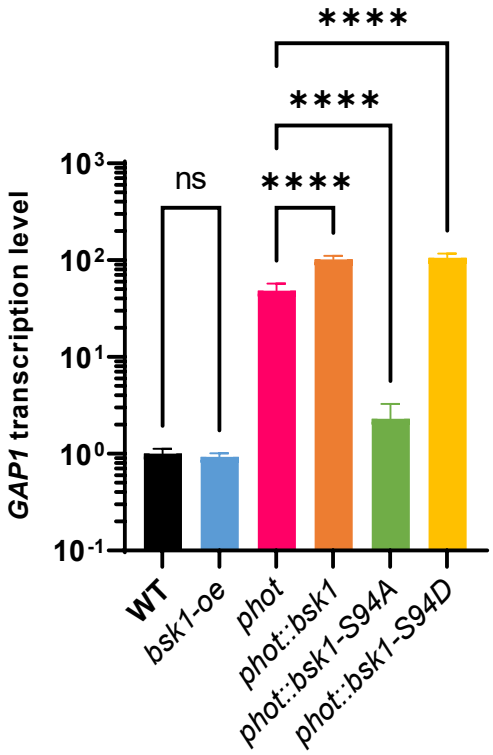


Figure 5.8. The GAP1 transcription level of WT, *bsk1-oe*, *phot*, *phot::bsk1*, *phot::bsk1^{S94A}*, *phot::bsk1^{S94D}* under continuous LL (15 µE) conditions in phototrophic HSM medium. Values relative to GBLP were normalized to wild-type samples (n = 3 biological samples, mean ± s.d.). The p-values for the comparisons indicated in the graph are based on ANOVA Dunnett's multiple comparisons test (****, P < 0.0001; ns, not significant).

However, *RWP5* and *GAPR4* transcription level was not affected in all the different BSK1 form overexpression lines (Fig. 5.9). These results indicated that BSK1 is involved in the PHOT-GAP1-starch pathway, and GAPR4 might not be the only TF regulator controlling GAP1.

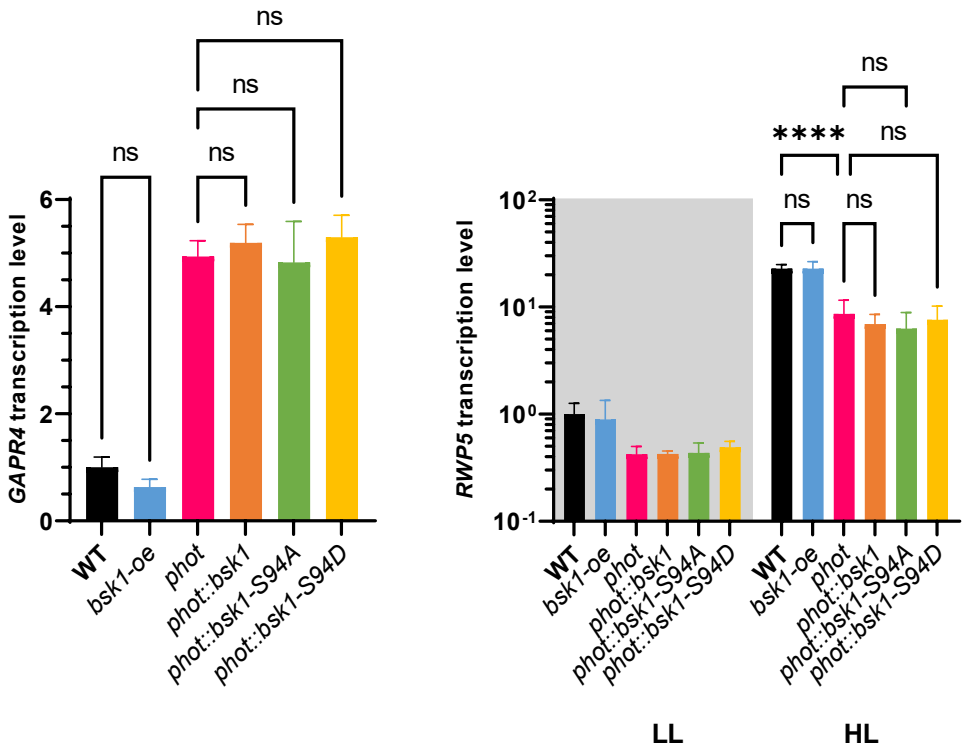


Figure 5.9. The *GAPR4* (Left)and *RWP5* (Right) transcription level of WT, *bsk1-oe*, *phot*, *phot::bsk1*, *phot::bsk1*^{S94A}, *phot:: bsk1*^{S94D} under continuous LL(15 µE) conditions and HL (300 µE) conditions in phototrophic HSM medium. Values relative to GBLP were normalized to wild-type samples (n = 3 biological samples, mean ± s.d.). The p-values for the comparisons indicated in the graph are based on ANOVA Dunnett's multiple comparisons test (****, P < 0.0001; ns, not significant).

5.3 Conclusion

In this chapter, we reported a novel kinase protein - BSK1, which is dephosphorylated by blue light via PHOT, involved in starch metabolism in *Chlamydomonas*.

BSK1 is mainly found in the phosphorylated form in the *phot* mutant. Interestingly, overexpressing dephosphorylated form of BSK1 in *phot* could bring its starch content

back to WT level, while overexpressing the endogenous and phosphorylated form of BSK1 in *phot* could even bring starch content to a much higher level. However, overexpressing the endogenous form of BSK1 in WT could not bring the starch level to a higher level. It might be that BSK1 can be dephosphorylated normally due to the presence of PHOT in WT. Moreover, GAP1 is under the control of BSK1, while GAPR4 and RWP5 are not. This also confirmed the link between GAP1 and starch.

Chapter 6

6.1 Summary and future directions

This thesis reported for the first time the link between blue light photoreceptor PHOT and starch metabolism in *Chlamydomonas reinhardtii* and three novel signaling cascades behind this link.

In higher plants-*Arabidopsis,* blue light could stimulate stomatal opening via activating PHOT by inducing the starch degradation in the guard cells. In this study, we showed that PHOT deletion could lead to a high starch content in green algae-*Chlamydomonas* and suggested a novel role of PHOT in regulating carbon partitioning in green algae.

Kamiya et al. reported that blue light could induce starch breakdown in *Chlorella*[117]. However, starch content did not decrease in *Phaeodactylum* in blue light acclimation conditions[119]. In this study, RNA-seq and RT-qPCR analysis showed that the expression of genes related to starch synthesis or degradation is both up-regulated in *phot* mutant (Table 2.1 & 2.2, Fig. 3.1). As Fig. 2.6 showed, starch synthesis and degradation rate in WT and *phot* are similar, lt indicated that the high starch phenotype has no connection with the circadian rhythm. All these results indicated that PHOT is indirectly connected with starch metabolism, and blue light inhibits starch accumulation is not via a single mechanism- promoting starch degradation or starch synthesis in *Chlamydomonas*. By generating the mutant *pkin*, which could constantly express the activated form of the PHOT kinase domain, we further confirmed that enzymatically active kinase is necessary for the blue light to inhibit the starch accumulation via PHOT (Fig. 2.10).

In *Arabidopsis,* red and blue light could both induce the expression of *GAPa* and *GAPb*, although red light is far less efficient[199]. GAPA and GAPB are the subunits of chloroplastic GAPDH, which plays a central role in the Calvin cycle in *Arabidopsis*. In *Chlamydomonas,* GAP3 is located in the chloroplast and functions like GAPA/B of *Arabidopsis*[200]. However, we did not see a significant difference in the protein or transcription level of GAP3 between the phototropin mutant and its wild type. In contrast, we found that GAP1 is upregulated in both protein and mRNA levels by the deletion of phototropin (Fig. 3.2). Red light could highly induce the expression level

of *GAP*1. Moreover, the deletion of PHOT will bring the GAP1 expression level up close to the level of WT under red light conditions (Fig. 3.3). This indicated that, the induction of *GAP1* in red light is mainly due to the absence of blue light.

Although the exact function of GAP1 in *Chlamydomonas is* still unknown, we found that overexpressing GAP1 could increase the starch content (Fig. 3.7). This gave the first role of GAP1 in starch metabolism in *Chlamydomonas*: GAP1 is involved in the PHOT-dependent starch metabolism. Moreover, by using GRN-based predictions[63], we reported a novel TF regulator of *GAP1*- GAPR4, which is under the control of PHOT (Fig. 3.12). Overexpression of GAPR4 could lead to a higher GAP1 expression level in different light quality conditions and also a higher starch content than its WT (Fig. 3.13).

In addition, when we tried to generate PHOT complementation lines and GAP1-OE lines with venus-tag, we first used the CDS of each gene, but it failed or with a meager success rate. Baier reported that intron-containing algal transgenes could enhance the gene expression in *Chlamydomonas reinhardtii*[178,201]. In our case, after switching to the genomic sequence, the success rate increased at least 20 times (from 0.1% to 2%). We think it is the introns' role in gene expression regulation. However, due to the high GC content of the *Chlamydomonas* genome and the longer length of gDNA than CDS, some researchers still prefer using CDS to avoid costly and technically challenging. Emrich-Mills recently reported a recombineering pipeline that could be a good solution for getting gDNA sequence in *Chlamydomonas*[202].

Light quality has been proved to affect nitrogen metabolism in different organisms. Blue light could induce the enzyme activities of nitrogen metabolism involved in *Malus domestica*, such as glutamate synthase (GOGAT) and nitrate reductase (NR)[203]. In *Arabidopsis,* the light control of asparagine synthetase (ASN1) and glutamine synthetase (GLN2), which are involved in nitrogen assimilation, has been proved to operate via phytochrome and by blue light[204]. In *Chlamydomonas*, blue light could activate nitrate reductase in *vivo*[205-207]. Taking advantage of proteomics and RNAseq, we found that many genes or enzymes involved in nitrogen metabolism have been affected in *phot* mutant (Table 4.1 & 4.2, Fig. 4.1). Motivated by this, we found a novel starch metabolism regulator- RWP5. RWP5 could be activated by blue light via PHOT, and the deletion of RWP5 leads to higher starch content (Fig. 4.3 & 4.7).

Overexpression of RWP5 in *phot* mutant could partially bring down the high starch content of *phot* mutant (Fig. 4.9). This confirmed the role of RWP5 in starch metabolism in blue light PHOT-dependent starch metabolism.

Taking advantage of phosphoproteomics, we reported the first phototropin-dependent dephosphorylation protein - BSK1 in *Chlamydomonas*. BSK1 is an ortholog of *Arabidopsis* HT1 and CBCs, and it could be dephosphorylated by blue light (Fig. 5.2 & 5.3). This indicated its possible role in starch metabolism in *Chlamydomonas*. By overexpressing the dephosphorylated form of BSK1 and phosphorylated form of BSK1 in the *phot* mutant, the starch content could go back to the wild-type level or even go higher, respectively (Fig. 5.7). Interestingly, overexpression of BSK1 in *phot* mutant could lead to a higher starch content in white light conditions, but this does not happen in wild type (Fig. 5.7). This confirms that BSK1 dephosphorylation needs not only blue light but also phototropin. Furthermore, the increase in the proportion of BSK1 dephosphorylated form not only bring down the starch content but also bring down the *GAP1* transcription level (Fig. 5.7 & 5.8). This indicated that the *GAP1* transcription level, the dephosphorylated form of BSK1 and starch content are positively correlated. It also further confirmed our link between GAP1 and starch.

However, there is still some pending work that would further advance our understanding of the link between PHOT and starch:

1) What would be the starch levels in *phot gap1* double knockout mutant*?* During my Ph.D., we have tried to get *phot gap1* double knockout mutant by CRISPR-Cas9 but failed after screening thousands of transformants. This might indicate that GAP1 is essential for *Chlamydomonas* to stay alive. So instead of getting knockout mutant, getting knockdown mutant might be a better choice. (in progress)

2) Could GAPR4 directly bind to *GAP1*? This could be checked by doing ChIP-qPCR.

3) How does the TF RWP5 control the starch metabolism? Which gene or protein is directly under the control of RWP5? Due to the two bp mismatch in the venus-flag region of our plasmid, the *rwp5* complementation line we generated in this study has no tag. However, this could be checked by generating a new *rwp5* complementation line using RWP5-tag (e.g. VENUS, FLAG, HIS) construction and then proceeding with ChIP-seq.

4) In *phot::rwp5-C* lines, *RWP5* transcription levels are higher than in WT. But the starch content of *phot* could only be partially rescued and can not fully go back to WT level. Does this mean that other regulators are also under the control of PHOT and involved in nitrogen metabolism?

5) What is the localization of BSK1? This could be checked by confocal microscopy. (in progress)

6) BSK1 is closely related to AtHT1 and AtCBCS. Can BSK1 also interact with PHOT and be phosphorylated by PHOT in *Chlamydomonas?* This could be answered by doing Yeast-Two-Hybrid or in *vitro/vivo* pull-down assays. (in progress)

7) Does BSK1 also involve in PHOT-dependent qE regulation? (in progress)

8) How do the starch content and *GAP1* transcription level of different BSK1 form overexpression mutants change under different light quality conditions? (in progress)

9) Is BSK1 under the control of red light or CO_2 concentration? In *Arabidopsis,* HT1 responds to red light and low CO_2 concentration[172]. In our case, RHP1, a CO_2 channel protein that responds to high CO_2[208], is upregulated in phototropin mutant, and AtHT1 could respond to low CO_2. Does BSK1 phosphorylation also respond to CO_2 concentration (in progress)?

In summary, this study provided evidence and three pathways to describe how

phototropin controls starch metabolism and a comprehensive answer to how light quality affects the carbon partitioning in the model green alga *Chlamydomonas reinhardtii* (Fig. 6.1). In addition, microalgae biomass could be directly catalyzed into biofuels. The two most common types of biofuels are bioethanol and biodiesel, which are made from sugar/starch and lipid/fats, respectively. Our findings will also give some new ideas for improving microalgal bioethanol productivity.

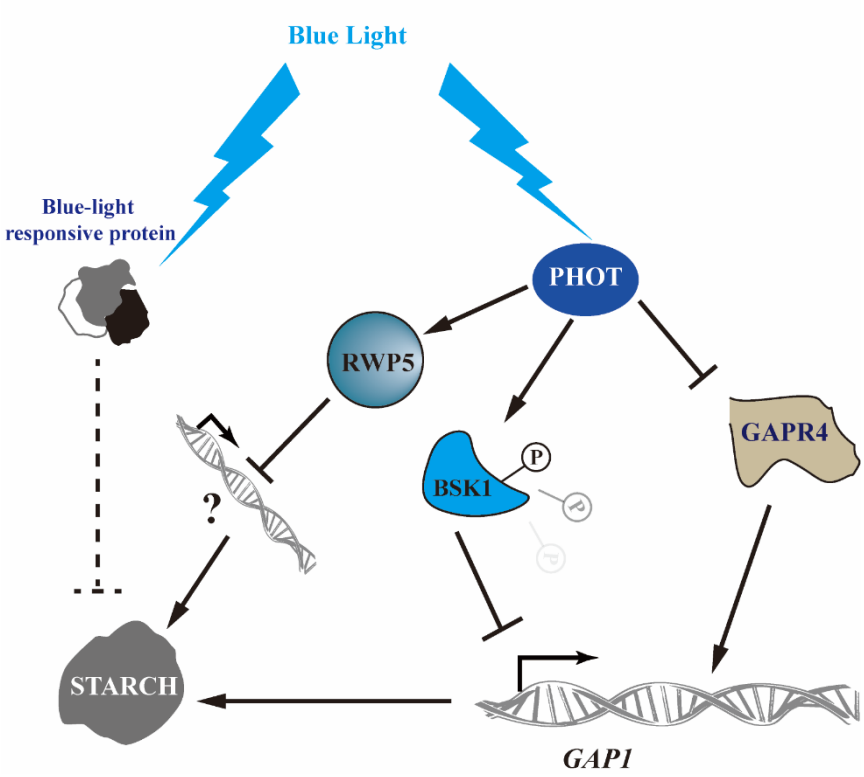


Figure. 6.1 Putative model of phototropin controls starch metabolism in *Chlamydomonas reinhardtii.*

6.2 Material and methods

6.2.1 Algal Strains

The following algal strains were obtained from the Chlamydomonas Resource Center and other labs (https://www.chlamycollection.org/) or generated in this study (Table 6.1). All the strains were maintained on solid TAP agar plates with or without appropriate antibiotics at 18℃ and 5 μmol/m^2/s.

Table 6.1. List of C. reinhardtii strains used in this study

Name	Description	Source/Reference
CC125	Wild type background for phot mutants	Chlamydomonas Resource Center
CC4533	Wild type background for CLiP mutants	Chlamydomonas Resource Center
myb	LMJ.RY0402.112219 mutant of Cre01.g034350	Chlamydomonas Resource Center
rwp5	LMJ.RY0402.184466 mutant of Cre06.g285600	Chlamydomonas Resource Center
phot	CRISPR/Cas9 mutant of PHOT in the *CC125* background;	Peter Hegemann's Lab[8]
phot-C	Full PHOT complemented line of *phot*	Generated in this study
phot-C2&C3	Full PHOT complemented line of *phot* (PHOT-VENUS-FLAG)	Generated in this study
pkin	Full PHOT kinase domain complemented line of *phot* (KINASE-BLE)	Generated in this study
rwp5-C	Full RWP5 complemented line of *rwp5*	Generated in this study
gap1oe-C1&C2&C3	GAP1-OE in *CC125* background (GAP1-VENUS-FLAG)	Generated in this study
gapr4-oe	GAPR4-OE in *CC125* background (GAPR4-VENUS-FLAG)	Generated in this study
bsk1-oe	BSK1-OE in *CC125* background (BSK1-VENUS-FLAG)	Generated in this study
bsk1 S94A-oe	BSK1(S94A)-OE in *CC125* background ($BSK1^{s94a}$-VENUS-FLAG)	Generated in this study
bsk1 S94D-oe	BSK1(S94D)-OE in *CC125* background ($BSK1^{s94d}$-VENUS-FLAG)	Generated in this study
phot::rwp5-C1&C2	RWP5-OE in *phot* background	Generated in this study
phot::bsk1	BSK1-OE in *phot* background (BSK1-VENUS-FLAG)	Generated in this study
phot::bsk1 S94A	phospho-site specific BSK1(S94A)-OE in *phot* background ($BSK1^{s94a}$-VENUS-FLAG)	Generated in this study
phot::bsk1 S94D	phospho-site specific BSK1(S94D)-OE in *phot* background ($BSK1^{s94d}$-VENUS-FLAG)	Generated in this study

6.2.2 Growth Conditions and Medium

All the experiments in this study were performed using standard (Tris Acetate

phosphate) TAP[209] medium or Sueoka's high salt medium (HSM)[210] medium.

For synchronized cultures, before the start of the experiment, cells were grown in 25 or 50ml liquid TAP medium in 150 ml or 250 ml flasks at 23 ℃, 120 rpm/min, 15 μmol/m^2/s to reach enough cell concentration. Then cells were allowed to grow in HSM medium for at least 5 days under the light/dark cycle conditions. (light/dark cycle 12h/12h: 12h dark and 12h light; Light: 50 μmol/m^2/s; Temperature, 18 ℃ in the dark, 23 ℃ in the light). Cells used for experiments were grown under the same conditions.

For non-synchronized culture, before the experiment, cells were grown in 25- or 50-ml liquid TAP medium in 150ml or 250ml flasks at 23 ℃, 120 rpm/min, 15 μmol/m^2/s. Cells used for experiments were grown under the same conditions. Low light (LL) was 50 μmol/m^2/s, while high light (HL) was 300μmol/m^2/s unless specified.

6.2.3 Cell count and chlorophyll measurement

For cell counts, 480 ml of algal culture was mixed with 20 μl of acetate solution, and cells were counted using Countess II Automated Cell Counter (Thermo Fisher Scientific, US).

Total chlorophyll (chlorophyll a+b) measurement was performed by using methanol extraction. 1 ml of collected culture was centrifugated at 11,000g for 5min and then the supernatant was discarded. Cell pellets were resuspended in 100% methanol and vortexed to extract the pigments. After 10 mins dark incubation, the chlorophyll concentration was determined using a UV–visible spectrophotometer (JASCO, Japan). The absorbance of the methanol extract was measured at 652, 655, 750 nm against a 100% methanol blank. The concentration of total chlorophyll (μg/ml) was calculated using the equation below[211].

$$\text{Chlorophyll } a \text{ (μg/ml)} = (16.29 * A665nm) - (8.54 * A652nm)$$
$$\text{Chlorophyll } b \text{ (μg/ml)} = (30.66 * A652nm) - (13.58 * A665nm)$$
$$\text{Chlorophyll } a + b = 22.12 * A652nm + 2.71 * A665nm$$

6.2.4 Fluorescence-based measurements

Fluorescence-based photosynthetic parameters were measured with a pulse modulated amplitude fluorimeter (MAXI-IMAGING-PAM, HeinzWaltz GmbH, Germany). Prior to the onset of the measurements, cells were acclimated to darkness for 15 min.

Chlorophyll fluorescence was recorded under different intensities of actinic light. Two programs were used for record in this study.

Program 1: starting with measurements in the dark (indicated as D below the x-axis of the graphs), followed by measurements at 21 μmol photons m^{-2} s^{-1} (indicated as L1 below the x-axis of the graphs) and 336 μmol photons m^{-2} s^{-1} (indicated as L2 below the x-axis of the graphs) and finishing with measurements of fluorescence relaxation in the dark.

Program 2: starting with measurements in the dark (indicated as D below the x-axis of the graphs), followed by measurements at 336 μmol photons m^{-2} s^{-1} (indicated as L2 below the x-axis of the graphs) and finishing with measurements of fluorescence relaxation in the dark.

YII was calculated as (Fm′ − F)/Fm′, NPQ was calculated as (Fm − Fm′)/Fm′ [212]. Fm is the maximal fluorescence yield in dark-adapted cells, F and Fm′ is the fluorescence yield in steady-state light and after a saturating pulse in the actinic light, respectively.

6.2.5 Transmission electron microscopy (TEM)

For transmission electron microscopy, the preparation of samples for TEM follows the previous protocol[213]. Sample cells were fixed in 0.1 M acodylate buffer (Sigma-Aldrich), pH 7.4, containing 2.5% glutaraldehyde (TAAB), 2% formaldehyde (Polysciences) for 1 h at room temperature. After the dehydration steps, the cells were infiltrated with ethanol/Epon resin mixture (2/3–1/3 for 1 h and 1/3–2/3 for 1 h) and finally embedded in Epon in a 60 °C oven for 48 h or longer. Ultrathin sections (60 nm) were prepared with a diamond knife on an UC6 Leica ultramicrotome and collected on 200 μM mesh nickel grids before examining on a JEOL 1200 EX electron microscope.

6.2.6 Starch and total carbohydrate content measurement

The starch content was determined by using Starch Assay Kit (Sigma-Aldrich, US) as described in the instruction. The result was calculated according to the standard curve made with glucose solution after starch digestion. Before starch digestion, samples were prepared as follows: 1 ml algal cells were pelleted by centrifugation at 11,000g for 5 min and then resuspended in 1ml 80% ethanol to incubate at 80℃ for 5 min. Another equal volume of 80% ethanol was added and the supernatant was discarded by centrifugation at 1,000g for 10 min. The previous step was repeated, the supernatant was carefully removed. After that there was 1 hour of waiting time to allow for the complete ethanol volatilization. Samples were now ready for starch digestion.

Total carbohydrate content was measured by the phenol–sulfuric acid method[182,214]. 0.1 ml of the sample was added to a 2 ml Eppendorf tube, spun down, and the supernatant was discarded. Then 0.1 ml of 5% phenol (v/v) solution and 0.1 ml of water were added and mixed. After mixing, 0.5 ml of concentrated sulfuric acid was added rapidly and mixed well by vortexing. After cooling down the samples for 10 min at room temperature to allow color development, 200 µl of the solution was pipetted into 96-well microplates (Greiner Bio-One, 655101) and the absorbance was measured at 488 nm with TECAN. The result was calculated based on a standard curve made with different concentrations of glucose solution (in the range from 0 µg/ml to 500 µg/ml).

6.2.7 Nucleic Acid Extraction

6.2.7.1 Genomic DNA extraction

Genomic DNA was extracted using Gentra Puregene Tissue Kit (Qiagen, Germany) following the manufacturer's online protocol (https://www.qiagen.com/cn/resources/resourcedetail?id=9a9d082d-26c5-40e1-8f7a-f6197d90ba91&lang=en).

For colony PCR, crude genomic DNA was extracted by using EDTA-method[215]. All the primers for colony PCR and screening are shown in Table 6.2 which were synthesized by Thermo-fisher or IDT.

Table 6.2. primers for colony PCR and screening in this study

Primer Name	Sense Primer
P1	GACGTTACAGCACACCCTTG
P2	GCACCAATCATGTCAAGCCT
P3	TTCCCTGCCGCTGCAACA
P4	AGTCAAGGTGGGTTCCAGTG
P5	GTTGCAAATGATCACAACGC
P6	CAAAGAGGAGGACGAGGATG
P7	CACATGTGTTGTCCTTTGGC

6.2.7.2 RNA extraction and cDNA synthesis

Total RNA was extracted using Trizol reagent (Invitrogen) or RNeasy Mini Kit (Qiagen). Two methods were used in this study described in follows:

Total RNA was extracted using Trizol reagent (Invitrogen, US). Around 10 million cells were pelleted by centrifugation at 4,000 g for 5 min. 1 ml of Trizol was added to the pellet, and cells were mixed by pipette up and down. Then cells were vortexed after adding 0.2 ml chloroform and incubated for 5 min at 25°C. The aqueous phase was aspirated from the Eppendorf tube after centrifugation at 13,400 g for 15 min at 4°C. The aspirated sample was precipitated with an equal volume of 100% isopropanol and incubated for 10 min at 25°C. Precipitated RNA was centrifuged at 13,400 g for 5 min at 4°C. The RNA pellet was washed with 75% ethanol, vortexed, and centrifuged at 13,400 g for 5 min in a refrigerated centrifuge. The supernatant was discarded, and the pellet was air-dried for 10 min. Isolated RNA was resuspended in 50 µl of RNAase and DNAase free water and quantified using Micro-Volume Full-Spectrum Fluorospectrometer (NanoDrop 3300, Thermo Scientific, US).

Total RNA was extracted using RNeasy Mini Kit (Qiagen, Germany) and treated with the RNase-Free DNase Set (Qiagen, Germany). 1 µg total RNA was reverse transcribed with oligo dT using Sensifast cDNA Synthesis kit (Meridian Bioscience, USA). qPCR reactions were performed and quantitated in a Bio-Rad CFX96 system using SsoAdvanced Universal SYBR Green Supermix (BioRad, US). The

housekeeping gene is G protein subunit-like protein (GBLP)[216]. The primers for RT-qPCR used in this study are shown in Table 6.3 which were synthesized by Thermo-fisher or IDT.

Table 6.3. RT-qPCR primers used in this study

Gene ID	Gene Name	Sense Primer	Anti-sense Primer
Cre08.g385500	*AMYA1*	GCATGGCAATGGCCGAGTA	AAAGCTTTCTCGTTGTGCCG
Cre06.g307150	*AMYB1*	GAATCGACCGCAGAATCGCA	GTGCATAGCCGTTATCGCTG
Cre07.g338451	*FBP1*	GTGTGTATGTTGGCTGTGGC	GAGACACTTACGGCGAGACA
Cre12.g510650	*FBP2*	GGGAGGCTGCATGGGATAC	GCAAATTGGTCGACAAGGCT
Cre01.g010900	*GAP3*	ATGGGACAAGGGACTTGCAG	GTGCACCTACGACTCTACCG
Cre07.g332300	*GWD2*	TTGAATGACGAGCACAGGCT	GGGCGTGAAAGCGTGAAAAT
Cre17.g698850	*ISA2*	CGAGTGCGAGTAAGACCTCG	CCACAACGTATGCAAACCCG
Cre12.g553250	*PFK2*	TGAAAGGCGGAGTTCTGTGT	TCAGAACGTGCGTGCTAGAG
Cre10.g450500	*PWD*	GCATCCACTTGGTCAGGGAA	CTACAGCGCGGATGCTATGA
Cre04.g215150	*SSS1*	TATGTCCGGCTCGATAGGGT	ATGACAACCCCACCACTGAC
Cre06.g282000	*STA3*	AGCCATCGACTACATCGAGC	CCCCTTGCCTCCTTTCTCTG
Cre03.g188250	*STA6*	GCACAGCTGCATAGAAGCAC	CTCGGTCCAAATCTCTGGGG
Cre02.g080200	*TRK1*	AAGAGCTCTGTCGGAACTGC	GCCGTTACAGGCGTTCAATC
Cre01.g032650	*TAL1*	GGTGGACAGCAGCCAACCT	TGGAGGCGAGCTTGATGTAG
Cre06.g283050	*LHCA1*	GCGAGGGACCAAGTCAATCA	CAGTTGTTAACACGCACGCA
Cre17.g720250	*LHCB4*	TCGGAGTTCGTGGTGATTGG	CTCCGAGCTAACCTCATCGG
Cre01.g010900	*GAP3*	ACTGAGGAGTGGACCTGTGA	CCATCGCCCTCCCCATTAAG
Cre07.g354200	*GAP2*	ATTGGGAAGTGCATACGGGG	GCACATGCGTGTCACAAACT
Cre12.g485150	*GAP1*	GATCGGCATCAACGGCTTCG	CGGCATCAATGAATGGGTCG
Cre12.g556600	*GAPN1*	GACAAGGAGTCCTACACGCT	AGGCCCGGCAAAGACAAATC
Cre01.g016600	*PSBS1*	TAAACCGTGTATTGGAACTCCG	CTCTGCACGCGGCGTGTT
Cre01.g016750	*PSBS2*	GGTTACGCACGTTATCACTTGT	CTCTCTGCACGCGTTGTACG
Cre08.g367500	*LHCSR3.1*	CACAACACCTTGATGCGAGATG	CCGTGTCTTGTCAGTCCCTG
Cre08.g367400	*LHCSR3.2*	TGTGAGGCACTCTGGTGAAG	CGCCTGTTGTCACCATCTTA
Cre08.g365900	*LHCSR1*	GAGTCTGAGATCACCCACGG	CCGATCTGCTGGAAGTGGTA
Cre01.g001600	*GAPR1*	CGGGGCAATTAGTGTCCAGA	ATTGCGGAAGCTTGGACTGA
Cre02.g118250	*GAPR2*	GAAGGATGCCGAGGACTGAG	GGGAATGCAGCCAACCCTAA
Cre06.g262800	*GAPR3*	ATCATGGCGAACATGCTGGA	CACGGTCAGGGACTTGGC
Cre07.g349152	*GAPR4*	CACGTGATAGTAAAGCGCGG	GACAGAGGCGTTGCGAATCT
Cre11.g480950	*GAPR5*	AAGCGCACCTGTTGCGTAAT	AAACAACAGGCCCAGAAAGC

Gene ID	Gene Name	Sense Primer	Anti-sense Primer
Cre13.g606750	*GAPR6*	GTAGGTAGTGACAGTGCGGG	GCCAGTCAGCGATGTCTCAAT
Cre16.g672300	*GAPR7*	GCATGCGTTTCTTCTGCTGT	GTGTTTAGGCCACGGAATGC
Cre06.g285600	*RWP5*	GGCGTAATGGACGCATGAAT	CCTCCGTGCCTGGAACATTA
Cre03.g199000	PHOT	GACGGACCTTCTTGGCTTCT	TTCAAGCAAAGACCAAGTCCT
Cre16.g659400	BSK1	GCTACTCCAACTTGGGCCTT	AGGGTGATTACAGCCCTTGC
Cre06.g278222	*GBLP*	TGGCTTTCTCGGTGGACAAC	CTCGCCAATGGTGTACTTGC

6.2.8 Gene cloning and Plasmid construction

6.2.8.1 Gene cloning and Plasmid construction

PHOT, GAP1, GAPR4, RWP5, BSK1, BSK1^{S94A}, and BSK1^{S94D} gDNA or CDS were generated by amplifying from gDNA or cDNA, or plasmid using different PCR enzymes.

PHOT CDS: The *PHOT* CDS was amplified by PCR from another plasmid which was preserved in our own lab by using primers G1 and G2 and KOD Hot Start DNA polymerase. The amplified fragment was digested by EcoRI and BamHI restriction enzymes and then was inserted into the EcoRI/BamHI site of vector phk330[60] to produce the transformation vector phk330-phot. Thus, the *PHOT* CDS is under the control of the Rubisco promoter, followed by a selectable marker of *ble* gene conferring resistance to zeocin. This plasmid will express PHOT protein and zeocin resistance protein separately in *vivo* because of the presence of the F2A sequence[217].

PHOT gDNA: The fragment 1 (1 bp to 3334 bp) and fragment 2 (3319 bp to 6536 bp) of *PHOT* gDNA were amplified from gDNA using the platinum polymerase and primer G3-G6. Two fragments together with HpaI digested pLM005 plasmid and assembled by Gibson assembly. The *PHOT* gDNA is under the control of the PSAD promoter, followed by the linked Venus-FLAG tag.

GAP1: GAP1 CDS was amplified by PCR using KOD Hot Start DNA polymerase and primers G7 and G8 from WT cDNA. GAP1 gDNA was amplified by PCR from WT gDNA using the platinum polymerase and primers G9 and G10. GAP1 CDS and GAP1 gDNA together with HpaI digested pLM005 or pRAM118 plasmid assembled by Gibson assembly.

GAPR4 gDNA: GAPR4 gDNA was amplified by PCR from WT gDNA using the platinum polymerase and primers G11 and G12. GAPR4 gDNA together with HpaI digested pLM005 plasmid and assembled by Gibson assembly.

RWP5: RWP5 gDNA was amplified by PCR using KOD Hot Start DNA Polymerase, Phusion HotstartII polymerase, and Platinum superfii DNA Polymerase. In particular, the fragment 1 from 1bp to 1710 bp of RWP5 was amplified from gDNA using the Platinum polymerase and primer G13 and G15; the fragment 2 from 1691 bp to 3411 bp of RWP5 was amplified from gDNA using Platinum and Phusion polymerase and primer G16 and G17; the fragment 3 from 3391 bp to 5233 bp of RWP5 was amplified from gDNA using Platinum and KOD polymerase and primers G18 and G14. Extension PCR was performed to ligate the three fragments. The details of the PCR program and conditions used for amplification were shown in Table 4.4.

BSK1: Full-length BSK1 gDNA was amplified from WT gDNA using the Platinum polymerase and primers G19 and G20. To induce point mutation, primers G21-G24 were used to get $BSK1^{S94A}$ and $BSK1^{S94D}$. The mutated base pair are underlined.

All the primers used in gene amplification were shown in Table 6.4.

Table 6.4 Primers for gene cloning is this study

Primer Name	Primer Sequence
G1	CCCGGATCCATGGCAGGGGTGC
G2	ttagaattctcagtagttgtcgaacgccg
G3	gctactcacaacaagcccagttATGGCAGGGGTGAGCAC
G4	gagccacccagatctccgttGTAGTTGTCGAACGCCGCG
G5	AATCACGTCTCCCACGTGTTG
G6	ATGGGGGCTGGGGTGTC
G7	GCAGGAGATTCGAGGTTATGGCCCTCCTCGCCC
G8	CCTTGGACACCATGTTCACCTTGACGGCGGCC
G9	gctactcacaacaagcccagttATGGCCCTCCTCGCCCA
G10	gagccacccagatctccgttCACCTTGACGGCGGC
G11	gctactcacaacaagcccagttATGGAGCTGGAGGACGCG
G12	gagccacccagatctccgttGTGGGCGCGCGG
G13	gctactcacaacaagcccagttATGCTAGACACTTTCGACAGCCT

Primer Name	Primer Sequence
G14	gagccacccagatctccgttATCCTCCTGCGGCACCT
G15	GCCGCCGCCGCCATCAGCAGCC
G16	CTGCTGATGGCGGCGGCGGCCT
G17	CCCTGTTGACCAAGGAGGCAGGT
G18	CTGCCTCCTTGGTCAACAGGGC
G19	gctactcacaacaagcccagttATGTCTTCTGGTCGGAAGCG
G20	gagccacccagatctccgttCTTCCCGCGCAGCTTGG
G21	TCGTGGGGggcAGATACACCATCAGCCCGC
G22	GTGTATCTgccCCCCACGAGGTATGCGTGG
G23	TCGTGGGGgtcAGATACACCATCAGCCCGC
G24	GTGTATCTgacCCCCACGAGGTATGCGTGG

6.2.8.2 Gibson Assembly

The Gibson assembly methods were modified from Gibson[177]. Fifteen fmol of each fragment were mixed together to a total volume of 5µl. and then 15µl isothermal Gibson assembly reaction mix were added and the resulting solution was mixed well by pipetting. After a 60 min-long incubation at 50°C the reaction products were ready to be used for bacteria transformation.

All PCR products were gel extracted and purified before Gibson assembly. The Gene of interest was mixed with a purified and linearized pLM005 or pRam118 plasmid, cloned in-frame with a C-terminal Venus-3Xflag tag. All the plasmid sequence has been confirmed by sequencing before using for transformation.

6.2.9 Bacterial Transformation

The bacterial transformation was performed using lab protocol as follows; 50µl of thawed competent bacteria and 1 - 5µl of DNA (usually 10pg - 10ng) were added to a pre-chilled Eppendorf tube and mixed gently. The suspensions were kept on ice for 30 min, and after that, the mixture was heated at 42°C for 45s. Following the heat-shock step, the tubes were kept on ice for another 2 min. Then add 1ml LB medium to the mixture and grow them in a shaking incubator for 1h at 37°C. Plate some or all the transformation onto a 10 cm LB agar plate containing the appropriate antibiotic. Incubate plates at 37°C overnight. For the plasmid size more than 8kbp, we used

Ecoli. DH10β, otherwise using DH5a.

6.2.10 Transformation of Chlamydomonas reinhardtii

The transformation was performed by electroporation, which follows the protocol of Zhang et al.[218] with modification. Cells for transformation were collected at 1-2h before the end of the light phase in a synchronized (12h light/ 12h dark) culture. Before using for transformation, 30 min of heat shock (40 ℃) was applied in a thermomixer with gentle mixing at 350 rpm. After 30 min recovery at RT, cells were ready for the transformation. For three reactions 11ng/kb linearized plasmid was mixed with 400 µl of 1.0 x10^7 *Chlamydomonas reinhardtii* cells/ml and electroporated in a volume of 125 ml in a 2-mm-gap electro cuvette using a NEPA21 square-pulse electroporator, using two poring pulses of 250 and 150 V for 8 msec each, and five transfer pulses of 50 msec each starting at 20 V with a "decay rate" of 40% (i.e., successive pulses of 20, 12, 7.2, 4.3, and 2.6 V). Electroporated cells were immediately transferred to a 15ml centrifugation tube containing 9ml TAP plus 40mM sucrose. After overnight incubation under low light, cells were collected by centrifugation and spread on TAP agar plates which contain the appropriate antibiotic (10µg/ml paromomycin or 7.5µg/ml zeocin or 10µg/ml hygromycin B). Transformants typically appear after 7 days.

6.2.11 Complementation strains screening

The plasmids were transformed into different strains using electroporation (described above), and then the putative antibiotic-resistant transformants were transferred into individual wells of a 96-well, flat-bottom transparent microplate; each well contained 250 µl of TAP medium. Cultures were grown for 3d without shaking under 15 µmol/m^2/s light, refreshed by replacing half of the culture with fresh medium, and were allowed to grow for an additional day. Transformants were then screened for Venus expression using a fluorescent microplate reader (using black 96-well plates). Parameters used: Venus (excitation 515/12 nm and emission 550/12 nm) and chlorophyll (excitation 440/9 nm and emission 680/20 nm). The fluorescence signal was normalized to the chlorophyll fluorescence signal. Colonies with high Venus/chlorophyll value were chosen as putative complemented strains. These putative positive transformants were further validated by western blotting using anti-

FLAG or RT-qPCR.

6.2.12 Google scholar web crawling

Source code for google scholar web crawling:

```
# coding:utf-8

"""
Google scholar
2021.01.05
"""
import random
import requests
import time
import re
from lxml import etree
from faker import Faker
import pandas as pd
import sys


class Scholar_Spider(object):
    def __init__(self):
        self.article_url =
'https://scholar.google.com/scholar?hl=zh-
CN&as_sdt=0%2C5&q=62&btnG=#'
        self.title_info_url =
'https://scholar.google.com/scholar?q=info:{}:scholar.google.com/&out
put=cite&scirp=0&hl=zh-CN'

    def get_data(self, url, wd):
        s = requests.session()
        s.verify = False
        faker = Faker()
        headers = {
            'user-agent': faker.user_agent(),
        }
        resp = s.get(url.format(wd), headers=headers)
        return resp.text

    def open_xlsx_read(self, file_name='workbook1.xlsx'):
        """
        读取 excel 文件
        :return: id 列表
        """
        data = pd.read_excel(file_name, keep_default_na=False)
        locusname_list = data['locusName'].values
        a_all_list = data['A_all'].values
        a_dc_list = data['A_DC'].values
        print(len(locusname_list) == len(a_all_list) ==
len(a_dc_list))
        name_dict = {}
        for i in range(len(locusname_list)):
            new_list = []
            new_list.append(a_all_list[i])
            new_list.append(a_dc_list[i])
            name_dict[locusname_list[i]] = new_list
```

```
48            print(len(name_dict))
49            return name_dict
50
51        def create_word_list(self, k, v):
52            """
53            创建关键词列表，进行优先级排序处理
54            :param k:
55            :param v:
56            :return:
57            """
58            word_list = []
59            word_list.append(k)
60            # 第二列
61            if len(v[0]) == 0:
62                pass
63            else:
64                if '(' in v[0]:
65                    # print('( 在v[0]中')
66                    r = re.findall('.*?\((.*?)\)', v[0])
67                    # print('提取括号里面的内容', r)
68                    if len(r) == 0:
69                        pass
70                    else:
71                        if ',' in r[0]:
72                            new_r = str(r[0]).split(',')
73                            for i in range(len(new_r)):
74
word_list.append(str(new_r[i]).replace(' ', ''))
75                        else:
76                            word_list.append(str(r[0]).replace(' ',
''))
77                    r = re.findall('(.*?)\(.*?\)', v[0])
78                    # print('提取括号外面的内容', r)
79                    if len(r) == 0:
80                        pass
81                    else:
82                        if ',' in r[0]:
83                            new_r = str(r[0]).split(',')
84                            for i in range(len(new_r)):
85
word_list.append(str(new_r[i]).replace(' ', ''))
86                        else:
87                            if len(r[0]) == 0:
88                                pass
89                            else:
90                                word_list.append(str(r[0]).replace('
', ''))
91                else:
92                    # print('( 不在v[0]中')
93                    if ',' in v[0]:
94                        new_r = str(v[0]).split(',')
95                        for i in range(len(new_r)):
96                            word_list.append(str(new_r[i]).replace('
', ''))
97                    else:
98                        if len(v[0]) == 0:
99                            pass
100                       else:
101                           word_list.append(str(v[0]).replace(' ',
''))
```

```
102             # 第三列
103             if len(v[1]) == 0:
104                 pass
105             else:
106                 if '(' in v[1]:
107                     # print('( 在v[1]中')
108                     r = re.findall('.*?\((.*?)\)', v[1])
109                     # print('提取括号里面的内容', r)
110                     if len(r) == 0:
111                         pass
112                     else:
113                         if ',' in r[0]:
114                             new_r = str(r[0]).split(',')
115                             for i in range(len(new_r)):
116 
word_list.append(str(new_r[i]).replace(' ', ''))
117                         else:
118                             word_list.append(str(r[0]).replace(' ',
''))
119                     r = re.findall('(.*?)\(.*?\)', v[1])
120                     # print('提取括号外面的内容', r)
121                     if len(r) == 0:
122                         pass
123                     else:
124                         if ',' in r[0]:
125                             new_r = str(r[0]).split(',')
126                             for i in range(len(new_r)):
127 
word_list.append(str(new_r[i]).replace(' ', ''))
128                         else:
129                             if len(r[0]) == 0:
130                                 pass
131                             else:
132                                 word_list.append(str(r[0]).replace('
', ''))
133                 else:
134                     # print('( 不在v[0]中')
135                     if ',' in v[1]:
136                         new_r = str(v[1]).split(',')
137                         for i in range(len(new_r)):
138                             word_list.append(str(new_r[i]).replace('
', ''))
139                     else:
140                         if len(v[1]) == 0:
141                             pass
142                         else:
143                             word_list.append(str(v[1]).replace(' ',
''))
144 
145             return word_list
146 
147         def get_title_data(self, cid_list):
148             # 还剩一个保存操作
149             save_str = ''
150             for cid in cid_list:
151                 html = self.get_data(self.title_info_url, cid)
152                 r2 = re.findall('请进行人机身份验证', html)
153                 if r2:
154                     print("出现人机验证，  程序退出")
```

```
155                 sys.exit()
156             ht2 = etree.HTML(html)
157             info_text =
ht2.xpath('/html/body/div[1]/table/tr[1]/td/div/text()')
158             if len(info_text) == 0:
159                 pass
160             else:
161                 # print(str(info_text[0]).replace('\n
', ''))
162                 save_str += str(info_text[0]).replace('\n
', '') + ', '
163             time.sleep(random.randint(1, 2))
164         return save_str
165
166     def save_to_csv(self, data, file_name='data.csv'):
167         with open(file_name, 'a', encoding='utf-8', newline='')
as fp:
168             fp.write(data)
169             fp.flush()
170
171     def run(self):
172         name_dict = self.open_xlsx_read()
173         for k, v in name_dict.items():
174             wd_str = k + ', ' + v[0] + ', ' + v[1] + ', '
175             print(k, '----', v[0], v[1])
176             # 创建关键词列表，按照顺序排列
177             word_list = self.create_word_list(k, v)
178             for wd in word_list:
179                 print(wd)
180                 html = self.get_data(self.article_url, wd)
181                 r = re.findall('与所有文章均不相符', html)
182                 r2 = re.findall('请进行人机身份验证', html)
183                 if r2:
184                     print("出现人机验证， 程序退出")
185                     sys.exit()
186                 if r:
187                     print("与所有文章均不相符, 下一个关键词")
188                     continue
189                 else:
190                     print('有数据，解析，保存')
191                     ht = etree.HTML(html)
192                     cid_list =
ht.xpath('/html/body/div/div[10]/div[2]/div[2]/div[2]/div/@data-cid')
193                     print(cid_list)
194                     title_str = self.get_title_data(cid_list)
195                     self.save_to_csv(wd_str + title_str + '\n')
196                     break
197             time.sleep(random.randint(1, 2))
198
199
200 if __name__ == '__main__':
201     s = Scholar_Spider()
202     s.run()
```

6.2.13 Immunoblotting

Protein samples of whole-cell extracts (0.5 µg chlorophyll or 10 µg protein- 20 µg protein in the case of PSBS detection) were loaded on 4-20% SDS-PAGE gels (Mini-PROTEAN TGX Precast Protein Gels, Bio-Rad) and blotted onto nitrocellulose membranes. Antisera against LHCSR1 (AS14 2819), LHCSR3 (AS14 2766), and ATPB (AS05 085) were from Agrisera (Vännäs, Sweden), Anti-FLAG was from ThermoFisher, US; previously described was antisera against *C. reinhardtii* PSBS (6). ATPB was used as a loading control. An anti-rabbit horseradish peroxidase-conjugated antiserum was used for detection. The blots were developed with the ECL detection reagent, and images of the blots were obtained using a CCD imager (ChemiDoc MP System, Bio-Rad) or ImageQuant^{TM} 800 (GE Healthcare, US).

6.2.14 Phylogenetic analysis of BSK1

To identify putative orthologs of the AtCBC1, AtCBC2 and AtHT1, the three full-length proteins were used as a query to run a BLASTp (Altshul, 1990) on the NCBI non-redundant sequence database[219].

The alignment of the 131 protein sequences was carried out at the amino acid level using Muscle[220]. During the process, alignments was curated using both the multiple alignment editing softwares ClustalX 2.0[221] and Jalview 2[222]. The phylogenetic relationship was inferred using the maximum likelihood method PhyML[223]. The evolutionary model selection was done with MEGA 7[224] under the Bayesian information criterion[225]. According to the model selection process, the best evolutionary was GTR+G+I[226]. The maximum likelihood tree was evaluated with aLRT[227], a non-parametric branch support based on a Shimodaira-Hasegawa-like procedure. Phylogenetic trees were visualized using FigTree (http://tree.bio.ed.ac.uk/software/figtree/).

6.2.15 Statistical analysis

All statistical tests were performed using the computing environment Prism 9 (Graphpad Software, LLC), with a significance level of 0.05. In order to confirm mRNA accumulation data to the distributional assumptions of ANOVA, i.e. the residuals should be normally distributed and variances should be equal among groups,

Two-Way Analysis of Variance was computed with log-transformed data Y=logX where X is mRNA accumulation[228]. P < 0.005, **, P < 0.01, ***, P < 0.001, ****, P < 0.0001.

6.3 Supplementary data

6.3.1 The map of plasmids used in this study

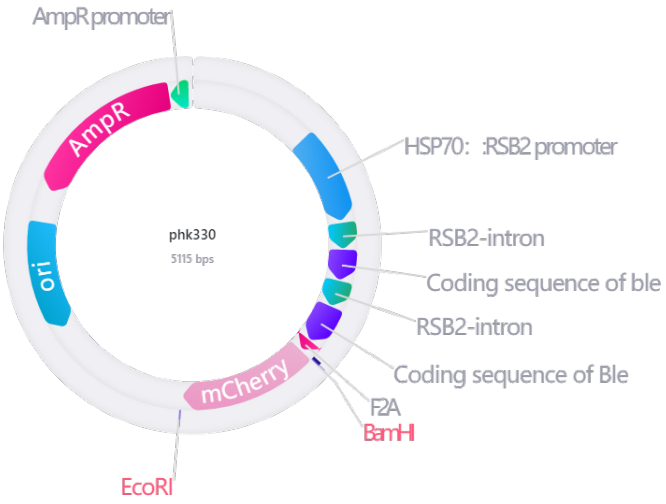


Figure. 6.2. pHK330 plasmid map

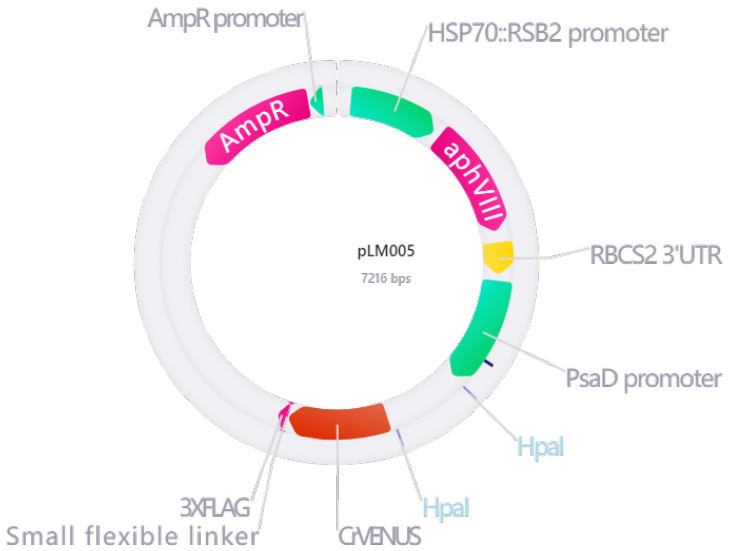


Figure. 6.3. pLM005 plasmid map

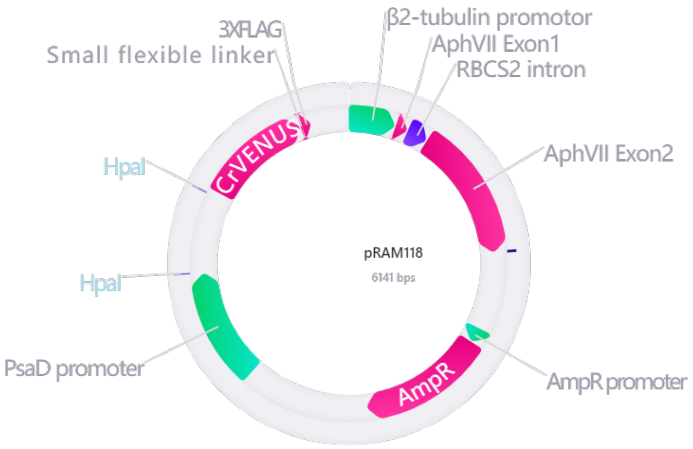


Figure. 6.4. Pram118 plasmid map

6.3.2 *rwp5* and *myb* colony PCR products sequence results.

```
MYB-5'          ACACGTCCGAGGACCCTGACGAGGACCTGCTGCCTGACGACCTGGATGACTCTCTGGACG
Sequencing      ATACGTACGAGGACCCTGACGAGGACATGGTGCCTGACGACCTGGATGACTCTCTGGACG
                * **** ******************** ** ******************************

MYB-5'          ACTCCA-CCACCAGCCGCACCGCACACACCGCAGCAGCGGCCGCGGCATCAGGCGGAGGA
Sequencing      ACTCCAACCACCAGCCGCACCGCACACACCGCAGCAGCGGCCGCGGCATCAGGCGGAGGA
                ****** ******************************************************

MYB-5'          AGAGGC---GTTG--------------------GACTCAGGAGCAACTAGTCTGCCAGAAG
Sequencing      AGAGGCCGGACTGCCAGCCTGACTCACTCAGTGACTCAGGAGCAACTAGTCTGCCAGAAG
                ******    **                 ****************************

MYB-5'          GCTCGACGTCAGTGCCCCGACGCCAGCGCGCCGTCTAGCGCGGGGCGTGTTGCAGCGGCAG
Sequencing      GCTCGACGTCAGTGCCCCGACGCCAGCGCGCCGTCTTGCGCGAAGCGT-TTCCAGGGCCAG
                ************************************ *****  **** ** *** * ***
```

Figure 6.5. Sequencing results using the Multiple Sequence Alignment by CLUSTALW-*myb*. CIB1 cassette was highlighted in yellow.

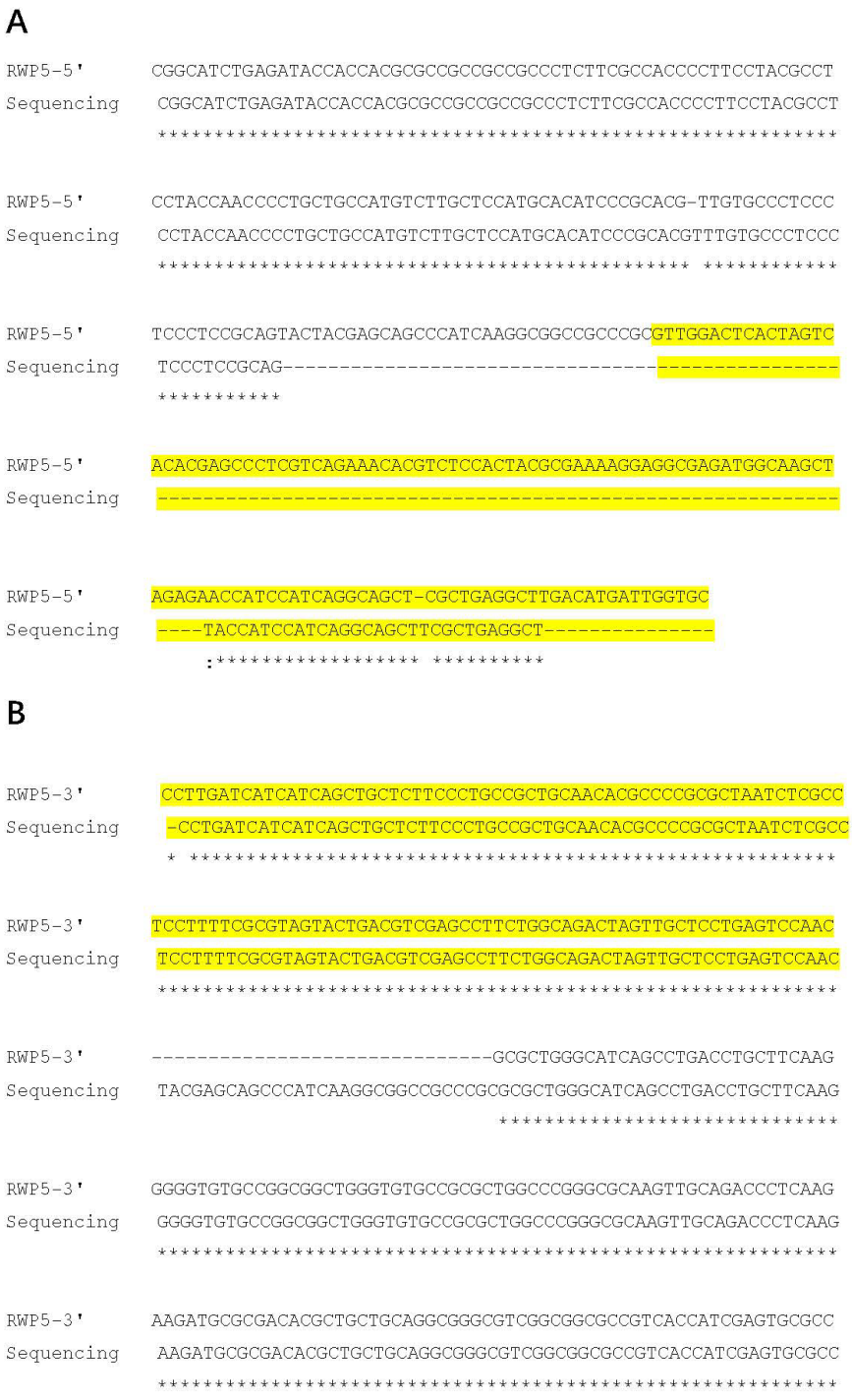


Figure 6.6. Sequencing results using the Multiple Sequence Alignment by CLUSTALW-*rwp5*. CIB1 cassette was highlighted in yellow.

6.3.3 Phylogenetic tree of BSK1

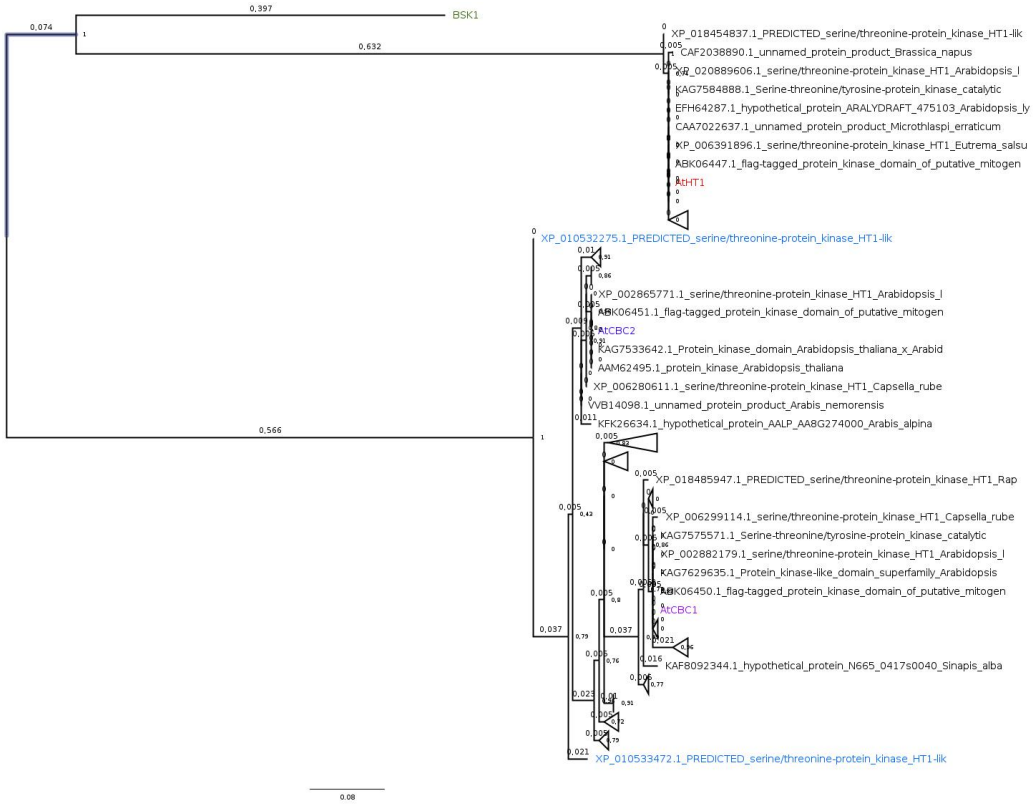


Figure 6.7. Full Phylogenetic tree of BSK1. (methods in 6.2.14, triangle nodes is the collapse of other proteins due to the paper size limitation**)**

6.4 List of manuscripts during Ph.D.

1) P. Redekop*, E. Sanz-Luque*, **Y. Yuan**, G. Villain, D. Petroutsos, A. R. Grossman. "Transcriptional regulation of photoprotection in dark-to-light transition - more than just a matter of excess light energy"
https://doi.org/10.1101/2021.10.23.463292

Science Advances, In press

2) A. Ruiz-Sola*, S. Flori*, **Y. Yuan***, G. Villain, E. Sanz-Luque, P. Redekop, R. Tokutsu, A. Tsichla, G. Allorent, M. Arend, F. Iacono, A. Kueken, G. Finazzi, M. Hippler, Z. Nikoloski, J. Minagawa, A. R. Grossman, D. Petroutsos. "Photoprotection is regulated by light independent CO2 availability"
https://doi.org/10.1101/2021.10.23.465040

Submitted to Nature Commuincations, co-first author

3) M. Arend*, **Y. Yuan***, N. Omranian, Z. Nikoloski, D. Petroutsos. "Widening the landscape of transcriptional regulation of algal photoprotection"
https://doi.org/10.1101/2022.02.25.482034

Under review in Nature Commuincations, co-first author

4) **Y. Yuan**, E. Werth M. Arend, A. Tsichla, G. Kepesidis, A. Ruiz-Sola, , N. Omranian, Z. Nikoloski, P. Hegemann, O. Bastien, L. Hicks, D. Petroutsos. "Phototropin controls carbon partitioning in the green microalgal *Chlamydomonas reinhardtii*"

Manuscript in preparation

5) F. Iacono, **Y. Yuan**, S. M. Esteves, A. Jadoul, G. Mossay, R. Houet, M. Hanikenne, T. Druet, D. Petroutsos, P. Cardol. "Link between carbon-concentrating mechanism and regulation of the trans-thylakoidal proton-motive force highlighted in the *Chlamydomonas reinhardtii* nigericin-insensitive field-isolated strain CC-2936"

Manuscript in preparation

6.5 Conferences attended during Ph.D.

I-BE-C 2021: The 1st International Bioenergy & Environment Congress (Web Congress), CEA Cadarache, France. 16th-18th February 2021. Poster Presentation, **Phototropin controls carbon Partitioning in the *Chlamydomonas reinhardtii*.** (Won the best poster prize-200euros)

Sfphi2021: Meeting of the French Photosynthesis Society (Web Congress), France. 8th-9th April 2021. Oral Presentation selected from abstract, **Phototropin controls carbon Partitioning in the *Chlamydomonas reinhardtii.***

ISPP2021: Plant Photobiology (Virtual). Cold Spring Harbor Laboratory, US. 22nd-25th July 2021. Oral Presentation selected from abstract, **Phototropin controls carbon Partitioning in the *Chlamydomonas reinhardtii.***

Chlamy 2020+1: The 19th International Conference on the Cell and Molecular Biology of Chlamydomonas, France. 29th August-3rd September 2021. Oral Presentation selected from abstract, **Phototropin controls carbon Partitioning in the *Chlamydomonas reinhardtii.***

Postdoc Round Table Seminar: Carnegie Institution for Science-Department of Plant Biology, Standford, CA, United States. 23rd March 2022. Oral Presentation, **PHOTOTROPIN links blue-light perception and starch accumulation in *Chlamydomonas.*** (Invited by Petra Redekop)

6.6 References

1 Bankaitis, V. A. The chlamydomonas sourcebook. *Cell* **61**, 559-560, doi:10.1016/0092-8674(90)90467-s (1990).

2 Lefebvre, P. A. & Silflow, C. D. Chlamydomonas: The Cell and Its Genomes. *Genetics* **151**, 9-14, doi:10.1093/genetics/151.1.9 (1999).

3 Merchant, S. S. *et al.* The Chlamydomonas genome reveals the evolution of key animal and plant functions. *Science* **318**, 245-250, doi:10.1126/science.1143609 (2007).

4 Popescu, C. E. & Lee, R. W. Mitochondrial genome sequence evolution in Chlamydomonas. *Genetics* **175**, 819-826, doi:10.1534/genetics.106.063156 (2007).

5 DeShaw, A. E., Figueroa-Martinez, F. & Reyes-Prieto, A. Complete chloroplast genomes of the Chlamydomonas reinhardtii nonphotosynthetic mutants CC-1375, CC-373, CC-4199, CC-2359 and CC-1051. *Mitochondrial DNA B Resour* **2**, 405-407, doi:10.1080/23802359.2017.1347838 (2017).

6 Shin, S. E. *et al.* CRISPR/Cas9-induced knockout and knock-in mutations in Chlamydomonas reinhardtii. *Sci Rep* **6**, 27810, doi:10.1038/srep27810 (2016).

7 Baek, K. *et al.* DNA-free two-gene knockout in Chlamydomonas reinhardtii via CRISPR-Cas9 ribonucleoproteins. *Sci Rep* **6**, 30620, doi:10.1038/srep30620 (2016).

8 Greiner, A. *et al.* Targeting of Photoreceptor Genes in Chlamydomonas reinhardtii via Zinc-Finger Nucleases and CRISPR/Cas9. *Plant Cell* **29**, 2498-2518, doi:10.1105/tpc.17.00659 (2017).

9 Kim, J., Lee, S., Baek, K. & Jin, E. Site-Specific Gene Knock-Out and On-Site Heterologous Gene Overexpression in Chlamydomonas reinhardtii via a CRISPR-Cas9-Mediated Knock-in Method. *Front Plant Sci* **11**, 306, doi:10.3389/fpls.2020.00306 (2020).

10 Bertalan, I. *et al.* A rapid, modular and marker-free chloroplast expression system for the green alga Chlamydomonas reinhardtii. *J Biotechnol* **195**, 60-66, doi:10.1016/j.jbiotec.2014.12.017 (2015).

11 Wannathong, T., Waterhouse, J. C., Young, R. E., Economou, C. K. & Purton, S. New tools for chloroplast genetic engineering allow the synthesis of human growth hormone in the green alga Chlamydomonas reinhardtii. *Appl Microbiol Biotechnol* **100**, 5467-5477, doi:10.1007/s00253-016-7354-6 (2016).

12 Sizova, I., Greiner, A., Awasthi, M., Kateriya, S. & Hegemann, P. Nuclear gene targeting in Chlamydomonas using engineered zinc-finger nucleases. *Plant J* **73**, 873-882, doi:10.1111/tpj.12066 (2013).

13 Kelterborn, S. *et al.* Gene Editing in Green Alga Chlamydomonas reinhardtii via CRISPR-Cas9 Ribonucleoproteins. *Methods Mol Biol* **2379**, 45-65, doi:10.1007/978-1-0716-1791-5_3 (2022).

14 Gao, H., Wang, Y., Fei, X., Wright, D. A. & Spalding, M. H. Expression activation and functional analysis of HLA3, a putative inorganic carbon transporter in Chlamydomonas reinhardtii. *Plant J* **82**, 1-11, doi:10.1111/tpj.12788 (2015).

15 Gao, H. *et al.* TALE activation of endogenous genes in Chlamydomonas

reinhardtii. *Algal Research* **5**, 52-60, doi:https://doi.org/10.1016/j.algal.2014.05.003 (2014).

16 Jiang, W., Brueggeman, A. J., Horken, K. M., Plucinak, T. M. & Weeks, D. P. Successful transient expression of Cas9 and single guide RNA genes in Chlamydomonas reinhardtii. *Eukaryot Cell* **13**, 1465-1469, doi:10.1128/EC.00213-14 (2014).

17 Wang, Q. *et al.* Photoactivation and inactivation of Arabidopsis cryptochrome 2. *Science* **354**, 343-347, doi:10.1126/science.aaf9030 (2016).

18 Li, X. *et al.* A genome-wide algal mutant library and functional screen identifies genes required for eukaryotic photosynthesis. *Nat Genet* **51**, 627-635, doi:10.1038/s41588-019-0370-6 (2019).

19 Petroutsos, D. *et al.* A blue-light photoreceptor mediates the feedback regulation of photosynthesis. *Nature* **537**, 563-566, doi:10.1038/nature19358 (2016).

20 Wakao, S. *et al.* Discovery of photosynthesis genes through whole-genome sequencing of acetate-requiring mutants of Chlamydomonas reinhardtii. *PLoS Genet* **17**, e1009725, doi:10.1371/journal.pgen.1009725 (2021).

21 Dent, R. M., Haglund, C. M., Chin, B. L., Kobayashi, M. C. & Niyogi, K. K. Functional genomics of eukaryotic photosynthesis using insertional mutagenesis of Chlamydomonas reinhardtii. *Plant Physiol* **137**, 545-556, doi:10.1104/pp.104.055244 (2005).

22 Mettler, T. *et al.* Systems Analysis of the Response of Photosynthesis, Metabolism, and Growth to an Increase in Irradiance in the Photosynthetic Model Organism Chlamydomonas reinhardtii. *Plant Cell* **26**, 2310-2350, doi:10.1105/tpc.114.124537 (2014).

23 Zhao, Q., Li, S., Shao, S., Wang, Z. & Pan, J. FLS2 is a CDK-like kinase that directly binds IFT70 and is required for proper ciliary disassembly in Chlamydomonas. *PLoS Genet* **16**, e1008561, doi:10.1371/journal.pgen.1008561 (2020).

24 Ruffer, U. & Nultsch, W. High-Speed Cinematographic Analysis of the Movement of Chlamydomonas. *Cell Motil Cytoskel* **5**, 251-263, doi:DOI 10.1002/cm.970050307 (1985).

25 Schmidt, J. A. & Eckert, R. Calcium couples flagellar reversal to photostimulation in Chlamydomonas reinhardtii. *Nature* **262**, 713-715, doi:10.1038/262713a0 (1976).

26 Segal, R. A., Huang, B., Ramanis, Z. & Luck, D. J. Mutant strains of Chlamydomonas reinhardtii that move backwards only. *J Cell Biol* **98**, 2026-2034, doi:10.1083/jcb.98.6.2026 (1984).

27 Kaye, Y. *et al.* The mitochondrial alternative oxidase from Chlamydomonas reinhardtii enables survival in high light. *J Biol Chem* **294**, 1380-1395, doi:10.1074/jbc.RA118.004667 (2019).

28 Rochaix, J. D. & Bassi, R. LHC-like proteins involved in stress responses and biogenesis/repair of the photosynthetic apparatus. *Biochem J* **476**, 581-593, doi:10.1042/BCJ20180718 (2019).

29 Garcia-Cerdan, J. G. *et al.* Chloroplast Sec14-like 1 (CPSFL1) is essential for normal chloroplast development and affects carotenoid accumulation in

Chlamydomonas. *Proc Natl Acad Sci U S A* **117**, 12452-12463, doi:10.1073/pnas.1916948117 (2020).

30 Xu, C., Moellering, E. R., Fan, J. & Benning, C. Mutation of a mitochondrial outer membrane protein affects chloroplast lipid biosynthesis. *Plant J* **54**, 163-175, doi:10.1111/j.1365-313X.2008.03417.x (2008).

31 Juppner, J. *et al.* Dynamics of lipids and metabolites during the cell cycle of Chlamydomonas reinhardtii. *Plant J* **92**, 331-343, doi:10.1111/tpj.13642 (2017).

32 Takeuchi, T. & Benning, C. Nitrogen-dependent coordination of cell cycle, quiescence and TAG accumulation in Chlamydomonas. *Biotechnol Biofuels* **12**, 292, doi:10.1186/s13068-019-1635-0 (2019).

33 Ball, S. G. & Morell, M. K. From bacterial glycogen to starch: understanding the biogenesis of the plant starch granule. *Annu Rev Plant Biol* **54**, 207-233, doi:10.1146/annurev.arplant.54.031902.134927 (2003).

34 Blaby, I. K. *et al.* Genome-wide analysis on Chlamydomonas reinhardtii reveals the impact of hydrogen peroxide on protein stress responses and overlap with other stress transcriptomes. *Plant J* **84**, 974-988, doi:10.1111/tpj.13053 (2015).

35 Wase, N., Black, P. N., Stanley, B. A. & DiRusso, C. C. Integrated quantitative analysis of nitrogen stress response in Chlamydomonas reinhardtii using metabolite and protein profiling. *J Proteome Res* **13**, 1373-1396, doi:10.1021/pr400952z (2014).

36 Colina, F., Carbo, M., Meijon, M., Canal, M. J. & Valledor, L. Low UV-C stress modulates Chlamydomonas reinhardtii biomass composition and oxidative stress response through proteomic and metabolomic changes involving novel signalers and effectors. *Biotechnol Biofuels* **13**, 110, doi:10.1186/s13068-020-01750-8 (2020).

37 Cross, F. R. & Umen, J. G. The Chlamydomonas cell cycle. *Plant J* **82**, 370-392, doi:10.1111/tpj.12795 (2015).

38 Torres-Romero, I. *et al.* Chlamydomonas cell cycle mutant crcdc5 over-accumulates starch and oil. *Biochimie* **169**, 54-61, doi:10.1016/j.biochi.2019.09.017 (2020).

39 Donnan, L., Carvill, E. P., Gilliland, T. J. & John, P. C. L. The Cell-Cycles of Chlamydomonas and Chlorella. *New Phytologist* **99**, 1-40, doi:DOI 10.1111/j.1469-8137.1985.tb03634.x (1985).

40 Huang, K. & Beck, C. F. Phototropin is the blue-light receptor that controls multiple steps in the sexual life cycle of the green alga Chlamydomonas reinhardtii. *Proc Natl Acad Sci U S A* **100**, 6269-6274, doi:10.1073/pnas.0931459100 (2003).

41 Ermilova, E. V., Zalutskaya, Z. M., Huang, K. & Beck, C. F. Phototropin plays a crucial role in controlling changes in chemotaxis during the initial phase of the sexual life cycle in Chlamydomonas. *Planta* **219**, 420-427, doi:10.1007/s00425-004-1241-6 (2004).

42 Zou, Y. *et al.* An Animal-Like Cryptochrome Controls the Chlamydomonas Sexual Cycle. *Plant Physiol* **174**, 1334-1347, doi:10.1104/pp.17.00493 (2017).

43 Pfennig, N. Photosynthetic bacteria. *Annu Rev Microbiol* **21**, 285-324, doi:10.1146/annurev.mi.21.100167.001441 (1967).

44 Nakamura, Y. *et al.* Complete genome structure of Gloeobacter violaceus PCC

7421, a cyanobacterium that lacks thylakoids. *DNA Res* **10**, 137-145, doi:10.1093/dnares/10.4.137 (2003).

45 Demmig-Adams, B. & Adams, W. W., 3rd. Harvesting sunlight safely. *Nature* **403**, 371, 373-374, doi:10.1038/35000315 (2000).

46 Li, Z., Wakao, S., Fischer, B. B. & Niyogi, K. K. Sensing and responding to excess light. *Annu Rev Plant Biol* **60**, 239-260, doi:10.1146/annurev.arplant.58.032806.103844 (2009).

47 Takahashi, S. & Badger, M. R. Photoprotection in plants: a new light on photosystem II damage. *Trends Plant Sci* **16**, 53-60, doi:10.1016/j.tplants.2010.10.001 (2011).

48 Schaller, K., David, R. & Uhl, R. How Chlamydomonas keeps track of the light once it has reached the right phototactic orientation. *Biophys J* **73**, 1562-1572, doi:10.1016/S0006-3495(97)78188-8 (1997).

49 Li, Z. *et al.* Lutein accumulation in the absence of zeaxanthin restores nonphotochemical quenching in the Arabidopsis thaliana npq1 mutant. *Plant Cell* **21**, 1798-1812, doi:10.1105/tpc.109.066571 (2009).

50 Havaux, M. & Niyogi, K. K. The violaxanthin cycle protects plants from photooxidative damage by more than one mechanism. *Proc Natl Acad Sci U S A* **96**, 8762-8767, doi:10.1073/pnas.96.15.8762 (1999).

51 Eberhard, S., Finazzi, G. & Wollman, F. A. The dynamics of photosynthesis. *Annu Rev Genet* **42**, 463-515, doi:10.1146/annurev.genet.42.110807.091452 (2008).

52 Niyogi, K. K., Li, X. P., Rosenberg, V. & Jung, H. S. Is PsbS the site of non-photochemical quenching in photosynthesis? *J Exp Bot* **56**, 375-382, doi:10.1093/jxb/eri056 (2005).

53 Külheim, C., Ågren, J. & Jansson, S. Rapid Regulation of Light Harvesting and Plant Fitness in the Field. *Science* **297**, 91-93, doi:10.1126/science.1072359 (2002).

54 Li, Z. *et al.* Evolution of an atypical de-epoxidase for photoprotection in the green lineage. *Nature Plants* **2**, 16140, doi:10.1038/nplants.2016.140 (2016).

55 Ledford, H. K. *et al.* Comparative profiling of lipid-soluble antioxidants and transcripts reveals two phases of photo-oxidative stress in a xanthophyll-deficient mutant of Chlamydomonas reinhardtii. *Molecular Genetics and Genomics* **272**, 470-479, doi:10.1007/s00438-004-1078-5 (2004).

56 Peers, G. *et al.* An ancient light-harvesting protein is critical for the regulation of algal photosynthesis. *Nature* **462**, 518-521, doi:10.1038/nature08587 (2009).

57 Dinc, E. *et al.* LHCSR1 induces a fast and reversible pH-dependent fluorescence quenching in LHCII in Chlamydomonas reinhardtii cells. *Proc Natl Acad Sci U S A* **113**, 7673-7678, doi:10.1073/pnas.1605380113 (2016).

58 Correa-Galvis, V. *et al.* Photosystem II Subunit PsbS Is Involved in the Induction of LHCSR Protein-dependent Energy Dissipation in Chlamydomonas reinhardtii. *J Biol Chem* **291**, 17478-17487, doi:10.1074/jbc.M116.737312 (2016).

59 Redekop, P. *et al.* PsbS contributes to photoprotection in Chlamydomonas reinhardtii independently of energy dissipation. *Biochim Biophys Acta Bioenerg* **1861**, 148183, doi:10.1016/j.bbabio.2020.148183 (2020).

60 Wang, Q., Peng, Z., Long, H., Deng, X. & Huang, K. Polyubiquitylation of alpha-tubulin at K304 is required for flagellar disassembly in Chlamydomonas. *Journal*

of cell science **132**, doi:10.1242/jcs.229047 (2019).
61 Allorent, G. & Petroutsos, D. Photoreceptor-dependent regulation of photoprotection. *Curr Opin Plant Biol* **37**, 102-108, doi:10.1016/j.pbi.2017.03.016 (2017).
62 Redekop, P. *et al.* Transcriptional regulation of photoprotection in dark-to-light transition- more than just a matter of excess light energy. *Sci Adv*, doi:In Press (2022).
63 Arend, M. *et al.* Widening the landscape of transcriptional regulation of algal photoprotection. *BioRxiv*, doi:10.1101/2022.02.25.482034 (2022).
64 Ruiz-Sola, M. Á. *et al. BioRxiv*, doi:10.1101/2021.10.23.465040 (2022).
65 Aihara, Y., Fujimura-Kamada, K., Yamasaki, T. & Minagawa, J. Algal photoprotection is regulated by the E3 ligase CUL4-DDB1(DET1). *Nat Plants* **5**, 34-40, doi:10.1038/s41477-018-0332-5 (2019).
66 Allorent, G. *et al.* UV-B photoreceptor-mediated protection of the photosynthetic machinery in Chlamydomonas reinhardtii. *Proc Natl Acad Sci U S A* **113**, 14864-14869, doi:10.1073/pnas.1607695114 (2016).
67 Tokutsu, R., Fujimura-Kamada, K., Matsuo, T., Yamasaki, T. & Minagawa, J. The CONSTANS flowering complex controls the protective response of photosynthesis in the green alga Chlamydomonas. *Nat Commun* **10**, 4099, doi:10.1038/s41467-019-11989-x (2019).
68 Reinfelder, J. R., Kraepiel, A. M. & Morel, F. M. Unicellular C4 photosynthesis in a marine diatom. *Nature* **407**, 996-999, doi:10.1038/35039612 (2000).
69 Kremer, B. P. C4-Metabolism in Marine Brown Macrophytic Algae. *Zeitschrift für Naturforschung C* **36**, 840-847, doi:10.1515/znc-1981-9-1024 (1981).
70 Liu, D. *et al.* Role of C4 carbon fixation in Ulva prolifera, the macroalga responsible for the world's largest green tides. *Commun Biol* **3**, 494, doi:10.1038/s42003-020-01225-4 (2020).
71 Yang, Y., Kwon, H. B., Peng, H. P. & Shih, M. C. Stress Responses and Metabolic Regulation of Glyceraldehyde-3-Phosphate Dehydrogenase Genes in Arabidopsis. *Plant Physiology* **101**, 209-216, doi:10.1104/pp.101.1.209 (1993).
72 Kirch, H.-H., Bartels, D., Wei, Y., Schnable, P. S. & Wood, A. J. The ALDH gene superfamily of Arabidopsis. *Trends in Plant Science* **9**, 371-377, doi:10.1016/j.tplants.2004.06.004 (2004).
73 Conley, T. R., Park, S. C., Kwon, H. B., Peng, H. P. & Shih, M. C. Characterization of cis-acting elements in light regulation of the nuclear gene encoding the A subunit of chloroplast isozymes of glyceraldehyde-3-phosphate dehydrogenase from Arabidopsis thaliana. *Mol Cell Biol* **14**, 2525-2533, doi:10.1128/mcb.14.4.2525-2533.1994 (1994).
74 Kwon, H. B. *et al.* Identification of a light-responsive region of the nuclear gene encoding the B subunit of chloroplast glyceraldehyde 3-phosphate dehydrogenase from Arabidopsis thaliana. *Plant Physiol* **105**, 357-367, doi:10.1104/pp.105.1.357 (1994).
75 Rius, S. P., Casati, P., Iglesias, A. A. & Gomez-Casati, D. F. Characterization of an Arabidopsis thaliana mutant lacking a cytosolic non-phosphorylating glyceraldehyde-3-phosphate dehydrogenase. *Plant Mol Biol* **61**, 945-957, doi:10.1007/s11103-006-0060-5 (2006).

76 Fromm, S., Senkler, J., Eubel, H., Peterhänsel, C. & Braun, H. P. Life without complex I: proteome analyses of an Arabidopsis mutant lacking the mitochondrial NADH dehydrogenase complex. *J Exp Bot* **67**, 3079-3093, doi:10.1093/jxb/erw165 (2016).

77 Senkler, J. *et al.* The mitochondrial complexome of Arabidopsis thaliana. *Plant J* **89**, 1079-1092, doi:10.1111/tpj.13448 (2017).

78 Nguyen-Kim, H. *et al.* Arabidopsis thaliana root cell wall proteomics: Increasing the proteome coverage using a combinatorial peptide ligand library and description of unexpected Hyp in peroxidase amino acid sequences. *PROTEOMICS* **16**, 491-503, doi:https://doi.org/10.1002/pmic.201500129 (2016).

79 Munoz-Bertomeu, J. *et al.* Plastidial glyceraldehyde-3-phosphate dehydrogenase deficiency leads to altered root development and affects the sugar and amino acid balance in Arabidopsis. *Plant Physiol* **151**, 541-558, doi:10.1104/pp.109.143701 (2009).

80 Tardif, M. *et al.* PredAlgo: a new subcellular localization prediction tool dedicated to green algae. *Mol Biol Evol* **29**, 3625-3639, doi:10.1093/molbev/mss178 (2012).

81 Kuken, A. *et al.* Effects of microcompartmentation on flux distribution and metabolic pools in Chlamydomonas reinhardtii chloroplasts. *Elife* **7**, doi:10.7554/eLife.37960 (2018).

82 Peden, E. A. *et al.* Identification of global ferredoxin interaction networks in Chlamydomonas reinhardtii. *J Biol Chem* **288**, 35192-35209, doi:10.1074/jbc.M113.483727 (2013).

83 Subramanian, V. *et al.* Ferredoxin5 Deletion Affects Metabolism of Algae during the Different Phases of Sulfur Deprivation. *Plant Physiol* **181**, 426-441, doi:10.1104/pp.19.00457 (2019).

84 Graciet, E., Mulliert, G., Lebreton, S. & Gontero, B. Involvement of two positively charged residues of Chlamydomonas reinhardtii glyceraldehyde-3-phosphate dehydrogenase in the assembly process of a bi-enzyme complex involved in CO2 assimilation. *Eur J Biochem* **271**, 4737-4744, doi:10.1111/j.1432-1033.2004.04437.x (2004).

85 Erales, J., Lignon, S. & Gontero, B. CP12 from Chlamydomonas reinhardtii, a permanent specific "chaperone-like" protein of glyceraldehyde-3-phosphate dehydrogenase. *J Biol Chem* **284**, 12735-12744, doi:10.1074/jbc.M808254200 (2009).

86 Mileo, E. *et al.* Dynamics of the intrinsically disordered protein CP12 in its association with GAPDH in the green alga Chlamydomonas reinhardtii: a fuzzy complex. *Mol Biosyst* **9**, 2869-2876, doi:10.1039/c3mb70190e (2013).

87 Lebreton, S., Andreescu, S., Graciet, E. & Gontero, B. Mapping of the interaction site of CP12 with glyceraldehyde-3-phosphate dehydrogenase from Chlamydomonas reinhardtii. Functional consequences for glyceraldehyde-3-phosphate dehydrogenase. *FEBS J* **273**, 3358-3369, doi:10.1111/j.1742-4658.2006.05342.x (2006).

88 Johnson, X. & Alric, J. Central carbon metabolism and electron transport in Chlamydomonas reinhardtii: metabolic constraints for carbon partitioning between oil and starch. *Eukaryot Cell* **12**, 776-793, doi:10.1128/EC.00318-12

(2013).
89 Alboresi, A. *et al.* Light Remodels Lipid Biosynthesis in Nannochloropsis gaditana by Modulating Carbon Partitioning between Organelles. *Plant Physiol* **171**, 2468-2482, doi:10.1104/pp.16.00599 (2016).
90 Dauvillée, D. *et al.* Genetic dissection of floridean starch synthesis in the cytosol of the model dinoflagellate Crypthecodinium cohnii. *Proc Natl Acad Sci U S A* **106**, 21126-21130, doi:10.1073/pnas.0907424106 (2009).
91 Markou, G., Angelidaki, I. & Georgakakis, D. Microalgal carbohydrates: an overview of the factors influencing carbohydrates production, and of main bioconversion technologies for production of biofuels. *Appl Microbiol Biotechnol* **96**, 631-645, doi:10.1007/s00253-012-4398-0 (2012).
92 Suzuki, E. & Suzuki, R. Variation of Storage Polysaccharides in Phototrophic Microorganisms. *Journal of Applied Glycoscience* **60**, 21-27, doi:10.5458/jag.jag.JAG-2012_016 (2013).
93 Luan, G., Zhang, S., Wang, M. & Lu, X. Progress and perspective on cyanobacterial glycogen metabolism engineering. *Biotechnol Adv* **37**, 771-786, doi:10.1016/j.biotechadv.2019.04.005 (2019).
94 Hirokawa, Y. *et al.* Structural and physiological studies on the storage beta-polyglucan of haptophyte Pleurochrysis haptonemofera. *Planta* **227**, 589-599, doi:10.1007/s00425-007-0641-9 (2008).
95 van de Wal, M. *et al.* Amylose is synthesized in vitro by extension of and cleavage from amylopectin. *J Biol Chem* **273**, 22232-22240, doi:10.1074/jbc.273.35.22232 (1998).
96 Ball, S. G. The intricate pathway of starch biosynthesis and degradation in the monocellular alga chlamydomonas reinhardtii. *Australian Journal of Chemistry* **55**, 49-59, doi:Unsp 004-9425/02/010049
10.1071/Ch02052 (2002).
97 Mouille, G. *et al.* Preamylopectin Processing: A Mandatory Step for Starch Biosynthesis in Plants. *Plant Cell* **8**, 1353-1366, doi:10.1105/tpc.8.8.1353 (1996).
98 Villarejo, A., Martinez, F., del Pino Plumed, M. & Ramazanov, Z. The induction of the CO2 concentrating mechanism in a starch-less mutant of Chlamydomonas reinhardtii. *Physiologia Plantarum* **98**, 798-802, doi:https://doi.org/10.1111/j.1399-3054.1996.tb06687.x (1996).
99 Itakura, A. K. *et al.* A Rubisco-binding protein is required for normal pyrenoid number and starch sheath morphology in Chlamydomonas reinhardtii. *Proc Natl Acad Sci U S A* **116**, 18445-18454, doi:10.1073/pnas.1904587116 (2019).
100 Meyer, M. T. *et al.* Assembly of the algal CO2-fixing organelle, the pyrenoid, is guided by a Rubisco-binding motif. *Sci Adv* **6**, doi:ARTN eabd2408
10.1126/sciadv.abd2408 (2020).
101 Toyokawa, C., Yamano, T. & Fukuzawa, H. Pyrenoid Starch Sheath Is Required for LCIB Localization and the CO2-Concentrating Mechanism in Green Algae. *Plant Physiol* **182**, 1883-1893, doi:10.1104/pp.19.01587 (2020).
102 Tamoi, M., Nagaoka, M., Miyagawa, Y. & Shigeoka, S. Contribution of fructose-1,6-bisphosphatase and sedoheptulose-1,7-bisphosphatase to the photosynthetic rate and carbon flow in the Calvin cycle in transgenic plants. *Plant Cell Physiol* **47**, 380-390, doi:10.1093/pcp/pcj004 (2006).

103 Ball, S. Vol. 7 549-567 (2006).
104 Busi, M. V., Barchiesi, J., Martin, M. & Gomez-Casati, D. F. Starch metabolism in green algae. *Starch-Starke* **66**, 28-40, doi:10.1002/star.201200211 (2014).
105 Ball, S. G. & Deschamps, P. in *The Chlamydomonas Sourcebook (Second Edition)* (eds Elizabeth H. Harris, David B. Stern, & George B. Witman) 1-40 (Academic Press, 2009).
106 Koo, K. M. *et al.* The Mechanism of Starch Over-Accumulation in Chlamydomonas reinhardtii High-Starch Mutants Identified by Comparative Transcriptome Analysis. *Front Microbiol* **8**, 858, doi:10.3389/fmicb.2017.00858 (2017).
107 Ran, W. *et al.* Storage of starch and lipids in microalgae: Biosynthesis and manipulation by nutrients. *Bioresour Technol* **291**, 121894, doi:10.1016/j.biortech.2019.121894 (2019).
108 Cheng, D. *et al.* Improving carbohydrate and starch accumulation in Chlorella sp. AE10 by a novel two-stage process with cell dilution. *Biotechnol Biofuels* **10**, 75, doi:10.1186/s13068-017-0753-9 (2017).
109 Blaby, I. K. *et al.* Systems-level analysis of nitrogen starvation-induced modifications of carbon metabolism in a Chlamydomonas reinhardtii starchless mutant. *Plant Cell* **25**, 4305-4323, doi:10.1105/tpc.113.117580 (2013).
110 Branyikova, I. *et al.* Microalgae--novel highly efficient starch producers. *Biotechnol Bioeng* **108**, 766-776, doi:10.1002/bit.23016 (2011).
111 Yao, C. H., Ai, J. N., Cao, X. P. & Xue, S. Characterization of cell growth and starch production in the marine green microalga Tetraselmis subcordiformis under extracellular phosphorus-deprived and sequentially phosphorus-replete conditions. *Appl Microbiol Biotechnol* **97**, 6099-6110, doi:10.1007/s00253-013-4983-x (2013).
112 Nakamura, Y. & Imamura, M. Change in properties of starch when photosynthesized at different temperatures in Chlorella vulgaris. *Plant Science Letters* **31**, 123-131, doi:https://doi.org/10.1016/0304-4211(83)90048-2 (1983).
113 Zachleder, V., Kawano, S. & Kuroiwa, T. Uncoupling of chloroplast reproductive events from cell cycle division processes by 5-fluorodeoxyuridine in the algaScenedesmus quadricauda. *Protoplasma* **192**, 228-234, doi:10.1007/BF01273894 (1996).
114 Yuan, Y., Li, X. & Zhao, Q. Enhancing growth and lipid productivity in Dunaliella salina under high light intensity and nitrogen limited conditions. *Bioresource Technology Reports* **7**, doi:10.1016/j.biteb.2019.100211 (2019).
115 Li, D., Yuan, Y., Cheng, D. & Zhao, Q. Effect of light quality on growth rate, carbohydrate accumulation, fatty acid profile and lutein biosynthesis of Chlorella sp. AE10. *Bioresour Technol* **291**, 121783, doi:10.1016/j.biortech.2019.121783 (2019).
116 Marchetti, J. *et al.* Effects of blue light on the biochemical composition and photosynthetic activity of Isochrysis sp. (T-iso). *J Appl Phycol* **25**, 109-119, doi:10.1007/s10811-012-9844-y (2012).
117 Kamiya, A. & Kowallik, W. Blue light-induced starch breakdown in Chlorella cells. *Plant Biol* **2**, 671-676 (1986).
118 Miyachi, S., Miyachi, S. & Kamiya, A. (Chlorella, 1978).

119 Jungandreas, A. *et al.* The acclimation of Phaeodactylum tricornutum to blue and red light does not influence the photosynthetic light reaction but strongly disturbs the carbon allocation pattern. *PLoS One* **9**, e99727, doi:10.1371/journal.pone.0099727 (2014).
120 Li, Z. *et al.* Identification and Biotechnical Potential of a Gcn5-Related N-Acetyltransferase Gene in Enhancing Microalgal Biomass and Starch Production. *Front Plant Sci* **11**, 544827, doi:10.3389/fpls.2020.544827 (2020).
121 Findinier, J. *et al.* Deletion of BSG1 in Chlamydomonas reinhardtii leads to abnormal starch granule size and morphology. *Sci Rep* **9**, 1990, doi:10.1038/s41598-019-39506-6 (2019).
122 Legris, M. *et al.* Phytochrome B integrates light and temperature signals in Arabidopsis. *Science* **354**, 897-900, doi:10.1126/science.aaf5656 (2016).
123 Chaves, I. *et al.* The Cryptochromes: Blue Light Photoreceptors in Plants and Animals. *Annual Review of Plant Biology* **62**, 335-364, doi:10.1146/annurev-arplant-042110-103759 (2011).
124 Fujii, Y. *et al.* Phototropin perceives temperature based on the lifetime of its photoactivated state. *Proc Natl Acad Sci U S A* **114**, 9206-9211, doi:10.1073/pnas.1704462114 (2017).
125 Jung, J. H. *et al.* Phytochromes function as thermosensors in Arabidopsis. *Science* **354**, 886-889, doi:10.1126/science.aaf6005 (2016).
126 Christie, J. M. Phototropin Blue-Light Receptors. *Annual Review of Plant Biology* **58**, 21-45, doi:10.1146/annurev.arplant.58.032806.103951 (2007).
127 Sharma, A. *et al.* UVR8 disrupts stabilisation of PIF5 by COP1 to inhibit plant stem elongation in sunlight. *Nat Commun* **10**, 4417, doi:10.1038/s41467-019-12369-1 (2019).
128 Huang, H. *et al.* PCH1 regulates light, temperature, and circadian signaling as a structural component of phytochrome B-photobodies in Arabidopsis</em&gt. *bioRxiv*, 566687, doi:10.1101/566687 (2019).
129 Kianianmomeni, A. & Hallmann, A. Algal photoreceptors: in vivo functions and potential applications. *Planta* **239**, 1-26, doi:10.1007/s00425-013-1962-5 (2014).
130 Im, C. S., Eberhard, S., Huang, K., Beck, C. F. & Grossman, A. R. Phototropin involvement in the expression of genes encoding chlorophyll and carotenoid biosynthesis enzymes and LHC apoproteins in Chlamydomonas reinhardtii. *Plant J* **48**, 1-16, doi:10.1111/j.1365-313X.2006.02852.x (2006).
131 Heijde, M. & Ulm, R. UV-B photoreceptor-mediated signalling in plants. *Trends in plant science* **17**, 230-237 (2012).
132 Petroutsos, D. in *Chlamydomonas: Biotechnology and Biomedicine Microbiology Monographs* (ed Michael Hippler) Ch. Chapter 1, 1-19 (Springer International Publishing, 2017).
133 Christie, J. M. *et al.* Arabidopsis NPH1: a flavoprotein with the properties of a photoreceptor for phototropism. *Science* **282**, 1698-1701, doi:10.1126/science.282.5394.1698 (1998).
134 Christie, J. M., Swartz, T. E., Bogomolni, R. A. & Briggs, W. R. Phototropin LOV domains exhibit distinct roles in regulating photoreceptor function. *Plant J* **32**, 205-219, doi:10.1046/j.1365-313x.2002.01415.x (2002).
135 Christie, J. M., Salomon, M., Nozue, K., Wada, M. & Briggs, W. R. LOV (light,

oxygen, or voltage) domains of the blue-light photoreceptor phototropin (nph1): binding sites for the chromophore flavin mononucleotide. *Proc Natl Acad Sci U S A* **96**, 8779-8783, doi:10.1073/pnas.96.15.8779 (1999).

136 Aihara, Y. *et al.* Mutations in N-terminal flanking region of blue light-sensing light-oxygen and voltage 2 (LOV2) domain disrupt its repressive activity on kinase domain in the Chlamydomonas phototropin. *J Biol Chem* **287**, 9901-9909, doi:10.1074/jbc.M111.324723 (2012).

137 Christie, J. M. Phototropin blue-light receptors. *Annu Rev Plant Biol* **58**, 21-45, doi:10.1146/annurev.arplant.58.032806.103951 (2007).

138 Sullivan, S. *et al.* Regulation of plant phototropic growth by NPH3/RPT2-like substrate phosphorylation and 14-3-3 binding. *Nat Commun* **12**, 6129, doi:10.1038/s41467-021-26333-5 (2021).

139 Sullivan, S., Thomson, C. E., Kaiserli, E. & Christie, J. M. Interaction specificity of Arabidopsis 14-3-3 proteins with phototropin receptor kinases. *Febs Lett* **583**, 2187-2193, doi:10.1016/j.febslet.2009.06.011 (2009).

140 Demarsy, E. *et al.* Phytochrome Kinase Substrate 4 is phosphorylated by the phototropin 1 photoreceptor. *EMBO J* **31**, 3457-3467, doi:10.1038/emboj.2012.186 (2012).

141 Christie, J. M. *et al.* phot1 inhibition of ABCB19 primes lateral auxin fluxes in the shoot apex required for phototropism. *PLoS Biol* **9**, e1001076, doi:10.1371/journal.pbio.1001076 (2011).

142 Inada, S., Ohgishi, M., Mayama, T., Okada, K. & Sakai, T. RPT2 is a signal transducer involved in phototropic response and stomatal opening by association with phototropin 1 in Arabidopsis thaliana. *Plant Cell* **16**, 887-896, doi:10.1105/tpc.019901 (2004).

143 Motchoulski, A. & Liscum, E. Arabidopsis NPH3: A NPH1 photoreceptor-interacting protein essential for phototropism. *Science* **286**, 961-964, doi:10.1126/science.286.5441.961 (1999).

144 Suetsugu, N. *et al.* RPT2/NCH1 subfamily of NPH3-like proteins is essential for the chloroplast accumulation response in land plants. *Proc Natl Acad Sci U S A* **113**, 10424-10429, doi:10.1073/pnas.1602151113 (2016).

145 Takemiya, A. *et al.* Phosphorylation of BLUS1 kinase by phototropins is a primary step in stomatal opening. *Nat Commun* **4**, 2094, doi:10.1038/ncomms3094 (2013).

146 Lariguet, P. *et al.* PHYTOCHROME KINASE SUBSTRATE 1 is a phototropin 1 binding protein required for phototropism. *Proc Natl Acad Sci U S A* **103**, 10134-10139, doi:10.1073/pnas.0603799103 (2006).

147 Kong, S. G. *et al.* Both phototropin 1 and 2 localize on the chloroplast outer membrane with distinct localization activity. *Plant Cell Physiol* **54**, 80-92, doi:10.1093/pcp/pcs151 (2013).

148 Han, I. S., Tseng, T. S., Eisinger, W. & Briggs, W. R. Phytochrome A regulates the intracellular distribution of phototropin 1-green fluorescent protein in Arabidopsis thaliana. *Plant Cell* **20**, 2835-2847, doi:10.1105/tpc.108.059915 (2008).

149 Kong, S.-G. *et al.* Blue light-induced association of phototropin 2 with the Golgi apparatus. *The Plant Journal* **45**, 994-1005, doi:https://doi.org/10.1111/j.1365-

313X.2006.02667.x (2006).
150 Sakamoto, K. & Briggs, W. R. Cellular and subcellular localization of phototropin 1. *Plant Cell* **14**, 1723-1735, doi:10.1105/tpc.003293 (2002).
151 de Carbonnel, M. *et al.* The Arabidopsis PHYTOCHROME KINASE SUBSTRATE2 protein is a phototropin signaling element that regulates leaf flattening and leaf positioning. *Plant Physiol* **152**, 1391-1405, doi:10.1104/pp.109.150441 (2010).
152 Yamauchi, S. *et al.* The Plasma Membrane H+-ATPase AHA1 Plays a Major Role in Stomatal Opening in Response to Blue Light. *Plant Physiol* **171**, 2731-2743, doi:10.1104/pp.16.01581 (2016).
153 Hosotani, S. *et al.* A BLUS1 kinase signal and a decrease in intercellular CO2 concentration are necessary for stomatal opening in response to blue light. *Plant Cell* **33**, 1813-1827, doi:10.1093/plcell/koab067 (2021).
154 Hiyama, A. *et al.* Blue light and CO2 signals converge to regulate light-induced stomatal opening. *Nat Commun* **8**, 1284, doi:10.1038/s41467-017-01237-5 (2017).
155 Horrer, D. *et al.* Blue Light Induces a Distinct Starch Degradation Pathway in Guard Cells for Stomatal Opening. *Curr Biol* **26**, 362-370, doi:10.1016/j.cub.2015.12.036 (2016).
156 Sakai, T. *et al.* Arabidopsis nph1 and npl1: blue light receptors that mediate both phototropism and chloroplast relocation. *Proc Natl Acad Sci U S A* **98**, 6969-6974, doi:10.1073/pnas.101137598 (2001).
157 Briggs, W. R. & Christie, J. M. Phototropins 1 and 2: versatile plant blue-light receptors. *Trends Plant Sci.* **7**, 204-210 (2002).
158 Briggs, W. R. & Christie, J. M. Phototropins 1 and 2: versatile plant blue-light receptors. *Trends Plant Sci* **7**, 204-210, doi:10.1016/s1360-1385(02)02245-8 (2002).
159 Hermanowicz, P., Banas, A. K., Sztatelman, O., Gabrys, H. & Labuz, J. UV-B Induces Chloroplast Movements in a Phototropin-Dependent Manner. *Front Plant Sci* **10**, 1279, doi:10.3389/fpls.2019.01279 (2019).
160 Islam, M. S., Van Nguyen, T., Sakamoto, W. & Takagi, S. Phototropin- and photosynthesis-dependent mitochondrial positioning in Arabidopsis thaliana mesophyll cells. *J Integr Plant Biol* **62**, 1352-1371, doi:10.1111/jipb.12910 (2020).
161 Ishishita, K. *et al.* Phototropin2 Contributes to the Chloroplast Avoidance Response at the Chloroplast-Plasma Membrane Interface. *Plant Physiol* **183**, 304-316, doi:10.1104/pp.20.00059 (2020).
162 Kagawa, T. *et al.* Arabidopsis NPL1: a phototropin homolog controlling the chloroplast high-light avoidance response. *Science* **291**, 2138-2141, doi:10.1126/science.291.5511.2138 (2001).
163 Jarillo, J. A. *et al.* Phototropin-related NPL1 controls chloroplast relocation induced by blue light. *Nature* **410**, 952-954, doi:10.1038/35073622 (2001).
164 Folta, K. M. & Spalding, E. P. Unexpected roles for cryptochrome 2 and phototropin revealed by high-resolution analysis of blue light-mediated hypocotyl growth inhibition. *Plant J* **26**, 471-478, doi:10.1046/j.1365-313x.2001.01038.x (2001).
165 Folta, K. M. & Kaufman, L. S. Phototropin 1 is required for high-fluence blue-light-mediated mRNA destabilization. *Plant Mol Biol* **51**, 609-618,

doi:10.1023/a:1022393406204 (2003).
166 Huang, K., Kunkel, T. & Beck, C. F. Localization of the blue-light receptor phototropin to the flagella of the green alga Chlamydomonas reinhardtii. *Molecular biology of the cell* **15**, 3605-3614, doi:10.1091/mbc.e04-01-0010 (2004).
167 Huang, K., Merkle, T. & Beck, C. F. Isolation and characterization of a Chlamydomonas gene that encodes a putative blue-light photoreceptor of the phototropin family. *Physiologia Plantarum* **115**, 613-622, doi:https://doi.org/10.1034/j.1399-3054.2002.1150416.x (2002).
168 Onodera, A. *et al.* Phototropin from Chlamydomonas reinhardtii is functional in Arabidopsis thaliana. *Plant Cell Physiol* **46**, 367-374, doi:10.1093/pcp/pci037 (2005).
169 Kinoshita, T. *et al.* phot1 and phot2 mediate blue light regulation of stomatal opening. *Nature* **414**, 656-660, doi:10.1038/414656a (2001).
170 Trippens, J. *et al.* Phototropin influence on eyespot development and regulation of phototactic behavior in Chlamydomonas reinhardtii. *Plant Cell* **24**, 4687-4702, doi:10.1105/tpc.112.103523 (2012).
171 Yang, D., Seaton, D. D., Krahmer, J. & Halliday, K. J. Photoreceptor effects on plant biomass, resource allocation, and metabolic state. *Proc Natl Acad Sci U S A* **113**, 7667-7672, doi:10.1073/pnas.1601309113 (2016).
172 Hashimoto, M. *et al.* Arabidopsis HT1 kinase controls stomatal movements in response to CO2. *Nat Cell Biol* **8**, 391-397, doi:10.1038/ncb1387 (2006).
173 Hashimoto-Sugimoto, M. *et al.* Dominant and recessive mutations in the Raf-like kinase HT1 gene completely disrupt stomatal responses to CO2 in Arabidopsis. *J Exp Bot* **67**, 3251-3261, doi:10.1093/jxb/erw134 (2016).
174 Inoue, S. I. *et al.* CIPK23 regulates blue light-dependent stomatal opening in Arabidopsis thaliana. *Plant J* **104**, 679-692, doi:10.1111/tpj.14955 (2020).
175 Tian, W. *et al.* A molecular pathway for CO(2) response in Arabidopsis guard cells. *Nat Commun* **6**, 6057, doi:10.1038/ncomms7057 (2015).
176 Mackinder, L. C. M. *et al.* A Spatial Interactome Reveals the Protein Organization of the Algal CO2-Concentrating Mechanism. *Cell* **171**, 133-147 e114, doi:10.1016/j.cell.2017.08.044 (2017).
177 Gibson, D. G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods* **6**, 343-345, doi:10.1038/nmeth.1318 (2009).
178 Baier, T., Wichmann, J., Kruse, O. & Lauersen, K. J. Intron-containing algal transgenes mediate efficient recombinant gene expression in the green microalga Chlamydomonas reinhardtii. *Nucleic Acids Res* **46**, 6909-6919, doi:10.1093/nar/gky532 (2018).
179 DuBois, M., Gilles, K. A., Hamilton, J. K., Rebers, P. A. & Smith, F. Colorimetric Method for Determination of Sugars and Related Substances. *Analytical Chemistry* **28**, 350-356, doi:10.1021/ac60111a017 (1956).
180 Dubois, M., Gilles, K., Hamilton, J. K., Rebers, P. A. & Smith, F. A Colorimetric Method for the Determination of Sugars. *Nature* **168**, 167-167, doi:10.1038/168167a0 (1951).
181 Yang, F., Xiang, W., Li, T. & Long, L. Transcriptome analysis for phosphorus starvation-induced lipid accumulation in Scenedesmus sp. *Sci Rep* **8**, 16420,

doi:10.1038/s41598-018-34650-x (2018).

182 Yuan, Y. *et al.* Enhancing Carbohydrate Productivity of Chlorella sp. AE10 in Semi-continuous Cultivation and Unraveling the Mechanism by Flow Cytometry. *Appl Biochem Biotechnol* **185**, 419-433, doi:10.1007/s12010-017-2667-1 (2018).

183 Dvir, I., Stark, A. H., Chayoth, R., Madar, Z. & Arad, S. M. Hypocholesterolemic effects of nutraceuticals produced from the red microalga Porphyridium sp in rats. *Nutrients* **1**, 156-167, doi:10.3390/nu1020156 (2009).

184 Hathwaik, L. T. *et al.* Transgressive, reiterative selection by continuous buoyant density gradient centrifugation of Dunaliella salina results in enhanced lipid and starch content. *Algal Research-Biomass Biofuels and Bioproducts* **9**, 194-203, doi:10.1016/j.algal.2015.03.009 (2015).

185 Zhou, Y. *et al.* Using flow cytometry to evaluate thermal extraction of EPS from Synechocystis sp PCC 6803. *Algal Research-Biomass Biofuels and Bioproducts* **20**, 276-281, doi:10.1016/j.algal.2016.10.024 (2016).

186 Kong, S. G. *et al.* The C-terminal kinase fragment of Arabidopsis phototropin 2 triggers constitutive phototropin responses. *Plant J* **51**, 862-873, doi:10.1111/j.1365-313X.2007.03187.x (2007).

187 Smith, R. T. & Gilmour, D. J. The influence of exogenous organic carbon assimilation and photoperiod on the carbon and lipid metabolism of Chlamydomonas reinhardtii. *Algal Research-Biomass Biofuels and Bioproducts* **31**, 122-137, doi:10.1016/j.algal.2018.01.020 (2018).

188 Onishi, M. & Pringle, J. R. Robust Transgene Expression from Bicistronic mRNA in the Green Alga Chlamydomonas reinhardtii. *G3 (Bethesda)* **6**, 4115-4125, doi:10.1534/g3.116.033035 (2016).

189 Zones, J. M., Blaby, I. K., Merchant, S. S. & Umen, J. G. High-Resolution Profiling of a Synchronized Diurnal Transcriptome from Chlamydomonas reinhardtii Reveals Continuous Cell and Metabolic Differentiation. *Plant Cell* **27**, 2743-2769, doi:10.1105/tpc.15.00498 (2015).

190 Strenkert, D. *et al.* Multiomics resolution of molecular events during a day in the life of Chlamydomonas. *Proc Natl Acad Sci U S A* **116**, 2374-2383, doi:10.1073/pnas.1815238116 (2019).

191 Riechmann, J. L. *et al.* Arabidopsis transcription factors: genome-wide comparative analysis among eukaryotes. *Science* **290**, 2105-2110, doi:10.1126/science.290.5499.2105 (2000).

192 Ledent, V. & Vervoort, M. The basic helix-loop-helix protein family: comparative genomics and phylogenetic analysis. *Genome Res* **11**, 754-770, doi:10.1101/gr.177001 (2001).

193 Carretero-Paulet, L. *et al.* Genome-wide classification and evolutionary analysis of the bHLH family of transcription factors in Arabidopsis, poplar, rice, moss, and algae. *Plant Physiol* **153**, 1398-1412, doi:10.1104/pp.110.153593 (2010).

194 Robinson, K. A. & Lopes, J. M. SURVEY AND SUMMARY: Saccharomyces cerevisiae basic helix-loop-helix proteins regulate diverse biological processes. *Nucleic Acids Res* **28**, 1499-1505, doi:10.1093/nar/28.7.1499 (2000).

195 Schmollinger, S. *et al.* Nitrogen-Sparing Mechanisms in Chlamydomonas Affect the Transcriptome, the Proteome, and Photosynthetic Metabolism. *Plant Cell* **26**, 1410-1435, doi:10.1105/tpc.113.122523 (2014).

196 Kono, A. & Spalding, M. H. LCI1, a Chlamydomonas reinhardtii plasma membrane protein, functions in active CO2 uptake under low CO2. *Plant J* **102**, 1127-1141, doi:10.1111/tpj.14761 (2020).

197 Choi, B. Y. *et al.* The Chlamydomonas transcription factor MYB1 mediates lipid accumulation under nitrogen depletion. *New Phytol*, doi:10.1111/nph.18141 (2022).

198 Perlaza, K. *et al.* The Mars1 kinase confers photoprotection through signaling in the chloroplast unfolded protein response. *Elife* **8**, doi:10.7554/eLife.49577 (2019).

199 Dewdney, J., Conley, T. R., Shih, M. C. & Goodman, H. M. Effects of Blue and Red Light on Expression of Nuclear Genes Encoding Chloroplast Glyceraldehyde-3-Phosphate Dehydrogenase of Arabidopsis thaliana. *Plant Physiology* **103**, 1115-1121, doi:10.1104/pp.103.4.1115 (1993).

200 Gontero, B., Avilan, L. & Lebreton, S. in *Annual Plant Reviews Volume 22: Control of Primary Metabolism in Plants* 187-218 (2006).

201 Baier, T., Jacobebbinghaus, N., Einhaus, A., Lauersen, K. J. & Kruse, O. Introns mediate post-transcriptional enhancement of nuclear gene expression in the green microalga Chlamydomonas reinhardtii. *PLoS Genet* **16**, e1008944, doi:10.1371/journal.pgen.1008944 (2020).

202 Emrich-Mills, T. Z. *et al.* A recombineering pipeline to clone large and complex genes in Chlamydomonas. *Plant Cell* **33**, 1161-1181, doi:10.1093/plcell/koab024 (2021).

203 Li, Z. *et al.* Analyses of the photosynthetic characteristics, chloroplast ultrastructure, and transcriptome of apple (Malus domestica) grown under red and blue lights. *BMC Plant Biology* **21**, 483, doi:10.1186/s12870-021-03262-5 (2021).

204 Thum, K. E., Shasha, D. E., Lejay, L. V. & Coruzzi, G. M. Light- and Carbon-Signaling Pathways. Modeling Circuits of Interactions. *Plant Physiology* **132**, 440-452, doi:10.1104/pp.103.022780 (2003).

205 Azuara, M. P. & Aparicio, P. J. In Vivo Blue-Light Activation of Chlamydomonas reinhardii Nitrate Reductase. *Plant Physiol* **71**, 286-290, doi:10.1104/pp.71.2.286 (1983).

206 Azuara, M. P. & Aparicio, P. J. Effects of light quality, CO2 tensions and NO3+− concentrations on the inorganic nitrogen metabolism of Chlamydomonas reinhardii. *Photosynthesis Research* **5**, 97-103, doi:10.1007/BF00028523 (1984).

207 Azuara, M. a. P. & Aparicio, P. J. Spectral Dependence of Photoregulation of Inorganic Nitrogen Metabolism in Chlamydomonas reinhardii 1. *Plant Physiology* **77**, 95-98, doi:10.1104/pp.77.1.95 (1985).

208 Soupene, E. *et al.* Rhesus expression in a green alga is regulated by CO2. *Proceedings of the National Academy of Sciences* **99**, 7769-7773, doi:10.1073/pnas.112225599 (2002).

209 Gorman, D. S. & Levine, R. P. Cytochrome f and plastocyanin: their sequence in the photosynthetic electron transport chain of Chlamydomonas reinhardi. *Proc Natl Acad Sci U S A* **54**, 1665-1669, doi:10.1073/pnas.54.6.1665 (1965).

210 Sueoka, N. Mitotic Replication of Deoxyribonucleic Acid in Chlamydomonas Reinhardi. *Proc Natl Acad Sci U S A* **46**, 83-91, doi:10.1073/pnas.46.1.83 (1960).

211 Porra, R. J., Thompson, W. A. & Kriedemann, P. E. Determination of accurate extinction coefficients and simultaneous equations for assaying chlorophylls a and b extracted with four different solvents: verification of the concentration of chlorophyll standards by atomic absorption spectroscopy. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **975**, 384-394, doi:10.1016/s0005-2728(89)80347-0 (1989).

212 Genty, B., Briantais, J.-M. & Baker, N. R. The relationship between the quantum yield of photosynthetic electron transport and quenching of chlorophyll fluorescence. *Biochimica et Biophysica Acta* **990**, 87-92 (1989).

213 Bailleul, B. *et al.* Energetic coupling between plastids and mitochondria drives CO2 assimilation in diatoms. *Nature* **524**, 366-369, doi:10.1038/nature14599 (2015).

214 Li, D., Wang, L., Zhao, Q., Wei, W. & Sun, Y. Improving high carbon dioxide tolerance and carbon dioxide fixation capability of Chlorella sp. by adaptive laboratory evolution. *Bioresour Technol* **185**, 269-275, doi:10.1016/j.biortech.2015.03.011 (2015).

215 Cao, M., Fu, Y., Guo, Y. & Pan, J. Chlamydomonas (Chlorophyceae) colony PCR. *Protoplasma* **235**, 107-110, doi:10.1007/s00709-009-0036-9 (2009).

216 Schloss, J. A. A Chlamydomonas gene encodes a G protein beta subunit-like polypeptide. *Molecular & general genetics : MGG* **221**, 443-452, doi:10.1007/BF00259410 (1990).

217 Rasala, B. A. *et al.* Robust expression and secretion of Xylanase1 in Chlamydomonas reinhardtii by fusion to a selection gene and processing with the FMDV 2A peptide. *PLoS One* **7**, e43349, doi:10.1371/journal.pone.0043349 (2012).

218 Zhang, R. *et al.* High-Throughput Genotyping of Green Algal Mutants Reveals Random Distribution of Mutagenic Insertion Sites and Endonucleolytic Cleavage of Transforming DNA. *Plant Cell* **26**, 1398-1409, doi:10.1105/tpc.114.124099 (2014).

219 Pruitt, K. D., Tatusova, T. & Maglott, D. R. NCBI Reference Sequence (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Res* **33**, D501-504, doi:10.1093/nar/gki025 (2005).

220 Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* **32**, 1792-1797, doi:10.1093/nar/gkh340 (2004).

221 Larkin, M. A. *et al.* Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947-2948, doi:10.1093/bioinformatics/btm404 (2007).

222 Waterhouse, A. M., Procter, J. B., Martin, D. M., Clamp, M. & Barton, G. J. Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189-1191, doi:10.1093/bioinformatics/btp033 (2009).

223 Guindon, S. *et al.* New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. *Systematic Biology* **59**, 307-321, doi:10.1093/sysbio/syq010 (2010).

224 Kumar, S., Stecher, G. & Tamura, K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol Biol Evol* **33**, 1870-1874, doi:10.1093/molbev/msw054 (2016).

225 Neath, A. A. & Cavanaugh, J. E. The Bayesian information criterion: background, derivation, and applications. *WIREs Computational Statistics* **4**, 199-203, doi:https://doi.org/10.1002/wics.199 (2012).

226 Nei, M. & Kumar, S. *Molecular evolution and phylogenetics*. (Oxford University Press, USA, 2000).

227 Anisimova, M. & Gascuel, O. Approximate likelihood-ratio test for branches: A fast, accurate, and powerful alternative. *Syst Biol* **55**, 539-552, doi:10.1080/10635150600755453 (2006).

228 Schlesselman, J. J. Data Transformation in Two-Way Analysis of Variance. *Journal of the American Statistical Association* **68**, 369-378, doi:10.2307/2284078 (1973).