Winter 12-15-2022

Mechanisms for High Light Tolerance in a Fast-Growing Cyanobacterium

Patricia Walker
Washington University in St. Louis

WASHINGTON UNIVERSITY IN ST. LOUIS
Division of Biology and Biomedical Sciences
Plant and Microbial Biosciences

Dissertation Examination Committee:
Himadri Pakrasi, Chair
Arpita Bose
Michael Caparon
Ram Dixit
Hani Zaher
Ru Zhang

Mechanisms for High Light Tolerance in a Fast-Growing Cyanobacterium

by

Patricia Lee Walker

A dissertation presented to
Washington University in St. Louis
in partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

December 2022
St. Louis, Missouri

Table of Contents

List of Figures

List of Tables

Acknowledgments

I would like to thank everyone who helped me get to this stage. While I cannot get to everyone who played a role, my special thanks go out to the following.

First, I am grateful to my advisor, Dr. Himadri B. Pakrasi, for welcoming me into his lab, providing me with various project opportunities, and creating a lab environment for me to grow in confidence and independence. I also thank my thesis committee: Dr. Ram Dixit, Dr. Ru Zhang, Dr. Arpita Bose, Dr. Michael Caparon, and Dr. Hani Zaher, for their support in my thesis project. And to Dr. Joe Jez, who first encouraged me to apply to the Plant and Microbiology Program and saw me through to the very end.

I thank my undergraduate scientific mentor, Dr. Sara Baer, who helped me grow through providing me with research opportunities as well as her time and patience, which I needed a lot.

I am indebted to my colleagues, particularly Dr. Virginia Johnson and Dr. Michelle Liberton, for their encouragement and assistance in editing my thesis. Thank you, Dr. Deng Liu, Dr. Anindita Bandyopadhyay, and Dr. Annesha Sengupta for answering my questions and being there to help me whenever needed. Thank you, Nancy Duan, for being such a wonderful and helpful person. And to the many other members, current and past: Dr. Sandeep Biswas, Dr. Kristen Wendt, Dr. Justin Ungerer, Dr. Po-Cheng Lin, Dr. Steven Hui-Yuan Chen, and so on. I will always cherish the lab's sense of community and traditions around travel and chocolate.

I could not have completed my thesis without the years of friendship and counsel from my cohort. Thank you, Maria Sorkin, Andrew Lin, Ginger Johnson, Jenny Shoots, Dennis Zhu, and Ryan Calcutt, for being there to celebrate the good times, but even more for helping me through the hardest of times. When I started six years ago, I had no idea how much I would, at times, struggle and question my choices. The friendship you provided made such a difference.

Thank you to the incredible and powerful women of my life who I have the privilege of calling my friends. Lisa Solinger, Lana Belkhadir, and Dr. Kathrine Kelton thank you for always rooting for me, believing in me, and despite the distance, you were there for me when I needed someone to vent to. Lana, you are an inspiration, a force, and a kind heart who kept me anchored throughout my years in graduate school. Michelle Mynx cultivated a loving and beautiful community like no other for me to escape to. And to Shayla Gunn for challenging me and sticking by my side for so many turbulent but wonderful years.

I thank my family, immediate and extended. Special thanks to my sister, nephew, and mother, who have been so understanding and supportive of me over the years. I am very proud to be the first Ph.D. in the family, and I could not have got here without the influence of my family. Finally, to my partner, statistics advisor, scientific writing critic, and number one fan, Dr. Marcos Barcellona. Thank you for the endless help, love, and encouragement, which has been especially meaningful while I write this dissertation.

I want to acknowledge the services I received during my graduate career. Being neurodivergent, the mental health support I received through student services was a game changer for me to grow and succeed. Last, I am thankful for the generous funding from the Bayer Graduate Research Fellowship.

Patricia Walker

Washington University in St. Louis

December 2022

Dedicated to my mentor, friend, and eco-warrior leader Craig Stettner,
a.k.a. Uncle Craigers.
Craig's biology courses, encouragement, environmental club adventures, and camping trips put me on the path that I followed to get to where I am today. He always believed in me and made sure that I knew that, so I know that he would be proud of me for completing this Ph.D. journey. I hope to one day inspire and make a difference in the lives of others the way Craig did for me, as well as so many others.

ABSTRACT OF THE DISSERTATION

Mechanisms of High Light Tolerance for a Fast-Growing Cyanobacterium

by

Patricia Walker

Doctor of Philosophy in Biology and Biomedical Sciences

Plant and Microbial Biology

Washington University in St. Louis, 2022

Himadri Pakrasi, Chair

Through oxygenic photosynthesis, cyanobacteria, algae, and plants convert light into chemical energy. However, highly energetic light often damages the photosynthetic apparatus, which can lead to the decline of photosynthetic activity (photoinhibition), and prolonged photoinhibition can lead to cell death. Cyanobacteria enact various protective mechanisms to mitigate photodamage, many of which have been conserved in plants. Repair of PSII reaction centers, decreased light absorption, quenching of excess absorbed energy, and biosynthesis of antioxidants all work to mitigate damage during high light. Due to the adjustments to photosynthesis, acclimation to high light requires remodeling of cellular metabolism and physiology. Therefore, phototrophs have evolved a diverse set of regulatory systems that efficiently detect increased light and enact cellular responses. During this study, I have endeavored to understand the unique features of the fast-growing cyanobacterium *Synechococcus elongatus* UTEX 2973 (*Synechococcus* 2973) that allow it to grow under extreme high light conditions.

In Chapter 1, I have reviewed the current understanding of regulatory mechanisms that control gene expression and physiology for high light response. The interconnectedness of light

response with nutrient metabolism demand is discussed and placed in the context of *Synechococcus* 2973. Chapter 2 takes a global look at the transcriptional changes that occur in two highly homologous *Synechococcus* strains during high light response and acclimation. Using this two-strain method, changes that are specific to *Synechococcus* 2973 were identified and discussed. Chapter 3 investigates a novel gene found to be involved in the high light tolerance of *Synechococcus* 2973.

In Chapter 4, conclusions, future directions, and additional data are discussed. In the Appendices, specific polymorphic loci that allow for natural competence in *Synechococcus* 7942 were identified and a naturally competent *Synechococcus* 2973 strain was generated. In summary, this dissertation describes the unique ways in which a fast-growing cyanobacteria acclimates to high light, reports the importance of several genes in high light tolerance, and characterizes a conserved phosphatase required for light response in cyanobacteria.

Chapter 1
Introduction: Regulation of high light response and tolerance in cyanobacteria

Chapter contributions:
All figures in this Chapter were generated by PLW.

1.1 Summary

In natural environments, light is one of the most dynamic parameters, fluctuating in intensity and spectral properties throughout the day. Cyanobacteria rely on light energy to perform oxygenic photosynthesis for chemical energy. However, light intensities beyond the photosynthetic capacity of an organism impairs photosynthesis and limits growth. To adapt to high-light conditions, cyanobacteria have developed multi-layered adaptive responses that rapidly modulate functions from gene expression to nutrient metabolism. Complex regulatory networks have evolved to coordinate the detection of upshifted light using molecular signals that connect light-stress sensors and transcriptional regulators to balance metabolic and energetic cellular systems. The recently characterized cyanobacterium *Synechococcus elongatus* UTEX 2973 is a suitable model for studying mechanisms of high-light response because of its fast growth rate combined with extreme high-light tolerance. This Chapter summarizes recent progress in our understanding of the regulation of high-light tolerance, including the role of transcription factors and circadian regulation in altering metabolism during high-light response. Finally, an overview of our understanding of high-light tolerance of *Synechococcus elongatus* UTEX 2973 is provided.

1.2 Introduction: Photoprotection and overcoming photoinhibition

Photosynthetic organisms utilize light energy to maintain physiological and biochemical functions. However, over-harvesting light energy by photosynthetic pigments can induce damage to the photosynthetic system resulting in photoinhibition and reduced growth. Therefore, the perception to changes in light and subsequent physiological responses are important adaptive responses that have been evolving at least since the emergence of cyanobacteria 1.9 billion years ago (Demoulin et al., 2019). High-light tolerance is typically framed around two universal Photosystem II (PSII) processes: prevention of damage and recovery. Photoinhibition occurs when the excess absorbed photon energy causing damage to the reaction center of PSII exceeds the rate of PSII repair (Aro et al., 1993, Hakala et al., 2006, Nishiyama et al., 2011, Nishiyama and Murata, 2014). Properly regulated stress responses are vital for cell survival under high light, as prolonged photoinhibition can trigger programmed cell death (Barteneva et al., 2022). In addition to light intensity and CO_2 availability, other environmental factors including temperature, high salt (Allakhverdiev and Murata, 2004), prior light intensity (Jimbo et al., 2019), and the period of excess light exposure contributes to the development and extent of photoinhibition. Resilience to photoinhibition differs among photosynthetic organisms. Many phototrophs experience a midday depression of photosynthesis while light intensity is highest (Fig. 1.1) (Koyama and Takemoto, 2014), although the high-light tolerant cyanobacterium *Synechococcus elongatus* UTEX 2973 grows optimally in high-light conditions of 1500-2400 μmol photons m^{-2} s^{-1} (Ungerer et al., 2018b). In comparison, the model species *Synechocystis* sp. PCC 6803 (hereafter, *Synechocystis)* is typically grown in 15-50 μmol photons

m^{-2} s^{-1} and 220-550 µmol photons m^{-2} s^{-1} is considered high light (Havaux et al., 2005, Zhang et al., 2021). Research has identified profound molecular and physiological changes in light-stressed cells that provide robust models for elucidating mechanisms of key cellular responses, including the molecular basis for PSII repair. An in-depth understanding of stress responses and their intrinsic regulatory mechanisms, as well as how these relate to physiological changes for high-light acclimation, is of considerable importance and will be reviewed in this Chapter.

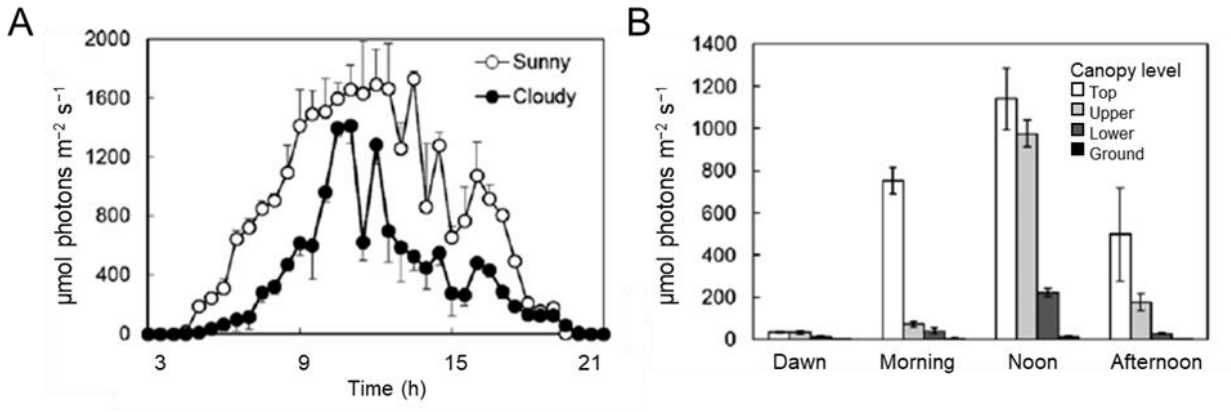

Figure 1.1: Environmental light intensity recordings (A) of the light in diurnal variation in cloudy and full sunlight, and (B) at different levels of the pine forest canopy. Presented as mean and standard deviations of three independent measurements taken in Norway, figure adapted from (Kovac et al., 2013).

Reactive oxygen species (ROS) including hydrogen peroxide (H_2O_2) and singlet oxygen accumulate in the photosynthetic machinery during photoinhibitory light conditions, inhibiting the synthesis of the D1 protein needed to repair photodamaged PSII (Fig. 1.2) (Nishiyama et al., 2005, Allakhverdiev and Murata, 2004, Nishiyama et al., 2011). Therefore, H_2O_2 is a key indicator of imbalance in photosynthetic components and a signal for photoprotective responses. Photoprotective mechanisms including excess energy dissipation through non-photochemical quenching (NPQ) (Kirilovsky, 2015), decreased light harvesting (by reduction in energy transfer

from antenna complex), and increased electron transfer out of PSII occur within seconds to minutes to protect PSII against damage. Recovery of PSII involves the removal of ROS by antioxidants like carotenoids that rapidly quench ROS near the photosynthetic complexes in the thylakoid membrane (Ksas et al., 2015) and the PSII reaction center protein D1 turnover (Jimbo et al., 2019, Izuhara et al., 2020). Additionally, high-light-inducible proteins (Hlips) are synthesized to help cells acclimate to high-light conditions (Shukla et al., 2018). Other high-light responses can include phototaxis and complementary chromatic acclimation, which are advantageous adaptations, but do not necessarily allow for high-light tolerance.

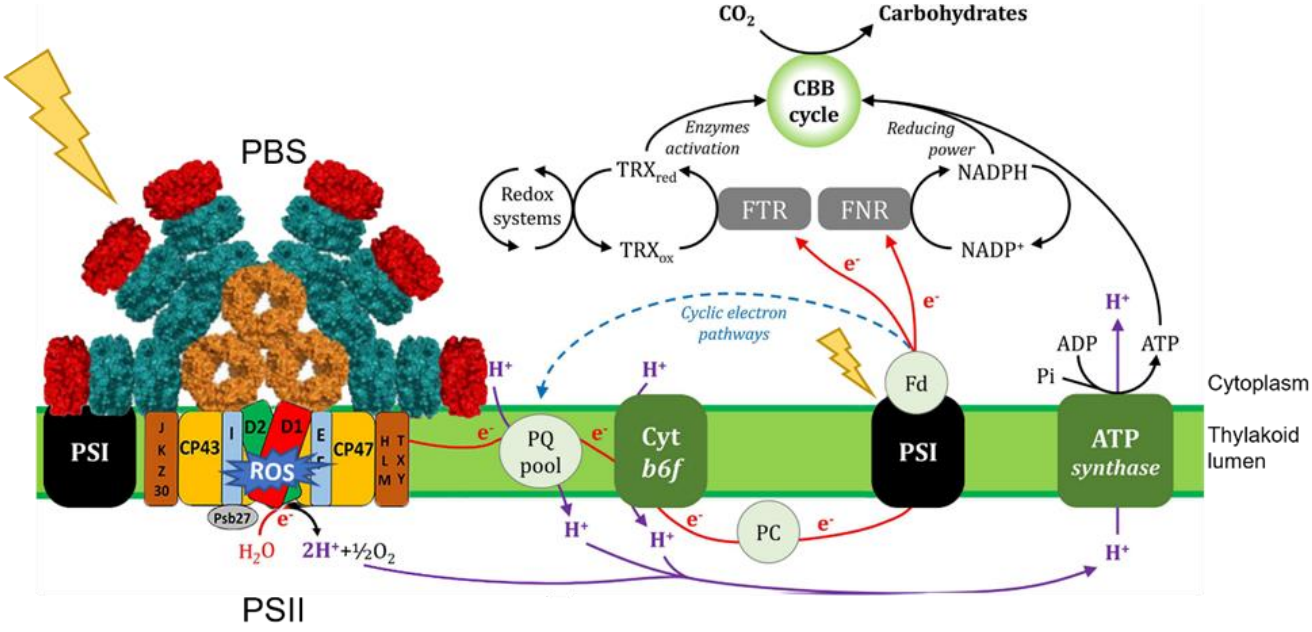

Figure 1.2: Schematic of the photosynthetic electron transport chain and connection to the Calvin-Benson-Bassham cycle (CBB) cycle. Diagram of the components of the PSII monomer protein and location of ROS damage to the PSII reaction center comprised of D1 and D2 proteins. The phycobilisome (PBS) antenna complex primarily associated with PSII also interacts with PSI in a low-light conformation. Light energy, indicated by lightning bolts, is absorbed by PBS and PSI. Photosynthetic electron transport and pathway diagram adapted from (Lima-Melo et al., 2021), PBS model adapted from (Blankenship, 2015), PSII structure diagram adapted from (Weisz et al., 2016). Linear electron transport is shown in red, cyclic electron transport is shown with a dashed blue arrow, and proton flux is shown in purple arrows. Color coding of PBS is as follows: phycoerythrin, red; phycocyanin, cyan; allophycocyanin, orange. ADP, adenosine diphosphate; ATP, adenosine triphosphate; CBB cycle, Calvin-Benson-Bassham cycle; Cyt *b6f*, cytochrome *b6f*; Fd, ferredoxin; FNR, ferredoxin:

$NADP^+$ oxidoreductase; FTR, ferredoxin: thioredoxin reductase; PBS, phycobilisome; $NADP^+$, oxidized nicotinamide adenine dinucleotide phosphate; NADPH, reduced nicotinamide adenine dinucleotide phosphate; PC, plastocyanin; Pi, inorganic phosphate; PQ, plastoquinone; PSI, photosystem I; PSII, photosystem II; TRX_{ox}, oxidized thioredoxin; TRX_{red}, reduced thioredoxin.

High-light tolerance is the ability to maintain photosynthetic activity during increasing light, which requires adjustments to photosynthetic machinery and metabolism. The composition of the photosynthetic apparatus and the rate of photosynthetic assimilatory reactions are optimized by long-term acclimation responses, which involve changes in gene expression, metabolic fluxes, and enzyme activity. The capacity to integrate light cues and respond appropriately is a highly developed feature of photosynthetic organisms. Physiological studies allude to the complex control of cyanobacterial light response by sensory proteins, histidine kinases, and response regulators that regulate transcription, physiology, and metabolism. High-light stress induces the expression and suppression of hundreds of genes. Novel genes involved in high-light response and acclimation have been identified and include *hltA* (Walker and Pakrasi, 2022), *slr0320*, *slr0151* (Yang et al., 2014), and *slr1916* (Yoshikawa et al., 2021). Recent advances in genome-wide experimental approaches have allowed for the expansion of regulatory mechanism characterization. Unraveling regulatory mechanisms that underly high-light tolerance remains a topic of interest and novel omics-based approaches allow the acquisition of new information. In this Chapter, I take a multi-level view to explore regulatory interactions between light-stress detection, regulation, and changes to physiology that assist in acclimation to high-light conditions and put into perspective how these metabolic and transcriptional changes assist in the acclimation to high-light conditions.

1.3 Light detection and signaling

Dynamic regulatory systems are required to respond to changes in light intensity of both anticipated diurnal changes and varying conditions due to weather, shading, or sunflecks (Fig. 1.1). A key to high-light response is the rapid detection of upshifted light and the initiation of high-light response signaling. Cyanobacteria absorb photons through bilin pigments in the phycobilisome (PBS) antenna and chlorophyll in the photosystems. Information about light intensity can be detected indirectly by by-products of photosynthesis including photosynthesis-induced ROS, changing redox potential due to buildup of electrochemical gradient across the thylakoid, and ATP levels. Light is also sensed directly with phytochromes, photoreceptor proteins that bind bilin chromophores to perceive light. Cyanobacteria and algae have phytochromes that detect a wide range of light wavelengths, and cyanobacteria have distinct bilin-binding photoreceptors called cyanobacteriochromes (CBCRs) (Enomoto et al., 2015, Ikeuchi and Ishizuka, 2008). Phytochromes have a bilin-binding GAF photosensory domain that allows them to detect changes in light. The signaling of phytochromes are believed to regulate aspects of growth, photosynthesis, phototaxis, and cell aggregation (Wiltbank and Kehoe, 2019).

Photoreceptor Proteins

Most characterized light-sensing proteins participate in two-component signal transduction pathways, consisting of a sensor kinase and response regulator protein. Two-component regulatory systems employ transient phosphorylation to convert an environmental stimulus into transcriptional regulation (Tuominen et al., 2008, Tuominen et al., 2006). The sensor kinase Cph1, found in *Synechocystis, is* a light-activated biliprotein photoreceptor that phosphorylates its response regulator Rcp1. This post-translational modification (PTM) enacts Rcp1 to modulate downstream signal transduction (Im et al., 2002). The CBCR RcaE (Regulator

of chromatic adaptation) (Kehoe and Grossman, 1996) also operates through a phosphorelay system to control PBS production (Terauchi et al., 2004). In addition to the two-component systems, partner-switching mechanisms where a phosphatase and a kinase antagonistically regulate the phosphorylation state of a target protein can also control downstream gene expression. Two GAF domain-containing phosphatase proteins, HltA from *Synechococcus* (Walker and Pakrasi, 2022) and *hmpU* from *Nostoc punctiforme* (Riley et al., 2018), were recently characterized. These phosphatase proteins are suspected to be sensing changes to light by binding bilin or sensing redox changes and regulating stress response genes through a partner-switching mechanism.

Redox signaling

During rapid influxes of light (Fig 1.1b), excess light energy is absorbed causing the over-reduction to the photosynthetic electron transport chain (ETC) (Fig. 1.2) and transiently changing the redox homeostasis. The disturbance to the photosynthetic redox balance triggers signaling cascades and light-stress response gene expression. Cyanobacteria encode five to six different thioredoxin (Trx) genes that are important for redox signaling. Light-dependent redox signals are transmitted through a pathway composed of ferredoxin (Fdx), and Fdx-thioredoxin reductase (FTR) or NADPH-thioredoxin reductase (NTR), which reduce Trx (Fig. 1.2). This pathway directly links light-induced changes in the redox state of the photosynthetic ETC to carbon metabolism (Schurmann and Buchanan, 2008).

Secondary Messengers

The transduction of environmental signals into biochemical processes through secondary messengers such as cyclic nucleotides (cAMP and cGMP) and nitrous oxide (NO) can be used to elicit stress response and physiological changes (Burlacot et al., 2020). The role of c-GMP and c-

AMP have been broadly studied regarding light quality, although not as much in light intensity. Light-controlled concentrations of secondary messengers such as cyclic-dimeric GMP (c-di-GMP) are involved in light signaling networks mediated by proteins with photosensory domains. In the filamentous cyanobacterium *Fremyella diplosiphon*, the c-di-GMP-specific phosphodiesterase proteins CdgA and CdgE bind c-di-GMP and adapt cells to light conditions by tuning the levels of photosynthetic pigments and inducing morphological changes (Agostoni et al., 2013). In Anabaena sp. PCC 7120 (hereafter Anabaena) grown in high-light conditions, c-di-GMP was found to regulate heterocyst differentiation (Neunuebel and Golden, 2008). In plants, light was shown to influence cAMP accumulation, and cAMP has been found to participate in photosynthetic acclimation by interacting with enzymes involved in the Calvin-Benson-Bassham cycle (CBB) and photorespiration (Donaldson et al., 2016).

The production of ROS and subsequent oxidative stress caused by core photosynthetic reactions are exacerbated during exposure to high heat, high light, and nutrient deprivation (Aro et al., 1993, Hakala et al., 2006, Nishiyama et al., 2011, Nishiyama and Murata, 2014). Photosynthetic organisms have developed numerous mechanisms of ROS detoxification to circumvent damage, as well as an array of regulatory proteins that utilize ROS as a secondary messenger. Because oxidative stress originates from several environmental stresses, redundancy of stress response pathways exists. Unique systems to respond to light-induced oxidative stress may not be necessary as the general systems still serve the same purpose, to decrease ROS production, scavenge existing ROS, and fix the damage that occurred (Aro et al., 1993).

H_2O_2 is rapidly produced during photosynthetic redox imbalance and is used by cells as a key stress signal that triggers a series of signaling cascades leading to photosynthetic redox balance and various adaptive responses. Detection of H_2O_2 in cyanobacteria includes a peroxide-

sensing repressor (PerR), transcriptional regulators (OxyR and RexT), histidine kinases (hik33 and hik34) as well as many putative components (Latifi et al., 2009) that induce transcription of detoxification genes to maintain redox homeostasis. During oxidative stress conditions such as high light, H_2O_2 inhibits RexT, a transcriptional repressor of thioredoxin A2 (*trxA2*), leading to the expression of *trxA2* (Ehira and Ohmori, 2012). Under low-light conditions, the highly conserved regulator protein PedR represses *ndhD2* and *rpe* genes and promotes *ntr* and *ftr-v* genes. Upon a shift to high light, the FTR and NTR pathways reduce Trx proteins (Fig. 1.2). Reduced Trx targets PedR and transiently inhibits its regulatory function within minutes of high-light exposure (Horiuchi et al., 2010). Further progress in our understanding of pathways involved in the light-induced ROS signaling network may further elucidate links to central metabolism regulation.

1.4 Photoprotective mechanisms

Because photoprotective mechanisms are essential for survival under high light, they are discussed briefly in this section. The processes that regulate these photoprotective mechanisms are discussed in depth in the following sections. Remodeling of the photosynthetic apparatus is a key photoprotective mechanism when dealing with high-light conditions. While the photosynthetic core complexes are conserved throughout photosynthetic organisms, the antenna complexes that harvest light for the photosynthetic machinery are different between eukaryotic plants and green algae, and prokaryotic cyanobacteria (Fig. 1.2). During high light, the light-harvesting systems that capture photons may rearrange within the thylakoid membrane (Casella et al., 2017), and PBSs are downsized or decoupled from PSII to decrease energy transfer to the reaction centers and prevent photodamage (Stoitchkova et al., 2007) (Tamary et al., 2012).

Alternative electron transport pathways are photoprotective mechanisms for the dissipation of excess absorbed light energy. Cyclic electron transfer (CET) around PSI (Battchikova et al., 2011) involves NADPH dehydrogenase-type I (NDH-1) complexes (Peltier et al., 2016). The NDH-1 complexes play different but important roles in high-light acclimation, which are discussed further in the following sections. Other important photoprotective mechanisms that utilize flavodiiron proteins (*flv*), the orange carotenoid protein (OCP), Hlips, and IsiA are induced by the absorption of excess photons during high light, nitrogen starvation, and low temperature stresses. When activated by high light, the OCP interacts with the PBS core to induce NPQ, which is the main mechanism of NPQ in most cyanobacteria (Sluchanko et al., 2018). Similarly, Hlips bind to chlorophyll and β-carotene, predominantly in PSII, and participate in the dissipation of excess energy absorbed by chlorophylls (Shukla et al., 2018, Bhaya et al., 2002, Latifi et al., 2009). IsiA can form a supercomplex with PSI and provide protection from photooxidative stress (Havaux et al., 2005, Ihalainen et al., 2005, Chen et al., 2021). Flv proteins participate in the photoreduction of ROS to alleviate the excess reduction in the photosystems (Bersanini et al., 2014, Santana-Sanchez et al., 2019). Finally, while the mechanisms of state transitions are highly debated, it is generally accepted that state transitions optimize the energy distribution between the photosystems, involving lower coupling of PSI in state 1 and higher coupling in state 2 (Fig. 1.2) (Cogdell and Gardiner, 2015).

1.5 Regulatory systems for high light response

Given the detrimental effects high light can have on photosynthesis, it should not be surprising that high light triggers widespread changes to gene expression. To downregulate the photosynthetic apparatus, remodel the carbon/nitrogen (C/N) balance, synthesize detoxification

proteins, and modify nutrient assimilation pathways, different levels of regulatory mechanisms take place. Light and oxidative-stress responsive transcription factors induce and repress large sets of genes. In addition, regulators of circadian genes that bring together the global processes of photosynthesis, C/N metabolism, and growth are also modulated in response to changes in light intensity. There are also many sigma factors and noncoding RNAs that are involved in light response regulation. In this section, we take a global view of the regulatory systems that transiently or sustainably modify global functions in the cell to acclimate to high-light intensity.

1.5.1 Transcription factors and regulatory cascades

Several transcription factors (TFs) are induced by stress signals to regulate the expression of stress-related genes. Gene targets of the TFs RpaB, FurA, IdiB, Rre1, NtcA, and NblR are differentially expressed in response to high light (Riediger et al., 2019). Specifically, the Regulator of PBS associated B (RpaB) TF is known to be essential for high-light response in cyanobacteria. Several other TFs such as the global regulator of nitrogen homeostasis, NtcA, are also involved in multiple abiotic stress responses. These light-responsive TFs are required to remodel cellular flux to align with changes relating directly to photosynthesis. These TFs also influence other high-light responsive regulatory genes, including *hik33, nusG, sigD*, *sigA3, rpoB, rpoC2, rpoA5, sigG*, *rpoA,* and *pedR* (Seki et al., 2007, Imamura et al., 2003). The large number of light-responsive regulatory factors is indicative of the dependence of TFs on additional regulators for proper high-light acclimation. In this section, key TFs involved in light stress response in cyanobacteria and their roles in enhanced light tolerance are reviewed.

RpaB is an essential regulator of high-light stress response in cyanobacteria. RpaB binds to and represses more than 150 promoters involved in diverse photosynthetic functions,

including photoprotection, CET, and photorespiration, making it indispensable in the regulation of the transcriptional network of high-light responsive genes (Riediger et al., 2019, Yasuda et al., 2020). RpaB is regulated by dual-redox-related signals relayed through a PTM phosphorylation by the sensor kinase NblS (also known as Hik33) (Riediger et al., 2019), and was recently found to be reduced by TrxA (Kato et al., 2022, Ibrahim et al., 2022). Having the RpaB gene-repression system under PTM regulation may be advantageous in keeping high-light response off until it is needed. Upon exposure to high light, RpaB is dephosphorylated and releases repression of target genes, including its paralog transcriptional regulator gene *ssrA*. NblS/Hik33 has also been found to be activated through interactions with SipA (NblS-interacting protein A) to promote the phosphorylation of RpaB and SsrA (Salinas et al., 2007, Kato et al., 2011). NblS/Hik33 shows a preference for phosphorylation of SsrA that may mechanistically promote continued expression of high-light responsive genes. RpaB targets also include the high-light induced genes *hliA*, *isiA*, *nblA*, *ssrA*, and *rpoD3* (Yasuda et al., 2020), some of which overlap with the TFs NtcA and Fur (Riediger et al., 2019). Additionally, RpaB participates in crosstalk with cyanobacterial circadian-regulated genes and can influence the activity of the master circadian clock TF RpaA (Espinosa et al., 2015), discussed further in the next section.

Oxidative defense regulatory proteins detect ROS and initiate signaling cascades, many of which have been conserved throughout aerobic organisms. The Fur-like PedR protein is a redox-activated transcriptional regulator, which interacts with Trxs to relay photosynthetic stress information to control oxidative response gene expression (Horiuchi et al., 2010). PerR regulates the transcription of peroxiredoxin antioxidant enzymes (Takeda et al., 2004). H_2O_2 responsive transcriptional regulator RexT upregulates TrxA2 thioredoxin to maintain cellular redox

homeostasis (Li et al., 2022, Ehira and Ohmori, 2012), and OxyR activates the transcription of detox genes including catalases, peroxidases, and thiol reductases (Pedre et al., 2018).

TFs associated with nitrogen and iron starvation are also active during high-light acclimation. During high-light exposure, PBSs are degraded, a process regulated by NtcA. NtcA promotes *nblA* expression in response to high light (Giner-Lamia et al., 2017). NtcA and coactivator PipX are global regulators in nitrogen control. PipX interacts with both NtcA and the PII protein, key regulators that integrate N/C balance signals with energy status (Espinosa et al., 2006). In Anabaena, the global nitrogen regulator *ntcA* is reported to play a role in the regulation of pigment synthesis and ROS production under iron stress conditions likely in response to oxidative stress (Kaushik et al., 2017). It is clear that nitrogen regulators are important factors for high-light acclimation, but the mechanisms that activate these TFs during high-light response have yet to be elucidated.

The transcriptional regulator Iron deficiency-induced protein B (IdiB) and its regulon that includes *idiA* and *idiC* may be important for long-term high-light acclimation. IdiA has been found to associate with PSII (Tolle et al., 2002) and may play a role in protecting the acceptor side of PSII under oxidative stress (Exss-Sonne et al., 2000). The IdiA protein was found in higher amounts under high and fluctuating light (Mustila et al., 2021). Redox-dependent TF Sll1961 from *Synechocystis,* previously found to regulate photosystem stoichiometry under high light (Fujimori et al., 2005), has also been shown to be reduced by TrxM (Kujirai et al., 2018). However, the regulatory network of *sll1961* has not yet been identified and may differ based on environmental stress experienced. Additional TFs implicated in the high-light response are still being identified and their mechanisms are being characterized.

1.5.2 Circadian regulation

Light cues derived from the photosynthetic ETC entrain the circadian clock (Mackey et al., 2011). The circadian clock inputs information to regulatory pathways to tune physiology based on anticipated daily environmental patterns. Therefore, at certain times of the day, cells are more primed for high-light acclimation. Changes in light intensity affect the expression of hundreds of genes enhanced or repressed by the circadian clock, many of which participate in key metabolic steps (Markson et al., 2013a, Piechura et al., 2017). Circadian-regulated genes with peak expression at subjective dawn are involved in carbon fixation, light-dependent photosynthesis, and growth processes, while dusk genes participate in functions such as glycogen catabolism (Diamond et al., 2015, Gutu and O'Shea, 2013). The interconnectedness of circadian regulation and light response has been demonstrated in *Synechococcus* (Piechura et al., 2017), probably facilitated by the transcriptional regulators RpaA, RpaB, and Rre1, which are integral to circadian clock function and abiotic stress acclimation.

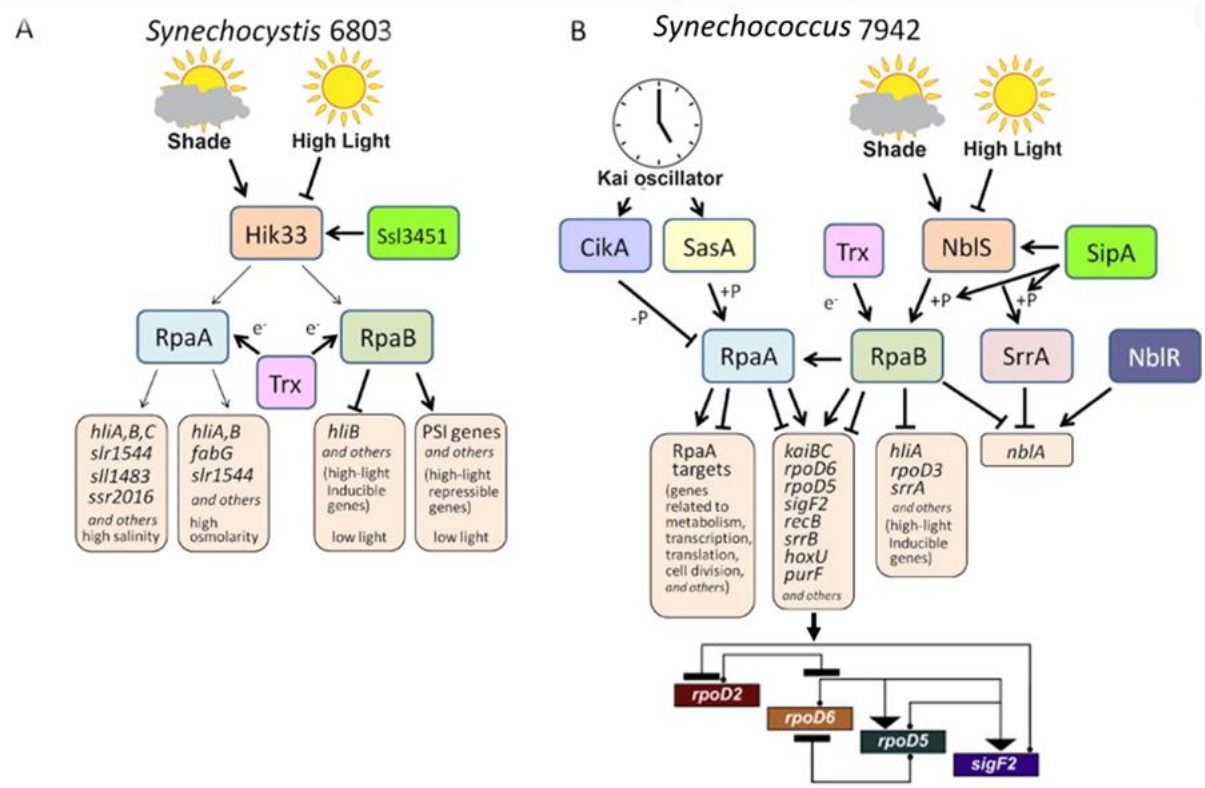

Figure 1.3: Diagram of the circadian clock and post-translational modifications that regulate circadian gene expression and sigma factor cascade in two model cyanobacterial species (a) *Synechocystis* and (b) *Synechococcus*. Circadian regulation diagram adapted from (Wilde and Hihara, 2016). Diagram of the sigma factor cascade adapted from (Fleming and O'Shea, 2018). Thin arrows indicate the putative interaction inferred from DNA microarray analysis. + P, phosphorylation; − P, dephosphorylation; e^-, reduction of thiols. Arrows indicate activation and bar-headed lines indicate inhibition. Ortholog proteins are colored the same in (a) and (b).

RpaA is considered to be the master regulator of circadian gene expression (Wilde and Hihara, 2016). The sensor kinase SasA interacts with the core circadian clock proteins and phosphorylates RpaA to mediate a circadian timing signal necessary for the transcriptional oscillation of clock-regulated genes (Fig. 1.3) (Fleming and O'Shea, 2018). The RpaA paralog RpaB also influences circadian target gene expression in a light-responsive manner. Circadian genes regulated by RpaB and RpaA are transiently repressed or activated in response to light upshift without altering global circadian oscillations (Piechura et al., 2017). RpaA, RpaB, and Rre1 have been shown to interact with TrxA *in vitro* (Kadowaki et al., 2015, Ibrahim et al.,

2022), which may indicate a redox response mechanism. RpaB is rapidly dephosphorylated following high-light exposure, leading to the expression of high-light responsive genes and influences on RpaA (Hanaoka and Tanaka, 2008, Espinosa et al., 2015, Yasuda et al., 2020). The general functions of RpaA and RpaB are conserved in *Synechocystis* and *Synechococcus,* although the sets of circadian genes that they regulate differ slightly (Fig. 1.3). For instance, in *Synechocystis* the *hli* genes are targets of both RpaA and RpaB, while they are targets of only RpaB in *Synechococcus.* Also, in *Synechocystis* RpaA has been reported to be involved in the regulation of monomeric PSI accumulation and *psbA* under high light (Majeed et al., 2012). It was recently suggested that RpaA may inhibit RpaB binding or compete for target gene binding under reduced plastoquinone (PQ) conditions (Kato et al., 2022). This regulation of RpaB by RpaA could result in large responses in the *Synechococcus* strains with different RpaA alleles during high-light acclimation.

1.5.3 Non-coding RNAs

Regulation at both the transcriptional and translational levels is important during early high-light response. Non-coding RNAs are often involved in regulatory mechanisms during environmental stress acclimation, and there are hundreds of potential regulatory small RNAs (sRNAs) in a cyanobacterium (Kopf and Hess, 2015). In *Synechocystis* there is an inverse relationship in the expression of *apcABC* and *ocp* genes. This is believed to be due to a regulatory PTM mediated by the sRNA, ApcZ, from the allophycocyanin operon of the PBS that inhibits the translation of OCP mRNA (Zhan et al., 2021). ApcZ functions in PTM to regulate photosynthetic processes in response to cellular physiology. During stress conditions, the expression of PBS genes is downregulated; ApcZ levels also decrease, allowing for an increase in OCP protein concentrations. ApcZ is also predicted to interact with *nblB2* and the response

regulator *pmgA* (Zhan et al., 2021). High-light activated kinase Hik26 phosphorylates a response regulator, possibly PmgA (Yoshikawa, 2021), although the exact regulator has not yet been identified. PmgA and NblB2 may play roles in PBS turnover (Dolganov and Grossman, 1999) and PSI/Chl suppression during exposure to high light (Hihara et al., 1998).

The PsrR1 sRNA regulates essential steps in high-light acclimation by targeting mRNAs of photosynthetic components, including PSI, PSII, PBS, and chlorophyll metabolism (Georg et al., 2014). Expression of *psrR* is dependent on RpaB (Riediger et al., 2019), creating a feed-forward loop between PsrR1 and RpaB to control light acclimation (Tan et al., 2018). PsrR1 represses the translation of PBS genes including *cpcA, cpcB, apcE,* and *apcF* (Nagarajan et al., 2014, Kirst et al., 2014), allowing for improved photosynthetic efficiency during stress (Joseph et al., 2014). The downregulation of *apcE* may function in promoting biomass and glycogen accumulation (Joseph et al., 2014), which is found in high concentrations in the high-light tolerant *Synechococcus* 2973 (Ungerer et al., 2018b). Cyanobacteria synthesize and accumulate glycogen during periods of light when energy production is high, which is later broken down in the dark for carbon and energy. Glycogen production may serve to balance the cellular redox state during high photosynthetic activity.

Antisense RNAs have been found to control the concentrations of photosynthetic components through control of target mRNA degradation. The *isiABC* operon is upregulated in response to high light (Yoshikawa et al., 2021) and may be regulated by several non-coding RNAs. IsrR is an antisense RNA from the *isiA* non-coding strand that controls IsiA concentrations through degradation control (Duhring et al., 2006), and its expression is dependent on RpaB (Riediger et al., 2019). IsrR, together with As1-flv4, regulates the flv2-4 operon (Zhang et al., 2012). Three sRNAs within the *isiA* operon were recently identified as

being related to oxidative stress and are predicted to be involved in the regulation of *isiA*, *cpcD*, *psaB*, *psaA*, and *gidA* by sequestering their respective mRNA (Cheng et al., 2020). Other antisense RNAs stabilize target mRNAs such as *rblR*, *psbA2R*, and *psbA3R*, which target *rbcL*, *psbA2*, and *psbA3* mRNAs respectively (Hu et al., 2017, Eisenhut et al., 2012). Relatively little is known about the non-coding RNAs, although our understanding of their role in regulating environmental stress response is growing.

1.5.4 Gene expression networks

To maintain homeostasis during high-light conditions, cells employ intricate regulatory networks coordinated by several transcription and sigma factors. Transcriptome studies have identified similar subsets of high-light induced or repressed genes, revealing portions of a complex transcriptional network that regulate core processes in response to high light (Ogawa et al., 2018a, Tan et al., 2018). In general, more genes are repressed than induced in response to high-light conditions. Gene expression analysis of *Synechocystis and Synechococcus* found that under high light there was downregulation of genes encoding phycobilisome, PSI, pilus synthesis (*pilA2*), pyridine nucleotide transhydrogenase (*pntA* and *pntB*), sigma factors (*sigA* and *sigF*), and upregulation of *psbA* for PSII repair (Ogawa et al., 2018a, Tan et al., 2018). Many of the genes in the light response network overlap with other stress response gene networks, indicating that regulators of other stress response pathways can act together with light-responsive regulators. Exactly how many light-responsive genes contribute to the light-stress response or how they fit into the light-stress response network remains to be elucidated.

Upregulation of oxidative stress protective mechanisms includes the accumulation of peroxidases (*aphC* and *gpx2*), and carotenoids (Izuhara et al., 2020). *Synechocystis* encodes many ROS-scavenging enzymes, including superoxide dismutase (sodB), catalase-peroxidase

(katG), peroxiredoxins (Prxs) and glutathione peroxidases (Bernroitner et al., 2009). In addition to cleaning up ROS and acting as a molecular chaperone, the Trx-dependent Prx proteins act as activators of cell signaling (Hosoya-Matsuda et al., 2005, Perez-Perez et al., 2009). As discussed in previous sections, Trxs are important for light signaling (Fig. 1.3) and act on a variety of regulators of central metabolic enzymes. TrxB and TrxQ reduce 2-Cys Prx and are involved in high-light related oxidative stress response (Perez-Perez et al., 2009). TrxA and CP12 are involved in the regulation of CBB cycle proteins, which produce changes in electron flux (Fig. 1.2) (Mallen-Ponce et al., 2021, Lindahl and Florencio, 2003).

Chaperone Proteins

Light-induced DNA and protein damage creates a demand for protease and chaperone activity. Chaperone proteins such as heat shock proteins (HSPs) are involved in folding newly synthesized proteins, stabilizing, or promoting protein degradation. Some HSPs have been suggested to function in protecting phycocyanin and the PBS complex, as in well as the regulation of tetrapyrrole biosynthesis (Watanabe et al., 2007). High light induces the expression of many of the HSPs including *groESL*, *clpB*, *dnaK2*, *hspA*, and *htpG* (Hihara et al., 2001). Analysis of *Synechocystis* grown under intense high light (1100 and 1300 μmol m^{-2} s^{-1}) found that genes encoding HSPs (*hsp, htpG, and dnaJ*) were the only category over-represented in upregulated genes (Ogawa et al., 2018a). The reduced state of the PQ appears to act as a key signal that affects the transcription of *hspA*, *groESL*, and *ocp* (Muramatsu and Hihara, 2012). Diverse regulatory mechanisms control the expression of HSPs, including two-component regulatory systems and sigma factors (Tuominen et al., 2008, Tuominen et al., 2006). Hik34 phosphorylates the regulator Rre1 to induce a signaling module that positively regulates heat and light-responsive genes including the sigma factor gene *rpoD2* and chaperone genes such as

dnaK2, *groESL*, *groEL2*, *hspA*, and *htpG* (Kobayashi et al., 2017). Additionally, several regulatory sequences have been found to be involved in heat and/or light-mediated regulation of the *groESL1* operon; however, differences in conservation and mutant studies suggest that the regulatory mechanisms of *groEL* expression differs among cyanobacterial species (Saito et al., 2020, Kojima and Nakamoto, 2007). In *Synechocystis*, GroEL protein chaperones stabilize the HrcA repressor protein, a negative regulator of the expression of *groEL*, *dnaJ*, and *dnaK*. During heat- and/or light-stress conditions, GroEL stops stabilizing HrcA, allowing it to unfold and resulting in the expression of the *groEL* operon (Nakamoto et al., 2003).

D1 turnover

High-light intensity response relies heavily on PSII repair, indicated by the upregulation of PSII D1 forms II and III (*psbAII-AIII*) and D1 protease (*ftsH*) genes (Jimbo et al., 2019) (Hihara et al., 2001, Tan et al., 2018). Another indicator of PSII turnover is the elongation factor EF-Tu, of which levels increased three-fold after light upshift, while ribosome protein levels did not change (Jimbo et al., 2019). Recently, *slr0320* from *Synechocystis* was found to be involved in optimizing PSII activity and electron transfer during high-light acclimation (Zhang et al., 2021). Future studies of these and other novel high-light response components may reveal mechanisms of regulation and acclimatization. Many TFs are potential candidates as light stress response regulators. Further studies are required to characterize and integrate these components into our knowledge of light response regulatory networks.

Sigma Factors

Many sigma factors have been suggested to be involved in high-light response, although not many have been fully characterized, and sigma factor regulons may alter depending on the stress condition that activates them. Group 2 sigma factors are non-essential under general

growth conditions but regulate global metabolic responses during various environmental stresses. All group 2 sigma factors play some role in high-light response, although only *rpoD3* (also called *sigD*) and *sigB* are consistently induced in response to high-light stress (Imamura et al., 2003, Zhang et al., 2021, Turunen et al., 2022) and are essential for acclimation to high light (Seki et al., 2007, Pollari et al., 2009). During growth in 750 μmol photons m^{-2} s^{-1} light the RpoD3/SigD regulon consists of ~62 genes, involved in the photosynthetic ETC, pigment synthesis, purine biosynthesis genes, and two-component response regulation (Turunen et al., 2022). When experiencing high heat, the RpoD3/SigD regulon contains a different and much smaller set of genes (Turunen et al., 2022). The expression of *rpoD3/sigD* is induced by RpaB, while activation of the sigma factor activity may rely on the redox state of the photosynthetic ETC (Seki et al., 2007). The role of SigB in high-light response is less clear although its regulon includes a few genes important for light response such as *trxA, ftsH* (involved in PSII turnover), and *psb28-2* (PSII subunit) (Turunen et al., 2022).

1.6 Metabolic remodeling

Nutrient assimilation processes require energy produced from photosynthesis and photosynthetic protein complexes require nutrients such as iron, S, N, and C. Therefore, cross-integration of regulation is required to balance the supply of electrons from photosynthetic processes with the demands of cellular metabolism. Under normal-light conditions, C and N metabolism act as sinks for reducing power from the photosynthetic ETC. Decreased energy production due to photoinhibition and downregulation of light absorption affect the ability of cells to fix CO_2. *Synechocystis* downregulates CO_2-concentrating mechanism proteins, which trigger an integrated homeostatic response and downregulates N-transport and assimilation

(Singh et al., 2008). This is further modulated during long-term acclimation when *Synechocystis* uses nitrogen-assimilation pathways as an electron sink (Mustila et al., 2021). Light intensity impacts the regulation of the CBB cycle, the primary metabolic pathway for CO_2 fixation (Fig. 2). The flux of carbon through the CBB cycle is regulated both at the transcriptional level (Kusian and Bowien, 1997) as well as the post-translational level through redox control (Michelet et al., 2013). A higher flux of metabolites may provide a sink for carbon assimilation (consuming trioses phosphate, ATP, and NADPH) from photosynthetic activity (Ibrahim et al., 2010). In plants, a higher C/N balance has been observed to be important for promoting the accumulation of flavonoids (Deng et al., 2019), suggesting that nitrogen supply modulates flavonoid and carotenoid biosynthesis.

High light commonly leads to the downregulation of nitrogen uptake-related genes and upregulation of CCM-related (Mustila et al., 2021). Central C- and N-metabolism are tightly linked through metabolites, energy requirements, and crosstalk to ensure balance across the cell. Changes to C/N balance are also observed under high-light conditions including increased expression of inorganic carbon (C_i) uptake (*ndhF3* operon, *cmp* operon, *sbtAB*) and repression of nitrogen uptake (*amt1*, *amt2*, *nrtB*, *urtB)*. Levels of 2-oxoglutarate (2-OG) from the tricarboxylic acid (TCA) cycle are detected by the PII protein as a signal of C/N status, and the detection of ADP by the PII protein indicates the ADP/ATP ratios (Zeth et al., 2014). PII relays low nitrogen information by promoting CO_2 fixation and preventing the nitrogen TF NtcA from forming a protein complex with PipX (Espinosa et al., 2006). In addition to 2-OG, glutamine also serves as an effector molecule making glutamine synthetase another key target for controlling N-assimilation. Interestingly, a recent study demonstrated that H_2O_2 interferes with sensing the C/N status by affecting the 2-OG levels (Robles-Rengel et al., 2019). In *Synechocystis* mRNA and

protein level accumulation of N-uptake and assimilation components were observed during high-light pulse and low CO_2 conditions (Mustila et al., 2021). Similar results were also observed in *Synechococcus* 2973 (Tan et al., 2018).

Some evidence has pointed to C- and N-assimilation regulatory proteins such as the PII protein (Fokina et al., 2010) and cyAbrB transcriptional regulators (Ishii and Hihara, 2008) as playing important roles during high-light acclimation. The CyAbrB2 regulatory protein negatively regulates the norB and hox operon and is a positive regulator of several nitrogen-related genes such as *urtA*, *amt1*, *narB*, and the *nrt* operon (Ishii and Hihara, 2008), as well as CCM-related genes (*cmpABCB*, *sbtAB* and *ndhF3/D3/cupA* operons) and complement CmpR and NdhR transcription factors (Orf et al., 2016). It was recently suggested that flv proteins may be involved in crosstalk with N-assimilation (Mustila et al., 2021) similar to the nitric oxide (NO) quenching function of Flv proteins in anaerobic microbes and *Chlamydomonas reinhardtii* (Burlacot et al., 2020). NO, a secondary messenger, is known to regulate TF in other phototrophs (Mengel et al., 2013) and could act as a signaling molecule in cyanobacteria to link the photosynthetic apparatus and N-metabolism. In plants, increased ROS level promotes the production of reactive nitrogen species including NO, which signals oxidative stress (Saxon, 2016).

Photosynthetic complexes are distributed through the thylakoid membrane in different ways, which may indicate different functional mechanisms for the photosynthetic ETC. The respiratory complex NDH-1 reorganizes in the thylakoid membrane under different light intensities (Casella et al., 2017), which could serve to alter electron flow pathways. NDH-1 can form a complex with PSI for efficient CET (Gao et al., 2016), allowing for higher ATP production required for C_i uptake and CCMs. The cyanobacterial NDH-1 complexes not only

participate in light reactions but also in C-assimilation. NDH-1_1 and NDH-1_2 function in respiration while NDH-1_3 and NDH-1_4 function in CCMs (Peltier et al., 2016). NDH-1_4 in particular appears to be important for high-light acclimation, as it was upregulated during high-light growth (Mustila et al., 2021). The carbon concentration regulator CcmR (also called NdhR) acts as a repressor of NDH-1 and C_i uptake genes (*ndhF3, ndhD3, cupA, bicA, sbtA,* and *ccmR*) and activates transcription of the *cmpABCD* operon (Woodger et al., 2007). C_i limitation is signaled by the metabolic intermediates α-ketoglutarate (α-KG) and $NADP^+$ and results in the PII protein binding to CcmR to enhance target gene repression (Daley et al., 2012). Ribulose1,5-bisphosphate (RuBP) and 2-phosphoglycolate (2-PG) act as co-activators of CmpR (Daley et al., 2012). Taken together, the NDH-1 complexes contribute to higher ATP levels and concentrated CO_2 levels, allowing for a high-flux towards CO_2 fixation and decreasing redox imbalance in the thylakoid.

In addition to PSI, PBS, and chemotaxis-related genes, sigma factors *sigA* and *sigF* as well as the pyridine nucleotide transhydrogenase genes *pntA* and *pntB* were negatively regulated in *Synechocystis* grown under intense high light (Ogawa et al., 2018a). In *Synechococcus,* PntAB is also downregulated in response to high light (Piechura et al., 2017). Mutants in PntA exhibit impaired ATP production (Tyystjarvi and Aro, 1996), state transitions, and C-assimilation (Kamarainen et al., 2017). PntAB keeps the intracellular $NADP^+$/NADPH ratios low to provide reductive power for C-assimilation and other metabolic reactions, such as amino acid and fatty acid biosynthesis or glutathione (GSH) reduction (Kamarainen et al., 2017). GSH is significant in high-light acclimatization because it reduces oxidative damage, together with other compounds such as ascorbate and TRX (Narainsamy et al., 2016). Photosynthetic production of NADPH increases in high light (Ogawa et al., 2018a) since more reducing equivalents are

produced during photosynthesis. $NADP^+$: NADPH balance may be important for adaptation to high-light conditions, considering that NADPH is also required to detoxify ROS (Diamond et al., 2017). Additionally, these genes may be regulated by *sigF* or *sigA* (Ogawa et al., 2018a). Disruption of *sigF*, known to be involved in phototaxis, led to changes in the expression of carotenoids and C-metabolism genes, as well as pigment release, which may protect against high light damage (Flores et al., 2019).

1.7 High-light tolerant strain *Synechococcus* 2973

In the 1950s, a mixed culture containing *Synechococcus* sp. PCC 6301 (or UTEX 625) originated from a creek in Texas and was deposited in several culture collections (Kratz and Myers, 1955). Recently, due to its fast-growth ability, *Synechococcus elongatus* UTEX 2973 was isolated from that mixed culture and found to be remarkably similar to the model cyanobacterium *Synechococcus* 7942 (Yu et al., 2015b) (Fig. 1.4). These strains share an average nucleotide identity of 99.9% and were first identified to have 55 genetic differences: 53 single nucleotide polymorphisms (SNPs), a 7-kb deletion, and a 177-kb inversion (Yu et al., 2015b). This total was later reduced to 51 SNPs as RpaA-Q121 and Aas-L95 were not SNPs but rather the result of sequencing a mutant strain of *Synechococcus* 7942 (Adomako et al., 2022). Despite these strains being nearly genetically identical, they have uniquely divergent phenotypes, with *Synechococcus* 2973 being tolerant to extreme high-light intensities (Fig. 1.4).

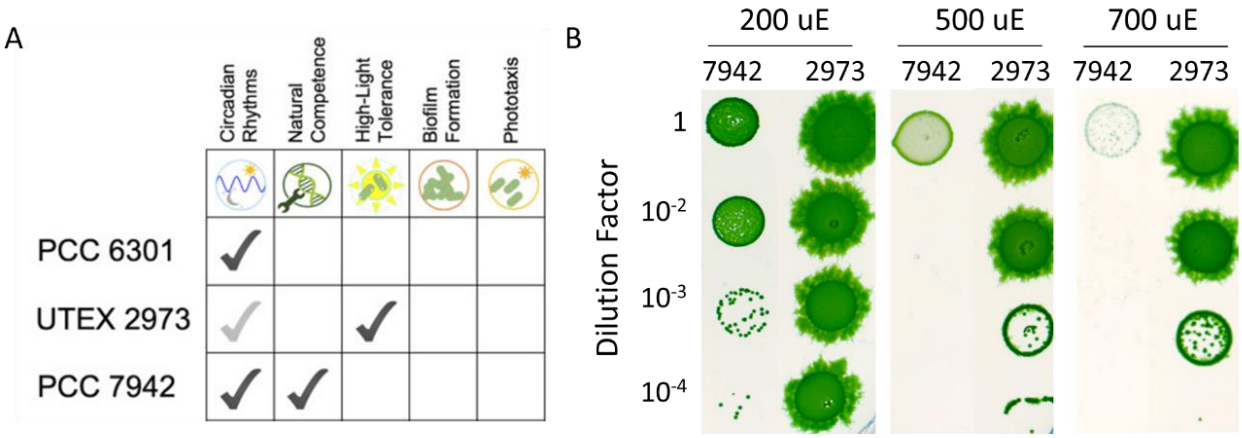

Figure 1.4: Divergent phenotypes of *Synechococcus* genotypes. (a) Combinations of phenotypes by strain (gray checkmark represents presumed circadian rhythms but not yet tested for in UTEX 2973). (Adomako et al., 2022) (b) Dilution spot plates comparing growth of *Synechococcus* 7942 and *Synechococcus* 2973 under three increasing light intensities.

One way in which *Synechococcus* 2973 achieves fast growth in high-light intensity is through more efficient photosynthetic electron flux. *Synechococcus* 2973 has been shown to have increased titers of cytochrome *f*, plastocyanin, and PSI, which lead to improved kinetics of the photosynthetic ETC (Ungerer et al., 2018b). In addition, the polymorphic alleles of ATP synthase subunit *atpA* and NAD^+ kinase *ppnK* result in increased production of $NADP^+$ and ATP (Ungerer et al., 2018c). The increased energy supply allows for a higher C-fixation rate supporting faster growth (Ungerer et al., 2018c). This is supported by a transcriptional study that found that *Synechococcus* 2973 upregulates C_i transport genes and enhances biomass and glycogen accumulation in high light (Tan et al., 2018). The more efficient electron flow allows for higher utilization of light energy, therefore increasing the high-light tolerance of *Synechococcus* 2973. Finally, a third allele encoding *rpaA* is required for fast growth in 900 µmol photons m^{-2} s^{-1} light. The *rpaA* allele contains differences both in the coding region and in the promoter region of the gene, and when the *Synechococcus* 2973 allele was introduced into *Synechococcus* 7942, the 7-bp deletion in the promoter resulted in decreased *rpaA* expression, as

well as many perplexing changes in the transcriptome (Ungerer et al., 2018c). However, this study lacked biological replicates, leaving questions about the pleiotropic effects caused by the RpaA allele unanswered.

High light may also damage cyanobacterial membranes. Genes involved in fatty acid biosynthesis (*accA*, *accC*, *fabH*, and *ado*) were upregulated during exposure to high light in *Synechococcus* 2973 (Tan et al., 2018). *Synechococcus* strains were originally thought to contain two SNPs in the long-chain-fatty-acid CoA ligase gene *aas* (aka *fadD*) (Yu et al., 2015b), although one of which (P295L) was found to be a sequencing error (Adomako et al., 2022). Aas is involved in free fatty acid recycling and activates the remodeling of membrane lipids for high-light acclimation in *Synechococcus* 7942 (Takatani et al., 2015). In both *Synechococcus* and *Synechocystis* inactivation of *aas* led to unstable PSII and increased photoinhibition under high light (Takatani et al., 2015, Jimbo et al., 2021). Increased thylakoid membrane fluidity more easily allows for PSII repair and reduces the harm caused by H_2O_2 (Yamamoto, 2016). Therefore, it may be that the polymorphic Aas protein could affect fatty acid biosynthesis and may indicate a difference in the membrane composition between the *Synechococcus* strains that affects the high-light tolerance.

Two other alleles were observed to weakly affect growth rate, although the growth rate range was not significantly different from the wild type. These SNPs were encoded by the genes chorismite mutase *aroH*, and DNA-directed RNA polymerase β subunit *rpoB* (Ungerer et al., 2018c). AroH, a component of the shikimate pathway, catalyzes the reaction to convert chorismite to prephenate, which is later used to produce phenylalanine and tyrosine, and RpoB is involved in transcription elongation. Considering the importance of elongation factors in PSII turnover is it feasible for the SNP of *rpoB* to be a contributor to the high-light tolerance of

Synechococcus 2973. In the transcriptome of *Synechococcus* 2973 under high light, the observed upregulation of ribosomal subunit genes suggested increased translation (Tan et al., 2018). Based on what we know about the regulation of high-light tolerance, a few other SNP-encoding genes could also be candidates for interacting with light-response pathways including molecular chaperone *dnaK1*, and a recently identified sigma factor interacting protein *sinA. sinA* is proposed to facilitate the activity of heat-shock responsive sigma factors during heat stress (Hasegawa et al., 2020). In *Escherichia coli* the Hsp70 protein encoded by *dnaK* acts as an organizer in the chaperone network by stabilizing proteins for folding by GroEL (Calloni et al., 2012). These genes could be involved in sigma factor activity and protein turnover providing additional tuning of light response.

Despite what we know, there are still many questions regarding the high-light tolerance of *Synechococcus* 2973. While much of our understanding of high-light acclimation has previously focused on the repair of PSII, there is not much evidence that this is a key mechanism in *Synechococcus* 2973. Rather, the strain may have reduced levels of PSII damage by having a more efficient ETC, although this requires more in-depth analysis. *Synechococcus elongatus* are some of the few cyanobacteria that do not encode an OCP, leaving the question of how they quench excess absorbed energy. Finally, given the requirement of the *rpaA* allele to confer fast growth under high-light conditions, it seems that other mechanisms beyond the photosynthetic ETC are at play that allow optimal growth in such high-light conditions. Therefore, the complex interaction of the circadian clock and high-light tolerance remains a topic of interest.

1.8 References

Adomako, M., Ernst, D., Simkovsky, R., Chao, Y.Y., Wang, J., Fang, M., Bouchier, C., Lopez-Igual, R., Mazel, D., Gugger, M., *et al.* (2022). Comparative Genomics of *Synechococcus elongatus* Explains the Phenotypic Diversity of the Strains. mBio *13*, e0086222.

Agostoni, M., Koestler, B.J., Waters, C.M., Williams, B.L., and Montgomery, B.L. (2013). Occurrence of cyclic di-GMP-modulating output domains in cyanobacteria: an illuminating perspective. mBio *4*.

Allakhverdiev, S.I., and Murata, N. (2004). Environmental stress inhibits the synthesis de novo of proteins involved in the photodamage-repair cycle of Photosystem II in *Synechocystis* sp. PCC 6803. Biochim Biophys Acta *1657*, 23-32.

Aro, E.M., Virgin, I., and Andersson, B. (1993). Photoinhibition of Photosystem II. Inactivation, protein damage and turnover. Biochim Biophys Acta *1143*, 113-134.

Barteneva, N.S., Meirkhanova, A., Malashenkov, D., and Vorobjev, I.A. (2022). To Die or Not to Die-Regulated Cell Death and Survival in Cyanobacteria. Microorganisms *10*.

Battchikova, N., Wei, L., Du, L., Bersanini, L., Aro, E.M., and Ma, W. (2011). Identification of novel Ssl0352 protein (NdhS), essential for efficient operation of cyclic electron transport around photosystem I, in NADPH:plastoquinone oxidoreductase (NDH-1) complexes of *Synechocystis* sp. PCC 6803. J Biol Chem *286*, 36992-37001.

Bernroitner, M., Zamocky, M., Furtmuller, P.G., Peschek, G.A., and Obinger, C. (2009). Occurrence, phylogeny, structure, and function of catalases and peroxidases in cyanobacteria. J Exp Bot *60*, 423-440.

Bersanini, L., Battchikova, N., Jokel, M., Rehman, A., Vass, I., Allahverdiyeva, Y., and Aro, E.M. (2014). Flavodiiron protein Flv2/Flv4-related photoprotective mechanism dissipates excitation pressure of PSII in cooperation with phycobilisomes in Cyanobacteria. Plant Physiol *164*, 805-818.

Bhaya, D., Dufresne, A., Vaulot, D., and Grossman, A. (2002). Analysis of the hli gene family in marine and freshwater cyanobacteria. FEMS Microbiol Lett *215*, 209-219.

Blankenship, R.E. (2015). Structural and functional dynamics of photosynthetic antenna complexes. Proc Natl Acad Sci U S A *112*, 13751-13752.

Burlacot, A., Richaud, P., Gosset, A., Li-Beisson, Y., and Peltier, G. (2020). Algal photosynthesis converts nitric oxide into nitrous oxide. Proc Natl Acad Sci U S A *117*, 2704-2709.

Calloni, G., Chen, T., Schermann, S.M., Chang, H.C., Genevaux, P., Agostini, F., Tartaglia, G.G., Hayer-Hartl, M., and Hartl, F.U. (2012). DnaK functions as a central hub in the E. coli chaperone network. Cell Rep *1*, 251-264.

Casella, S., Huang, F., Mason, D., Zhao, G.Y., Johnson, G.N., Mullineaux, C.W., and Liu, L.N. (2017). Dissecting the Native Architecture and Dynamics of Cyanobacterial Photosynthetic Machinery. Mol Plant *10*, 1434-1448.

Chen, H.S., Niedzwiedzki, D.M., Bandyopadhyay, A., Biswas, S., and Pakrasi, H.B. (2021). A Novel Mode of Photoprotection Mediated by a Cysteine Residue in the Chlorophyll Protein IsiA. mBio *12*.

Cheng, Y., Zhang, T., Wang, L., and Chen, W. (2020). Transcriptome Analysis Reveals IsiA-Regulatory Mechanisms Underlying Iron Depletion and Oxidative-Stress Acclimation in *Synechocystis* sp. Strain PCC 6803. Appl Environ Microbiol *86*.

Cogdell, R.J., and Gardiner, A.T. (2015). Activated OCP unlocks nonphotochemical quenching in cyanobacteria. Proc Natl Acad Sci U S A *112*, 12547-12548.

Daley, S.M., Kappell, A.D., Carrick, M.J., and Burnap, R.L. (2012). Regulation of the cyanobacterial CO2-concentrating mechanism involves internal sensing of NADP+ and alpha-ketogutarate levels by transcription factor CcmR. PLoS One *7*, e41286.

Demoulin, C.F., Lara, Y.J., Cornet, L., Francois, C., Baurain, D., Wilmotte, A., and Javaux, E.J. (2019). Cyanobacteria evolution: Insight from the fossil record. Free Radic Biol Med *140*, 206-223.

Deng, B., Li, Y., Lei, G., and Liu, G. (2019). Effects of nitrogen availability on mineral nutrient balance and flavonoid accumulation in Cyclocarya paliurus. Plant Physiol Biochem *135*, 111-118.

Diamond, S., Jun, D., Rubin, B.E., and Golden, S.S. (2015). The circadian oscillator in *Synechococcus elongatus* controls metabolite partitioning during diurnal growth. Proc Natl Acad Sci U S A *112*, E1916-1925.

Diamond, S., Rubin, B.E., Shultzaberger, R.K., Chen, Y., Barber, C.D., and Golden, S.S. (2017). Redox crisis underlies conditional light-dark lethality in cyanobacterial mutants that lack the circadian regulator, RpaA. Proc Natl Acad Sci U S A *114*, E580-E589.

Dolganov, N., and Grossman, A.R. (1999). A polypeptide with similarity to phycocyanin alpha-subunit phycocyanobilin lyase involved in degradation of phycobilisomes. J Bacteriol *181*, 610-617.

Donaldson, L., Meier, S., and Gehring, C. (2016). The arabidopsis cyclic nucleotide interactome. Cell Commun Signal *14*, 10.

Duhring, U., Axmann, I.M., Hess, W.R., and Wilde, A. (2006). An internal antisense RNA regulates expression of the photosynthesis gene isiA. Proc Natl Acad Sci U S A *103*, 7054-7058.

Ehira, S., and Ohmori, M. (2012). The redox-sensing transcriptional regulator RexT controls expression of thioredoxin A2 in the cyanobacterium Anabaena sp. strain PCC 7120. J Biol Chem *287*, 40433-40440.

Eisenhut, M., Georg, J., Klahn, S., Sakurai, I., Mustila, H., Zhang, P., Hess, W.R., and Aro, E.M. (2012). The antisense RNA As1_flv4 in the Cyanobacterium *Synechocystis* sp. PCC 6803 prevents premature expression of the flv4-2 operon upon shift in inorganic carbon supply. J Biol Chem *287*, 33153-33162.

Enomoto, G., Ni Ni, W., Narikawa, R., and Ikeuchi, M. (2015). Three cyanobacteriochromes work together to form a light color-sensitive input system for c-di-GMP signaling of cell aggregation. Proc Natl Acad Sci U S A *112*, 8082-8087.

Espinosa, J., Boyd, J.S., Cantos, R., Salinas, P., Golden, S.S., and Contreras, A. (2015). Cross-talk and regulatory interactions between the essential response regulator RpaB and cyanobacterial circadian clock output. Proc Natl Acad Sci U S A *112*, 2198-2203.

Espinosa, J., Forchhammer, K., Burillo, S., and Contreras, A. (2006). Interaction network in cyanobacterial nitrogen regulation: PipX, a protein that interacts in a 2-oxoglutarate dependent manner with PII and NtcA. Mol Microbiol *61*, 457-469.

Exss-Sonne, P., Tolle, J., Bader, K.P., Pistorius, E.K., and Michel, K.P. (2000). The IdiA protein of *Synechococcus* sp. PCC 7942 functions in protecting the acceptor side of Photosystem II under oxidative stress. Photosynth Res *63*, 145-157.

Fleming, K.E., and O'Shea, E.K. (2018). An RpaA-Dependent Sigma Factor Cascade Sets the Timing of Circadian Transcriptional Rhythms in *Synechococcus elongatus*. Cell Rep *25*, 2937-2945 e2933.

Flores, C., Santos, M., Pereira, S.B., Mota, R., Rossi, F., De Philippis, R., Couto, N., Karunakaran, E., Wright, P.C., Oliveira, P., *et al.* (2019). The alternative sigma factor SigF is a key player in the control of secretion mechanisms in *Synechocystis* sp. PCC 6803. Environ Microbiol *21*, 343-359.

Fokina, O., Chellamuthu, V.R., Forchhammer, K., and Zeth, K. (2010). Mechanism of 2-oxoglutarate signaling by the *Synechococcus elongatus* PII signal transduction protein. Proc Natl Acad Sci U S A *107*, 19760-19765.

Fujimori, T., Higuchi, M., Sato, H., Aiba, H., Muramatsu, M., Hihara, Y., and Sonoike, K. (2005). The mutant of sll1961, which encodes a putative transcriptional regulator, has a defect in regulation of photosystem stoichiometry in the cyanobacterium *Synechocystis* sp. PCC 6803. Plant Physiol *139*, 408-416.

Gao, F., Zhao, J., Chen, L., Battchikova, N., Ran, Z., Aro, E.M., Ogawa, T., and Ma, W. (2016). The NDH-1L-PSI Supercomplex Is Important for Efficient Cyclic Electron Transport in Cyanobacteria. Plant Physiol *172*, 1451-1464.

Georg, J., Dienst, D., Schurgers, N., Wallner, T., Kopp, D., Stazic, D., Kuchmina, E., Klahn, S., Lokstein, H., Hess, W.R., *et al.* (2014). The small regulatory RNA SyR1/PsrR1 controls photosynthetic functions in cyanobacteria. Plant Cell *26*, 3661-3679.

Giner-Lamia, J., Robles-Rengel, R., Hernandez-Prieto, M.A., Muro-Pastor, M.I., Florencio, F.J., and Futschik, M.E. (2017). Identification of the direct regulon of NtcA during early acclimation to nitrogen starvation in the cyanobacterium *Synechocystis* sp. PCC 6803. Nucleic Acids Res *45*, 11800-11820.

Gutu, A., and O'Shea, E.K. (2013). Two antagonistic clock-regulated histidine kinases time the activation of circadian gene expression. Mol Cell *50*, 288-294.

Hakala, M., Rantamaki, S., Puputti, E.M., Tyystjarvi, T., and Tyystjarvi, E. (2006). Photoinhibition of manganese enzymes: insights into the mechanism of photosystem II photoinhibition. J Exp Bot *57*, 1809-1816.

Hanaoka, M., Takai, N., Hosokawa, N., Fujiwara, M., Akimoto, Y., Kobori, N., Iwasaki, H., Kondo, T., and Tanaka, K. (2012). RpaB, another response regulator operating circadian clock-dependent transcriptional regulation in *Synechococcus elongatus* PCC 7942. J Biol Chem *287*, 26321-26327.

Hanaoka, M., and Tanaka, K. (2008). Dynamics of RpaB-promoter interaction during high light stress, revealed by chromatin immunoprecipitation (ChIP) analysis in *Synechococcus elongatus* PCC 7942. Plant J *56*, 327-335.

Hasegawa, H., Tsurumaki, T., Kobayashi, I., Imamura, S., and Tanaka, K. (2020). Identification and analysis of a principal sigma factor interacting protein SinA, essential for growth at high temperatures in a cyanobacterium *Synechococcus elongatus* PCC 7942. J Gen Appl Microbiol *66*, 66-72.

Havaux, M., Guedeney, G., Hagemann, M., Yeremenko, N., Matthijs, H.C., and Jeanjean, R. (2005). The chlorophyll-binding protein IsiA is inducible by high light and protects the cyanobacterium *Synechocystis* PCC6803 from photooxidative stress. FEBS Lett *579*, 2289-2293.

Hihara, Y., Kamei, A., Kanehisa, M., Kaplan, A., and Ikeuchi, M. (2001). DNA microarray analysis of cyanobacterial gene expression during acclimation to high light. Plant Cell *13*, 793-806.

Hihara, Y., Sonoike, K., and Ikeuchi, M. (1998). A novel gene, pmgA, specifically regulates photosystem stoichiometry in the cyanobacterium *Synechocystis* species PCC 6803 in response to high light. Plant Physiol *117*, 1205-1216.

Horiuchi, M., Nakamura, K., Kojima, K., Nishiyama, Y., Hatakeyama, W., Hisabori, T., and Hihara, Y. (2010). The PedR transcriptional regulator interacts with thioredoxin to connect photosynthesis with gene expression in cyanobacteria. Biochem J *431*, 135-140.

Hosoya-Matsuda, N., Motohashi, K., Yoshimura, H., Nozaki, A., Inoue, K., Ohmori, M., and Hisabori, T. (2005). Anti-oxidative stress system in cyanobacteria. Significance of type II peroxiredoxin and the role of 1-Cys peroxiredoxin in *Synechocystis* sp. strain PCC 6803. J Biol Chem *280*, 840-846.

Hu, J., Li, T., Xu, W., Zhan, J., Chen, H., He, C., and Wang, Q. (2017). Small Antisense RNA RblR Positively Regulates RuBisCo in *Synechocystis* sp. PCC 6803. Front Microbiol *8*, 231.

Ibrahim, I.M., Rowden, S.J.L., Cramer, W.A., Howe, C.J., and Puthiyaveetil, S. (2022). Thiol redox switches regulate the oligomeric state of cyanobacterial Rre1, RpaA and RpaB response regulators. FEBS Lett *596*, 1533-1543.

Ibrahim, M.H., Jaafar, H.Z., Rahmat, A., and Rahman, Z.A. (2010). The relationship between phenolics and flavonoids production with total non structural carbohydrate and photosynthetic rate in Labisia pumila Benth. under high CO2 and nitrogen fertilization. Molecules *16*, 162-174.

Ihalainen, J.A., D'Haene, S., Yeremenko, N., van Roon, H., Arteni, A.A., Boekema, E.J., van Grondelle, R., Matthijs, H.C., and Dekker, J.P. (2005). Aggregates of the chlorophyll-binding protein IsiA (CP43') dissipate energy in cyanobacteria. Biochemistry *44*, 10846-10853.

Ikeuchi, M., and Ishizuka, T. (2008). Cyanobacteriochromes: a new superfamily of tetrapyrrole-binding photoreceptors in cyanobacteria. Photochem Photobiol Sci *7*, 1159-1167.

Im, Y.J., Rho, S.H., Park, C.M., Yang, S.S., Kang, J.G., Lee, J.Y., Song, P.S., and Eom, S.H. (2002). Crystal structure of a cyanobacterial phytochrome response regulator. Protein Sci *11*, 614-624.

Imamura, S., Asayama, M., Takahashi, H., Tanaka, K., Takahashi, H., and Shirai, M. (2003). Antagonistic dark/light-induced SigB/SigD, group 2 sigma factors, expression through redox potential and their roles in cyanobacteria. FEBS Lett *554*, 357-362.

Ishii, A., and Hihara, Y. (2008). An AbrB-like transcriptional regulator, Sll0822, is essential for the activation of nitrogen-regulated genes in *Synechocystis* sp. PCC 6803. Plant Physiol *148*, 660-670.

Izuhara, T., Kaihatsu, I., Jimbo, H., Takaichi, S., and Nishiyama, Y. (2020). Elevated Levels of Specific Carotenoids During Acclimation to Strong Light Protect the Repair of Photosystem II in *Synechocystis* sp. PCC 6803. Front Plant Sci *11*, 1030.

Jimbo, H., Izuhara, T., Hihara, Y., Hisabori, T., and Nishiyama, Y. (2019). Light-inducible expression of translation factor EF-Tu during acclimation to strong light enhances the repair of photosystem II. Proc Natl Acad Sci U S A *116*, 21268-21273.

Jimbo, H., Yuasa, K., Takagi, K., Hirashima, T., Keta, S., Aichi, M., and Wada, H. (2021). Specific Incorporation of Polyunsaturated Fatty Acids into the sn-2 Position of

Phosphatidylglycerol Accelerates Photodamage to Photosystem II under Strong Light. Int J Mol Sci *22*.

Joseph, A., Aikawa, S., Sasaki, K., Matsuda, F., Hasunuma, T., and Kondo, A. (2014). Increased biomass production and glycogen accumulation in apcE gene deleted *Synechocystis* sp. PCC 6803. AMB Express *4*, 17.

Joshua, S., and Mullineaux, C.W. (2005). The rpaC gene product regulates phycobilisome-photosystem II interaction in cyanobacteria. Biochim Biophys Acta *1709*, 58-68.

Kadowaki, T., Nishiyama, Y., Hisabori, T., and Hihara, Y. (2015). Identification of OmpR-family response regulators interacting with thioredoxin in the Cyanobacterium *Synechocystis* sp. PCC 6803. PLoS One *10*, e0119107.

Kamarainen, J., Huokko, T., Kreula, S., Jones, P.R., Aro, E.M., and Kallio, P. (2017). Pyridine nucleotide transhydrogenase PntAB is essential for optimal growth and photosynthetic integrity under low-light mixotrophic conditions in *Synechocystis* sp. PCC 6803. New Phytol *214*, 194-204.

Kato, H., Kubo, T., Hayashi, M., Kobayashi, I., Yagasaki, T., Chibazakura, T., Watanabe, S., and Yoshikawa, H. (2011). Interactions between histidine kinase NblS and the response regulators RpaB and SrrA are involved in the bleaching process of the cyanobacterium *Synechococcus elongatus* PCC 7942. Plant Cell Physiol *52*, 2115-2122.

Kato, N., Iwata, K., Kadowaki, T., Sonoike, K., and Hihara, Y. (2022). Dual Redox Regulation of the DNA-Binding Activity of the Response Regulator RpaB in the Cyanobacterium *Synechocystis* sp. PCC 6803. Plant Cell Physiol *63*, 1078-1090.

Kaushik, M.S., Srivastava, M., Singh, A., and Mishra, A.K. (2017). Impairment of ntcA gene revealed its role in regulating iron homeostasis, ROS production and cellular phenotype under iron deficiency in cyanobacterium Anabaena sp. PCC 7120. World J Microbiol Biotechnol *33*, 158.

Kehoe, D.M., and Grossman, A.R. (1996). Similarity of a chromatic adaptation sensor to phytochrome and ethylene receptors. Science *273*, 1409-1412.

Kirilovsky, D. (2015). Modulating energy arriving at photochemical reaction centers: orange carotenoid protein-related photoprotection and state transitions. Photosynth Res *126*, 3-17.

Kirst, H., Formighieri, C., and Melis, A. (2014). Maximizing photosynthetic efficiency and culture productivity in cyanobacteria upon minimizing the phycobilisome light-harvesting antenna size. Biochim Biophys Acta *1837*, 1653-1664.

Kobayashi, I., Watanabe, S., Kanesaki, Y., Shimada, T., Yoshikawa, H., and Tanaka, K. (2017). Conserved two-component Hik34-Rre1 module directly activates heat-stress inducible transcription of major chaperone and other genes in *Synechococcus elongatus* PCC 7942. Mol Microbiol *104*, 260-277.

Kojima, K., and Nakamoto, H. (2007). A novel light- and heat-responsive regulation of the groE transcription in the absence of HrcA or CIRCE in cyanobacteria. FEBS Lett *581*, 1871-1880.

Kopf, M., and Hess, W.R. (2015). Regulatory RNAs in photosynthetic cyanobacteria. FEMS Microbiol Rev *39*, 301-315.

Kovac, D., Malenovsky, Z., Urban, O., Spunda, V., Kalina, J., Ac, A., Kaplan, V., and Hanus, J. (2013). Response of green reflectance continuum removal index to the xanthophyll de-epoxidation cycle in Norway spruce needles. J Exp Bot *64*, 1817-1827.

Koyama, K., and Takemoto, S. (2014). Morning reduction of photosynthetic capacity before midday depression. Sci Rep *4*, 4389.

Kratz, W.A., and Myers, J. (1955). Photosynthesis and Respiration of Three Blue-Green Algae. Plant Physiol *30*, 275-280.

Ksas, B., Becuwe, N., Chevalier, A., and Havaux, M. (2015). Plant tolerance to excess light energy and photooxidative damage relies on plastoquinone biosynthesis. Sci Rep *5*, 10919.

Kujirai, J., Nanba, S., Kadowaki, T., Oka, Y., Nishiyama, Y., Hayashi, Y., Arai, M., and Hihara, Y. (2018). Interaction of the GntR-family transcription factor Sll1961 with thioredoxin in the cyanobacterium *Synechocystis* sp. PCC 6803. Sci Rep *8*, 6666.

Kusian, B., and Bowien, B. (1997). Organization and regulation of cbb CO2 assimilation genes in autotrophic bacteria. FEMS Microbiol Rev *21*, 135-155.

Latifi, A., Ruiz, M., and Zhang, C.C. (2009). Oxidative stress in cyanobacteria. FEMS Microbiol Rev *33*, 258-278.

Li, B., Jo, M., Liu, J., Tian, J., Canfield, R., and Bridwell-Rabb, J. (2022). Structural and mechanistic basis for redox sensing by the cyanobacterial transcription regulator RexT. Commun Biol *5*, 275.

Lima-Melo, Y., Kilic, M., Aro, E.M., and Gollan, P.J. (2021). Photosystem I Inhibition, Protection and Signalling: Knowns and Unknowns. Front Plant Sci *12*, 791124.

Lindahl, M., and Florencio, F.J. (2003). Thioredoxin-linked processes in cyanobacteria are as numerous as in chloroplasts, but targets are different. Proc Natl Acad Sci U S A *100*, 16107-16112.

Mackey, S.R., Golden, S.S., and Ditty, J.L. (2011). The itty-bitty time machine genetics of the cyanobacterial circadian clock. Adv Genet *74*, 13-53.

Majeed, W., Zhang, Y., Xue, Y., Ranade, S., Blue, R.N., Wang, Q., and He, Q. (2012). RpaA regulates the accumulation of monomeric photosystem I and PsbA under high light conditions in *Synechocystis* sp. PCC 6803. PLoS One *7*, e45139.

Mallen-Ponce, M.J., Huertas, M.J., Sanchez-Riego, A.M., and Florencio, F.J. (2021). Depletion of m-type thioredoxin impairs photosynthesis, carbon fixation, and oxidative stress in cyanobacteria. Plant Physiol *187*, 1325-1340.

Markson, J.S., Piechura, J.R., Puszynska, A.M., and O'Shea, E.K. (2013). Circadian control of global gene expression by the cyanobacterial master regulator RpaA. Cell *155*, 1396-1408.

Mengel, A., Chaki, M., Shekariesfahlan, A., and Lindermayr, C. (2013). Effect of nitric oxide on gene transcription - S-nitrosylation of nuclear proteins. Front Plant Sci *4*, 293.

Michelet, L., Zaffagnini, M., Morisse, S., Sparla, F., Perez-Perez, M.E., Francia, F., Danon, A., Marchand, C.H., Fermani, S., Trost, P., *et al.* (2013). Redox regulation of the Calvin-Benson cycle: something old, something new. Front Plant Sci *4*, 470.

Muramatsu, M., and Hihara, Y. (2012). Acclimation to high-light conditions in cyanobacteria: from gene expression to physiological responses. J Plant Res *125*, 11-39.

Mustila, H., Muth-Pawlak, D., Aro, E.M., and Allahverdiyeva, Y. (2021). Global proteomic response of unicellular cyanobacterium *Synechocystis* sp. PCC 6803 to fluctuating light upon CO2 step-down. Physiol Plant *173*, 305-320.

Nagarajan, A., Page, L.E., Liberton, M., and Pakrasi, H.B. (2014). Consequences of Decreased Light Harvesting Capability on Photosystem II Function in *Synechocystis* sp. PCC 6803. Life (Basel) *4*, 903-914.

Nakamoto, H., Suzuki, M., and Kojima, K. (2003). Targeted inactivation of the hrcA repressor gene in cyanobacteria. FEBS Lett *549*, 57-62.

Narainsamy, K., Farci, S., Braun, E., Junot, C., Cassier-Chauvat, C., and Chauvat, F. (2016). Oxidative-stress detoxification and signalling in cyanobacteria: the crucial glutathione synthesis pathway supports the production of ergothioneine and ophthalmate. Mol Microbiol *100*, 15-24.

Neunuebel, M.R., and Golden, J.W. (2008). The Anabaena sp. strain PCC 7120 gene all2874 encodes a diguanylate cyclase and is required for normal heterocyst development under high-light growth conditions. J Bacteriol *190*, 6829-6836.

Nishiyama, Y., Allakhverdiev, S.I., and Murata, N. (2005). Inhibition of the repair of photosystem II by oxidative stress in cyanobacteria. Photosynth Res *84*, 1-7.

Nishiyama, Y., Allakhverdiev, S.I., and Murata, N. (2011). Protein synthesis is the primary target of reactive oxygen species in the photoinhibition of photosystem II. Physiol Plant *142*, 35-46.

Nishiyama, Y., and Murata, N. (2014). Revised scheme for the mechanism of photoinhibition and its application to enhance the abiotic stress tolerance of the photosynthetic machinery. Appl Microbiol Biotechnol *98*, 8777-8796.

Ogawa, K., Yoshikawa, K., Matsuda, F., Toya, Y., and Shimizu, H. (2018). Transcriptome analysis of the cyanobacterium *Synechocystis* sp. PCC 6803 and mechanisms of photoinhibition tolerance under extreme high light conditions. J Biosci Bioeng.

Orf, I., Schwarz, D., Kaplan, A., Kopka, J., Hess, W.R., Hagemann, M., and Klahn, S. (2016). CyAbrB2 Contributes to the Transcriptional Regulation of Low CO2 Acclimation in *Synechocystis* sp. PCC 6803. Plant Cell Physiol *57*, 2232-2243.

Pedre, B., Young, D., Charlier, D., Mourenza, A., Rosado, L.A., Marcos-Pascual, L., Wahni, K., Martens, E., A, G.d.l.R., Belousov, V.V., *et al.* (2018). Structural snapshots of OxyR reveal the peroxidatic mechanism of H2O2 sensing. Proc Natl Acad Sci U S A *115*, E11623-E11632.

Peltier, G., Aro, E.M., and Shikanai, T. (2016). NDH-1 and NDH-2 Plastoquinone Reductases in Oxygenic Photosynthesis. Annu Rev Plant Biol *67*, 55-80.

Perez-Perez, M.E., Mata-Cabana, A., Sanchez-Riego, A.M., Lindahl, M., and Florencio, F.J. (2009). A comprehensive analysis of the peroxiredoxin reduction system in the Cyanobacterium *Synechocystis* sp. strain PCC 6803 reveals that all five peroxiredoxins are thioredoxin dependent. J Bacteriol *191*, 7477-7489.

Piechura, J.R., Amarnath, K., and O'Shea, E.K. (2017). Natural changes in light interact with circadian regulation at promoters to control gene expression in cyanobacteria. Elife *6*.

Pollari, M., Ruotsalainen, V., Rantamaki, S., Tyystjarvi, E., and Tyystjarvi, T. (2009). Simultaneous inactivation of sigma factors B and D interferes with light acclimation of the cyanobacterium *Synechocystis* sp. strain PCC 6803. J Bacteriol *191*, 3992-4001.

Riediger, M., Kadowaki, T., Nagayama, R., Georg, J., Hihara, Y., and Hess, W.R. (2019). Biocomputational Analyses and Experimental Validation Identify the Regulon Controlled by the Redox-Responsive Transcription Factor RpaB. iScience *15*, 316-331.

Riley, K.W., Gonzalez, A., and Risser, D.D. (2018). A partner-switching regulatory system controls hormogonium development in the filamentous cyanobacterium *Nostoc punctiforme*. Mol Microbiol *109*, 555-569.

Robles-Rengel, R., Florencio, F.J., and Muro-Pastor, M.I. (2019). Redox interference in nitrogen status via oxidative stress is mediated by 2-oxoglutarate in cyanobacteria. New Phytol *224*, 216-228.

Saito, M., Watanabe, S., Nimura-Matsune, K., Yoshikawa, H., and Nakamoto, H. (2020). Regulation of the groESL1 transcription by the HrcA repressor and a novel transcription factor Orf7.5 in the cyanobacterium *Synechococcus elongatus* PCC7942. J Gen Appl Microbiol *66*, 85-92.

Salinas, P., Ruiz, D., Cantos, R., Lopez-Redondo, M.L., Marina, A., and Contreras, A. (2007). The regulatory factor SipA provides a link between NblS and NblR signal transduction

pathways in the cyanobacterium *Synechococcus* sp. PCC 7942. Mol Microbiol *66*, 1607-1619.

Santana-Sanchez, A., Solymosi, D., Mustila, H., Bersanini, L., Aro, E.M., and Allahverdiyeva, Y. (2019). Flavodiiron proteins 1-to-4 function in versatile combinations in O2 photoreduction in cyanobacteria. Elife *8*.

Schurmann, P., and Buchanan, B.B. (2008). The ferredoxin/thioredoxin system of oxygenic photosynthesis. Antioxid Redox Signal *10*, 1235-1274.

Seki, A., Hanaoka, M., Akimoto, Y., Masuda, S., Iwasaki, H., and Tanaka, K. (2007). Induction of a group 2 sigma factor, RPOD3, by high light and the underlying mechanism in *Synechococcus elongatus* PCC 7942. J Biol Chem *282*, 36887-36894.

Shukla, M.K., Llansola-Portoles, M.J., Tichy, M., Pascal, A.A., Robert, B., and Sobotka, R. (2018). Binding of pigments to the cyanobacterial high-light-inducible protein HliC. Photosynth Res *137*, 29-39.

Singh, A.K., Elvitigala, T., Bhattacharyya-Pakrasi, M., Aurora, R., Ghosh, B., and Pakrasi, H.B. (2008). Integration of carbon and nitrogen metabolism with energy production is crucial to light acclimation in the cyanobacterium *Synechocystis*. Plant Physiol *148*, 467-478.

Sluchanko, N.N., Slonimskiy, Y.B., Shirshin, E.A., Moldenhauer, M., Friedrich, T., and Maksimov, E.G. (2018). OCP-FRP protein complex topologies suggest a mechanism for controlling high light tolerance in cyanobacteria. Nat Commun *9*, 3869.

Stoitchkova, K., Zsiros, O., Javorfi, T., Pali, T., Andreeva, A., Gombos, Z., and Garab, G. (2007). Heat- and light-induced reorganizations in the phycobilisome antenna of *Synechocystis* sp. PCC 6803. Thermo-optic effect. Biochim Biophys Acta *1767*, 750-756.

Takatani, N., Use, K., Kato, A., Ikeda, K., Kojima, K., Aichi, M., Maeda, S., and Omata, T. (2015). Essential Role of Acyl-ACP Synthetase in Acclimation of the Cyanobacterium *Synechococcus elongatus* Strain PCC 7942 to High-Light Conditions. Plant Cell Physiol *56*, 1608-1615.

Takeda, K., Nishiyama, Y., Yoda, K., Watanabe, T., Nimura-Matsune, K., Mura, K., Tokue, C., Katoh, T., Kawasaki, S., and Niimura, Y. (2004). Distribution of Prx-linked hydroperoxide reductase activity among microorganisms. Biosci Biotechnol Biochem *68*, 20-27.

Tamary, E., Kiss, V., Nevo, R., Adam, Z., Bernat, G., Rexroth, S., Rogner, M., and Reich, Z. (2012). Structural and functional alterations of cyanobacterial phycobilisomes induced by high-light stress. Biochim Biophys Acta *1817*, 319-327.

Tan, X., Hou, S., Song, K., Georg, J., Klahn, S., Lu, X., and Hess, W.R. (2018). The primary transcriptome of the fast-growing cyanobacterium *Synechococcus elongatus* UTEX 2973. Biotechnol Biofuels *11*, 218.

Terauchi, K., Montgomery, B.L., Grossman, A.R., Lagarias, J.C., and Kehoe, D.M. (2004). RcaE is a complementary chromatic adaptation photoreceptor required for green and red light responsiveness. Mol Microbiol *51*, 567-577.

Tolle, J., Michel, K.P., Kruip, J., Kahmann, U., Preisfeld, A., and Pistorius, E.K. (2002). Localization and function of the IdiA homologue Slr1295 in the cyanobacterium *Synechocystis* sp. strain PCC 6803. Microbiology (Reading) *148*, 3293-3305.

Tuominen, I., Pollari, M., Aguirre von Wobeser, E., Tyystjarvi, E., Ibelings, B.W., Matthijs, H.C., and Tyystjarvi, T. (2008). Sigma factor SigC is required for heat acclimation of the cyanobacterium *Synechocystis* sp. strain PCC 6803. FEBS Lett *582*, 346-350.

Tuominen, I., Pollari, M., Tyystjarvi, E., and Tyystjarvi, T. (2006). The SigB sigma factor mediates high-temperature responses in the cyanobacterium *Synechocystis* sp. PCC6803. FEBS Lett *580*, 319-323.

Turunen, O., Koskinen, S., Kurkela, J., Karhuvaara, O., Hakkila, K., and Tyystjarvi, T. (2022). Roles of Close Homologues SigB and SigD in Heat and High Light Acclimation of the Cyanobacterium *Synechocystis* sp. PCC 6803. Life (Basel) *12*.

Tyystjarvi, E., and Aro, E.M. (1996). The rate constant of photoinhibition, measured in lincomycin-treated leaves, is directly proportional to light intensity. Proc Natl Acad Sci U S A *93*, 2213-2218.

Ungerer, J., Lin, P.C., Chen, H.Y., and Pakrasi, H.B. (2018a). Adjustments to Photosystem Stoichiometry and Electron Transfer Proteins Are Key to the Remarkably Fast Growth of the Cyanobacterium *Synechococcus elongatus* UTEX 2973. MBio *9*.

Ungerer, J., Wendt, K.E., Hendry, J.I., Maranas, C.D., and Pakrasi, H.B. (2018b). Comparative genomics reveals the molecular determinants of rapid growth of the cyanobacterium *Synechococcus elongatus* UTEX 2973. Proc Natl Acad Sci U S A.

Walker, P.L., and Pakrasi, H.B. (2022). A Ubiquitously Conserved Cyanobacterial Protein Phosphatase Essential for High Light Tolerance in a Fast-Growing Cyanobacterium. Microbiol Spectr *10*, e0100822.

Watanabe, S., Kobayashi, T., Saito, M., Sato, M., Nimura-Matsune, K., Chibazakura, T., Taketani, S., Nakamoto, H., and Yoshikawa, H. (2007). Studies on the role of HtpG in the tetrapyrrole biosynthesis pathway of the cyanobacterium *Synechococcus elongatus* PCC 7942. Biochem Biophys Res Commun *352*, 36-41.

Weisz, D.A., Gross, M.L., and Pakrasi, H.B. (2016). The Use of Advanced Mass Spectrometry to Dissect the Life-Cycle of Photosystem II. Front Plant Sci *7*, 617.

Wilde, A., and Hihara, Y. (2016). Transcriptional and posttranscriptional regulation of cyanobacterial photosynthesis. Biochim Biophys Acta *1857*, 296-308.

Wiltbank, L.B., and Kehoe, D.M. (2019). Diverse light responses of cyanobacteria mediated by phytochrome superfamily photoreceptors. Nat Rev Microbiol *17*, 37-50.

Woodger, F.J., Bryant, D.A., and Price, G.D. (2007). Transcriptional regulation of the CO2-concentrating mechanism in a euryhaline, coastal marine cyanobacterium, *Synechococcus* sp. Strain PCC 7002: role of NdhR/CcmR. J Bacteriol *189*, 3335-3347.

Yamamoto, Y. (2016). Quality Control of Photosystem II: The Mechanisms for Avoidance and Tolerance of Light and Heat Stresses are Closely Linked to Membrane Fluidity of the Thylakoids. Front Plant Sci *7*, 1136.

Yang, H., Liao, L., Bo, T., Zhao, L., Sun, X., Lu, X., Norling, B., and Huang, F. (2014). Slr0151 in *Synechocystis* sp. PCC 6803 is required for efficient repair of photosystem II under high-light condition. J Integr Plant Biol *56*, 1136-1150.

Yasuda, A., Inami, D., and Hanaoka, M. (2020). RpaB, an essential response regulator for high-light stress, is extensively involved in transcriptional regulation under light-intensity upshift conditions in *Synechococcus elongatus* PCC 7942. J Gen Appl Microbiol *66*, 73-79.

Yoshikawa, K., Ogawa, K., Toya, Y., Akimoto, S., Matsuda, F., and Shimizu, H. (2021). Mutations in hik26 and slr1916 lead to high-light stress tolerance in *Synechocystis* sp. PCC6803. Commun Biol *4*, 343.

Yu, J., Liberton, M., Cliften, P.F., Head, R.D., Jacobs, J.M., Smith, R.D., Koppenaal, D.W., Brand, J.J., and Pakrasi, H.B. (2015). *Synechococcus elongatus* UTEX 2973, a fast growing cyanobacterial chassis for biosynthesis using light and CO(2). Sci Rep *5*, 8132.

Zeth, K., Fokina, O., and Forchhammer, K. (2014). Structural basis and target-specific modulation of ADP sensing by the *Synechococcus elongatus* PII signaling protein. J Biol Chem *289*, 8960-8972.

Zhan, J., Steglich, C., Scholz, I., Hess, W.R., and Kirilovsky, D. (2021). Inverse regulation of light harvesting and photoprotection is mediated by a 3'-end-derived sRNA in cyanobacteria. Plant Cell *33*, 358-380.

Zhang, H., Ge, H., Zhang, Y., Wang, Y., and Zhang, P. (2021). Slr0320 Is Crucial for Optimal Function of Photosystem II during High Light Acclimation in *Synechocystis* sp. PCC 6803. Life (Basel) *11*.

Zhang, P., Eisenhut, M., Brandt, A.M., Carmel, D., Silen, H.M., Vass, I., Allahverdiyeva, Y., Salminen, T.A., and Aro, E.M. (2012). Operon flv4-flv2 provides cyanobacterial photosystem II with flexibility of electron transfer. Plant Cell *24*, 1952-1971.

Chapter 2: *Synechococcus elongatus* genotypes differentially regulate C/N balance and circadian gene expression during high-light acclimation

Chapter contributions: All data, tables, and figures in this Chapter were generated by PLW

2.1 Summary

Two strains of the freshwater cyanobacterium *Synechococcus elongatus*, PCC 7942 (*S.* 7942) and UTEX 2973 (*S.* 2973), have emerged as model systems due to their intricate circadian clock biology and the availability of genetic tools. Additionally, *S.* 2973 has recently grown in popularity specifically for its robust growth in high temperature and light. We set out to identify the differences in global cellular regulation between these strains that account for the dramatic difference in high-light acclimation. Using RNA sequencing, we compared the global transcriptomes of closely related strains *S.* 7942 and *S.* 2973 under high-light conditions. Our data provide insight into specialized high-light adaptations of *S.* 2973 that allow it to thrive under extreme high-light intensity. We observed differences in the expression of genes related to nitrogen assimilation, ion transport, amino acid synthesis, photosynthesis, and oxidation reduction between the two strains. The RpaA protein, the master regulator of the circadian clock alleles, contains differences in both the coding and promoter sequence regions. Analysis revealed that the RpaA allele of *S.* 2973 is downregulated under high light in *S.* 2973, unlike in *S.* 7942, leading to large changes in the expression of circadian genes and global metabolism. We placed these differing processes into a cellular context to understand the driving factors in the differential high-light tolerance of *S.* 7942 and *S.* 2973.

2.2 Introduction

The two cyanobacterial strains *Synechococcus elongatus* UTEX 2973 and *Synechococcus elongatus* PCC 7942 are genetically similar, sharing an average nucleotide identity of 99.9%, but have notable differing phenotypes. These strains were first reported to contain a total of 55 genetic differences but were later found to have a total 53 SNPs and indels (Yu et al., 2015b, Adomako et al., 2022). The high degree of homology of these strains provides a platform for studying the genetic factors responsible for the phenotypic differences. Previous studies found *S.* 2973 grows faster at higher light intensities (Ungerer et al., 2018c, Yu et al., 2015b) and displays higher photosynthetic rates (Ungerer et al., 2018b). The faster growth rate has been attributed to alleles of three genes, ATP synthase, NAD+ kinase, and RpaA, the master regulator of the circadian clock (Ungerer et al., 2018c). The increase in photosynthetic efficiency was attributed to the increase in photosystem I (PSI), cytochrome b_6f, and plastocyanin content in *S.* 2973 (Ungerer et al., 2018b). This was corroborated by a high resolution atomic force microscopy study of *S.* 2973, which found an enrichment of PSI complexes and type-1 NADH dehydrogenase-like complexes in high light-adapted thylakoid membranes (Zhao et al., 2022).

Cyanobacteria have a circadian rhythm that anticipates light and dark and regulates gene expression in an oscillatory fashion (Ito et al., 2009, Hosokawa et al., 2011). RpaA is known to connect environmental responses to physiology by controlling the global oscillation of gene expression (Ito et al., 2009). Additionally, fluctuating light has been shown to alter the expression of most circadian genes (Piechura et al., 2017). The *S.* 7942 RpaA allele is important for natural transformation and rapid growth (Wendt et al., 2021, Ungerer et al., 2018c). The effect of RpaA alleles on high-light tolerance or regulation of circadian genes has not yet been explored. RpaA regulates a large set of genes including interdependent sigma factors that

feedback to the Kai clock, and therefore we predicted that the RpaA allele may potentially lead to global transcriptional differences. Although changes in RpaA likely affect regulation of photosynthetic and carbon metabolism genes, how those processes are differentially regulated in response to increased light makes the high-light tolerant *S.* 2973 an intriguing area of investigation.

Transcriptomic studies of *Synechococcus elongatus* have previously investigated dark adaptation, heat response, circadian oscillations, nitrogen starvation, and fluctuating light (Ito et al., 2009, Piechura et al., 2017, Hosokawa et al., 2011, Tan et al., 2018, Choi et al., 2016). To our knowledge, only one study has examined the transcriptome response of *S.* 2973 to high light (1000 μmol photons m^{-2} s^{-1}) (Tan et al., 2018). Due to the nature of the primary study, they did not utilize the optimal light condition for fast-growth (l500-2500 μmol photons m^{-2} s^{-1}) or perform a time-series analysis. Our study is unique in that we directly compare *S.* 2973 and *S.* 7942 during early response and acclimation to high light. Previous studies comparing *S.* 7942 and *S.* 2973 have typically constructed mutants with individual allele reversions to identify the minimal changes for certain phenotypes (Walker and Pakrasi, 2022, Ungerer et al., 2018c, Wendt et al., 2021). In this study, we consider all the genome variance through comparative global transcriptomics to take a holistic look at the differences in high-light response and determine key factors that lead to distinctive high-light growth phenotypes.

Following high-light exposure, both strains experienced an immediate drop in photosynthetic efficiency, although the photosynthetic efficiency of *S.* 7942 failed to recover during prolonged exposure to high-light intensity. Our transcriptome analysis aligned with previous reports on light stress response but also differed in the activity of many of the circadian genes known to peak at subjective dawn. This difference in regulation of circadian genes

including interconnected regulators could have led to higher carbon fixation, and less reactive chemical species generation. Altogether, these results suggest that *S.* 2973 undergoes broad transcriptome changes initiated by the RpaA allele to acclimate to extreme high-light conditions.

2.3 Results

2.3.1 Changes in photosynthetic efficiency from high light in *Synechococcus elongatus* strains

To directly compare the high light (HL) transcriptome response in *S.* 7942 and *S.* 2973, it was necessary to apply the same experimental conditions for both strains. To determine key times in the different HL responses of *S.* 7942 and *S.* 2973, growth rate and PSII efficiency were measured throughout growth under 1500 μmol photons m^{-2} s^{-1} light at 38°C in BG11 media. We found that the maximum photochemical efficiency of PSII (Fv/Fm) decreased dramatically over the first hour of HL exposure in both strains. Photochemical efficiency gradually recovered to pre-HL levels after two hours in *S.* 2973, but never fully recovered in *S.* 7942 (Fig. 2.1a). We chose a HL, 1500 μmol photons m^{-2} s^{-1}, intensity that allowed for optimal growth of *S.* 2973 without preventing the growth of *S.* 7942. Using chloramphenicol, an inhibitor of protein synthesis, to examine PSII photoinhibition, we observed a negligible rate of photodamage to PSII in the lower light intensity of 200 μmol photons m^{-2} s^{-1} that still allowed for optimal growth of *S.* 7942 (Fig. 2.1c and 2.1f). The inhibition of D1 protein synthesis experiments showed that *S.* 7942 underwent more damage to PSII and relied more heavily on D1 synthesis to recover. Taken together, this indicated that the key differences in the HL tolerance of *S.* 2973 occurred in the early times after HL intensity increase, and that HL inhibited the photosynthetic activity of *S.* 7942 to a greater and longer-lasting effect than *S.* 2973. Therefore, we isolated RNA from cells

during exposure to HL after 30 min, to represent the initial stress phase, and 4 h, to represent the acclimated phase.

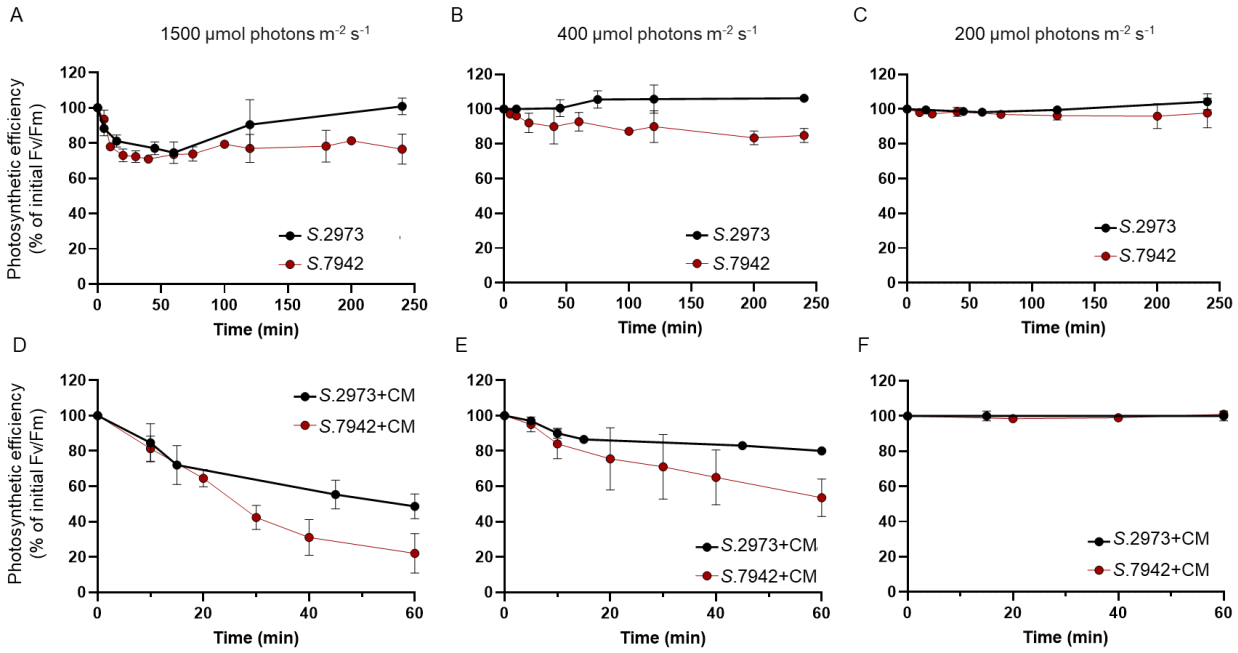

Figure 2.1: Photosynthetic efficiency of *S. elongatus* strains under high-, medium-, and low-light conditions (d-f) with and (a-c) without chloramphenicol (CM). The Fv/Fm ratios are expressed as a percentage of the control (measurements before applying light treatment). (d-f) CM was added to inhibit PSII repair through D1 protein synthesis, and a decrease in the Fv/Fm ratio indicates damage to PSII. Values represent means +/- SD of results from three independent experiments.

2.3.2 *S.* 2973 undergoes greater transcriptome changes during HL response

To determine the effect of the 53 genomic differences on the transcriptome during HL acclimation, we compared genome-wide gene expression at 30 min and 4 h after HL treatment. Significantly differentially expressed genes (DEG) were determined using a $\log_2$FoldChange ($\log_2$FC) > 2, and an adjusted p-value < 0.05 as selection criteria. Differential expression (DE) analysis revealed a larger percentage of the *S.* 2973 genome underwent transcriptional changes compared to *S.* 7942 in response to HL across all time points (Fig. 2.2). This difference was particularly strong at the 30 min time point in which 250 genes (9.4%) of the *S.* 7942 genome

was DE compared to 780 genes (29.3%) of the *S.* 2973 genome (Fig. 2.2). 70% (114 genes) of the DEGs from *S.* 7942 were also DE in *S.* 2973 and 31% (608 genes) of DEGs from *S.* 2973 were DE in *S.* 7492 (Fig. 2.2). This data indicated that the *S.* 2973 transcriptome undergoes fast and large changes in response light upshift, suggesting that the genetic background of *S.* 2973 influences the light response expression dynamics of about 20% of the genome.

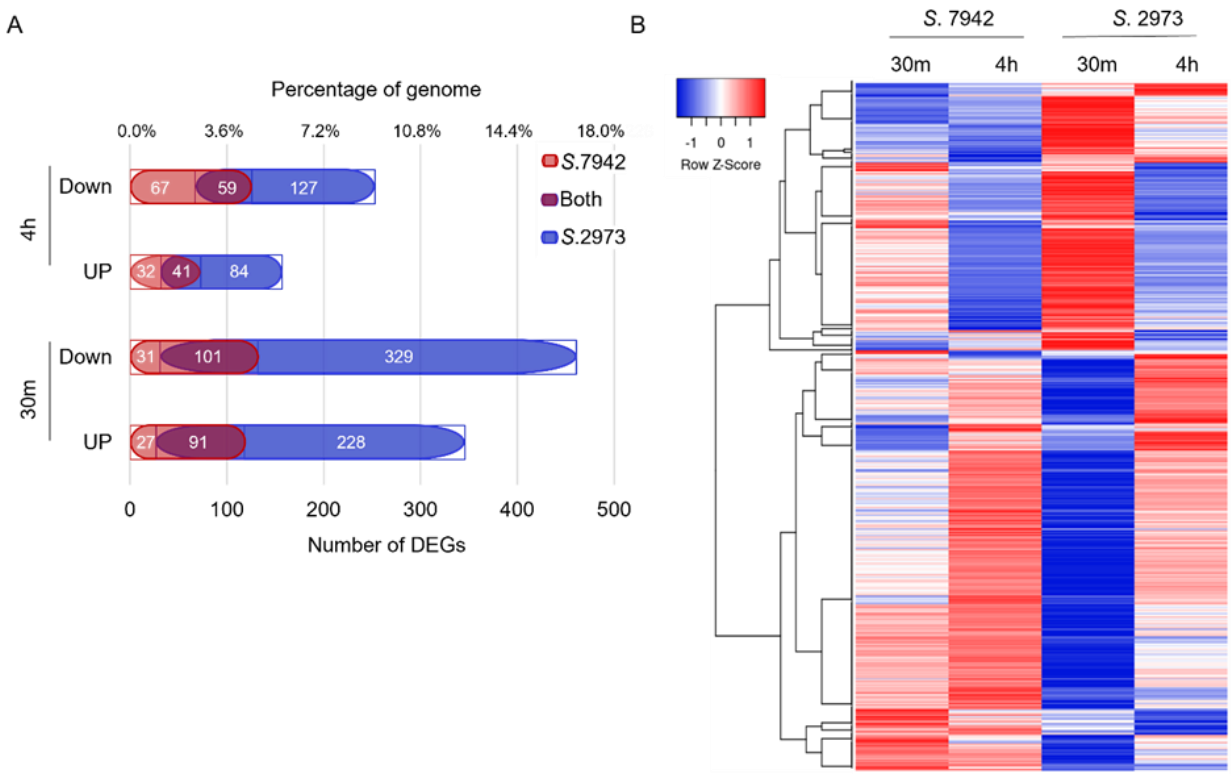

Figure 2.2: Up- and down-regulated DEGs ($log_2FC > 2$; $P<0.05$) at 30 min and 4 h. (a) Venn diagrams shown inside bar plots representing the number of DEGs of each strain. DEGs unique to *S.* 7942 are colored red; DEGs unique to *S.* 2973 are colored blue, and genes that were found to be DE in both strains are shown in purple. (b) Heat map and dendrogram of the DEG. Constructed using Spearman rank correlation with average linkage clustering.

Half of the 33 most differentially regulated genes (Table 1.1) have been reported in previous studies as part of the general HL response, including *hliA, sigA3, irpA-B, isiA-B, cpcC1-2,* and *psaB* (Yasuda et al., 2020, Tan et al., 2018). The putative Mg^{2+} transporter *srpB* was only upregulated in *S.* 2973, suggesting increased Mg^{2+} uptake may be uniquely important

for *S.* 2973. Overall, we find more overlap between the strains in the top significantly upregulated genes compared to downregulated genes, and the downregulation of unique genes may be important for HL adaptation of *S.* 2973 (Fig. 2.2a).

Table 1.1: The top ten differentially up and down regulated genes. Gene product and their log_2FC value, totaling 33 genes. The top five up and down log_2FC values for each category are in bold.

Locus	Name	Product	7942:30m	7942: 4h	2973: 30m	2973: 4h
Synpcc7942_1050	*cpcC2*	phycobilisome rod linker polypeptide	**-3.45**		**-7.38**	
Synpcc7942_1049	*cpcC1*	phycobilisome rod linker polypeptide	-3.04		**-6.84**	
Synpcc7942_1051	*cpcD*	phycocyanin linker protein 9K	**-3.71**		**-6.76**	
Synpcc7942_0551		conserved hypothetical protein	-1.83	-1.35	**-5.98**	-2.56
Synpcc7942_1048	*cpcA1*	phycocyanin, alpha subunit	-2.60		**-4.82**	
Synpcc7942_2048	*psaB*	photosystem I core protein	**-3.28**		-4.58	
Synpcc7942_0418		conserved hypothetical protein	**-3.61**		-4.09	-1.12
Synpcc7942_1656	*katG*	catalase/peroxidase monomer	-1.41		-3.96	**-3.28**
Synpcc7942_B2638	*anL11*	conserved hypothetical protein			-3.68	**-4.18**
Synpcc7942_1655		hypothetical protein	-1.59	**-3.56**	-3.63	**-4.50**
Synpcc7942_0417		AAA -type ATPase, SpoVK/Ycf46/Vps4 family	**-3.22**		-3.49	-1.15
Synpcc7942_RS00440		hypothetical protein			-3.21	**-3.13**
Synpcc7942_1570	*pacS*	cation translocating P-type ATPase	-1.77	-2.20	-2.83	**-3.16**
Synpcc7942_R0049	*tRNA-Lys1*	tRNA-Lys		**-4.47**	1.11	
Synpcc7942_1389	*psbA2*	photosystem II D1 protein	**2.56**	2.31	1.13	1.92
Synpcc7942_B2646	*anL22*	two-component sensor histidine kinase		**-4.22**	1.57	
Synpcc7942_0672	*sigA3*	group2 RNA polymerase sigma factor SigD	**3.51**	2.92	2.02	2.74
Synpcc7942_1997	*hliA*	high light-inducible protein	**3.32**		3.78	1.58
Synpcc7942_0243	*hliC*	high light inducible polypeptide	**3.79**	1.45	**4.22**	2.18
Synpcc7942_B2621	*srpB*	MgtC/SapB family protein			**4.58**	4.03
Synpcc7942_1488	*cmpA*	bicarbonate-binding protein			**4.69**	
Synpcc7942_B2657	*anL03*	hypothetical protein			**4.89**	
Synpcc7942_1120		high light inducible protein	**5.33**	1.60	**5.96**	1.99
Synpcc7942_0834		Alpha/beta hydrolase		**-3.48**		
Synpcc7942_B2645	*anL21*	conserved hypothetical protein		**-5.02**		1.22
Synpcc7942_1939	*gap3*	glyceraldehyde-3-phosphate dehydrogenase	1.32	**3.61**		1.85
Synpcc7942_2173	*idiC*	Iron deficiency-induced protein C		1.99		3.71
Synpcc7942_2175	*idiA*	Fe3+ ABC transporter substrate-binding protein		**4.55**		5.39
Synpcc7942_1461	*irpB*	thiol oxidoreductase-like		**4.31**		**6.87**
Synpcc7942_1462	*irpA*	iron regulated protein A		**4.49**		**7.00**
Synpcc7942_1542	*isiA*	iron-stress chlorophyll-binding protein		2.88		**7.65**
Synpcc7942_1541	*isiB*	Flavodoxin, long chain		2.80		**7.81**
Synpcc7942_1607	*somA(2)*	probable porin		**6.43**		**8.80**

2.3.3 Gene Ontology (GO) Enrichment of the DEGs

GO annotation and pathway enrichment analysis was carried out to classify functions of DEGs and categorized as either early (30 min) or late (4 hr), depending on the time of DE, as

well as a few sustained (30 min and 4 hr) terms. As expected, most GO-terms were overrepresented in both strains including, several key mechanism terms related to the central dogma, photosynthesis, and macromolecule biosynthesis (Hood et al., 2016) (Fig. 2.3). The early downregulated GO terms common in both strains included terms related to light absorption such as photosynthesis and protein-chromophore linkage. The analysis of the GO terms revealed a few important gene categories in *S.* 2973 and *S.* 7942 during HL acclimation. Nitrogen cycle metabolic process and cellular nitrogen compound metabolic process terms were only overrepresented in early response of *S.* 2973. Generation of precursor energy and metabolites, and tetrapyrrole binding were downregulated early in both strains but upregulated late in *S.* 2973 (Fig. 2.3).

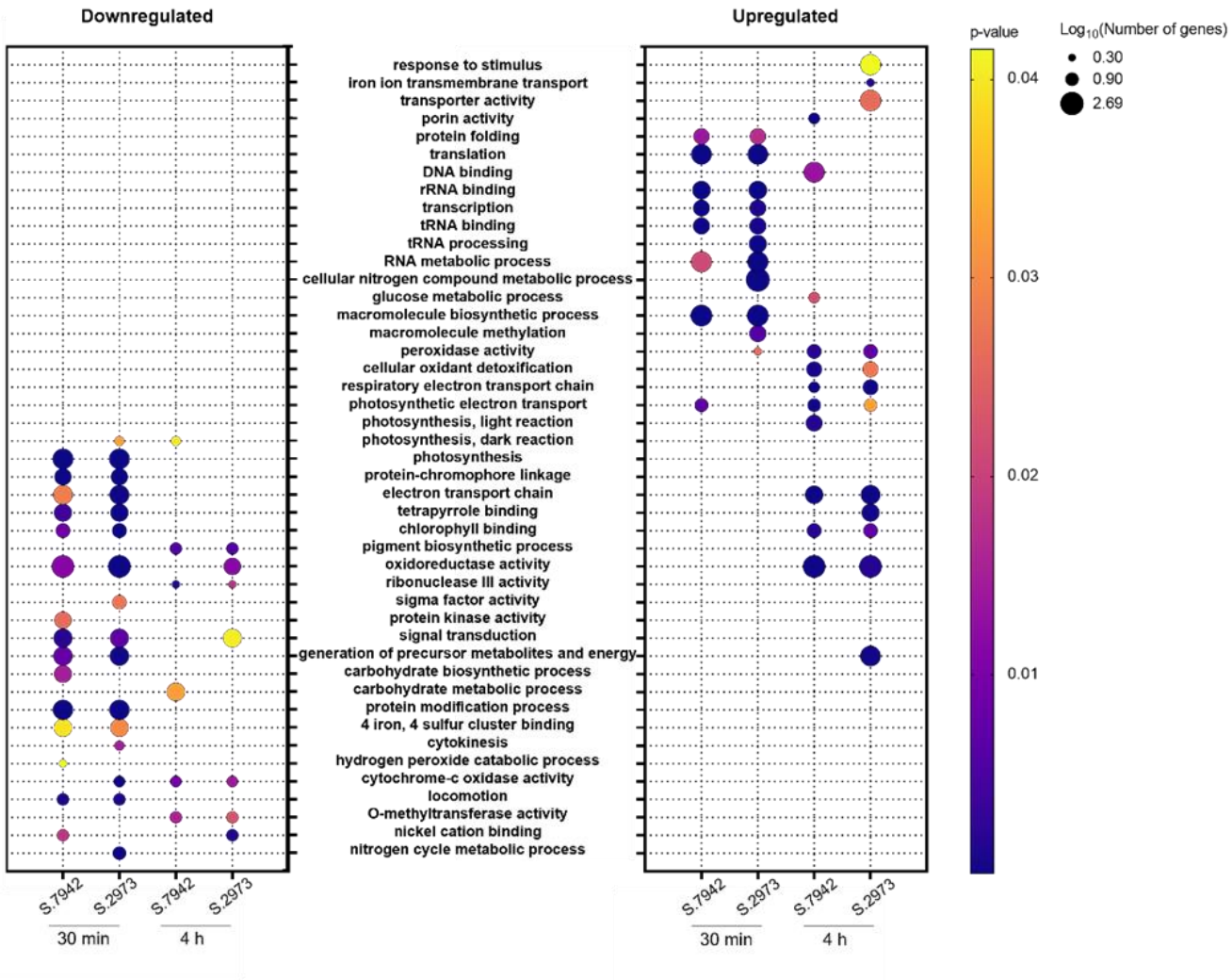

Figure 2.3: Over-represented GO terms (molecular function and biological processes) of DEGs. P-value represents the significance of term enrichment. Log_{10} (number of genes) is the log_{10} of the number of genes that belong to the corresponding GO term. Terms used: nitrogen cycle metabolic process (GO:0071941), nickel cation binding (GO:0016151), O-methyltransferase activity (GO:0008171), locomotion (GO:0040011), cytochrome-c oxidase activity (GO:0004129), hydrogen peroxide catabolic process (GO:0042744), cytokinesis (GO:0000910), 4 iron, 4 sulfur cluster binding (GO:0051539), protein modification process (GO:0036211), carbohydrate metabolic process (GO:0005975), generation of precursor metabolites and energy (GO:0006091), signal transduction (GO:0007165), protein kinase activity (GO:0004672), sigma factor activity (GO:0016987), ribonuclease III activity (GO:0004525), oxidoreductase activity (GO:0016491), pigment biosynthetic process (GO:0046148), chlorophyll binding (GO:0016168), tetrapyrrole binding (GO:0046906), electron transport chain (GO:0022900), protein-chromophore linkage (GO:0018298), photosynthesis (GO:0015979), photosynthesis, dark reaction (GO:0019685), photosynthesis, light reaction (GO:0019684), photosynthetic electron transport (GO:0009767), respiratory electron transport chain (GO:0022904), cellular oxidant detoxification (GO:0098869), peroxidase activity (GO:0004601), macromolecule methylation (GO:0043414), macromolecule biosynthetic process (GO:0009059), glucose metabolic process (GO:0006006), cellular nitrogen compound metabolic

process (GO:0034641), RNA metabolic process (GO:0016070), tRNA processing (GO:0008033), tRNA binding (GO:0000049), transcription (GO:0006351), rRNA binding (GO:0019843), DNA binding (GO:0003677), translation (GO:0006412), protein folding (GO:0006457), porin activity (GO:0015288), transporter activity (GO:0005215), iron ion transmembrane transport (GO:0034755), response to stimulus (GO:0050896), negative regulation of circadian rhythm (GO:0042754).

2.3.4 Differences in fold-changes of photosynthetic component genes expression

Despite the high photosynthetic rate of *S.* 2973 attributed to high levels of PSI and photosynthetic electron carriers (Ungerer et al., 2018c), we found that during HL acclimation most of the PSII, PBS, and PSI components were down-regulated three-fold more in *S.* 2973 than in *S.* 7942 (Fig. 2.3, Fig. 2.4). As expected, phycobiliprotein genes were in the top-most downregulated genes for both strains (Table 1.1). Phycobilisome (PBS) linker genes *apcE* and *cpcG* were downregulated early solely in *S.* 2973 and the timing of downregulation of the phycocyanin and phycocyanin linker proteins (*cpcA, cpcB1, cpcC*) differed (Fig. 2.4). We observed greater early downregulation of PSII in *S.* 2973, which included the SNP encoding the *psbD2* gene. The expression of most cytochrome genes varied considerably between the strains, the majority of which were downregulated around four-fold in *S.* 2973 (Fig. 2.4). As expected, *psbA1,* which encodes the PSII D1 protein, was downregulated in both strains, although forms *psbA2* and *psbA3* were two of the few genes which were upregulated to a greater extent in *S.* 7942 (Table 1.1). This higher demand for D1 synthesis in *S.* 7942 aligned with our data that suggested a greater dependence on D1 turnover in *S.* 7942 in HL conditions (Fig. 2.1b).

2.3.5 Key differences in carbon/nitrogen balance genes

S. 2973 was enriched for multiple terms relating to nitrogen and *S.* 7942 was enriched for multiple terms relating to carbohydrate metabolism (Fig. 2.3), indicating that there may be a difference in the carbon/nitrogen (C/N) balance regulation. The PII protein senses signals from

the central carbon metabolism and regulates nitrogen transporters by acting on the transcription factors NtcA, PipX, and PlmA (Forcada-Nadal et al., 2018, Watzer et al., 2019). We observed significant differences in the expression of carbon and nitrogen transporters. Bicarbonate transporters *sbtA* and *cmpA-D* were highly upregulated in *S.* 2973, while the PII-regulated NO_3-transporters *nrtA-D* and *amt1* were downregulated early in *S.* 2973 (Fig. 2.4). Additionally, nitrate transport/cyanate degradation genes *cynA-B* were expressed in opposing fashions between the strains (Fig. 2.4). Downregulation of nitrogen transport genes in *S.* 2973 may prevent accumulation of toxic nitric oxide (NO) species or be due to a reduced ATP supply hindering the function of ABC-type transporters (Verma and Prasad, 2021). We observed that three SNP-containing genes (*gss, trpE, glsF*) and DE genes (*glnA, nirA, gshB, gst3, chaC, gpx1*) involving glutamate and glutathione metabolism (Fig. 2.4). Glutamate synthesis incorporates 2-OG, acts as an indicator to PII of the carbon/nitrogen balance (Watzer et al., 2019). The differing activity of the polymorphic proteins may lead to divergent accumulations of 2-OG, which triggers different C/N balance responses.

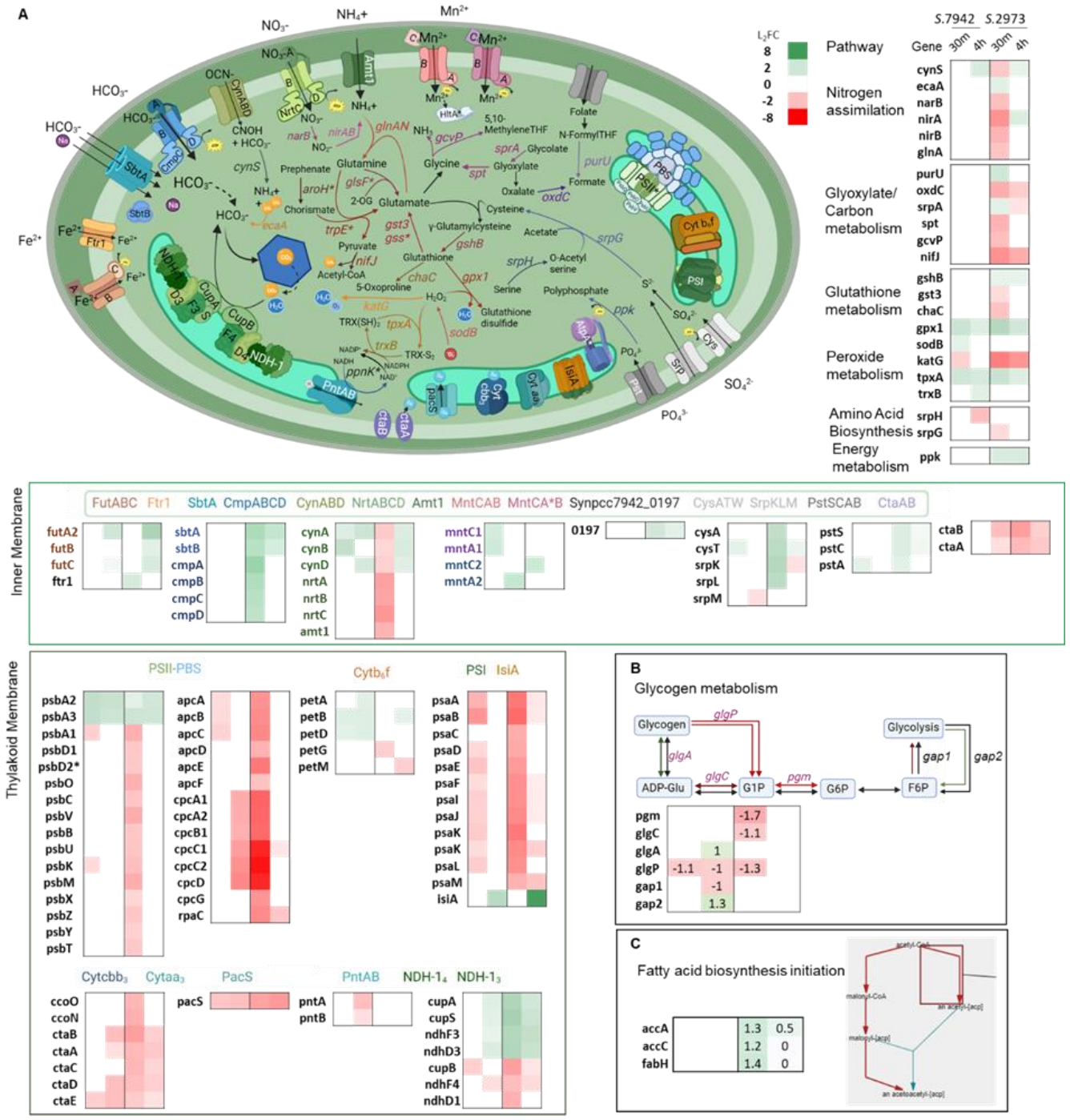

Figure 2.4: Schematic representation of differentially regulated transporters and metabolism with respective log$_2$FC comparison. (a) MntCAB and FutABC transport Mn^{2+} and Fe (Sharon et al., 2014), respectively, from the periplasm into the cytoplasm, indicated by black arrows. Bicarbonate transporters SbtA and CmpCAB are shown in blue. Translocation of proton across the thylakoid membrane and simultaneous reduction of $NADP^+$ by PntAB shown in blue (Kamarainen et al., 2017)Gene expressions for the corresponding protein complexes represented in the inner and thylakoid membranes are shown below the cell, with the order of the complex names respective to the order of the complexes clockwise around the cell. Pathways shown in the

cytoplasm correspond with genes to the right of the cell, span nitrogen assimilation to amino acid metabolism, and carbon metabolism. The asterisks indicate polymorphic genes. TFH: tetrahydrofolate. Arrows and associated enzymes are color coded. (b) Glycogen metabolism pathway comparison in the strains. Green arrow indicates upregulation, red indicates downregulation. The left or bottom arrow represents expression of gene in *S.* 7942 and right or top arrow represents expression of gene in *S.* 2973.

2.3.6 The role of variant genes in differing transcriptome response

We hypothesized that the 53 genome differences led to the profound discrepancies in the transcriptome responses observed between the strains. To assess this, we identified which polymorphic genes were DE. Three of the six genes that are deleted in the small plasmid of *S.* 2973 sustained downregulated in *S.* 7942, and nine of 19 SNP-containing alleles were DE (Table 1.2). *rpoB, somA* (2), and *pyrG* were upregulated in both strains, while *mntA2* was only upregulated in *S.* 7942. *hltA, psbD1, sinA, pilN,* and *rpaA* were downregulated only in *S. 2*973. We previously found that the introduction of the *S.* 7942 alleles of *hltA, pyrG,* and *sinA* into *S.* 2973 led to a loss of HL tolerance (Walker and Pakrasi, 2022). Four of the variant genes (*rpaA, sinA, rpoB, rrlA*) are involved in transcriptional regulation. *rpaA* was of particular interest as this master transcriptional regulator controls the expression of 170 genes, including itself and *hltA* (Table 1.2) (Walker and Pakrasi, 2022). This early downregulation of the RpaA target *hltA* and late dysregulation of *rpaA* were both unique to *S.* 2973 and indicated that the different *rpaA* alleles led to changes in the expression of its regulon.

We also observed polymorphic genes that were not DE but shared enriched GO terms such as nitrogen compound metabolic processes (Fig. 2.3). Three of the polymorphic genes are involved in L-glutamate synthesis pathways: the SNP encoding *glsF, trpE,* and the indel encoding glutathione S-transferase (Fig. 2.4a). Interestingly, *gshB* was upregulated in *S.* 2973 and enzymes which utilize glutathione as a substrate (*gst3* and *chaC*) were downregulated in *S.* 2973 (Fig 2.4a). Together this suggested reduced oxidative damage in *S.* 2973 through a higher

flux of glutathione, an antioxidant required for environmental acclimation, and subsequent reduction of peroxides (Cameron and Pakrasi, 2010).

Table 1.2: log_2FC comparison of polymorphic protein-coding genes.

Symbol	Product	2973 Locus ID (M744_)	7942 Locus ID (Synpcc7942_)	7942: 30m	7942: 4h	2973: 30m	2973: 4h	Variation type
mntA2	Manganese ABC transporter ATP-binding protein	3335	2574		2.10			SNP
rpoB	DNA-directed RNA polymerase, beta subunit	8615	1522	1.43		1.68		SNP
pyrG	CTP synthase	6650	1954	1.22		1.56		SNP
hltA	PP2C-type phosphatase	3855	2473			-1.41		SNP
psbD1	D2 protein	13540	655			-1.72		SNP
rpaA	Response regulator for circadian clock	2605	95				-1.07	SNP
somA(2)	Porin; major outer membrane protein	9095	1607	6.40		8.80		Silent
sinA	Sigma factor interacting protein	705	452			-1.07		SNP
pilN	type IV pilus assembly protein	3975	2452			-1.62		SNP
atpA	ATP synthase F_0F_1 subunit alpha	1335	336					SNP
PpnK	inorganic polyphosphate/ATP-NAD kinase	4780	2304					SNP
	hypothetical protein	5865	2103					SNP
dnaK	Molecular chaperone	6025	2073					SNP
	Hydrolase	6570	1971					SNP
aroH	Chorismate mutase	6850	1915					SNP
trpE	Anthranilate synthase, component I	11685	1003					SNP
fadD	Long-chain-fatty-acid CoA ligase	12130	918					SNP
glsF	Glutamate synthase	12285	890					SNP
rrlA	23S ribosomal RNA	3320	R0005					2 SNPs
apt	Adenine phosphoribosytransferase	3965	2454					243 bp deletion
Coq4	Ubiquinone biosynthesis protein	-	768	-1.6	-2.0	-	-	deletion
	hypothetical protein	-	769	-1.9	-2.0	-	-	deletion
	5'-nucleotidase	-	770	-2.12	-2.0	-	-	deletion
	glutathione S-transferase	-	771			-	-	deletion
	CpeT/CpcT family DUF1001	-	772			-	-	deletion

2.3.7 RpaA alleles cause dramatic differences in genome-wide transcript levels

The KaiABC-based circadian clock determines genome-wide transcript level circadian oscillations (Hosokawa et al., 2011), and establishes RpaA activity through antagonistic clock proteins SasA and CikA (Gutu and O'Shea, 2013). RpaA is known to control primary metabolism (Markson et al., 2013a) and the expression of RpaA and RpaB-regulated genes in response to changing light intensity (Piechura et al., 2017, Iijima et al., 2015). We observed differential regulation of core clock genes *rpaA*, *sasA*, and *kaiC* solely in *S.* 2973. Additionally,

clock-associated genes *crm* and *lalA*, paralog of *labA*) (Adomako et al., 2022) were downregulated four-fold early in *S.* 2973.

The RpaA regulon is believed to consist of 103 to 170 genes (Piechura et al., 2017, Markson et al., 2013a). Of the 170 targets, we found that 60 genes (58%) were DE in one or both strains, 28% of the RpaA regulon was DE in *S.* 7942, while 37% were DE in *S.* 2973 (Fig. 2.5a). Most circadian genes that peak at dusk are dependent on RpaA activity and decrease during HL exposure (Piechura et al., 2017). We found 116 of 187 dusk genes are DE after HL exposure, with almost three times the number of early DEGs in *S.* 2973 compared to *S.* 7942. Dawn gene expression showed a similar difference with 40% of dawn genes DE in *S.* 2973 while only 16% were DE in *S.* 7942 (Fig. 2.5c). Unexpectedly, a greater proportion of dawn genes were downregulated than upregulated in *S.* 2973 (Fig. 2.5c, Supp. Table 2.1). RpaA-dependent sigma factors repress dawn genes and activate the transcription of non-RpaA target dusk genes (Markson et al., 2013a). We hypothesized that the RpaA alleles differentially induce σ factor signaling cascades that would result in these differences in the regulation of circadian genes.

Four sigma factors (*rpoD2*, *rpoD3*, *rpoD4*, and *rpoD5/sigC*) act as secondary regulators of RpaA and make up the complex RpaA-dependent transcriptional regulatory loop that controls circadian genes (Fleming and O'Shea, 2018, Markson et al., 2013a). RpaA regulates the mRNA levels of these sigma factor proteins, with *sigF2* and *rpoD5* found to peak at dusk, the final peaks in the regulatory cascade (Fleming and O'Shea, 2018). Both *sigF2* and *rpoD5* were downregulated early by a 1.7 $\log_2$FC early in *S.* 2973 (Fig 2.5b). Because of the intertwined nature of the σ factor cascade, many of the genes associated with each sigma factor overlap. In analysis of DEGs associated with each sigma factor, the largest set of DEGs were associated with *rpoD5* (Fig. 2.5b). Interestingly, *rpoA5*, a target of both RpaA and RpaB, was among the

most downregulated genes in *S.* 7942 (log_2FC -3), whereas it was only downregulated in *S.* 2973 by -1.4 log_2FC (Table 1.1). These differences in regulation of RpaA-dependent genes likely impact the complex bidirectional relationship of the circadian clock, affecting cellular metabolic processes that alter cell physiology according to the increased light environment. In addition to controlling circadian regulatory elements, RpaA is required for the regulation of cell division and the coordination of central carbon metabolism (Dong et al., 2010, Puszynska and O'Shea, 2017). RpaA target genes associated with carbon catabolism (*glgP* and *gnd*), glycogen synthesis (*pgm1* and *glgC*) (Fig. 2.4b), and carbohydrate metabolism genes (*oxdC* and *fbp*) were all downregulated in *S.* 2973 (Puszynska and O'Shea, 2017, Markson et al., 2013a). These results suggest that the *S.* 2973 *rpaA* allele regulates carbon metabolism differently than *S.* 7942 in response to HL, which may impact the C/N balance.

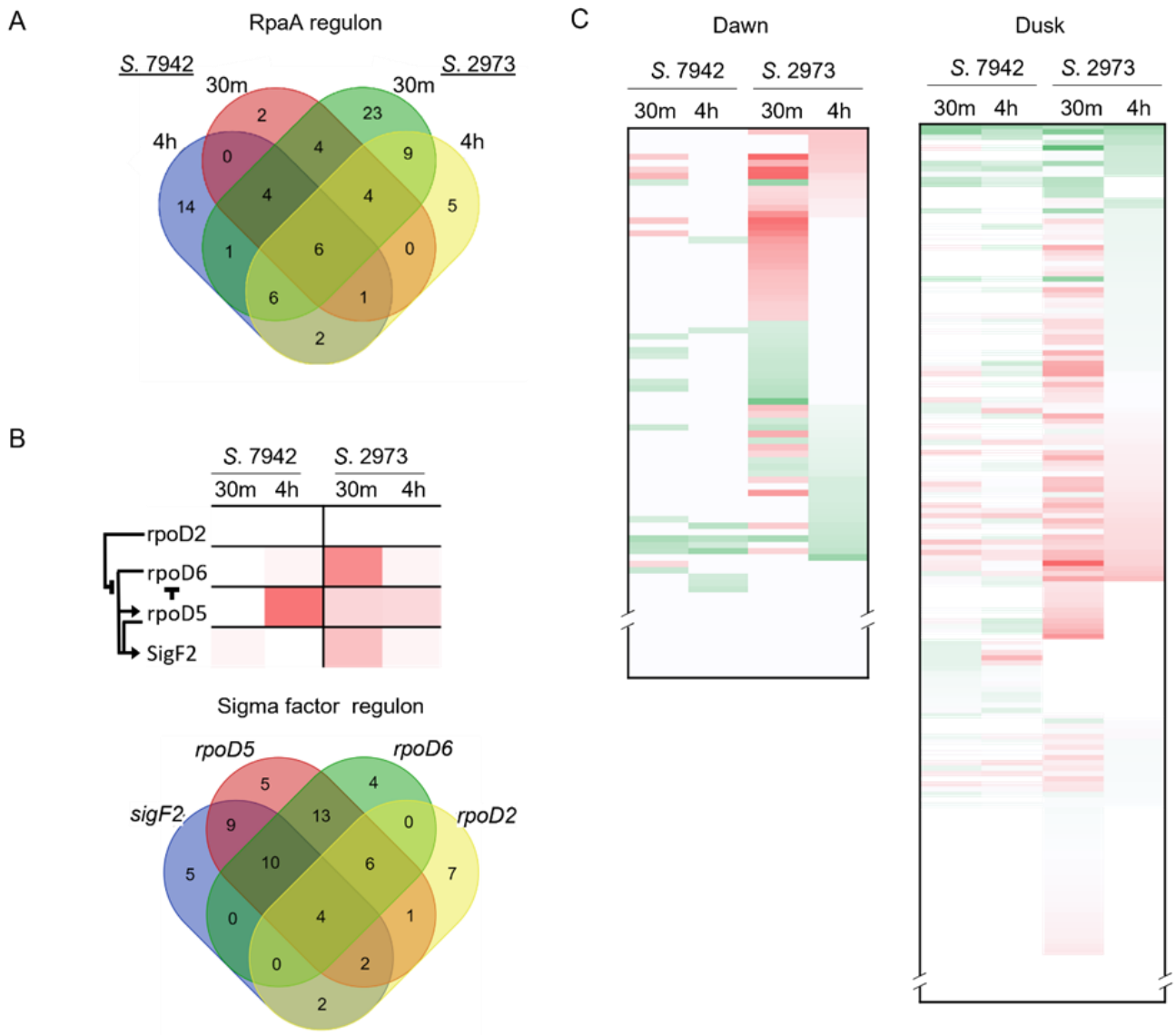

Figure 2.5: Comparison of RpaA regulon and circadian genes. (a) Venn diagram representing the number of genes from *S.* 7942 and *S.* 2973 in the RpaA regulon that were DE, from each condition. (b) Sigma factor cascade (21) and Venn diagram of DE circadian genes for each corresponding sigma factor(s). (c) Heatmap of dawn and dusk circadian genes.

2.3.8 Clustering analysis reveals altered gene expression profiles between strains

To further investigate the contributions of RpaA on enriched pathways and identify transcriptional signatures associated with each strain, we performed a K-means cluster analysis on 476 DEGs ($\log_2$FC ≥ 1.5; P≤0.05), which revealed eight clusters with distinct expression patterns (Fig. 2.6). The largest of these were cluster 2 with 118 genes and cluster 4 with 154

genes, and the expression profiles of these clusters contrast with each other in *S.* 2973. The most downregulated genes were in cluster 1 in *S.* 2973 and in cluster 6 in *S.* 7942. We focused our analysis on clusters that contain different expression patterns between the strains to identify unique light adaptation mechanisms in *S.* 2973. Two distinctive clusters that showed different gene expression patterns were clusters 4 and 6 (Fig. 2.6b). Cluster 4 was enriched in NtcA regulated genes (including regulators *ntcA* and *ntcB*) that are involved in nitrogen assimilation and metabolism, as well as enriched in PSII and thylakoid membrane genes (Fig. 2.6c). More than half of cluster 4 genes were circadian, or clock regulated, in total 26% were dusk genes and 22% were genes regulated directly by RpaA or RpaA-dependent sigma factors (Fig. 2.6d). Cluster 4 included two RpaA-dependent sigma factors, *sigF2* and *rpoD6,* the clock-associated *lalA* (Fleming and O'Shea, 2018). 16% of cluster 4 genes were associated with *rpoD5* (in cluster 6) (Fig. 2.5b and 6d). The sustained downregulation of *rpoD5* in *S.* 2973 may have allowed the release of inhibition on *rpoD6,* resulting in the late expression recovery seen in cluster 4. Cluster 6 was dominantly regulated by RpaA and was enriched for tRNAs and ncRNA metabolism (Fig. 2.6). This cluster included a two-component sensor histidine kinase (Synpcc7942_B2646), PII interacting protein *pipX*, and *rpoD5*. Cluster 6 may represent early response genes that connect the clock and the RpaA allele to the wide net of downstream changes that allow for HL tolerance in *S.* 2973.

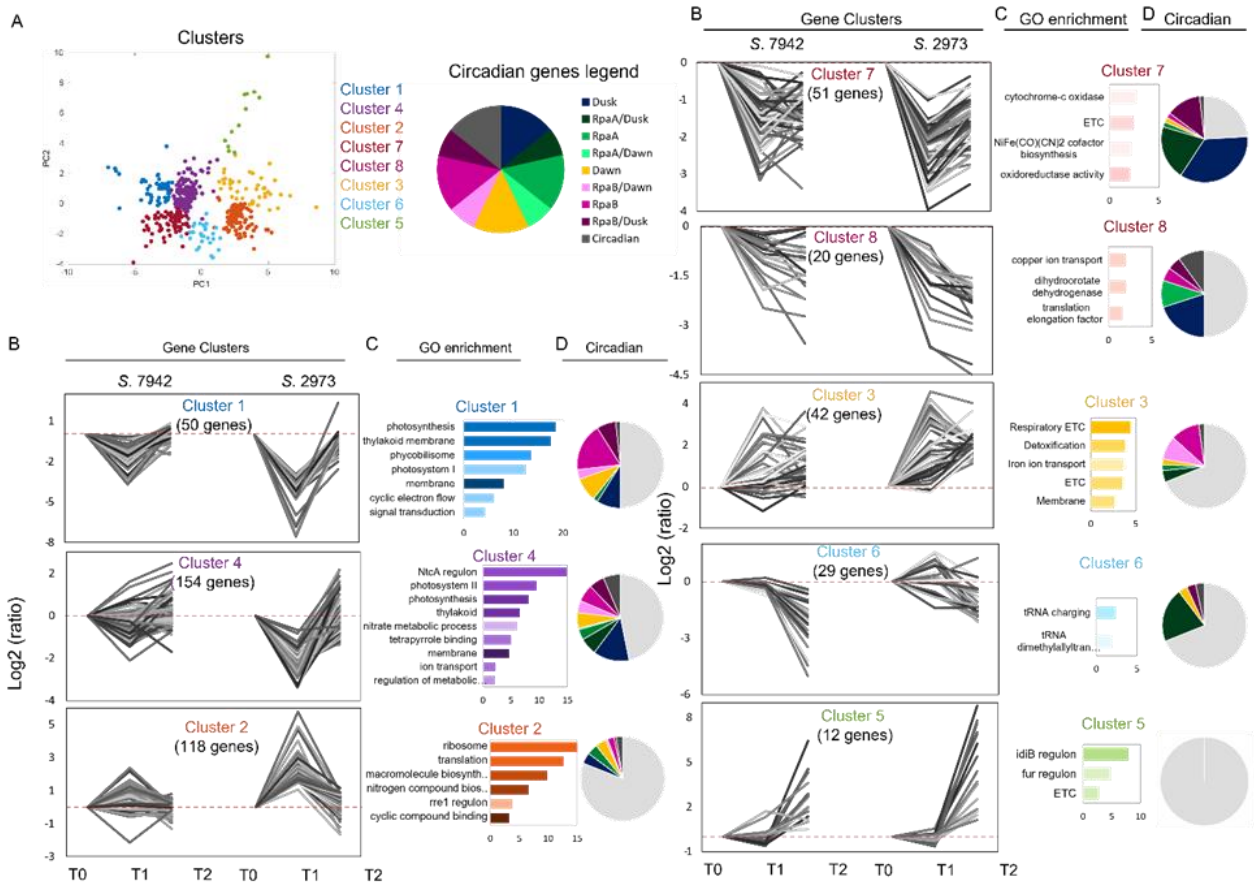

Figure 2.6: Expression dynamics and characteristics of clustered genes with differing responses to HL. (a) K-means clustering into 8 major clusters. Colors indicate the grouping of genes in each cluster. (b) Gene signatures and expression patterns of DEGs for each cluster. The x-axis represents times of HL treatment T0: 0 min, T1: 30 min, T2: 4 h and the y-axis represents log_2FC values. Grey scale lines indicate expression trends for individual genes, dotted red line indicates log_2FC of 0. (c) GO analyses of each gene cluster. Significance values shown across x-asix, darker colors represent a larger number of genes. (d) Ratio of circadian and/or RpaA/RpaB regulated genes out of total number of genes per cluster.

2.3.9 Roles of RpaB and NtcA in HL tolerance

The RpaB transcription factor controls genes involved in photoprotection, cyclic electron flow, and state transitions (Riediger et al., 2019) and is widely known to be involved in light stress regulation (Yasuda et al., 2020) (Fig. 2.6b and 6d). RpaB is activated by HL responsive Hik33 (also known as NblS or DspA) (Seki et al., 2007), and exhibits cross regulation with NtcA and FurA (Riediger et al., 2019). Of its 135 possible gene targets (Piechura et al., 2017), 92 (68% of the RpaB regulon) were found to be DE in our data (Table S2.1, Fig. 2.6), 9 of which were in

the top 10 most up- or downregulated (Table 1.1). As expected, RpaB targets nblA and hliA were upregulated in both strains after 30 min, although hliA upregulation was sustained in S. 2973. These genes are known to be transiently induced, suggesting this upregulation in S. 2973 may be a new homeostasis level for HL acclimation.

NtcA, the global nitrogen transcription factor, emerged as a significant player in the different transcriptome responses of *S.* 7942 and *S.* 2973. Cluster 4 was highly enriched for NtcA regulon genes, most of which were downregulated early in *S.* 2973, while in *S.* 7942 targets were not DE or in the case of the *cyn* operon sustained upregulation (Fig. 2.4 and 2.6). After four hours in *S.* 29734, most NtcA targets returned to control expression levels, although some gene targets (*nirA, cynS, cynD,* and *cynB*) were upregulated (cluster 4) (Fig. 2.4 and 2.6). NtcA mediates an intricate regulatory network of other transcription factors, including the response regulator NrrA (also known as Rre37) (Liu and Yang, 2014), SigE, and *glnB*, which encodes the PII protein. We found that *nrrA* and *glnB* were downregulated early in *S.* 2973. PII and PipX form a protein complex under low C/N and ATP/ADP ratios, which activates the *plmA* transcriptional regulator (Labella et al., 2016), and we found that *plmA* was DE only in *S.* 2973. These data indicated that the nitrogen stress pathway induced transcriptional changes in *S.* 2973. The connection of NtcA and HL response may be due to the overlap of oxidative stress, which has been observed to alter the expression of genes related to nitrogen metabolism through NtcA (Robles-Rengel et al., 2019). Downregulation of many nitrogen assimilation genes may also be a consequence of C/N balance. *S.* 2973 upregulated C metabolism is detected by the PII protein which controls transcription factors NtcA, PipX, and PlmA to alter nitrogen assimilation activity.

2.4 Discussion

This study captured the transcriptome landscapes of two *Synechococcus elongatus* strains during HL acclimation to identify key genetic and transcriptional differences that account for the HL tolerance of *S.* 2973. We found that key differences for the divergence in HL tolerance between *S.* 7942 and *S.* 2973 occurred within the first hour of HL intensity upshift, as indicated by changes in photosynthetic efficiency (Fig. 2.1a and 2.1d) and in global transcriptional activity (Fig. 2.2). Specifically, we observed that in the first hour of HL, *S.* 7942 underwent greater damage to PSII (Fig. 2.1d) and relied more heavily on D1 protein turnover (Table 1.1). We found that clusters of genes are dysregulated to different degrees between the strains. Cluster 2, which was enriched for nitrogen compound metabolism, cluster 4, which was enriched for regulation of metabolic activity, and cluster 7, which was enriched for cellular respiration, all underwent greater DE in *S.* 2973 compared to *S.* 7942 (Fig. 2.6). These differences likely affect the metabolic activity of cells, causing the difference in growth rate under HL conditions (Fig. 2.1). There may also be a relationship between the faster growth rate of *S.* 2973 and the higher expression levels observed, as the reverse is true for *Synechococcus* cells grown in the dark that undergo gene expression depressions (Hood et al., 2016). In clustering analysis, clusters 1 and 2 (which together comprise 17% of the DEG) exhibited the same expression pattern but were expressed at least two-fold higher in *S.* 2973, reflecting the overall trend of larger gene expression changes in *S.* 2973. The large differences in gene expression may be due to alleles of transcriptional regulation genes *rpoB, rrlA* (23S ribosomal RNA), and *rpaA.* The larger and greater fold changes were likely driven by different activation dynamics that allow for *S.* 2973 to acclimate faster to HL. This study is the first RNA-seq dataset analyzing *S.* 2973 under its optimal fast-growth conditions.

We observe a complex interlaced response involving iron stress, oxidative stress, nitrogen stress, and light stress response pathways. Through pathway analysis we observed upregulation of bicarbonate transport and CO_2 fixation genes in *S.* 2973 (Fig. 2.4), which aligns with previous studies of moderate HL acclimated *S.* 2973 that found up-regulation of central metabolic pathway genes (Tan et al., 2018). The low-CO_2-responsive regulator CmpR induces the *cmp* operon upon low-to-high light transition (Takahashi et al., 2004), which may be more active in *S.* 2973 than *S.* 7942. Additionally, we observed large expression changes to genes involving nitrogen metabolism (Fig. 2.3, 2.4, and 2.6). Nitrate assimilation genes *nirA-B,* nitrate transporter genes *nrtA-D,* nitrilase (Synpcc7942_0839), and ammonium transport gene *amt1* (Watzer et al., 2019) are all significantly downregulated only in S. 2973. This suggests that the cells experienced nitrogen stress, likely in response to reactive nitrogen species and for carbon/nitrogen balance due to the strong demand for CO_2 under HL. The difference seen in nitrogen regulation may be due to the SNP encoding *glsF*. *glsF* catalyzes the reductive synthesis of L-glutamate from 2-OG and L-glutamine. Levels of 2-OG in the cell act as an indicator of the C/N status detected by PII, which regulates nitrogen transporters (Amt, NrtABCD) and co-actives the transcription factor NtcA. NtcA-regulated targets underwent large differences between the strains, such as glutamine synthase inactivation factor *gifB* (Synpcc7942_2529), which was upregulated eight-fold in *S.* 2973. Additionally, AbrB-like transcriptional regulator (Synpcc7942_1969), which has been found to regulate nitrogen uptake and carbon metabolism in other cyanobacteria (Lieman-Hurwitz, 2009), sustained upregulation in *S.* 2973 but was not DE in *S.* 7942 (Table S2.2). The altered expression of nitrogen metabolism genes may also be due to less ROS accumulating in *S.* 2973, as hydrogen peroxide can decrease concentrations of 2-OG (Robles-Rengel et al., 2019). This is reflected by the downregulation of the oxidative stress

responsive catalase-peroxidase encoded by *katG*, which showed sustained downregulation in *S.* 2973 but was only downregulated early in *S.* 7942 (Fig. 2.4).

Photosynthetic functions were downregulated in the early response of both strains but upregulated after 4 h in *S.* 2973, reflecting its acclimation to the HL. Unexpectedly, early downregulation of Cluster 1 included cyclic electron flow (CEF), which has been shown in *S.* 2973 to increase as an alternative electron pathway for energy dissipation needed during over-saturating light exposure (Ungerer et al., 2018b). Genes encoding PSII, PBS, and PSI proteins were downregulated in the early HL response of both strains (Fig 2.4 and Fig. 2.6), which differs from previous transcriptome analysis of *S.* 2973 in HL that saw upregulation of PSII genes (Tan et al., 2018). However, the metalloprotease FtsH was only upregulated in *S.* 7942, which suggests D1 degradation was not upregulated in *S.* 2973 (Sacharz et al., 2015). This aligns with our data (Fig. 2.1b) showing that PSII does not incur as much damage under HL in *S.* 2973 as S. 7942. Enrichment for ribosomal transcripts was sustained in *S.* 2973 but only present during early HL response in *S.* 7942. This suggests an acceleration of translation in *S.* 2973, in agreement with previous transcriptome analysis (Tan et al., 2018). RpaC, a proposed component controlling PBS-PSII stability (Joshua and Mullineaux, 2005), was highly downregulated only in *S.* 2973 (Fig. 2.4).

Of the 90 regulatory elements (transcription factors, sigma factors, RNA polymerases, and putative regulatory proteins) that we identified in the *Synechococcus* genome, 21 were DE in *S.* 7942 while 36 were DE in *S.* 2973 (Table S2.2). DE regulatory factors included Fur, IdiB, Rre1, NtcA, RpaA, and RpaB. NtcA is also partly responsible for *rpoD4* regulation (Imamura and Asayama, 2009). The two sigma factors *rpoD4* and *rpoD6* were downregulated, 8-fold and 6-fold respectively, in *S.* 2973 but were not DE in *S.* 7942. These regulatory elements were

therefore partially responsible for the transcriptional differences observed between *S.*2973 and *S.*7942 (Fig. 2.6).

We observed at least a three-fold down-regulation of 37% of dusk genes in S. 2973 and 27% of dusk genes in *S.* 7942 (Table S2.1). This is comparable to HL-pulse response results found in a circadian regulation study (Piechura et al., 2017). We hypothesized that the difference in circadian gene regulation was due to the RpaA allele. Changes in light intensity transiently alter the action of RpaA through the mediation of RpaB (Piechura et al., 2017, Espinosa et al., 2015) and are predicted to transiently stall the progression of the RpaA-dependent σ factor cascade, thus affecting the regulation of many circadian-controlled genes (Fleming and O'Shea, 2018). Clustering analysis allowed us to examine and compare the contribution of regulatory factors to the HL transcriptome response. As expected, the dynamics of transcriptional regulators RpaA and RpaA-dependent sigma factors varied between the two strains. *rpoD2* represses the activity of the other sigma factors and was the only RpaA-dependent σ factor that was not significantly down-regulated in *S.* 2973. Overall, our results suggested that the *rpaA* allele in *S.* 2973 orchestrated large transcriptome changes that gave rise to HL tolerance. This study provides a deeper understanding of the interconnected relationship between circadian gene regulators and how small changes in the RpaA transcription factor can lead to global transcriptome changes.

RpaB is a benchmark of the HL response whose gene targets are activated immediately following HL (Hanaoka and Tanaka, 2008, Riediger et al., 2019, Yasuda et al., 2020, Espinosa et al., 2015). Some RpaB targets are transiently expressed with peak mRNA levels after 15 min that then are degraded after 30 min (Hanaoka and Tanaka, 2008). Key transient RpaB targets *hliA* and *rpoD3* (Seki et al., 2007) were upregulated about six-fold in both strains in our data set,

confirming that the 30 min time point was able to capture transient high light response. In *Synechocystis* 6803, RpaB has been found to coregulate genes with other transcription factors such as Fur and NtcA (Riediger et al., 2019). RpaA and RpaB participate in interdependent regulatory mechanisms, and RpaB regulation of RpaA has been shown to occur through coregulation of gene targets that take part in activating RpaA (Espinosa et al., 2015). In summary, it appears that multiple and complex regulatory interactions predominantly originating from the RpaA allele in *S.* 2973 cooperate to better acclimate cells to HL conditions.

The process of Mn^{2+} uptake is light-dependent, while regulated Mn^{2+} import occurs through ABC-type transporter MntCAB and FutABC (Eisenhut, 2019, Sharon et al., 2014). Two operons encode MntCAB in *Synechococcus*, with a SNP in *mntA2* in *S.* 2973. The expression of the Mn transporters varied greatly between the strains, likely due to the shortened *mntA2* in *S.* 2973. In *S.* 7942 *mntA1* and *mntC1* were upregulated early while *mntA2* was upregulated late (Fig. 2.4).

We previously identified that the *S.* 2973 allele of *hltA,* a RpaA-regulated manganese-dependent protein phosphatase, was essential for HL tolerance (Walker and Pakrasi, 2022). In our data RpaA was downregulated in *S.* 2973 and not *S.* 7942. To our knowledge this is the first report of the difference in RpaA expression between the strains. The predicted interacting partners of *hltA* were also DE following HL. The expression pattern of HltA's predicted target, an anti-anti-σ factor regulator, was opposite of *hltA* in *S.* 2973, and the predicted cognate kinase was downregulated within 30 min of HL. The pathway of *hltA* activity has yet to be identified, although from the transcriptome data it shares expression patterns with genes involved in energy production (*ndbC* and *glcE*), cell wall biogenesis (*corA, ntcB,* Synpcc7942_0047), and transcription (*srpG,* Synpcc7942_0837). Most of the predicted interacting genes with *hltA* are

putative factors, and future studies of *hltA* will elucidate the pathway and may further connect circadian regulation to light sensing and response.

A single SNP is present in the β subunit of the RNA polymerase encoded by *rpoB*, we found *rpoB* to be upregulated 3-fold after 30 minutes in both strains. We suspect that the polymorphic RNA polymerase subunit may cause changes that result in global differences. Mutations to RpoB have been seen to impact promoter binding, elongation, and survival, and mutation to *rpoB* has previously been reported in *Synechococcus elongatus* PCC 11801 to possibly confer increased tolerance to solvent stresses (Srivastava et al., 2021). Additionally, the 23S ribosomal RNA encodes two SNPs. Exactly what differences seen in the HL response can be accounted for by these different alleles is unknown and is an interesting topic for further investigation.

The group 2 sigma factors are required for HL tolerance in *Synechocystis* 6803 (Koskinen et al., 2016, Turunen et al., 2022). The conserved, HL-responsive *rpoD3* showed sustained upregulation in both strains (Seki et al., 2007). We find that the M-type group 2 sigma factors *rpoD4* and *rpoD6* are highly downregulated specifically in *S.* 2973 (Imamura and Asayama, 2009). *rpoD4* has been observed to be downregulated in *Prochlorococcus marinus* PCC9511 cultures exposed to HL and UV, and it was suggested that *rpoD4* may regulate genes expressed during the light-dark transition (Kolowrat et al., 2010). Group 3 sigma factors SigF and SigF2 (SigJ) were also downregulated in *S.* 2973. SigF2 was found to function in photoprotection in Anabaena *sp*. PCC 7120 (Srivastava et al., 2020). The extra-cytoplasmic function sigma factor *sigG* was downregulated in only *S.* 7942. Not much is known about the ECF sigma factors, which may be controlled by post-translational mechanisms such as phosphorylation-based partner switching (Srivastava et al., 2020).

Finally, a limitation arose from mismatched or missing locus identification between the strains, which made it difficult to identify orthologous genes. Difficult to compare plasmid encoded genes. To overcome this issue in addition to aligning *S.* 2973 transcripts to their respective genome, we also ran an alignment and DEG pipeline using the *S.* 7942 reference genome. The pUH24 small plasmid in *S.* 2973 was first reported to contain 16 of the 55 genomic differences (Yu et al., 2015), though recent resequencing reported that many of the SNPs were the result of sequencing error (Adomako et al., 2022). Data provided here and accumulating omics studies will help resolve the discrepancies observed in transcriptome datasets and in genome annotations, which will assist in future *Synechococcus* comparative analysis studies.

2.5 Materials and Methods

Cells and culture conditions

Synechococcus elongatus UTEX 2973 and *Synechococcus elongatus* PCC 7942 were grown in BG11 media at 38°C with 1% CO_2 bubbling, as described in Ungerer et al. (Ungerer et al., 2018c), unless otherwise indicated. For both photosynthetic efficiency and RNA-seq experiments the following culturing steps were followed. 50 mL cultures were grown to midlogarithmic growth phase (optical density at 730 nm of 0.6–0.8) in an MC-1000 Multicultivator Bioreactor (Photon Systems Instruments, Czech Republic) at 38°C under normal light (NL, 200 μmol photons m^{-2} s^{-1}) and diluted with fresh BG11 medium to an OD_{730} of ~0.3 before beginning exposure to light treatments.

Cultures for RNA-seq were kept at NL (200 μmol photons m^{-2} s^{-1}) or transferred to HL (1,500 μmol photons m^{-2} s^{-1}) for 30 minutes and 4 hours at which times they were promptly collected for RNA isolation.

Measurement of photosynthetic activity

To determine the crucial time points of high light stress on cell physiology, cultures were exposed to NL (200 μmol photons m^{-2} s^{-1}), ML (400 μmol photons m^{-2} s^{-1}), and HL (1,500 μmol photons m^{-2} s^{-1}) with and without 200 μg/ml chloramphenicol. The fluorescence was measured every five to 60 minutes for up to 4 hours using a double-fluorometer, FL-200 (Photon Systems Instruments, Brno, Czech Republic). 1 mL samples were dark adapted for 3 minutes before measurement. The instrument contains red LEDs for both actinic (20-s) and measuring (2.5-s) flashes and was used in the time range of 100 μs to 100 s. The photosynthetic maximum quantum yield of PSII parameter was calculated as the equation of $F_v/F_m = (F_m–F_0)/F_m$, where F_0 is the minimum fluorescence, F_v is the variable fluorescence, and F_m is the maximum fluorescence.

RNA isolation and purification

Fifty mL cultures for each biological replicate (n=3) of control, 30 minutes, and 4 hours of high light treated for both *Synechococcus elongatus* UTEX 2973 and *Synechococcus elongatus* PCC 7942 were spun down and immediately frozen to -80°C. Total RNA was extracted by using RNAwiz reagent (Ambion) according to the manufacturer's instructions and treated with RNase-free DNase I to eliminate contaminating chromosomal DNA. RNA quality and quantity were evaluated using a NanoDrop Spectrophotometer ND-1000 instrument (NanoDrop Technologies). Additional concentration measurements were collected with Qubit Fluorometer (Invitrogen).

Library construction and transcriptome sequencing

Ribosomal depletion, library preparation, and RNA-sequencing were performed by the Genome Technology Access Center (GTAC) at Washington University in St. Louis. Library preparation was performed with 1 μg of total RNA, and integrity was determined using an Agilent bioanalyzer (Invitrogen). RNAseq library preparation was completed with FastSelect 5S/16S/23S rRNA depletion (Bacteria) kits (Qiagen). mRNA was then fragmented and reverse

transcribed to yield cDNA using SuperScript III RT enzyme (Life Technologies) and random hexamers. A second-strand reaction was performed to yield ds-cDNA. cDNA was blunt ended, had an A base added to the 3′ ends, and then had Illumina sequencing adapters ligated to the ends. Ligated fragments were then amplified for 15 cycles using primers incorporating unique index tags. All libraries were then sequenced on an Illumina NovaSeq6000 S4 XP platform with 2x150 targeting 50M paired-end reads.

Transcriptome analysis

The Prokseq RNA-Seq data analysis package for Prokaryotes was used for raw fastq file data processing and quality control (Firoj Mahmud et al., 2020). Processed paired-end reads were aligned using Bowtie to the reference genomes of *Synechococcus elongatus* PCC 7942 (*Synechococcus*_elongatus_pcc_7942_gca_000012525.ASM1252v1.51) from Ensembl (bacteria.ensembl.org) and *Synechococcus elongatus* UTEX 2973 (GCF_000817325.1_ASM81732v1) from NCBI.

To validate the alignment results, two pipelines were used to align and count differentially expressed genes (DEGs) for each strain. As a secondary validation RNA-seq reads were additionally processes by GTAC at Washington University in St. Louis. Briefly, RNA-seq reads were then aligned to the Ensembl 76 assembly with STAR 2.5.1a. To find critical genes, the raw counts were variance stabilized with the R/Bioconductor package DESeq211 and was then analyzed via weighted gene correlation network analysis with the R/Bioconductor package WGCNA12.

Differential gene expression was determined using DESeq2. The differentially expressed genes (DEGs) were filtered based on a minimum log2FoldChange = 1.5 and P-value <0.05. Gene annotation involved using Blast2GO (blast2go.com) and Biocyc (Biocyc.org). Heatmaps were

generated using Heatmapper (heatmapper.ca/expression) and GO and pathway enrichment from Biocyc (Biocyc.org) and KEGG pathway enrichment analysis (kegg.jp).

Clustering analysis

The 475 genes found to be DE (based on a minimum log_2FoldChange = 1.5 and P-value <0.05) in at least one time point were grouped according to their expression patterns by K-means clustering analysis. K-means clustering analysis was performed in MatLab (MathWorks Ltd., Cambridge, UK) using the pairwise similarity of each pair of samples determined by the Spearman correlation coefficient and hierarchical clustering of expression with default settings.

2.6 Supporting information

Table S2.1: DEGs that exhibit circadian mRNA expression or are regulated by a circadian component and its cluster.

Locus	7942: 30m	7942: 4h	2973: 30m	2973: 4h	RpaA regulon	RpaB regulon	Dusk / Dawn	Circadian	Cluster
Synpcc7942_0569	-0.40	0.49	-4.56	-0.17			Dusk	yes	1
Synpcc7942_0855	-0.69	0.07	-3.87	-0.67		RpaB	Dawn	yes	1
Synpcc7942_0858	-2.09	-0.11	-4.36	-0.86			Dawn	yes	1
Synpcc7942_1042	-2.19	-0.03	-4.15	-1.24		RpaB	Dusk	yes	1
Synpcc7942_1108	-1.08	0.82	-3.85	-0.96			Dusk	yes	1
Synpcc7942_1234	-0.68	0.22	-3.03	-1.11				yes	1
Synpcc7942_1446	-1.67	-0.44	-3.24	-1.29			Dusk	yes	1
Synpcc7942_1987	-0.68	-0.36	-3.62	-1.74	rpaA	RpaB	Dusk	yes	1
Synpcc7942_2352	-1.00	0.07	-3.29	-0.42		RpaB	Dusk	yes	1
Synpcc7942_2499	-1.29	-0.49	-3.59	-0.92			Dusk	yes	1
Synpcc7942_0920	-1.60	0.42	-4.16	0.18		RpaB	Dawn		1
Synpcc7942_0551	-1.83	-1.35	-5.98	-2.56		RpaB	Dusk		1
Synpcc7942_2049	-2.38	0.18	-4.17	-0.71		RpaB			1
Synpcc7942_0327	-1.07	0.25	-3.37	-0.43		RpaB			1
Synpcc7942_0328	-0.91	0.37	-4.18	-0.22		RpaB			1
Synpcc7942_1047	-2.44	0.13	-4.60	-0.49		RpaB			1
Synpcc7942_1322	-1.85	0.21	-3.13	0.17		RpaB			1
Synpcc7942_2030	-0.75	0.26	-3.31	-0.47		RpaB			1
Synpcc7942_2343	-1.46	0.29	-3.44	-0.62		RpaB			1
Synpcc7942_0857	-1.46	-0.19	-4.46	-1.31			Dawn		1
Synpcc7942_0856	-1.21	-0.25	-4.27	-1.23			Dawn		1
Synpcc7942_1236	-0.72	-0.14	-3.20	-0.39			Dawn		1

Synpcc7942_2342	-1.56	0.24	-3.42	-0.35			Dawn		1
Synpcc7942_0407	-1.06	-0.18	-3.61	-1.46			Dusk		1
Synpcc7942_2498	-1.68	-0.65	-4.04	-0.98			Dusk		1
Synpcc7942_1419	-1.06	-0.35	-3.23	-0.82			Dusk		1
Synpcc7942_0859	-1.78	0.13	-3.64	-0.61			Dusk		1
Synpcc7942_1630	-1.59	0.49	-3.08	-0.35			Dusk		1
Synpcc7942_0325	-1.36	0.27	-3.70	-0.30			Dusk		1
Synpcc7942_0019	1.18	-0.25	3.43	-0.74			Dusk	yes	2
Synpcc7942_0020	1.02	-0.25	4.03	-1.55				yes	2
Synpcc7942_0829	0.17	0.01	3.41	0.18			Dawn	yes	2
Synpcc7942_1262	0.56	-0.24	1.63	-0.16				yes	2
Synpcc7942_1290	-2.16	-0.20	5.77	0.89			Dusk	yes	2
Synpcc7942_1482	0.26	-0.51	1.54	-0.23				yes	2
Synpcc7942_1313	1.14	0.09	1.53	0.31	rpaA				2
Synpcc7942_1328	0.23	0.00	1.60	0.87	rpaA				2
Synpcc7942_1552	1.07	0.22	1.68	0.35	rpaA				2
Synpcc7942_2119	0.60	-0.76	1.76	-0.69	rpaA				2
Synpcc7942_2524	0.98	-0.38	1.60	-0.29	rpaA				2
Synpcc7942_2088	0.27	0.09	1.80	0.47		RpaB	Dawn		2
Synpcc7942_2127	1.10	0.31	2.48	0.76		RpaB	Dusk		2
Synpcc7942_0882	0.23	-0.66	1.61	0.35		RpaB			2
Synpcc7942_0885	1.22	0.37	1.52	0.24		RpaB			2
Synpcc7942_1596	0.30	-0.44	1.59	0.01		RpaB			2
Synpcc7942_2409	1.16	-0.96	2.59	-0.69			Dawn		2
Synpcc7942_2232	1.22	0.04	1.61	0.18			Dawn		2
Synpcc7942_0632	1.48	0.42	1.74	0.18			Dawn		2
Synpcc7942_0631	1.30	0.49	1.62	0.43			Dawn		2
Synpcc7942_1124	0.70	-1.31	2.58	-0.50			Dusk		2
Synpcc7942_2203	1.19	-0.11	1.81	-0.20			Dusk		2
Synpcc7942_1523	1.63	0.23	1.75	0.17			Dusk		2
Synpcc7942_0672	3.51	2.92	2.02	2.74		RpaB	Dusk	yes	3
Synpcc7942_1214	2.17	1.37	3.04	1.92				yes	3
Synpcc7942_2309	2.10	1.72	2.20	1.41		RpaB	Dawn	yes	3
Synpcc7942_1548	1.70	1.14	-0.27	1.45	rpaA	RpaB	Dawn		3
Synpcc7942_0243	3.79	1.45	4.22	2.18	rpaA	RpaB	Dusk		3
Synpcc7942_1475	-0.22	0.81	3.69	2.24	rpaA				3
Synpcc7942_0197	0.67	0.72	1.85	1.13		RpaB			3
Synpcc7942_0364	-0.02	0.85	0.32	1.54		RpaB			3
Synpcc7942_1120	5.33	1.60	5.96	1.99		RpaB			3
Synpcc7942_1840	1.67	1.24	1.55	1.45		RpaB			3
Synpcc7942_1997	3.32	0.44	3.78	1.58		RpaB			3
Synpcc7942_1628	0.11	0.99	-0.14	2.53			Dawn		3
Synpcc7942_0023	-0.29	0.53	-1.99	0.61				yes	4

Synpcc7942_0096	-0.75	-0.22	-2.31	-0.56		RpaB	Dusk	yes	4
Synpcc7942_0278	-0.52	-0.33	-1.68	-0.92				yes	4
Synpcc7942_0370	-1.24	-0.56	-2.62	-0.16	rpaA		Dusk	yes	4
Synpcc7942_0383	-0.14	-0.03	-2.45	-1.09			Dusk	yes	4
Synpcc7942_0422	-0.80	-0.41	-2.07	-1.03			Dusk	yes	4
Synpcc7942_0568	-0.28	-0.10	-2.23	0.32			Dusk	yes	4
Synpcc7942_0595	-0.30	0.11	-1.70	-0.40		RpaB	Dusk	yes	4
Synpcc7942_0599	-0.40	-0.28	-1.69	-0.70			Dusk	yes	4
Synpcc7942_0630	-0.09	1.23	-2.21	0.80				yes	4
Synpcc7942_0655	0.03	0.38	-1.72	0.05				yes	4
Synpcc7942_0699	-0.34	0.72	-2.54	-0.83			Dusk	yes	4
Synpcc7942_0797	-1.07	0.43	-1.56	-0.56	rpaA	RpaB	Dusk	yes	4
Synpcc7942_1135	-0.27	0.03	-1.87	0.41				yes	4
Synpcc7942_1397	-0.64	0.04	-2.57	-0.88		RpaB	Dusk	yes	4
Synpcc7942_1402	-0.17	-0.03	-2.02	-1.25			Dusk	yes	4
Synpcc7942_1506	-0.56	-0.11	-1.96	-0.66	rpaA		Dusk	yes	4
Synpcc7942_1549	-0.71	-0.04	-2.24	-0.85				yes	4
Synpcc7942_1557	-0.14	-0.56	-2.48	-0.89	rpaA	RpaB	Dusk	yes	4
Synpcc7942_1637	0.14	0.42	-1.67	0.05				yes	4
Synpcc7942_1646	-0.14	0.21	-1.63	-0.16	rpaA	RpaB	Dusk	yes	4
Synpcc7942_1757	-0.51	-0.49	-2.32	-0.87	rpaA		Dusk	yes	4
Synpcc7942_1776	0.21	-0.21	-1.31	-1.79		RpaB	Dusk	yes	4
Synpcc7942_1784	-0.26	-0.13	-1.71	-0.83	rpaA	RpaB	Dusk	yes	4
Synpcc7942_1976	-0.35	-0.29	-2.43	-0.64	rpaA		Dusk	yes	4
Synpcc7942_2033	-0.44	0.09	-2.10	-0.06				yes	4
Synpcc7942_2059	-0.12	0.44	-2.72	-0.18	rpaA		Dusk	yes	4
Synpcc7942_2080	-0.13	-0.23	-1.55	0.66			Dusk	yes	4
Synpcc7942_2160	0.01	0.27	-2.34	-0.08		RpaB	Dawn	yes	4
Synpcc7942_2187	-0.71	-0.39	-2.27	-0.95			Dusk	yes	4
Synpcc7942_2200	-0.53	-0.25	-1.56	-0.83			Dusk	yes	4
Synpcc7942_2306	-0.17	-0.15	-1.57	-1.38	rpaA		Dusk	yes	4
Synpcc7942_2404	-0.53	-0.09	-2.46	-1.05			Dusk	yes	4
Synpcc7942_2412	0.04	0.70	-1.58	0.03				yes	4
Synpcc7942_2447	-0.49	0.46	-2.77	0.44				yes	4
Synpcc7942_2500	-1.03	-0.38	-2.34	-0.85			Dusk	yes	4
Synpcc7942_2602	-0.13	-0.57	-2.25	-1.30			Dusk	yes	4
Synpcc7942_2609	-0.39	0.32	-2.10	0.22			Dusk	yes	4
Synpcc7942_0157	-0.07	0.11	-1.95	-0.61	rpaA	RpaB			4
Synpcc7942_0891	-0.82	0.83	-2.46	1.36	rpaA	RpaB			4
Synpcc7942_2605	-0.47	-0.12	-1.56	-1.24	rpaA	RpaB			4
Synpcc7942_0156	0.02	0.86	-1.73	0.58	rpaA		Dawn		4
Synpcc7942_0013	-0.52	1.53	-0.68	-0.51	rpaA				4
Synpcc7942_0342	-1.40	0.81	-1.55	1.33	rpaA				4

Synpcc7942_2473	-0.28	-0.48	-1.41	-0.53	rpaA				4
Synpcc7942_0259	0.09	1.05	-2.84	0.09		RpaB	Dawn		4
Synpcc7942_0697	-0.25	0.51	-1.72	0.08		RpaB	Dawn		4
Synpcc7942_1254	-0.55	0.55	-2.16	-0.53		RpaB	Dawn		4
Synpcc7942_2010	-0.13	0.12	-2.56	-0.45		RpaB	Dawn		4
Synpcc7942_2156	0.19	0.82	-1.80	0.36		RpaB	Dawn		4
Synpcc7942_2158	-0.14	0.45	-1.86	-0.01		RpaB	Dawn		4
Synpcc7942_0240	0.11	0.52	-2.32	-0.04		RpaB	Dusk		4
Synpcc7942_0130	-0.11	-0.11	-1.94	-1.10		RpaB			4
Synpcc7942_0415	-0.34	-0.04	-1.97	-0.87		RpaB			4
Synpcc7942_0424	-1.46	0.95	-2.47	-0.01		RpaB			4
Synpcc7942_0470	-0.84	-0.12	-1.92	-0.57		RpaB			4
Synpcc7942_0759	-0.55	0.05	-1.60	-0.25		RpaB			4
Synpcc7942_1002	-1.21	0.33	-2.66	0.10		RpaB			4
Synpcc7942_1858	-0.12	0.25	-2.35	-0.08		RpaB			4
Synpcc7942_1143	-0.38	0.06	-2.14	-1.24			Dawn		4
Synpcc7942_2503	-0.14	-0.02	-1.66	-0.93			Dawn		4
Synpcc7942_1882	-0.51	-0.03	-1.79	-0.40			Dawn		4
Synpcc7942_0294	-0.54	0.75	-2.42	0.03			Dawn		4
Synpcc7942_1304	0.27	0.33	-1.50	0.03			Dawn		4
Synpcc7942_0656	-0.05	0.37	-1.78	0.13			Dawn		4
Synpcc7942_0442	0.01	0.73	-2.34	0.50			Dawn		4
Synpcc7942_1032	0.05	0.50	-2.93	1.03			Dawn		4
Synpcc7942_2105	0.78	1.70	-1.43	1.29			Dawn		4
Synpcc7942_2106	1.61	2.47	-1.24	1.58			Dawn		4
Synpcc7942_1125	-0.29	-0.13	-2.11	-0.74			Dusk		4
Synpcc7942_2453	-0.31	-0.04	-1.50	-0.43			Dusk		4
Synpcc7942_0593	0.12	0.11	-1.66	-0.40			Dusk		4
Synpcc7942_0307	-1.38	-0.29	-2.04	-0.30			Dusk		4
Synpcc7942_1460	-0.65	0.40	-1.77	-0.25			Dusk		4
Synpcc7942_0168	-0.39	0.42	-1.70	-0.16			Dusk		4
Synpcc7942_1535	-0.33	0.03	-1.54	-0.13			Dusk		4
Synpcc7942_1827	-0.47	0.29	-1.89	-0.10			Dusk		4
Synpcc7942_2452	-0.23	0.56	-1.62	0.02			Dusk		4
Synpcc7942_0199	-0.33	-0.17	-1.97	0.06			Dusk		4
Synpcc7942_0201	-0.29	-0.29	-2.22	0.33			Dusk		4
Synpcc7942_0834	-0.23	-3.48	0.12	-0.43	rpaA		Dusk	yes	6
Synpcc7942_1612	-0.24	-2.14	0.05	0.08	rpaA		Dusk	yes	6
Synpcc7942_1849	-0.25	-2.82	-1.42	-1.37	rpaA	RpaB	Dusk	yes	6
Synpcc7942_2267	-0.05	-0.97	-0.73	-1.54	rpaA		Dusk	yes	6
Synpcc7942_2386	-1.60	-1.80	-0.15	-1.13				yes	6
Synpcc7942_2590	-0.09	-1.81	0.10	-0.57	rpaA		Dusk	yes	6
Synpcc7942_2591	-0.23	-2.24	0.82	0.23	rpaA		Dusk		6

Synpcc7942_2120	-0.25	-2.61	1.51	-1.88	rpaA				6
Synpcc7942_2061	-0.04	-0.67	-0.69	-1.52			Dawn		6
Synpcc7942_0069	-2.84	-2.07	-3.30	-2.52			Dusk	yes	7
Synpcc7942_0100	-1.92	-1.29	-3.47	-2.82	rpaA	RpaB	Dusk	yes	7
Synpcc7942_0215	-0.27	-0.75	-1.18	-1.72		RpaB	Dusk	yes	7
Synpcc7942_0236	-1.11	-0.30	-2.91	-2.33		RpaB	Dusk	yes	7
Synpcc7942_0312	-2.72	-1.77	-2.64	-1.73			Dusk	yes	7
Synpcc7942_0498	-1.62	-1.80	-2.13	-1.55				yes	7
Synpcc7942_0499	-1.76	-1.11	-3.23	-1.87			Dusk	yes	7
Synpcc7942_0573	-2.20	-0.76	-1.97	-1.30			Dusk	yes	7
Synpcc7942_0680	-0.90	-1.33	-0.73	-1.97			Dusk	yes	7
Synpcc7942_0703	-2.46	-1.27	-3.40	-2.11			Dusk	yes	7
Synpcc7942_0780	-2.07	-1.70	-1.96	-1.11			Dusk	yes	7
Synpcc7942_0781	-1.14	-0.28	-3.22	-2.26			Dusk	yes	7
Synpcc7942_0836	-0.62	-0.81	-1.32	-1.83				yes	7
Synpcc7942_1040	-3.02	-2.21	-2.61	-2.34			Dusk	yes	7
Synpcc7942_1227	-2.31	-1.31	-2.99	-1.86			Dusk	yes	7
Synpcc7942_1301	-1.94	-1.47	-1.40	-1.78				yes	7
Synpcc7942_1656	-1.41	-0.55	-3.96	-3.28	rpaA		Dusk	yes	7
Synpcc7942_2082	-0.22	-1.81	-2.49	-2.95			Dusk	yes	7
Synpcc7942_2301	-1.29	-0.61	-2.21	-1.75			Dusk	yes	7
Synpcc7942_2384	-1.63	-2.02	-3.39	-2.55			Dusk	yes	7
Synpcc7942_2388	-0.69	-0.71	-2.49	-1.51	rpaA		Dusk	yes	7
Synpcc7942_2458	-1.73	-1.48	-2.39	-2.36			Dusk	yes	7
Synpcc7942_2554	-1.92	-1.25	-2.53	-2.30	rpaA		Dusk	yes	7
Synpcc7942_1572	-0.32	-2.33	-1.83	-1.62	rpaA	RpaB	Dusk		7
Synpcc7942_2254	-1.44	-1.18	-1.68	-0.58	rpaA	RpaB	Dusk		7
Synpcc7942_0571	-1.20	-0.87	-1.62	-1.03	rpaA		Dusk		7
Synpcc7942_1573	-1.68	-2.22	-1.74	-1.07	rpaA		Dusk		7
Synpcc7942_2307	-0.52	-0.38	-2.67	-2.00	rpaA		Dusk		7
Synpcc7942_2387	-1.36	-1.42	-2.03	-1.12	rpaA		Dusk		7
Synpcc7942_2553	-1.85	-1.21	-1.69	-1.56	rpaA		Dusk		7
Synpcc7942_0433	-0.84	-0.93	-2.10	-2.75	rpaA				7
Synpcc7942_1574	-1.84	-1.47	-1.01	-0.37	rpaA				7
Synpcc7942_R0024	-0.53	-0.30	-0.98	-2.11	rpaA				7
Synpcc7942_0267	-1.14	-0.67	-1.88	-1.77		RpaB	Dusk		7
Synpcc7942_0309	-0.43	-1.10	-1.99	-1.08		RpaB	Dusk		7
Synpcc7942_0768	-1.60	-1.78	Deleted	Deleted		RpaB	Dusk		7
Synpcc7942_0416	-3.18	-1.17	-3.35	-2.02		RpaB			7
Synpcc7942_1148	-1.43	-2.21	-1.80	-2.02		RpaB			7
Synpcc7942_1674	-0.38	-0.54	-1.67	-1.55			Dawn		7
Synpcc7942_2540	-0.71	-0.97	-1.34	-2.32			Dusk		7
Synpcc7942_2385	-2.48	-2.73	-1.70	-2.17			Dusk		7

Synpcc7942_1420	-1.48	-1.33	-3.02	-1.82			Dusk		7
Synpcc7942_0704	-1.81	-0.89	-1.57	-1.53			Dusk		7
Synpcc7942_2603	-0.64	-1.65	-1.98	-1.44			Dusk		7
Synpcc7942_2367	-0.84	-0.65	-2.01	-1.30			Dusk		7
Synpcc7942_0769	-1.90	-2.08	Deleted	Deleted			Dusk		7
Synpcc7942_1445	-1.51	-0.59	-1.85	-0.87			Dusk		7
Synpcc7942_2604	-1.48	-1.63	-1.61	-0.85			Dusk		7

Table S2.2: DE transcriptional regulators (putative and characterized).

Locus	description	Gene	7942: 30m	7942: 4h	2973: 30m	2973: 4h
Synpcc7942_0672	Sigma Factor	*sigA3 rpoD3*	3.51	2.90	2.02	2.74
Synpcc7942_1523	DNA-directed RNAP subunit	*rpoC1*	1.63		1.75	
Synpcc7942_1522	DNA-directed RNAP subunit	*rpoB*	1.43		1.68	
Synpcc7942_2209	DNA-directed RNAP subunit	*rpoA*	1.42		1.48	
Synpcc7942_1524	DNA-directed RNAP subunit	*rpoC2*	1.30		1.49	
Synpcc7942_1979	Membrane protein-like		1.27			1.31
Synpcc7942_1923	Sigma-24 (FecI-like)		-1.04			
Synpcc7942_2466	Two component transcriptional regulator		-1.10			
Synpcc7942_2174	Transcriptional regulator	*idiB*		2.80		4.39
Synpcc7942_1969	Transcriptional regulator	*abrB*		1.20	1.48	1.08
Synpcc7942_0987	Ferric uptake regulation protein	*fur*		1.10		
Synpcc7942_1465	Transcriptional regulator			-1.00	1.49	
Synpcc7942_1849	Sigma Factor			-3.00	-1.42	-1.37
Synpcc7942_0569	Sigma Factor	*sigA4 rpoD4*			-4.56	
Synpcc7942_2305	Response regulator	*nblR*				
Synpcc7942_2022	Transcription termination factor	*nusA*			1.10	
Synpcc7942_2170	Putative fur-like					2.18
Synpcc7942_2416	Two component transcriptional regulator					1.18
Synpcc7942_2585	Transcriptional regulator					-1.04
Synpcc7942_0095	Two component transcriptional regulator	*rpaa*				-1.07
Synpcc7942_1510	Sigma Factor					-1.26
Synpcc7942_0688	Transcriptional regulator					-1.28
Synpcc7942_0090	Transcriptional regulator				-1.01	
Synpcc7942_1978	Thioredoxin	*trx*			-1.10	
Synpcc7942_0677	Transcriptional regulator				-1.13	-1.44
Synpcc7942_1056	Transcriptional regulator	*XRE-like*			-1.19	
Synpcc7942_0321	Nitrogen regulatory PII	*glnB*			-1.24	
Synpcc7942_0599	Transcriptional regulator				-1.69	
Synpcc7942_1784	Sigma Factor				-1.71	
Synpcc7942_1242	Probable nitrogen regulator	*ntcB*			-1.76	
Synpcc7942_1557	Sigma Factor	*rpoD6*			-2.48	
Synpcc7942_0127	Global nitrogen regulator	*ntcA*			-2.60	
Synpcc7942_1969	Transcriptional regulator	*abrB*			1.48	1.08

2.7 References

Adomako, M., Ernst, D., Simkovsky, R., Chao, Y. Y., Wang, J., Fang, M., . . . Golden, S. S. (2022). Comparative Genomics of *Synechococcus elongatus* Explains the Phenotypic Diversity of the Strains. *MBio, 13*(3), e0086222. doi:10.1128/mbio.00862-22

Cameron, J. C., & Pakrasi, H. B. (2010). Essential role of glutathione in acclimation to environmental and redox perturbations in the cyanobacterium *Synechocystis* sp. PCC 6803. *Plant Physiol, 154*(4), 1672-1685. doi:10.1104/pp.110.162990

Choi, S. Y., Park, B., Choi, I. G., Sim, S. J., Lee, S. M., Um, Y., & Woo, H. M. (2016). Transcriptome landscape of *Synechococcus elongatus* PCC 7942 for nitrogen starvation responses using RNA-seq. *Sci Rep, 6*, 30584. doi:10.1038/srep30584

Dong, G., Yang, Q., Wang, Q., Kim, Y. I., Wood, T. L., Osteryoung, K. W., . . . Golden, S. S. (2010). Elevated ATPase activity of KaiC applies a circadian checkpoint on cell division in *Synechococcus elongatus*. *Cell, 140*(4), 529-539. doi:10.1016/j.cell.2009.12.042

Eisenhut, M. (2019). Manganese Homeostasis in Cyanobacteria. *Plants (Basel), 9*(1). doi:10.3390/plants9010018

Espinosa, J., Boyd, J. S., Cantos, R., Salinas, P., Golden, S. S., & Contreras, A. (2015). Cross-talk and regulatory interactions between the essential response regulator RpaB and cyanobacterial circadian clock output. *Proc Natl Acad Sci U S A, 112*(7), 2198-2203. doi:10.1073/pnas.1424632112

Firoj Mahmud, A. K. M., Delhomme, N., Nandi, S., & Fallman, M. (2020). ProkSeq for complete analysis of RNA-Seq data from prokaryotes. *Bioinformatics*. doi:10.1093/bioinformatics/btaa1063

Fleming, K. E., & O'Shea, E. K. (2018). An RpaA-Dependent Sigma Factor Cascade Sets the Timing of Circadian Transcriptional Rhythms in *Synechococcus elongatus*. *Cell Rep, 25*(11), 2937-2945 e2933. doi:10.1016/j.celrep.2018.11.049

Forcada-Nadal, A., Llacer, J. L., Contreras, A., Marco-Marin, C., & Rubio, V. (2018). The PII-NAGK-PipX-NtcA Regulatory Axis of Cyanobacteria: A Tale of Changing Partners, Allosteric Effectors and Non-covalent Interactions. *Front Mol Biosci, 5*, 91. doi:10.3389/fmolb.2018.00091

Gutu, A., & O'Shea, E. K. (2013). Two antagonistic clock-regulated histidine kinases time the activation of circadian gene expression. *Mol Cell, 50*(2), 288-294. doi:10.1016/j.molcel.2013.02.022

Hanaoka, M., & Tanaka, K. (2008). Dynamics of RpaB-promoter interaction during high light stress, revealed by chromatin immunoprecipitation (ChIP) analysis in *Synechococcus elongatus* PCC 7942. *Plant J, 56*(2), 327-335. doi:10.1111/j.1365-313X.2008.03600.x

Hood, R. D., Higgins, S. A., Flamholz, A., Nichols, R. J., & Savage, D. F. (2016). The stringent response regulates adaptation to darkness in the cyanobacterium *Synechococcus elongatus*. *Proc Natl Acad Sci U S A, 113*(33), E4867-4876. doi:10.1073/pnas.1524915113

Hosokawa, N., Hatakeyama, T. S., Kojima, T., Kikuchi, Y., Ito, H., & Iwasaki, H. (2011). Circadian transcriptional regulation by the posttranslational oscillator without de novo clock gene expression in *Synechococcus*. *Proc Natl Acad Sci U S A, 108*(37), 15396-15401. doi:10.1073/pnas.1019612108

Ibrahim, I. M., Rowden, S. J. L., Cramer, W. A., Howe, C. J., & Puthiyaveetil, S. (2022). Thiol redox switches regulate the oligomeric state of cyanobacterial Rre1, RpaA and RpaB response regulators. *FEBS Lett, 596*(12), 1533-1543. doi:10.1002/1873-3468.14340

Iijima, H., Shirai, T., Okamoto, M., Kondo, A., Hirai, M. Y., & Osanai, T. (2015). Changes in primary metabolism under light and dark conditions in response to overproduction of a response regulator RpaA in the unicellular cyanobacterium *Synechocystis* sp. PCC 6803. *Front Microbiol, 6*, 888. doi:10.3389/fmicb.2015.00888

Imamura, S., & Asayama, M. (2009). Sigma factors for cyanobacterial transcription. *Gene Regul Syst Bio, 3*, 65-87. doi:10.4137/grsb.s2090

Ito, H., Mutsuda, M., Murayama, Y., Tomita, J., Hosokawa, N., Terauchi, K., . . . Iwasaki, H. (2009). Cyanobacterial daily life with Kai-based circadian and diurnal genome-wide transcriptional control in *Synechococcus elongatus*. *Proc Natl Acad Sci U S A, 106*(33), 14168-14173. doi:10.1073/pnas.0902587106

Joshua, S., & Mullineaux, C. W. (2005). The rpaC gene product regulates phycobilisome-photosystem II interaction in cyanobacteria. *Biochim Biophys Acta, 1709*(1), 58-68. doi:10.1016/j.bbabio.2005.06.005

Kadowaki, T., Nagayama, R., Georg, J., Nishiyama, Y., Wilde, A., Hess, W. R., & Hihara, Y. (2016). A Feed-Forward Loop Consisting of the Response Regulator RpaB and the Small RNA PsrR1 Controls Light Acclimation of Photosystem I Gene Expression in the Cyanobacterium *Synechocystis* sp. PCC 6803. *Plant Cell Physiol, 57*(4), 813-823. doi:10.1093/pcp/pcw028

Kamarainen, J., Huokko, T., Kreula, S., Jones, P. R., Aro, E. M., & Kallio, P. (2017). Pyridine nucleotide transhydrogenase PntAB is essential for optimal growth and photosynthetic integrity under low-light mixotrophic conditions in *Synechocystis* sp. PCC 6803. *New Phytol, 214*(1), 194-204. doi:10.1111/nph.14353

Kato, N., Iwata, K., Kadowaki, T., Sonoike, K., & Hihara, Y. (2022). Dual Redox Regulation of the DNA-Binding Activity of the Response Regulator RpaB in the Cyanobacterium *Synechocystis* sp. PCC 6803. *Plant Cell Physiol, 63*(8), 1078-1090. doi:10.1093/pcp/pcac079

Kolowrat, C., Partensky, F., Mella-Flores, D., Le Corguille, G., Boutte, C., Blot, N., . . . Garczarek, L. (2010). Ultraviolet stress delays chromosome replication in light/dark synchronized cells of the marine cyanobacterium Prochlorococcus marinus PCC9511. *BMC Microbiol, 10*, 204. doi:10.1186/1471-2180-10-204

Koskinen, S., Hakkila, K., Gunnelius, L., Kurkela, J., Wada, H., & Tyystjarvi, T. (2016). In vivo recruitment analysis and a mutant strain without any group 2 sigma factor reveal roles of different sigma factors in cyanobacteria. *Mol Microbiol, 99*(1), 43-54. doi:10.1111/mmi.13214

Labella, J. I., Obrebska, A., Espinosa, J., Salinas, P., Forcada-Nadal, A., Tremino, L., . . . Contreras, A. (2016). Expanding the Cyanobacterial Nitrogen Regulatory Network: The GntR-Like Regulator PlmA Interacts with the PII-PipX Complex. *Front Microbiol, 7*, 1677. doi:10.3389/fmicb.2016.01677

Liu, D., & Yang, C. (2014). The nitrogen-regulated response regulator NrrA controls cyanophycin synthesis and glycogen catabolism in the cyanobacterium *Synechocystis* sp. PCC 6803. *J Biol Chem, 289*(4), 2055-2071. doi:10.1074/jbc.M113.515270

Markson, J. S., Piechura, J. R., Puszynska, A. M., & O'Shea, E. K. (2013). Circadian control of global gene expression by the cyanobacterial master regulator RpaA. *Cell, 155*(6), 1396-1408. doi:10.1016/j.cell.2013.11.005

Piechura, J. R., Amarnath, K., & O'Shea, E. K. (2017). Natural changes in light interact with circadian regulation at promoters to control gene expression in cyanobacteria. *Elife, 6*. doi:10.7554/eLife.32032

Puszynska, A. M., & O'Shea, E. K. (2017). Switching of metabolic programs in response to light availability is an essential function of the cyanobacterial circadian output pathway. *Elife, 6*. doi:10.7554/eLife.23210

Riediger, M., Kadowaki, T., Nagayama, R., Georg, J., Hihara, Y., & Hess, W. R. (2019). Biocomputational Analyses and Experimental Validation Identify the Regulon Controlled by the Redox-Responsive Transcription Factor RpaB. *iScience, 15*, 316-331. doi:10.1016/j.isci.2019.04.033

Robles-Rengel, R., Florencio, F. J., & Muro-Pastor, M. I. (2019). Redox interference in nitrogen status via oxidative stress is mediated by 2-oxoglutarate in cyanobacteria. *New Phytol, 224*(1), 216-228. doi:10.1111/nph.15979

Sacharz, J., Bryan, S. J., Yu, J., Burroughs, N. J., Spence, E. M., Nixon, P. J., & Mullineaux, C. W. (2015). Sub-cellular location of FtsH proteases in the cyanobacterium *Synechocystis* sp. PCC 6803 suggests localised PSII repair zones in the thylakoid membranes. *Mol Microbiol, 96*(3), 448-462. doi:10.1111/mmi.12940

Seki, A., Hanaoka, M., Akimoto, Y., Masuda, S., Iwasaki, H., & Tanaka, K. (2007). Induction of a group 2 sigma factor, RPOD3, by high light and the underlying mechanism in

Synechococcus elongatus PCC 7942. *J Biol Chem, 282*(51), 36887-36894. doi:10.1074/jbc.M707582200

Sharon, S., Salomon, E., Kranzler, C., Lis, H., Lehmann, R., Georg, J., . . . Keren, N. (2014). The hierarchy of transition metal homeostasis: iron controls manganese accumulation in a unicellular cyanobacterium. *Biochim Biophys Acta, 1837*(12), 1990-1997. doi:10.1016/j.bbabio.2014.09.007

Srivastava, A., Summers, M. L., & Sobotka, R. (2020). Cyanobacterial sigma factors: Current and future applications for biotechnological advances. *Biotechnol Adv, 40*, 107517. doi:10.1016/j.biotechadv.2020.107517

Srivastava, V., Amanna, R., Rowden, S. J. L., Sengupta, S., Madhu, S., Howe, C. J., & Wangikar, P. P. (2021). Adaptive laboratory evolution of the fast-growing cyanobacterium *Synechococcus elongatus* PCC 11801 for improved solvent tolerance. *J Biosci Bioeng, 131*(5), 491-500. doi:10.1016/j.jbiosc.2020.11.012

Takahashi, Y., Yamaguchi, O., & Omata, T. (2004). Roles of CmpR, a LysR family transcriptional regulator, in acclimation of the cyanobacterium *Synechococcus* sp. strain PCC 7942 to low-CO(2) and high-light conditions. *Mol Microbiol, 52*(3), 837-845. doi:10.1111/j.1365-2958.2004.04021.x

Tan, X., Hou, S., Song, K., Georg, J., Klahn, S., Lu, X., & Hess, W. R. (2018). The primary transcriptome of the fast-growing cyanobacterium *Synechococcus elongatus* UTEX 2973. *Biotechnol Biofuels, 11*, 218. doi:10.1186/s13068-018-1215-8

Turunen, O., Koskinen, S., Kurkela, J., Karhuvaara, O., Hakkila, K., & Tyystjarvi, T. (2022). Roles of Close Homologues SigB and SigD in Heat and High Light Acclimation of the Cyanobacterium *Synechocystis* sp. PCC 6803. *Life (Basel), 12*(2). doi:10.3390/life12020162

Ungerer, J., Lin, P. C., Chen, H. Y., & Pakrasi, H. B. (2018). Adjustments to Photosystem Stoichiometry and Electron Transfer Proteins Are Key to the Remarkably Fast Growth of the Cyanobacterium *Synechococcus elongatus* UTEX 2973. *MBio, 9*(1). doi:10.1128/mBio.02327-17

Ungerer, J., Wendt, K. E., Hendry, J. I., Maranas, C. D., & Pakrasi, H. B. (2018). Comparative genomics reveals the molecular determinants of rapid growth of the cyanobacterium *Synechococcus elongatus* UTEX 2973. *Proc Natl Acad Sci U S A*. doi:10.1073/pnas.1814912115

Verma, N., & Prasad, S. M. (2021). Interplay of hydrogen peroxide and nitric oxide: systemic regulation of photosynthetic performance and nitrogen metabolism in cadmium challenged cyanobacteria. *Physiol Mol Biol Plants, 27*(10), 2181-2199. doi:10.1007/s12298-021-01083-2

Walker, P. L., & Pakrasi, H. B. (2022). A Ubiquitously Conserved Cyanobacterial Protein Phosphatase Essential for High Light Tolerance in a Fast-Growing Cyanobacterium. *Microbiol Spectr, 10*(4), e0100822. doi:10.1128/spectrum.01008-22

Watzer, B., Spat, P., Neumann, N., Koch, M., Sobotka, R., Macek, B., . . . Forchhammer, K. (2019). The Signal Transduction Protein PII Controls Ammonium, Nitrate and Urea Uptake in Cyanobacteria. *Front Microbiol, 10*, 1428. doi:10.3389/fmicb.2019.01428

Wendt, K. E., Walker, P., Sengupta, A., Ungerer, J., & Pakrasi, H. B. (2021). Engineering natural competence into the fast-growing cyanobacterium *Synechococcus elongatus* UTEX 2973. *Appl Environ Microbiol*, AEM0188221. doi:10.1128/AEM.01882-21

Yasuda, A., Inami, D., & Hanaoka, M. (2020). RpaB, an essential response regulator for high-light stress, is extensively involved in transcriptional regulation under light-intensity upshift conditions in *Synechococcus elongatus* PCC 7942. *J Gen Appl Microbiol, 66*(2), 73-79. doi:10.2323/jgam.2020.01.010

Yu, J., Liberton, M., Cliften, P. F., Head, R. D., Jacobs, J. M., Smith, R. D., . . . Pakrasi, H. B. (2015). *Synechococcus elongatus* UTEX 2973, a fast growing cyanobacterial chassis for biosynthesis using light and CO(2). *Sci Rep, 5*, 8132. doi:10.1038/srep08132

Zhao, L. S., Li, C. Y., Chen, X. L., Wang, Q., Zhang, Y. Z., & Liu, L. N. (2022). Native architecture and acclimation of photosynthetic membranes in a fast-growing cyanobacterium. *Plant Physiol*. doi:10.1093/plphys/kiac372

Chapter 3: A Ubiquitously Conserved Cyanobacterial Protein Phosphatase Essential for High Light Tolerance in a Fast-Growing Cyanobacterium

This Chapter was adapted, with permission from:
Walker PL and Pakrasi HB. (2022). A ubiquitously conserved cyanobacterial protein phosphatase essential for high light tolerance in a fast-growing cyanobacterium. Microbiology Spectrum. DOI: 10.1128/spectrum.01008-22.

Chapter contributions: All data presented was generated by PLW

3.1 Summary

Synechococcus elongatus UTEX 2973, the fastest-growing cyanobacterial strain known, optimally grows under extreme high-light (HL) intensities of 1500-2500 μmol photons m^{-2} s^{-1}, which is lethal to most other photosynthetic microbes. We leveraged the few genetic differences between *Synechococcus* 2973 and the HL-sensitive strain *Synechococcus elongatus* PCC 7942 to unravel factors essential for the high light tolerance. We identified a novel protein in *Synechococcus* 2973 that we have termed HltA for High light tolerance protein A. Using bioinformatic tools, we determined that HltA contains a functional PP2C-type protein phosphatase domain. Phylogenetic analysis showed that the PP2C domain belongs to the bacterial-specific Group II family and is closely related to the environmental stress response phosphatase RsbU. Additionally, we showed that unlike any previously described phosphatases, HltA contains a single N-terminal regulatory GAF domain. We found *hltA* to be ubiquitous throughout cyanobacteria, indicative of its potentially important role in the photosynthetic lifestyle of these oxygenic phototrophs. Mutations in the *hltA* gene resulted in severe defects specific to high light growth. These results provide evidence that *hltA* is a key factor in the tolerance of *Synechococcus* 2973 to high light and will open new insights into the mechanisms of cyanobacterial light stress response.

Cyanobacteria are a diverse group of photosynthetic prokaryotes. The cyanobacterium *Synechococcus* 2973 is a high light tolerant strain with industrial promise due to its fast growth under high light conditions and the availability of genetic modification tools. Currently, little is known about the high light tolerance mechanisms of *Synechococcus* 2973 and there are many unknowns overall regarding high light tolerance of cyanobacteria. In this study, a comparative genomic analysis of *Synechococcus* 2973 identified a single nucleotide polymorphism in a locus

encoding a serine phosphatase as a key factor for high light tolerance. This novel GAF-containing phosphatase was found to be the sole Group II metal-dependent protein phosphatase that is evolutionarily conserved throughout cyanobacteria. These results shed new light on the light response mechanisms of *Synechococcus* 2973, improving our understanding of environmental stress response. Additionally, this work will help facilitate the development of *Synechococcus* 2973 as an industrially useful organism.

3.2 Introduction

Photosynthesis requires light energy, but all photosynthetic organisms have a limited capacity for light utilization. As a result, excess absorbed light causes increased production of reactive oxygen species (ROS) within the photosynthetic machinery, leading to severe damage to photosystem II (PSII) (Aro et al., 1993). The accumulation of photodamage reduces photosynthetic capacity, referred to as photoinhibition. Maximum photosynthetic efficiency is achieved when all light energy absorbed is utilized for carbon assimilation. Under this optimal condition, the rate of PSII damage is lower than the rate of PSII repair and photoinhibition does not occur (Allakhverdiev et al., 2005, Aro et al., 1993). Because avoiding prolonged photoinhibition is vital to survival, organisms have evolved numerous mechanisms to detect changes in light intensity, prevent photodamage, and repair damage to PSII (Derks et al., 2015). Although the photosynthetic reaction centers are identical throughout oxygenic photosynthetic organisms, many cyanobacteria have developed adaptation strategies that allow them to survive under high-light (HL) conditions (Allakhverdiev et al., 2005, Aro et al., 1993, Bhaya et al., 2002).

An important mechanism to prevent excess absorbed energy from causing photoinhibition is to reduce light absorption by decreasing the amount of chlorophyll and

phycobilisome (PBS) and by lowering the PSI/PSII stoichiometry (Ogawa et al., 2018b, Latifi et al., 2009). Another initial response to prevent photodamage and acclimate to increased light is by removing excess energy through quenching. Non-photochemical quenching (NPQ) can occur through the orange carotenoid protein (OCP) as well as state transitions which redirect light energy between PSI and PSII (Derks et al., 2015). Quenching can also involve high-light-inducible proteins (HLIPs) (Latifi et al., 2009, Daddy et al., 2015). However, since photodamage to PSII occurs under over-saturating light, an essential method to overcome photoinhibition is to repair PSII (Daddy et al., 2015). ROS are inadvertently produced from the energy transfer in PSII, and the accumulation of ROS prevents PSII repair (Allakhverdiev et al., 2005, Kale et al., 2017). To allow for PSII repair, cyanobacteria have evolved a variety of antioxidant molecules such as carotenoids to eliminate ROS (Latifi et al., 2009).

Cyanobacteria have developed various regulatory mechanisms to respond to high-intensity light, with many of these acclimation mechanisms mediated at the transcriptional level. One aspect that is not well understood of the stress response pathways is how stress is detected. Some cyanobacteria sense light through GAF (for cGMP-specific phosphodiesterase, adenylyl cyclase, and FhlA) domain-containing photoreceptors that bind bilin pigments (Montgomery and Lagarias, 2002). The enrichment of these photoreceptors is critical for acclimation to HL through mediation of signaling pathways that regulate HLIP, sigma factors, and pigments (Cho et al., 2015, Ogawa et al., 2018b, Daddy et al., 2015). Having a variety of signal transduction networks allows for detection and response to environmental changes. Members of the Mn^{2+}/Mg^{2+}-dependent protein phosphatase family known as PPM or PP2C provide rapid and reversible post-translational modifications that play important roles in environmental response (Zhang and Shi, 2004). In gram-positive bacteria, environmental stress activates the stressosome complex that

leads to the activation of the PP2C family phosphatase RsbU, a regulator of sigma factor SigB, to bring about stress response that enhances survival (Yang et al., 1996, Pane-Farre, 2005 #4542, Pane-Farre et al., 2005). Further, regulation of phosphorylation levels of the cyanobacterial photosynthetic apparatus and metabolic enzymes was recently found to be pivotal for responding to environmental conditions (Toyoshima et al., 2020). Yet, little is known about the identity and physiological role of phosphatases in cyanobacteria.

Many of the known HL response mechanisms are accompanied by decreased growth rates. This highlights the need to study the adaptive mechanisms evolved by HL tolerant species which grow optimally under HL conditions. *Synechococcus elongatus* UTEX 2973 (*S.* 2973) is one of the fastest-growing cyanobacteria, growing optimally under extreme HL up to 2,400 μmol photons m^{-2} s^{-1} (Yu et al., 2015b, Ungerer et al., 2018b) with a high photosynthetic efficiency (Clark et al., 2018). Both *S.* 2973 and other HL tolerant species are found to upregulate PSI (Dobson et al., 2021, Ungerer et al., 2018b) as well as electron carriers in response to HL, thus increasing photosynthetic capacity. *S.* 2973 shares over 99% genome similarity with the model strain *Synechococcus elongatus* PCC 7942 (*S.* 7942), yet the strains have distinct HL responses. Similar HL conditions result in fast growth of *S.* 2793 but decreased viability of *S.* 7942. Intriguingly, these strains do not encode an OCP, so mechanisms of NPQ remain unknown. The total genetic differences between these strains are comprised of 53 single nucleotide polymorphisms (SNPs), one large inversion, and a small deletion (Yu et al., 2015b). We hypothesized that these genetic differences induce differential physiological responses that alter HL tolerance. Through the comparison of these two strains, the genetic markers for differential phenotypic traits can be identified. Using a CRISPR/Cas12a genome editing system, 36 mutants in *S.* 2973 containing substituted alleles from *S.* 7942 were constructed without antibiotic

markers (Ungerer and Pakrasi, 2016a). Previous analysis of these SNPs uncovered key genes for fast growth and natural competency (Ungerer et al., 2018c, Wendt et al., 2021).

In this work, we report on the finding of an uncharacterized protein, which we term HltA, that is vital for HL tolerance in *S.* 2973. Bioinformatic characterization highlighted a resemblance of the C-terminal domain with the PP2C family phosphatase domain of RsbU and showed that the PP2C domain retains the conserved residues and structure necessary for its function. We found that the N-terminal signaling domain does not have strong homologs and that this combination of domains is unique to HltA. We further showed that this protein is conserved throughout cyanobacteria. Moreover, inactivation of *hltA* led to impaired viability during HL growth. Thus, *hltA* has an essential role in cyanobacterial high-light tolerance.

3.3 Results

3.2.1 Identification of high light sensitive mutants

To elucidate alleles responsible for the differential light tolerance between *S.* 7942 and *S.* 2973, a series of previously generated *S.* 2973 CRISPR/Cas12 mutant lines (Ungerer et al., 2018c) were assayed for their HL tolerance. These mutant lines contained *S.* 2973 alleles modified to the *S.* 7942 versions and a *S.* 2973 small plasmid deletion strain. In total, 36 mutant strains (Table S3.1) were assayed for a loss of HL tolerance through dilution series spot plates and verified for HL_L (1500 μmol photons m^{-2} s^{-1}) sensitivity in liquid growth assays. The dilution spot plate assays were more sensitive at detecting slight changes to HL tolerance as there was no shading of cells allowing HL_P exposure to remain consistent. Because of the shading and movement that occurred in liquid cultures, a higher light intensity was used to test HL tolerance when assaying liquid cultures. HL screening of all mutants led to the identification of three alleles of the most significant SNPs that individually decreased viability under HL. The

corresponding genes encoded HltA, a hypothetical protein, and CTP synthetase (Fig 3.1a), respectively. In liquid cultures, all three mutants grew at similar rates under medium light (ML_L, 500 µmol photons m^{-2} s^{-1}) conditions (Fig 3.1b), suggesting that these mutations did not affect normal cell growth. In liquid culture, two of the three mutants grew significantly slower than the wild-type strain under HL (Fig. 3.1c). The SNP with the most significant effect was in the putative serine phosphatase locus M744_03855, which will be referred to as *hltA* (High light tolerance protein A). The *S.* 2973 *hltA* SNP mutant ($hltA_{7942}$) contains an Arg, as in *S.* 7942, in place of the native Cys, as the 35th residue of the encoded protein (Fig. 3.2a).

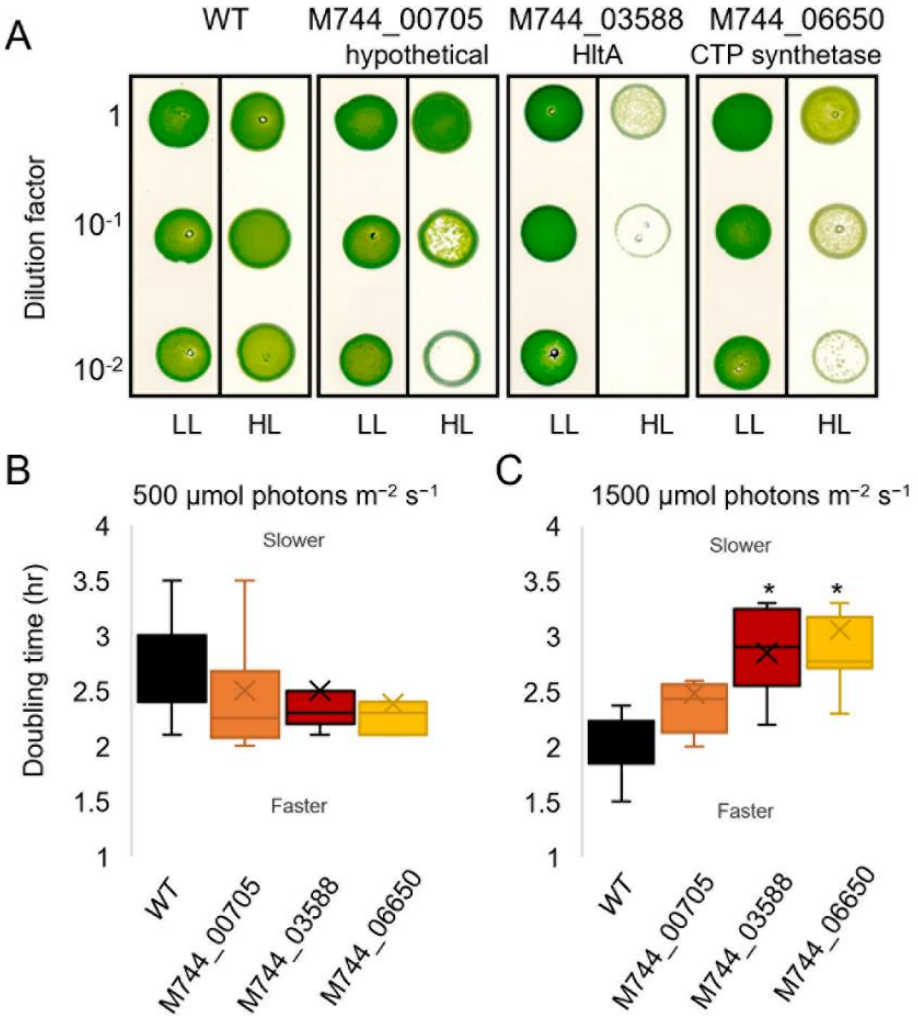

Figure 3.1 *S.* 2973 allele reversion mutants decrease high light tolerance. (A) High light sensitivity screening plate assay performed on wild type and single allele reversion mutants of select loci. The representative plate of n≥3 shown; HL_P, 500 μmol photons m^{-2} s^{-1}, and LL_P, 50 μmol photons m^{-2} s^{-1}. (B) Box plot representation of doubling times of wild-type *S.* 2973 and single SNP point mutants at ML_L, 500 μmol photons m^{-2} s^{-1} and (C) HL_L, 1500 μmol photons m^{-2} s^{-1}. Mean doubling time calculated from at least four experiments. The asterisk indicates p-value = 0.01-0.05, from Dunnett's multiple comparisons test for each strain compared to *S.* 2973. Box plots indicate median (line), mean (x), and first and fourth quartile (whiskers) of doubling times.

3.2.2 Sequence analysis of HltA

M744_03588 has been annotated as a guanylate cyclase (NCBI: AJD57037.1) as well as a serine phosphatase (Uniprot: Q8GIR9) in different databases, as previously noted (Yu et al., 2015b). To understand the nature of this protein, we conducted a series of analyses to examine its sequence features, dissect the domain components, and establish relationships with other well characterized proteins. The *hltA* gene is flanked by two ORFs that are transcribed in the same direction (Fig. 3.5a), although in a previous analysis of the transcriptional start site, the gene did not appear to be in an operon (Tan et al., 2018). The HltA protein has a predicted molecular mass of 51.1 kDa with 462 residues. Among them, residues 227-461 form a PP2C-family SpoIIE domain [Pfam: PF07228]. Remarkably, in contrast to any previously characterized PP2C, HltA also contains a GAF domain near the N terminus (residues 59-200) (Fig. 3.2a and b). A BLASTP homology search of HltA revealed that it shared the highest degree of similarity (>50%) to sequences in other cyanobacteria. To search for homologs outside of the cyanobacterial clade, a BLASTP search excluding cyanobacteria was conducted that yielded 38 uncharacterized bacterial homologs, with no hits from eukaryotic organisms (Table S1.2). The closest characterized sequence is the PP2C-family phosphatase RsbU from *Bacillus subtilis*, which shares 27.8% sequence identity. However, RsbU does not have a GAF domain. We next

modelled the structure using AlphaFold (Jumper et al., 2021), which predicts protein structures with high accuracy even without known similar structures (Fig. 3.2b).

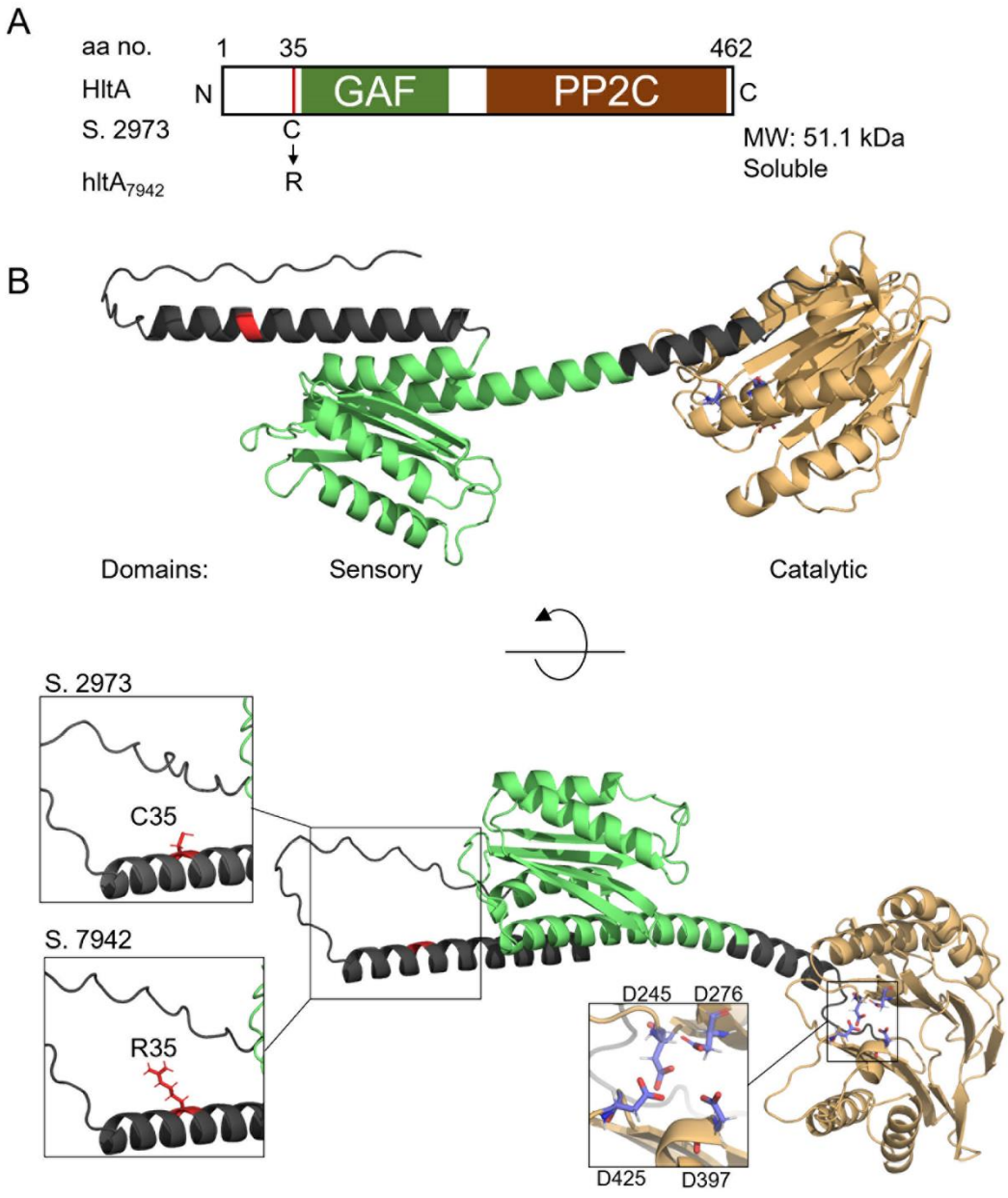

Figure 3.2 The modeled structure of the HltA protein. (A) Schematic diagram of the domain structure with SNP variation and characteristics shown below. (B) Ribbon diagram of the predicted structures of *Synechococcus* HltA protein, generated using AlphaFold (Jumper et al.,

2021). Location of the SNP shown in red, GAF domain shown in green, PP2C domain shown in light brown, active site Asp residues involved in metal binding shown in violet. Left insets highlight the Cys and Arg SNPs at residue 35, right inset shows the positioning of putative metal coordinating residues of the active site.

3.2.3 HltA is a PP2C-family phosphatase

Members of the PP2C family are characterized by the presence of 11 signature motifs that correspond to the structure and function of the catalytic domain (Bork et al., 1996). HltA has all four metal-binding Asp residues (Zhang and Shi, 2004, Shi, 2009) (Fig. 3.2b) within the 11 conserved motifs (Fig. 3.3a). The number of bound metal ions is a differentiating characteristic of the Group I and Group II PP2Cs. Members of Group I such as tPphA bind three metal ions, whereas SpoIIE, a Group II member, binds two metal ions (Su et al., 2011). We observed that HltA shares some important residues with tPphA, including a Met in motif 2 which supports the active site, and the second Asp in motif 8 involved in stabilizing metal binding (Fig. 3.3a). Notably, HltA belongs to Group II (Fig. 3.3b), and the absence of a Gly in motif 5 suggests that it, like other Group II members, cannot bind a third metal ion. The predicted structure of HltA displays a conserved composition of secondary structural elements composed of 10 β sheets and 5 α helices (Fig. 3.2b and Fig. 3.3a) observed in the active site of SpoIIE (Levdikov et al., 2012). The structure also revealed that HltA possesses an additional α helix (α0) preceding the PP2C core domain (Fig. 3.3a and c), a structural feature that may function as a regulatory module to control its phosphatase activity (Brody et al., 2009, Bradshaw et al., 2017).

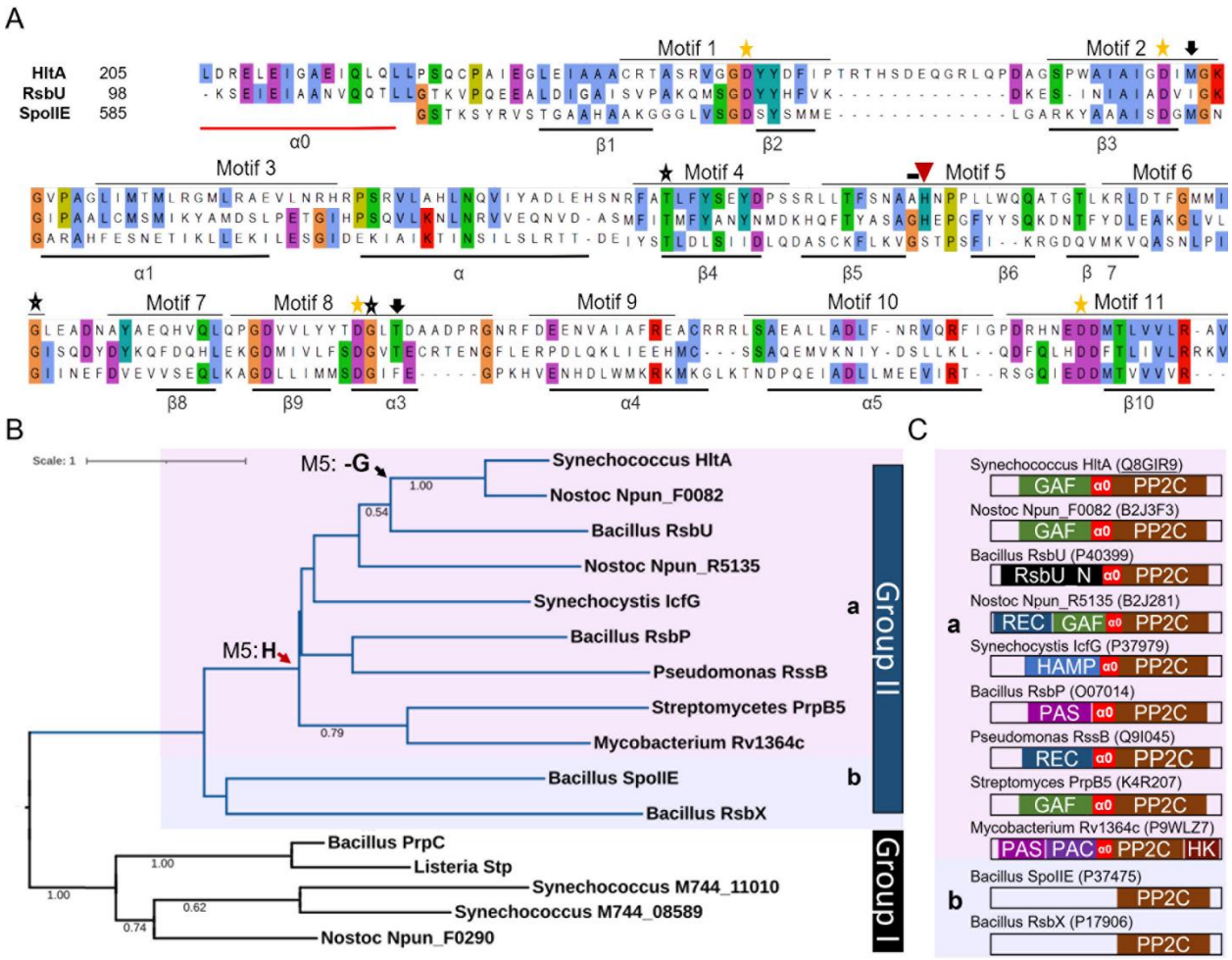

Figure 3.3 HltA is a Group IIa bacterial PP2C phosphatase. (A) Multiple sequence alignment of the 11 conserved PP2C motifs and preceding α0, of HltA from *S.* 2973, and RsbU and SpoIIE from *B. subtilis*. Stars indicate highly conserved residues across PP2Cs; yellow stars denote essential metal-binding residues. Black arrows indicate residues that may support catalytic activity, the red arrow indicates conserved residue in Group IIa sequences. Underlined sequences indicate predicted secondary structures of HltA, labeled according to structure of SpoIIE (Levdikov et al., 2012). Identical residues are highlighted. Secondary structures below the alignment represent properties from SpoIIE and are supported by PSIPRED structural predictions for HltA. (B) Maximum likelihood tree of *hltA* and bacterial PP2C family phosphatases. PP2C domain sequences were aligned with MUSCLE and used to infer the maximum likelihood protein tree. Branch color indicates PP2C group: pink highlight for Group IIa, blue highlight for Group IIb. The red arrow indicates conserved H in motif 5. The tree was generated using MEGA-X with LG+G+I model, supporting values from 1000 bootstrap greater than 0.5 are shown on branches. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. (C) Schematic representation of domain architectures (SMART and Pfam domains) of Group II PP2Cs, with Uniprot ID. Protein domains shown are GAF (SM00065), PP2C (SM00331), HAMP (SM00304), HatPase Histidine Kinase (HK, SM000387), Receiver (Rec, SM00448), PAS (SM000091), PAC (SM000086). α0 linker between sensory

domains and PP2C domain identified using Alphafold. Domain architectures are not drawn to scale.

To explore relationships between HltA and previously characterized proteins, we searched for structurally related proteins using JPRED4, which resulted in several hits in the PP2C family: RssB [PDB:3f7a] (e-value $4e^{-09}$), and Rv1364c [PDB:3ke6] (e-value $1e^{-06}$). We note that despite the significant degree of structural conservation, the level of sequence identity was low (~24.8%). To evaluate similarity of domains that may lead to functional inferences, PP2C domain sequences from *Synechococcus* 2973, *Synechocystis* 6803, *Nostoc punctiforme*, *Streptomyces sp.* and previously characterized PP2Cs were used to construct a phylogenetic tree. The maximum likelihood tree in Fig. 3.3b shows separation of the Group I eukaryotic-like PP2Cs that contain the motifs 5b and 5c and Group II bacterial-specific PP2Cs (Zhang and Shi, 2004), which reflects the functional differences between these groups. Within Group II, we observed two subclades, one of which, Group IIa, contains a known conserved His residue (H348) within motif 5 (Kerk et al., 2015), indicated by a red arrow in Fig. 3.3a and b. Additionally, we found that HltA and the HltA ortholog Npun_F0082 lack a Gly. The replacement of Gly with Ala (A347) in motif 5 was observed in all cyanobacterial HltA orthologs. As seen in the phylogenetic analysis, the closest characterized phosphatase to the PP2C domain of HltA is RsbU (Fig. 3.3b). Strains of *Synechococcus elongatus* encode three probable PP2C-type phosphatases, none of which have been previously characterized. The two other putative PP2C-type phosphatases belong to Group I and do not contain any sensory domain (Fig. 3.3b and c). Together, our analysis suggests that the HltA phosphatase domain likely functions similarly to RsbU and differs in functionality from the other PP2Cs in *S.* 2973.

3.2.4 HltA is ubiquitous in cyanobacteria

To evaluate the conservation and evolutionary relationship of HltA in cyanobacteria, we searched for *hltA* orthologs in 410 cyanobacterial genomes, available in the Integrated Microbial Genome database (IMG: https://img.jgi.doe.gov). We selected orthologs with sequence similarity spanning both domains, >50% identity, with a 0.01 e-value cut off. We identified orthologs in 400 genomes, including the non-photosynthetic UCYN-A and the most primitive cyanobacterium *Gloeobacter violaceous* PCC 7421. The ortholog sequences from the remaining 10 genomes mostly comprised of *Prochlorococcus* and *Synechococcus*. All genomes contained a single copy of the *hltA* ortholog. In two *Arthrospira* strains, *hltA* appeared to be misannotated as two separate genes, separating the GAF and PP2C domains with only one start and one stop codon between the two. We selected 75 ortholog sequences from morphological and ecologically diverse cyanobacterial strains for phylogenetic analysis to investigate the relationship of *hltA* orthologs across cyanobacteria. The distribution of the resulting *hltA* tree topology maintained a high degree of similarity with the core cyanobacterial phylogeny (Shih et al., 2013, Will et al., 2019, Moya et al., 2020) shown by the preservation of distinct clades A-H (Fig. 3.4a, and Fig. S3), with little divergence in clade distribution. The most notable topological difference is of clade C (Fig. 3.4b), with the presence of subclades c1 and c2 at the bottom of the tree suggesting that these sequences are the most divergent ones. Clade C is predominantly composed of *Prochlorococcus* and *Synechococcus,* which are the fastest evolving lineages and have undergone significant genome reductions (Will et al., 2019, Moya et al., 2020). Additionally, a single species, *Geitlerinema sp.* PCC 7407, did not group together with its proposed clade (clade D).

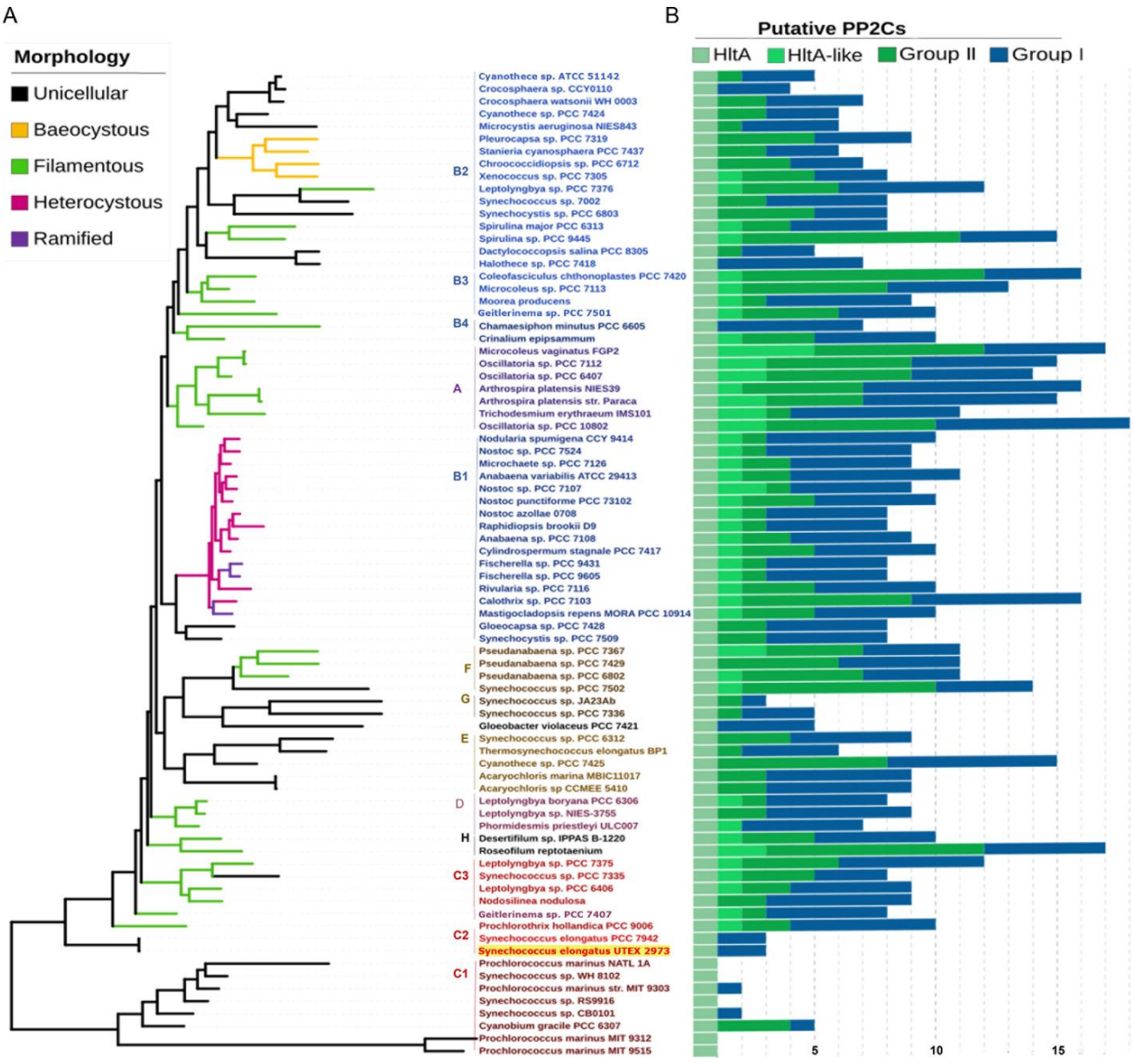

Figure 3.4 HltA is conserved throughout cyanobacteria. (A) Maximum likelihood phylogeny of *hltA* across cyanobacteria and distribution of additional HltA-like proteins. Branches are color-coded according to the morphological trait. Taxa names are colored by phylogenetic species clades (Shih et al., 2013). (B) The number of putative PP2C genes present in each genome is divided into four categories. Group II PP2Cs in green, with HltA orthologs, HltA-like sequences (which include PP2C sequences with at least one GAF domain) subcategories, and Group I PP2Cs in blue. Sequence homology search was carried out using the IMG database and aligned with MUSCLE with ICFG and RsbU outgroups (not shown). The tree was created using MEGAX software, bootstrapped 1000 replicates. Outgroups and bootstrap values are shown in Fig. S3.

We have hypothesized that the aberrant *hltA* clades could be attributed to Group II PP2C duplication events that could allow for sequence variability to occur. To assess this, we identified the number of putative PP2Cs in each genome and categorized these sequences based on Group I or Group II features, and more specifically Group II sequences encoding one or more GAF domains. In this analysis, we observed that HltA is the only PP2C-family phosphatase with orthologs in all cyanobacteria, shown by the light green *hltA* bar in all genomes (Fig. 3.4b). Interestingly, the possible Group II duplication events did not coincide with the aberrant clades, but rather with cellular morphologies. Unicellular species in subclades c1 and c2 predominantly encode a single copy of a Group II PP2C, the HltA ortholog. In contrast, subclades c3 and D contain filamentous species that encode multiple PP2C proteins with several copies of PP2Cs containing GAF domains (Fig. 3.4b). Filamentous and heterocystous cyanobacteria have a greater number of Group II PP2C genes, due to genome expansion events (Moya et al., 2020) and sensory domain recruitment to PP2Cs that possibly assisted in adaptation to different environments (Zhang and Shi, 2004). Rather, the divergent *hltA* clades associated closer with genome complexity. Clades c1, b2, A and *Geitlerinema sp.* PCC 7407 of clade D all have higher genome complexity, measured by sequence compositional complexity, which resulted from trends towards driven progressive evolution (Moya et al., 2020). The overall conservation of clades and presence of *hltA* throughout cyanobacteria suggest that its function is necessary and was established early in cyanobacterial evolution.

3.2.5 HltA is essential for the high-light tolerance of *S.* 2973

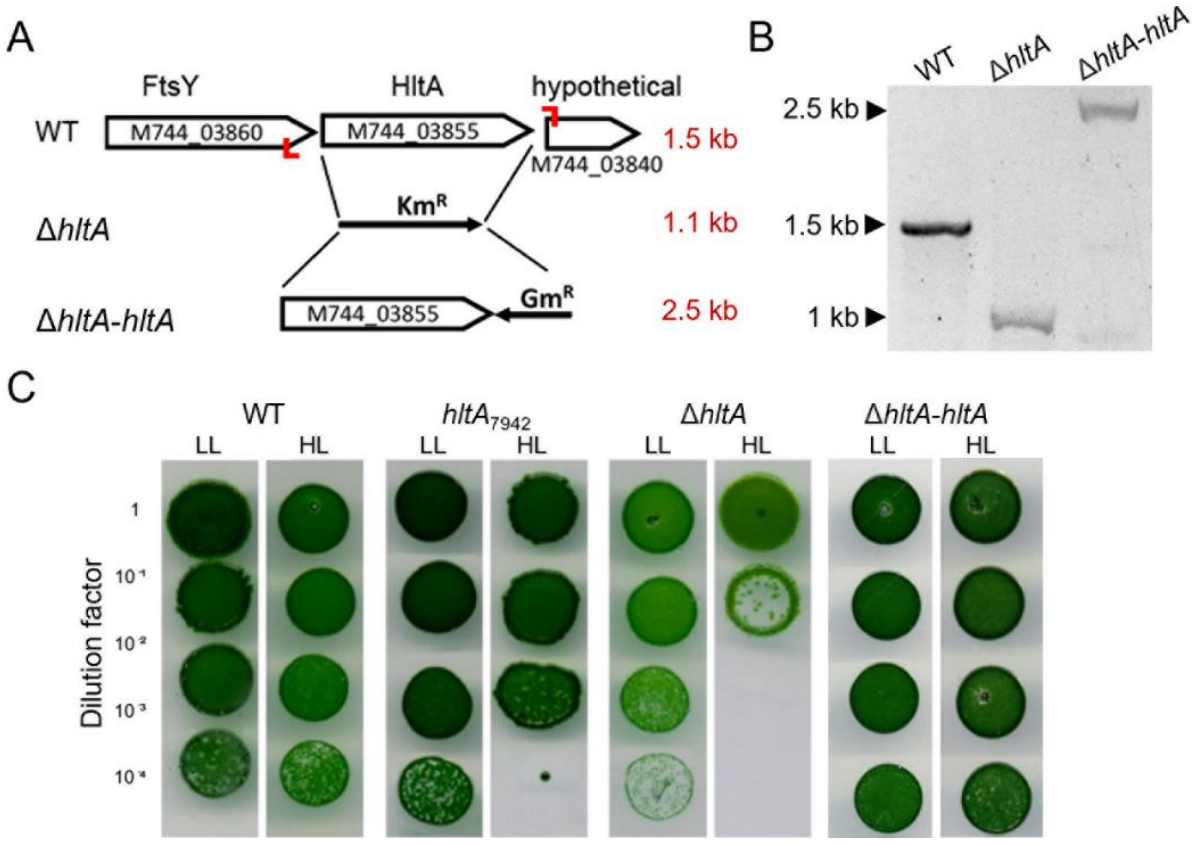

Figure 3.5 Mutations to hltA cause sensitivity to high light in S. 2973. (A) Organization of hltA and flanking DNA region containing M744_03860 (FtsY), M744_03855 (HltA), and M744_03840 (hypothetical protein), the position of the inserted kanamycin resistance cassette (KmR) to generate the ΔhltA strain, and replacement of KmR with hltA and gentamycin resistance cassette (GmR) to generate the complementation strain ΔhltA-hltA. Red bars indicate the location of primers used in panel B; red text indicates the size of the PCR product. (B) Deletion and complementation of hltA in the ΔhltA and ΔhltA-hltA strains were tested by PCR. (C) Growth of WT and hltA mutants from HL dilution spot plate assays. Strains were grown on plates with their respective antibiotic resistance. HLP (500 µmol photons m−2 s−1), LLP (50 µmol photons m−2 s−1).

To verify that HltA is important for HL growth, we engineered a *hltA* deletion strain by replacing the full-length *hltA* with a Kan^R cassette (Fig. 3.5a). Since *S.* 2973 maintains multiple genome copies, the replacement of *hltA* was checked with PCR primers set outside the homology arms (Table S3) and growth on antibiotic plates. The *hltA* mutant obtained after transformation was partially segregated and the wild-type gene could still be detected by PCR. Repeated

subculturing under higher antibiotic concentrations was performed, resulting in the *hltA* deletion mutation (Fig. 3.5b). This deleted *hltA* mutant strain (Δ*hltA*) exhibited HL sensitivity more severe than the $hltA_{7942}$ mutant (Fig. 3.5c). To confirm that the sensitivity to HL was caused by the *hltA* mutation, we introduced the full-length *hltA* gene back into the Δ*hltA* strain (Fig. 3.5a) and examined the complemented strain for the restoration of HL_P tolerance. As shown in Fig. 3.5c, the complemented strain has regained its HL tolerance, pointing to the critical nature of this gene for HL adaptation.

To further understand the mechanisms of HltA, we assessed photosynthesis parameters of the mutant in response to HL treatment. Wild-type and *hltA* mutant cultures subjected to ML_L and HL_L treatments were monitored for differences in pigment levels recorded by the absorption spectra, and for changes in energy transfer between PBS and photosystems. The absorbance and fluorescence emission spectra exhibited no significant difference between the mutants and wild type strains (Fig. S3.1a-c. Additionally, quantum yield efficiency of PSII showed no difference between these strains in either light condition (Fig. S1d). These results indicate that HltA is not directly involved in regulating photosynthetic capacity.

3.2.6 HltA is specific to high light stress response

To determine whether HltA affects more general stress pathways, we challenged the wild type and *hltA* mutants with various stress conditions. The mutants did not exhibit growth defects as compared to the wild type under various nutrient-deprived or high salinity conditions (Fig. S3.2). All studies were performed at 38°C for of the comparison with *S.* 7942, which grows best at this temperature, although the optimum fast-growth temperature for *S.* 2973 is 42°C (Ungerer et al., 2018b). Thus, we examined growth at 42°C under HL_L, ML_L, and LL_L (50 μmol photons m^{-2} s^{-1}) conditions. Under HL_L, both mutants exhibited acute growth defects, with the Δ*hltA*

strain losing viability, while under ML_L only the Δ*hltA* mutant exhibited minor sensitivity (Fig. 3.6a and b). Since photoinhibition still occurs at the ML_L condition, growth at 42°C under a LL_L condition was necessary to isolate the effects of the increased temperature. Under LL_L, we observed no differences in growth between the *hltA*$_{7942}$ and wild-type strains, whereas the Δ*hltA* strain did not exhibit growth defects for the first two days of treatment (Fig. 3.6a and c). Since high temperature amplified the detrimental effects of the HL but was not harmful to the *hltA* mutants on their own, our analysis demonstrated that HltA is primarily involved in alleviating high-light stress.

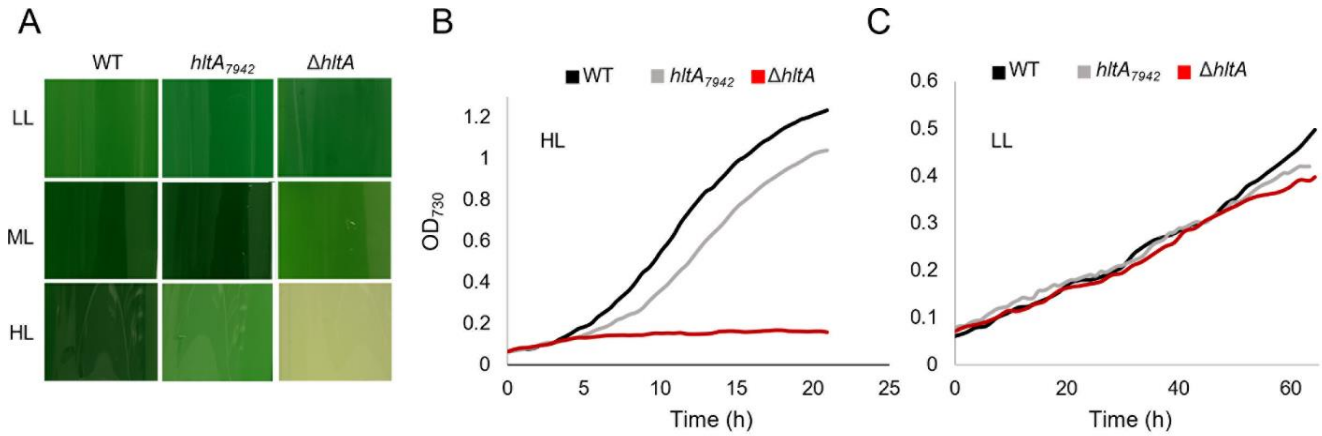

Figure 3.6 *hltA* mutant sensitivity is high-light specific. (A) Phenotype of wild-type and hltA strains, grown at 42°C under different light conditions for 24 hours. Cultures were grown with 1% CO2. (B) Growth curve of wild-type and mutant strains grown under HL_L (1500 µmol photons m-2 s-1) and 42°C. (C) Growth of wild-type and mutant strains under LL_L (50 µmol photons m-2 s-1) and 42°C. Data shown represents three replicates.

3.4 Discussion

S. 2973 optimally grows at some of the highest light conditions recorded for photosynthetic organisms (Ungerer et al., 2018c). Despite the few genetic differences between *S.* 2973 and the HL sensitive *S.* 7942, only a few SNPs had been investigated for links to phenotypic differences. We set out to identify polymorphic alleles in *S.* 2973 to discover

unexplored mechanisms of HL tolerance. We identified *hltA* to be essential for HL growth in the fast-growing *S.* 2973, as demonstrated by decreased colony formation and cell viability upon mutations to *hltA* (Fig. 3.1 and 3.5). In addition to these severe growth defects there were also slight defects observed under low-light conditions (Fig. 3.5c and 3.6). Given our finding that HltA is present in all cyanobacteria, including HL sensitive species, it is evident that HltA is necessary for tolerance to light at all intensities. Further, this decrease in growth was light specific and was not observed in any other general stress conditions. Therefore, we suggest *hltA* may be a key component to light response mechanisms specific to cyanobacteria.

Cyanobacterial phosphatases have been widely left unstudied in contrast with their cognate protein kinases, despite the observation that most proteins detected under different light conditions in cyanobacteria are phosphorylated on Ser or Thr residues (Zhang and Shi, 2004). The RsbU phosphatase regulates alternative sigma factor B activity, which controls over 150 stress response genes (Rodriguez Ayala et al., 2020). The similarity between *hltA* and *rsbU* (Fig. 3.3) suggests that *hltA* may regulate sigma factor activity. While neither *Synechococcus elongatus* strain encodes homologs of the RsbU protein target RsbV, analysis of the *hltA* neighborhood did not reveal any target candidates. Homologs to the stressosome are not present in *Synechococcus* and *hltA* does not share identity with the N-terminal RsbU N domain that detects input from the stressosome pathway. Therefore, it is likely that HltA is part of a distinct environmental response pathway with regulation through the GAF domain. GAF domains are extremely widespread, known to bind to small molecules like cyclic nucleotides or chromophores, and take part in signal transduction pathways (Rumyantsev et al., 2015, Montgomery and Lagarias, 2002). GAF domains are used by phytochromes to sense light through bound bilin pigments (Yoshihara et al., 2000, Aravind and Ponting, 1997), and are

involved in plastoquinone redox sensing that can signal photosynthetic imbalances (Ibrahim et al., 2020). Additionally, it is evident that the GAF domain of HltA is important, as demonstrated from the widespread conservation of the GAF-PP2C fusion that we observed throughout cyanobacteria. Analysis of the signals that activate the HltA GAF domain would provide insights into the regulation of HltA activity.

hltA expression has been found to exhibit weak circadian oscillation (Ito et al., 2009) and was identified as a target of the master regulator of the circadian clock, RpaA (Markson et al., 2013a). Therefore, it is plausible that *hltA* functions as a part of the circadian clock. To test circadian growth, cultures were entrained to a 12h light/ 12h dark cycle for three cycles, then exposed to constant light to observe the free running rhythm growth. We found that *hltA* mutations did not affect growth either during light/dark entrainment or following continuous light growth (Fig. S2). While *hltA* may not contribute to fitness under light-dark conditions, the transcriptional data does suggest its expression to be a function of both time-of-day and light intensity-dependent response (Piechura et al., 2017). The homology to RsbU supports a role for HltA in a partner-switching system that regulates stress response gene expression through an alternative sigma factor. Given the various evidence, we present a speculative model (Fig. 3.7) in which RpaA controls *hltA* expression (Markson et al., 2013a), with HltA then sensing changes to light conditions via its N-terminal GAF domain. The light-induced activation stimulates dephosphorylation of the target, possibly the putative sigma factor B antagonist M744_05770, which shares 28% identity with RsbV. The HltA target can then bind to and sequester the anti-sigma factor kinase, freeing the alternative sigma factor to promote transcription of light stress response genes. In addition, activated HltA may indirectly block its expression in a negative feedback loop. This is supported by data showing the *hltA* gene product increases under shade

conditions but decreases following high light pulse (Piechura et al., 2017), as HltA is no longer needed. The model suggested here outlines directions for future studies investigating the partners and mechanisms of HltA as well as its role in the circadian clock.

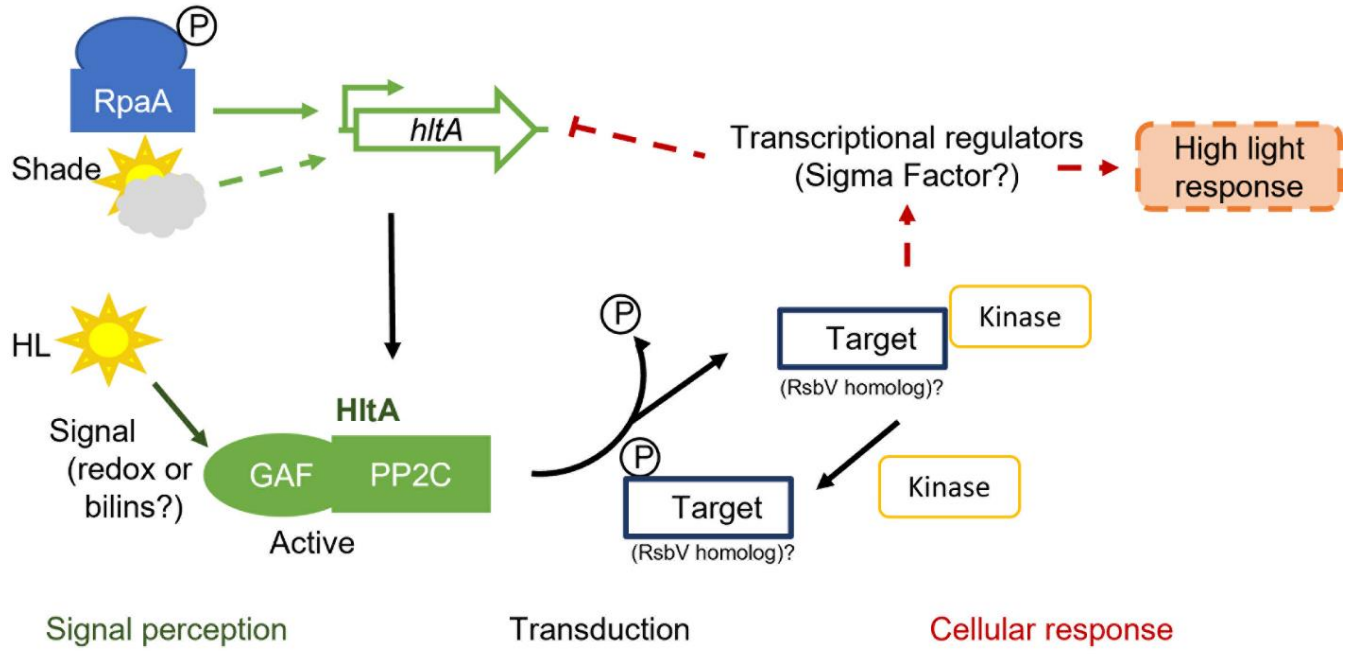

Figure 3.7: A model proposing regulation and function of HltA for *Synechococcus*. Phosphorylated RpaA controls hltA expression at dusk. Increases in light are detected by the GAF domain of HltA, which activates the HltA protein. Activated HltA de-phosphorylates a target, possibly the RsbV homolog M744_05770 (NCBI: WP_011244281), which controls downstream regulation leading to high light response through unknown mechanisms, indirectly repressing expression of the hltA gene. An unknown kinase works in opposition to HltA, phosphorylating the HltA target. Phosphorylation events are indicated by "P". Arrows indicate activation, T-bar represents inhibition. Solid and dashed lines indicate direct and indirect or unknown regulation, respectively.

We show that HltA belongs to the Group Inia phosphatases, a group that contains few characterized proteins (Fig. 3.3). Most PP2Cs have not been characterized and many mechanisms of regulation are unknown (Shi, 2009). The majority of crystalized bacterial PP2Cs, such as tPphA and MtPstP (Su et al., 2011), belong to Group I, which contain a third metal-

binding site and lack regulatory domains. In contrast, many Group II members are regulated by a diverse set of N-terminal domains (Shi, 2009). A PP2C with a lone N-terminal GAF domain has yet to be characterized. Further functional characterization of *hltA* is a promising avenue for advancing our understanding of Group II protein phosphatases.

The observation that PP2Cs are not found in all bacteria, together with the presence of the eukaryotic-like Group I PP2Cs in bacteria, led to the hypothesis that the PP2C family arose in eukaryotes and was acquired by certain prokaryotes through horizontal gene transfer (HGT) (Zhang and Shi, 2004). This notion was challenged through phylogenetic analysis that suggested that PP2Cs originated in bacteria before inheritance of PP2Cs by eukaryotes, followed by eukaryotic-like Group I PP2Cs moving back into bacteria via HGT (Kerk et al., 2015). Our phylogenetic analysis suggests *hltA* was fixed in the cyanobacterial genome before cyanobacterial clade divergence and therefore preceded eukaryotic PP2Cs. This is supported by the finding that *hltA* is part of the core gene set in many cyanobacterial genomes (Shi and Falkowski, 2008, Larsson et al., 2011). Genes involved in adaptation to different light intensities are believed to have been fixed in the core genome before cyanobacterial clades diverged (Kettler et al., 2007). Transcriptomic evidence from *Nodularin spumigena* shows *hltA* ortholog expression clustered with genes that responded differently to environmental variables and showed enrichment for functional categories: environmental adaptation and transcription (Asplund-Samuelsson et al., 2016). Together these observations suggest that HltA is a core component of an environmental adaptation process conserved in cyanobacteria.

Based on the observations in this study, we have shown that HltA is a novel factor that facilitates a light-stress response. The single SNP adaptation to this protein in *S.* 2973 allows this strain to grow optimally under HL intensity. This SNP proceeds the predicted start of the N-

terminal GAF domain, which may have been co-opted as a light-recognition module early during the evolution of cyanobacteria. Our study shows that *hltA* is maintained throughout cyanobacteria and presents an opportunity for further exploration of its physiological role in light response. *Synechococcus* encodes a single putative RsbV homolog (M744_05770) and 10 serine protein kinases (7 of which are putative). Functional studies to elucidate the partners of HltA in *Synechococcus* will lead to a greater understanding of high light tolerance and post-translational modification networks in cyanobacteria.

3.5 Materials and methods

Cell growth and treatment

The wild-type *Synechococcus elongatus* UTEX 2973 and constructed mutant strains in this study (listed in Table SA1) were grown as described in (Ungerer et al., 2018c) unless otherwise indicated. Cultures were maintained on BG11 agar plates at 38°C, 150 µmol photons m^{-2} s^{-1}. To determine the effect of high-light stress conditions on gene expression, cells were grown to mid-logarithmic growth phase (optical density at 730 nm of 0.6-0.8) in an MC-1000 multicultivator (Photon Systems Instruments, Czech Republic), and diluted with fresh medium to an OD 730 of ~0.3 before exposure to high light (1500 µmol photons m^{-2} s^{-1}) for 30 minutes, 1 hour, and 2 hours, when cells were harvested. Growth experiments were repeated at least three times to confirm the growth patterns. Doubling times were calculated by fitting an exponential curve to the logarithmic section of the growth data (typically OD_{720} of <0.3) and using the slope, *m*, as *K′* ($y = ke^{mx}$). Doubling times were then calculated as ln(2)/*K′*. Mean doubling times were compared to *S.* 2973 using a One-Way ANOVA and Dunnett's multiple comparisons test.

Spot plate assays

Plates spotted with cultures with a starting OD (730) of 0.3- 0.35, in 5µl spots, and grown for 4 days on BG11 plates, under low light (50 µmol photons m^{-2} s^{-1}) or high light (500 µmol photons m^{-2} s^{-1}) at 38 °C with 0.6% CO_2 in a Caron plant growth chamber (Caron Products & services, Inc. Marietta, OH).

Structural prediction

Three-dimensional protein structural predictions were derived using AlphaFold (Jumper et al., 2021), accessed using AlphaFold Colab (https://colab.research.google.com/github/deepmind/alphafold/blob/main/notebooks/AlphaFold.ipynb). Predicted secondary structures were also checked using HHpred, which queries the Pfam, PDB, and SCOP databases. Structural visualization and alignments used the PyMOL Molecular Graphics System (Schrödinger, LLC).

Sequence analysis of *hltA* and genomic context

Cyanobacterial genomes with finished and permanent draft genomes were compiled in the IMG database. *hltA* ortholog sequences were identified from the genome cart using the full-length M744_03855 sequence as our query. One representative gene sequence containing homology to both domains (>70% coverage on query gene) was selected per genome. The obtained candidate *hltA* homologs from each genome were further inspected for Interpro and Pfam conserved domain architecture (GAF and PP2C). Additional BLAST searches were executed using the standard parameters on NCBI for any cyanobacterial genomes of interest not found on IMG. Search for non-cyanobacterial homologs was performed with NCBI BLASTP search excluding cyanobacteria with an >70% query coverage cut off, to cover homology to both domains.

Sequence alignment and phylogenetic analysis of HltA

Genomes for phylogenetic analysis were selected from (Will et al., 2019), *hltA* orthologs from these genomes were aligned by multiple sequence alignment using MUSCLE with default settings. The phylogenetic tree was constructed in MEGAX by maximum likelihood method and LG model (Le and Gascuel, 2008). The tree with the highest log-likelihood was selected. Initial trees for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the JTT model and then selecting the topology with superior log likelihood value. A discrete Gamma distribution was used to model evolutionary rate differences among sites. The rate variation model allowed for some sites to be evolutionarily invariable. The tree was drawn to scale with branch lengths measured in the number of substitutions per site. Evolutionary analyses were conducted in MEGAX (Kumar et al., 2018). The final dendrogram was visualized on the Interactive Tree of Life (iTOL) (Letunic and Bork, 2021). Conserved domains were identified using SMART (Simple Modular Architecture Research Tool).

To identify putative PP2C genes, UniProt (http://www.uniprot.org/), and JGI Genome Portal (http://genome.jgi.doe.gov/) were used to search against cyanobacterial genomes used in the phylogenetic analysis. Putative PP2C gene counts were searched individually using JGI pfam search using keywords: Protein phosphatase 2C-like, SpoIIE, and validated through searches on UniProt.

Construction of *hltA* deletion and complementation strains

For the deletion of the *hltA* gene, *KanR* from PVZ321 and PUC backbone from PAM3103 (Wu et al., 2007) were amplified for Gibson assembly was performed using primers (Table S3.3). Upstream and downstream genes of *hltA* were amplified from genomic DNA using

primers (Fragment 2.FOR and REV, and Fragment 4.FOR and REV). Gibson assembled fragments (pUC and kanR) to construct the depletion strain (Table S3.3).

For the complementation of the *hltA* gene, the entire coding region of M744_03855 was amplified by PCR using the *hltA* forward and reverse primers (Table S3.3). UP, PUC backbone, and DS were amplified from the deletion plasmid (primers HltA_DS and HltA_US) and assembled with the full-length *hltA* and gentamicin resistance cassette.

The resulting exiting vector plasmids were checked by sequencing and used to conjugate into *Synechococcus* 2973, and ex-conjugants were selected with 40 μg/ml of kanamycin or 20 μg/ml gent. The transformants were obtained and passed several times on BG11 plates supplemented with their appropriate antibiotic (gent concentration of 5-10 μg/mL, kanamycin 10-20 μg/mL) to achieve segregation. Mutants were verified by amplifying a region of chromosomal DNA with primers located outside of the homology region of the editing plasmids. Confirmed depletion mutants were used to conjugate with the complementation plasmid and selected for with gentamicin. All the primers used for cloning and plasmid construction are listed in Table S3.3.

Functional validation of wild type and *hltA* mutants under various stress conditions

Liquid culture assays were carried out to ascertain the function of the mutated *hltA* and wild-type gene under various abiotic stress treatments (salt, nutrient, heat). Wild-type and mutant *S.* 2973 were grown for 4 days in BG11 media. Cultures were diluted and incubated at 38°C for up to four days under various abiotic stress conditions (high salt, and nutrient deprivation). For 42°C growth, cultures were diluted and allowed to adapt to Multicultivator conditions to an absorbance of (OD_{720}) 0.5-0.6, then diluted once again to (OD_{720}) 0.1 before incubating at

42°C with bubbling CO_2 and various light intensities. Photos were taken after treatment. Stress tolerance was determined with respect to control cultures.

Room temperature fluorescence kinetics

The fluorescence parameter F_v/F_m (maximum quantum yield of PSII) was calculated as the equation of $F_v/F_m = (F_m - F_0)/F_m$, where F_0 is the minimum fluorescence, F_v is the variable fluorescence, and Fm is the maximum fluorescence. Cultures were dark adapted for 3 minutes at room temperature before measured using a double-modulation fluorescence fluorometer, FL-200 (Photon Systems Instruments, Brno, Czech Republic). The instrument contained red LEDs for both actinic (20-μs) and measuring (2.5-μs) flashes and was used in the time range of 100 μs to 10 s.

Fluorescence and absorption spectroscopy

The fluorescence emission spectra of phycobilins and chlorophyll from whole cells of each strain were measured at 77 K on a Fluoromax-2 fluorometer (Jobin Yvon, Longjumeau, France). Excitation occurred at 580 nm and 435nm respectively, and fluorescence emission was recorded between 600 nm and 750 nm and normalized to the readings at 730 nm. Whole-cell absorbance was measured on an Olis DW-2000 spectrophotometer, and data were analyzed with Olis Globalworks software (On-Line Instrument Systems, Bogart, GA). All spectra were normalized at 730 nm to correct for differences in light scattering.

3.6 Supporting information

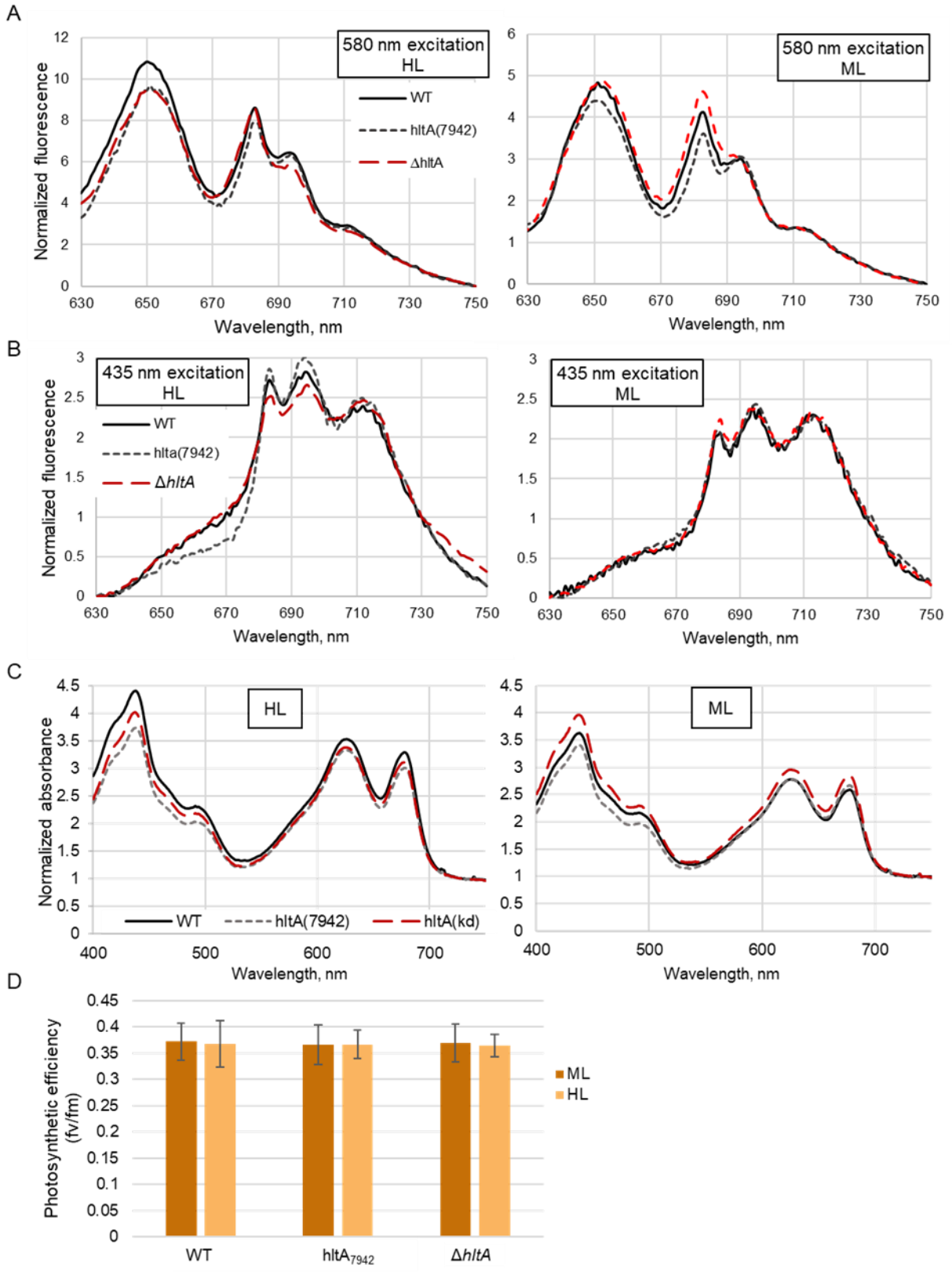

Figure S3.1: Photosynthetic phenotype and spectra analysis of *hltA* mutant under increasing light. (A-B) Low temperature (77K) fluorescence spectra of WT and *hltA* mutants cultivated under ML_L (500 µmol photons m^{-2} s^{-1}) and HL_L (1500 µmol photons m^{-2} s^{-1}) for 4 hours, normalized to 720 nm. (A) Excitation at 580 nm for PBS, peaks 650-660 nm indicate energy transfer to phycocyanin, 685 nm relates to PSII, peak 695 nm relates to PSI. (B) Excitation of chlorophyll at 435 nm show PSI corresponding to 715 nm, and PSII to 685 nm and 695 nm peaks. (C) Comparison of whole-cell absorbance spectra grown in ML_L or HL_L. To facilitate comparisons, spectra were normalized to 720 nm. Spectra shown are representative data from three biological replicates. (D) Flash induced chlorophyll fluorescence decay kinetics analysis for WT and *hltA* mutants. Bars represent mean Fv/Fm ratio ± standard deviation from three independent replicates.

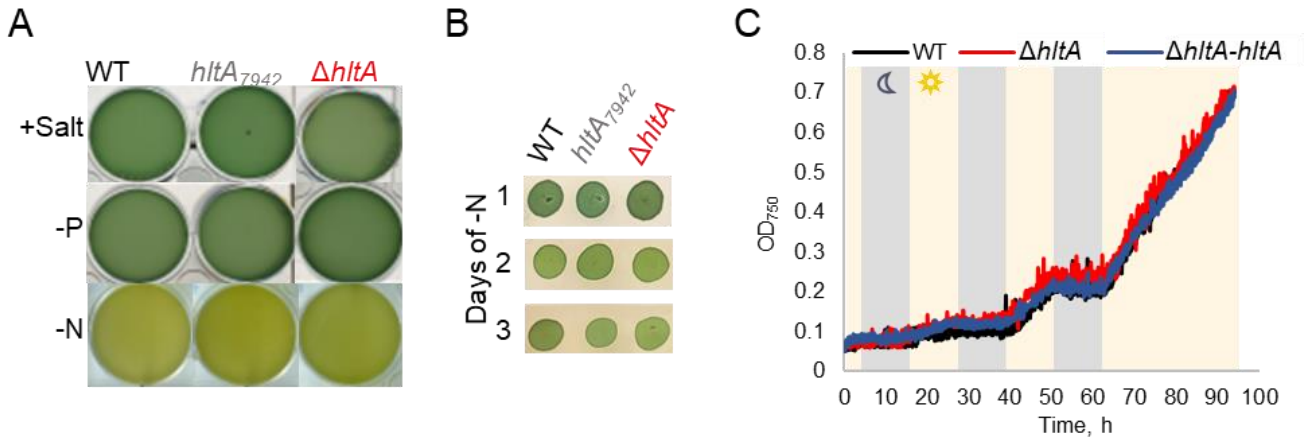

Figure S3.2: Mutations to HltA do not affect general stress growth or recovery. Non-light environmental stress analysis of *hltA* mutants. (A) Phenotype of cultures after 4 days of high salt, nitrogen starvation, and phosphorus starvation, grown in 6-well plates. (B) Recovery on BG11 plates after 1-3 days of nitrogen-depletion stress. (C) Growth curves of wild-type and *hltA* strains in 12 hr light/12 hr dark conditions followed by constant light. Light conditions were 75 µmol photons m^{-2} s^{-1}, with bubbling 1% CO_2. (A-B) Experiments were performed at low light with ambient air.

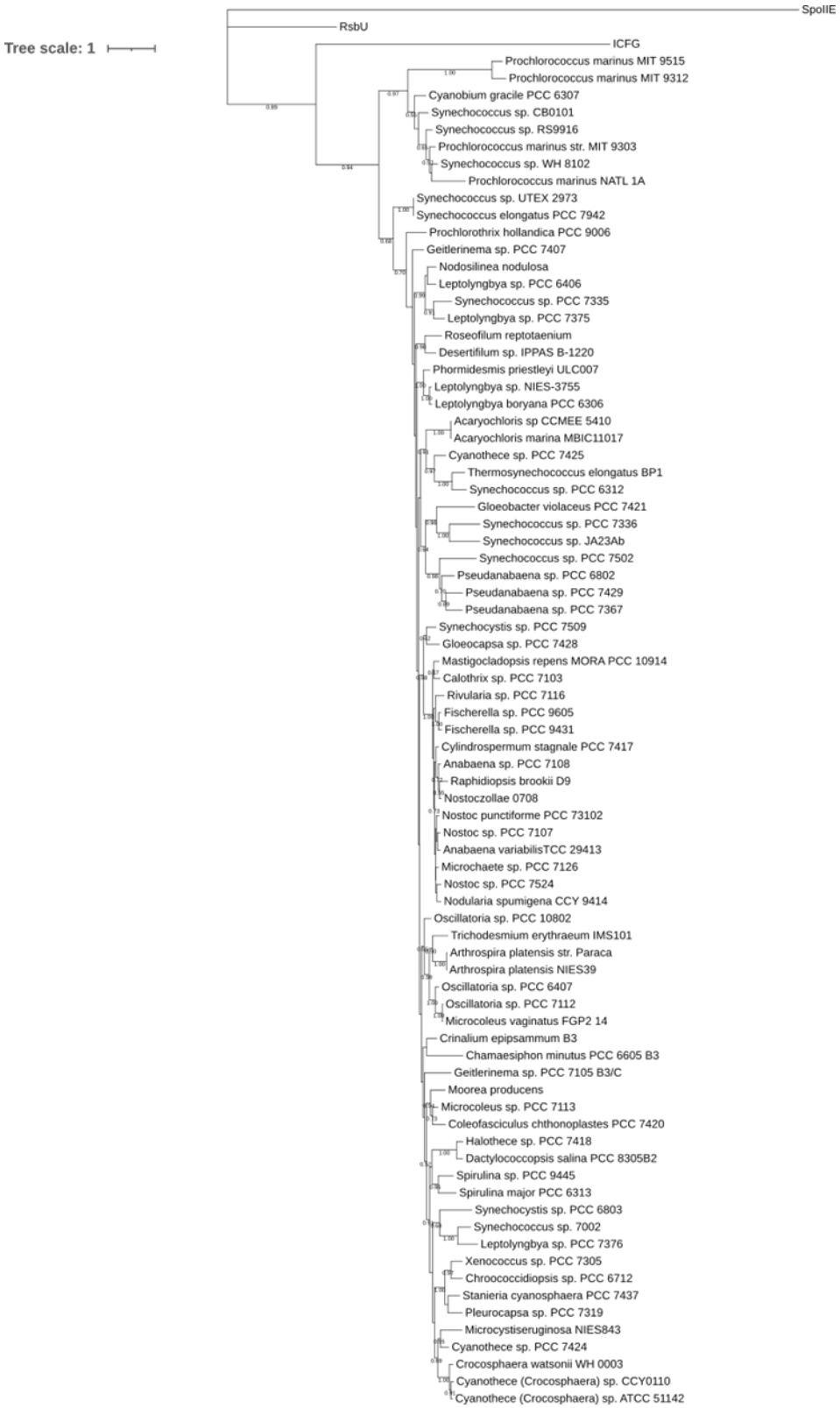

Figure S3.3: Cyanobacterial tree with bootstrap values (n=1000) and outgroups.

Table S3.1: Strains used in this study

Nucleotide position	Locus tag	Gene function/ Annotation	Strain Name	Source
2973-WT	—	Control		Yu, Liberton et al. 2015 (Yu et al., 2015b)
pUH24	—	Cure 2973 of plasmid	JU161	Ungerer, 2018 (Ungerer et al., 2018c)
126938	M744_00705	Hypothetical protein	Ju162	Ungerer, 2018 (Ungerer et al., 2018c)
236706	M744_1335	ATP synthase F0F1 alpha	Ju129	Ungerer, 2018 (Ungerer et al., 2018c)
474883	—	Noncoding 101 bp upstream of *rpaA*	JU212	Ungerer, 2018 (2)
475352	M744_2605	*rpaA*, SNP I, circadian response regulator	Ju144	Ungerer, 2018 (2)
475390	M744_2605	*rpaA*, SNP II, circadian response regulator	Ju144	Ungerer, 2018 (2)
610804	M744_3335	Manganese ABC transporter ATP-binding	JU177	Ungerer, 2018 (2)
705129	M744_03855	Guanylate cyclase/ Serine phosphatase/ HltA	$hltA_{7942}$	Ungerer, 2018 (2)
891346	M744_4780	PpnK – NAD+ kinase	JU175	Ungerer, 2018 (2)
1042862	—	Noncoding, 362 bp upstream of *ycfA*	JU214	Ungerer, 2018 (2)
1080351	M744_5865	Hypothetical protein	Ju165	Ungerer, 2018 (2)
1113358	M744_6025	Molecular chaperone DnaK	Ju145	Ungerer, 2018 (2)
1222741	M744_6570	Hydrolase	Ju146	Ungerer, 2018 (2)
1237531	M744_6650	CTP synthetase	Ju179	Ungerer, 2018 (2)
1238113	M744_6650	CTP synthetase	Ju181	Ungerer, 2018 (2)
1273424	M744_6850	Chorismate mutase	JU155	Ungerer, 2018 (2)
1533430	—	Noncoding, 120 bp upstream of COP23	JU216	Ungerer, 2018 (2)
1619501	M744_8615	DNA-directed RNA polymerase β subunit	JU127	Ungerer, 2018 (2)
1718274	M744_11685	Anthranilate synthase	JU154	Ungerer, 2018 (2)

1741647	—	Noncoding, downstream of two genes	JU218	Ungerer, 2018 (2)
2139224	—	Noncoding, 130 bp upstream of histone like protein	JU220	Ungerer, 2018 (2)

Table S3.2: HltA homologs in non-cyanobacteria

Scientific Name	Metagenome?	Max Score	Total Score	Query Cover	E value	Per. ident	Acc. Len	Accession
Tatlockia sp.	yes	514	514	94%	5.00E-177	60.5	458	MBA2750123.1
Chloroflexaceae bacterium	yes	493	493	93%	1.00E-168	57.08	474	NJL83976.1
Propionibacteriaceae bacterium	yes	428	428	93%	4.00E-143	50.68	469	MAR52998.1
Gemmatimonadaceae bacterium	yes	414	414	93%	5.00E-138	52.18	436	MBC8122912.1
Anaerolineae bacterium	yes	411	411	94%	5.00E-137	51.26	428	MBC7881039.1
Acidobacteriaceae bacterium	yes	212	212	93%	2.00E-56	33.48	879	NUQ27114.1
Acidobacteria bacterium	yes	208	208	93%	4.00E-55	33.04	879	MBW8747969.1
Granulicella sp. S190	-	207	207	91%	7.00E-55	34.38	874	WP_158945906.1
Terriglobus albidus	-	206	206	93%	2.00E-54	32.6	879	WP_147650223.1
Acidobacteria bacterium	yes	202	202	92%	4.00E-54	31.47	628	MBV9957074.1
Acidobacteria bacterium 13_1_40CM_4_58_4	yes	197	197	90%	6.00E-54	33.41	427	OLC97355.1
Bryobacterales bacterium	yes	204	204	92%	8.00E-54	33.48	862	MBL8228477.1
Acidobacteria bacterium	yes	198	198	92%	5.00E-53	31.92	581	PYS47485.1
Edaphobacter lichenicola	-	196	196	91%	8.00E-51	32.74	824	WP_179636654.1
Chloroflexi bacterium	yes	193	193	92%	9.00E-51	32.65	656	MBM3135102.1
Edaphobacter lichenicola	-	196	196	91%	1.00E-50	32.59	874	MBB5329045.1
Edaphobacter lichenicola	-	195	195	91%	2.00E-50	30.49	876	WP_183789934.1
Edaphobacter lichenicola	-	195	195	91%	2.00E-50	33.63	876	WP_179581468.1
Edaphobacter modestus	-	195	195	94%	2.00E-50	31.42	874	WP_130418875.1
Edaphobacter lichenicola	-	194	194	91%	3.00E-50	32.88	882	WP_183812128.1
Acidobacteriia bacterium	yes	194	194	92%	5.00E-50	32.68	869	MBZ5535594.1
Acidobacteriia bacterium	yes	193	193	94%	8.00E-50	31.6	877	MBZ5640185.1

Candidatus Poribacteria bacterium	yes	186	186	93%	1.00E-49	30.09	450	**RKY01516.1**
Edaphobacter lichenicola	-	192	192	91%	3.00E-49	32.81	877	**WP_183767659.1**
Edaphobacter aggregans	-	191	191	91%	7.00E-49	31.77	877	**WP_051978591.1**
Acidobacteriia bacterium	yes	190	190	92%	1.00E-48	32.28	850	**MBZ5724199.1**
Edaphobacter sp. 4G125	yes	189	189	91%	2.00E-48	32.09	879	**WP_186695529.1**
Vicinamibacteria bacterium	yes	188	188	92%	6.00E-48	32.09	870	**MBK5255026.1**
Edaphobacter aggregans	-	187	187	91%	1.00E-47	30.79	865	**WP_125486341.1**
Acidobacteriaceae bacterium	yes	186	186	92%	1.00E-47	32.23	781	**MBV8895472.1**
Chloroflexi bacterium	yes	180	180	89%	1.00E-47	32.03	421	**MBU0492686.1**
Geobacter argillaceus	yes	186	186	85%	1.00E-47	30.75	727	**TWJ14039.1**

Table S3.3: Primer sequences used in this study

Deletion	
Frag1_R_KanR	gctgcccgtcctctttgaacttttgctttgccacggaac
Frag1_F_KanR	cgatccatcgtctcgttagaaaaactcatcgagcatcaaatgaaactgc
Frag_2_F_HltA_DS	gatgagtttttctaacgagacgatggatcggtg
Frag2_R_HltA_DS	ctgtgactggtgagttgcgttcaaagccgagtaatgc
Frag3_R_PUC	tcggctttgaacgcaactcaccagtcacagaaaagcat
Frag3_F_PUC	cctcgattgtctcaagatcctgcctcgcgcg
Frag4_F_HltA_US	gcgcgaggcaggatcttgagacaatcgaggcactgttgctac
Frag4_R_HltA_US	caaagcaaaagttcaaagaggacgggcagct
Complementation	
Gm_Rev_Comp	caccgatccatcgtctcgctcgaattgacataagcctgttcgg
SL3360_FOR	cgggccgtttgaaccaggcatcaaataaaacgaaag
HltA_UP_Comp	gtcgtcataagaggacgggcaagaggacgggcagct
HltA_DS_Comp	caggcttatgtcaattcgagcgagacgatggatcggtgag
HltA_for_Comp	ccgtcctcttatgacgacgctcccctattt
HltA_Rev_p2103	tttgatgcctggttcaaacggcccgcagg
Sequencing	
Chr_HltA_UP_seq	taaccgttgggaacttttctcc
Chr_HltA_DS_seq	tctacagcagatcgctcacc

3.7 References

Allakhverdiev, S. I., Nishiyama, Y., Takahashi, S., Miyairi, S., Suzuki, I., & Murata, N. (2005). Systematic analysis of the relation of electron transport and ATP synthesis to the photodamage and repair of photosystem II in Synechocystis. *Plant Physiol, 137*(1), 263-273. doi:10.1104/pp.104.054478

Aravind, L., & Ponting, C. P. (1997). The GAF domain: an evolutionary link between diverse phototransducing proteins. *Trends Biochem Sci, 22*(12), 458-459. doi:10.1016/s0968-0004(97)01148-1

Aro, E. M., Virgin, I., & Andersson, B. (1993). Photoinhibition of Photosystem II. Inactivation, protein damage and turnover. *Biochim Biophys Acta, 1143*(2), 113-134.

Asplund-Samuelsson, J., Sundh, J., Dupont, C. L., Allen, A. E., McCrow, J. P., Celepli, N. A., . . . Ekman, M. (2016). Diversity and Expression of Bacterial Metacaspases in an Aquatic Ecosystem. *Front Microbiol, 7*, 1043. doi:10.3389/fmicb.2016.01043

Beauchamp, J. M., Leveque, R. M., Dawid, S., & DiRita, V. J. (2017). Methylation-dependent DNA discrimination in natural transformation of Campylobacter jejuni. *Proceedings of the National Academy of Sciences, 114*(38), E8053-E8061.

Bhaya, D., Bianco, N. R., Bryant, D., & Grossman, A. (2000). Type IV pilus biogenesis and motility in the cyanobacterium Synechocystis sp. PCC6803. *Molecular microbiology, 37*(4), 941-951.

Bhaya, D., Dufresne, A., Vaulot, D., & Grossman, A. (2002). Analysis of the hli gene family in marine and freshwater cyanobacteria. *FEMS Microbiol Lett, 215*(2), 209-219.

Bolivar, F., Rodriguez, R. L., Greene, P. J., Betlach, M. C., Heyneker, H. L., Boyer, H. W., . . . Falkow, S. (1977). Construction and characterization of new cloning vehicle. II. A multipurpose cloning system. *Gene, 2*(2), 95-113.

Bork, H. R., Heinritz, G. n., Wiessner, R., & Deutsche Gesellschaft für Geographie. (1996). *Raumentwicklung und Umweltverträglichkeit*. Stuttgart: F. Steiner.

Bradshaw, N., Levdikov, V. M., Zimanyi, C. M., Gaudet, R., Wilkinson, A. J., & Losick, R. (2017). A widespread family of serine/threonine protein phosphatases shares a common regulatory switch with proteasomal proteases. *Elife, 6*. doi:10.7554/eLife.26111

Brody, M. S., Stewart, V., & Price, C. W. (2009). Bypass suppression analysis maps the signalling pathway within a multidomain protein: the RsbP energy stress phosphatase 2C from Bacillus subtilis. *Mol Microbiol, 72*(5), 1221-1234. doi:10.1111/j.1365-2958.2009.06722.x

Carbonnelle, E., Helaine, S., Nassif, X., & Pelicic, V. (2006). A systematic genetic analysis in Neisseria meningitidis defines the Pil proteins required for assembly, functionality, stabilization and export of type IV pili. *Molecular microbiology, 61*(6), 1510-1522.

Chang, Y.-W., Rettberg, L. A., Treuner-Lange, A., Iwasa, J., Søgaard-Andersen, L., & Jensen, G. J. (2016). Architecture of the type IVa pilus machine. *Science, 351*(6278), aad2001.

Chen, I., & Dubnau, D. (2004). DNA uptake during bacterial transformation. *Nature Reviews Microbiology, 2*(3), 241-249.

Cho, S. M., Jeoung, S. C., Song, J. Y., Kupriyanova, E. V., Pronina, N. A., Lee, B. W., . . . Park, Y. I. (2015). Genomic Survey and Biochemical Analysis of Recombinant Candidate Cyanobacteriochromes Reveals Enrichment for Near UV/Violet Sensors in the Halotolerant and Alkaliphilic Cyanobacterium Microcoleus IPPAS B353. *J Biol Chem, 290*(47), 28502-28514. doi:10.1074/jbc.M115.669150

Clark, R. L., McGinley, L. L., Purdy, H. M., Korosh, T. C., Reed, J. L., Root, T. W., & Pfleger, B. F. (2018). Light-optimized growth of cyanobacterial cultures: Growth phases and productivity of biomass and secreted molecules in light-limited batch growth. *Metab Eng, 47*, 230-242. doi:10.1016/j.ymben.2018.03.017

Collier, J. L., & Grossman, A. (1994). A small polypeptide triggers complete degradation of light-harvesting phycobiliproteins in nutrient-deprived cyanobacteria. *The EMBO journal, 13*(5), 1039-1047.

Daddy, S., Zhan, J., Jantaro, S., He, C., He, Q., & Wang, Q. (2015). A novel high light-inducible carotenoid-binding protein complex in the thylakoid membranes of Synechocystis PCC 6803. *Sci Rep, 5*, 9480. doi:10.1038/srep09480

Deatherage, D. E., & Barrick, J. E. (2014). Identification of mutations in laboratory-evolved microbes from next-generation sequencing data using breseq. In *Engineering and analyzing multicellular systems* (pp. 165-188): Springer.

Derks, A., Schaven, K., & Bruce, D. (2015). Diverse mechanisms for photoprotection in photosynthesis. Dynamic regulation of photosystem II excitation in response to rapid environmental change. *Biochim Biophys Acta, 1847*(4-5), 468-485. doi:10.1016/j.bbabio.2015.02.008

Dobson, Z., Ahad, S., Vanlandingham, J., Toporik, H., Vaughn, N., Vaughn, M., . . . Mazor, Y. (2021). The structure of photosystem I from a high-light-tolerant cyanobacteria. *Elife, 10*. doi:10.7554/eLife.67518

Golden, S. S., Brusslan, J., & Haselkorn, R. (1986). Expression of a family of psbA genes encoding a photosystem II polypeptide in the cyanobacterium Anacystis nidulans R2. *The EMBO journal, 5*(11), 2789.

Golden, S. S., & Sherman, L. A. (1984). Optimal conditions for genetic transformation of the cyanobacterium Anacystis nidulans R2. *Journal of bacteriology, 158*(1), 36-42.

Hamilton, H. L., & Dillard, J. P. (2006). Natural transformation of Neisseria gonorrhoeae: from DNA donation to homologous recombination. *Molecular microbiology, 59*(2), 376-385.

Ibrahim, I. M., Wu, H., Ezhov, R., Kayanja, G. E., Zakharov, S. D., Du, Y., . . . Puthiyaveetil, S. (2020). An evolutionarily conserved iron-sulfur cluster underlies redox sensory function of the Chloroplast Sensor Kinase. *Commun Biol, 3*(1), 13. doi:10.1038/s42003-019-0728-4

Ito, H., Mutsuda, M., Murayama, Y., Tomita, J., Hosokawa, N., Terauchi, K., . . . Iwasaki, H. (2009). Cyanobacterial daily life with Kai-based circadian and diurnal genome-wide transcriptional control in Synechococcus elongatus. *Proc Natl Acad Sci U S A, 106*(33), 14168-14173. doi:10.1073/pnas.0902587106

Jaiswal, D., Sengupta, A., Sohoni, S., Sengupta, S., Phadnavis, A. G., Pakrasi, H. B., & Wangikar, P. P. (2018). Genome Features and Biochemical Characteristics of a Robust, Fast Growing and Naturally Transformable Cyanobacterium Synechococcus elongatus PCC 11801 Isolated from India. *Scientific Reports, 8*(1), 16632.

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., . . . Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature, 596*(7873), 583-589. doi:10.1038/s41586-021-03819-2

Kale, R., Hebert, A. E., Frankel, L. K., Sallans, L., Bricker, T. M., & Pospisil, P. (2017). Amino acid oxidation of the D1 and D2 proteins by oxygen radicals during photoinhibition of Photosystem II. *Proc Natl Acad Sci U S A, 114*(11), 2988-2993. doi:10.1073/pnas.1618922114

Kerk, D., Silver, D., Uhrig, R. G., & Moorhead, G. B. (2015). "PP2C7s", Genes Most Highly Elaborated in Photosynthetic Organisms, Reveal the Bacterial Origin and Stepwise Evolution of PPM/PP2C Protein Phosphatases. *PLoS One, 10*(8), e0132863. doi:10.1371/journal.pone.0132863

Kettler, G. C., Martiny, A. C., Huang, K., Zucker, J., Coleman, M. L., Rodrigue, S., . . . Chisholm, S. W. (2007). Patterns and implications of gene gain and loss in the evolution of Prochlorococcus. *PLoS Genet, 3*(12), e231. doi:10.1371/journal.pgen.0030231

Kufryk, G. I., Sachet, M., Schmetterer, G., & Vermaas, W. F. (2002). Transformation of the cyanobacterium Synechocystis sp. PCC 6803 as a tool for genetic mapping: optimization of efficiency. *FEMS Microbiology letters, 206*(2), 215-219.

Kumar, S., Stecher, G., Li, M., Knyaz, C., & Tamura, K. (2018). MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol Biol Evol, 35*(6), 1547-1549. doi:10.1093/molbev/msy096

Larsson, J., Nylander, J. A., & Bergman, B. (2011). Genome fluctuations in cyanobacteria reflect evolutionary, developmental and adaptive traits. *BMC Evol Biol, 11*, 187. doi:10.1186/1471-2148-11-187

Latifi, A., Ruiz, M., & Zhang, C. C. (2009). Oxidative stress in cyanobacteria. *FEMS Microbiol Rev, 33*(2), 258-278. doi:10.1111/j.1574-6976.2008.00134.x

Le, S. Q., & Gascuel, O. (2008). An improved general amino acid replacement matrix. *Mol Biol Evol, 25*(7), 1307-1320. doi:10.1093/molbev/msn067

Letunic, I., & Bork, P. (2021). Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res, 49*(W1), W293-W296. doi:10.1093/nar/gkab301

Levdikov, V. M., Blagova, E. V., Rawlings, A. E., Jameson, K., Tunaley, J., Hart, D. J., . . . Wilkinson, A. J. (2012). Structure of the phosphatase domain of the cell fate determinant SpoIIE from Bacillus subtilis. *J Mol Biol, 415*(2), 343-358. doi:10.1016/j.jmb.2011.11.017

Li, S., Sun, T., Xu, C., Chen, L., & Zhang, W. (2018). Development and optimization of genetic toolboxes for a fast-growing cyanobacterium Synechococcus elongatus UTEX 2973. *Metabolic engineering*.

Livak, K. J., & Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2− ΔΔCT method. *methods, 25*(4), 402-408.

Markson, J. S., Piechura, J. R., Puszynska, A. M., & O'Shea, E. K. (2013). Circadian control of global gene expression by the cyanobacterial master regulator RpaA. *Cell, 155*(6), 1396-1408. doi:10.1016/j.cell.2013.11.005

Markson, J. S., Piechura, J. R., Puszynska, A. M., & O'Shea, E. K. (2013). Circadian control of global gene expression by the cyanobacterial master regulator RpaA. *Cell, 155*(6), 1396-1408.

Montgomery, B. L., & Lagarias, J. C. (2002). Phytochrome ancestry: sensors of bilins and light. *Trends Plant Sci, 7*(8), 357-366. doi:10.1016/s1360-1385(02)02304-x

Moya, A., Oliver, J. L., Verdu, M., Delaye, L., Arnau, V., Bernaola-Galvan, P., . . . Roman-Roldan, R. (2020). Driven progressive evolution of genome sequence complexity in Cyanobacteria. *Sci Rep, 10*(1), 19073. doi:10.1038/s41598-020-76014-4

Nagarajan, A., Winter, R., Eaton-Rye, J., & Burnap, R. (2011). A synthetic DNA and fusion PCR approach to the ectopic expression of high levels of the D1 protein of photosystem II in Synechocystis sp. PCC 6803. *Journal of Photochemistry and Photobiology B: Biology, 104*(1), 212-219.

Nies, F., Mielke, M., Pochert, J., & Lamparter, T. (2020). Natural transformation of the filamentous cyanobacterium Phormidium lacuna. *PloS one, 15*(6), e0234440.

Ogawa, K., Yoshikawa, K., Matsuda, F., Toya, Y., & Shimizu, H. (2018). Transcriptome analysis of the cyanobacterium Synechocystis sp. PCC 6803 and mechanisms of

photoinhibition tolerance under extreme high light conditions. *J Biosci Bioeng, 126*(5), 596-602. doi:10.1016/j.jbiosc.2018.05.015

Onai, K., Morishita, M., Kaneko, T., Tabata, S., & Ishiura, M. (2004). Natural transformation of the thermophilic cyanobacterium Thermosynechococcus elongatus BP-1: a simple and efficient method for gene transfer. *Molecular Genetics and Genomics, 271*(1), 50-59.

Pakrasi, H. B., & Wendt, K. E. (2019). Genomics approaches to deciphering natural transformation in cyanobacteria. *Frontiers in microbiology, 10*, 1259.

Pane-Farre, J., Lewis, R. J., & Stulke, J. (2005). The RsbRST stress module in bacteria: a signalling system that may interact with different output modules. *J Mol Microbiol Biotechnol, 9*(2), 65-76. doi:10.1159/000088837

Piechura, J. R., Amarnath, K., & O'Shea, E. K. (2017). Natural changes in light interact with circadian regulation at promoters to control gene expression in cyanobacteria. *Elife, 6*. doi:10.7554/eLife.32032

Piepenbrink, K. H., Maldarelli, G. A., De La Peña, C. F. M., Dingle, T. C., Mulvey, G. L., Lee, A., . . . Sundberg, E. J. (2015). Structural and evolutionary analyses show unique stabilization strategies in the type IV pili of Clostridium difficile. *Structure, 23*(2), 385-396.

Porter, R. D. (1986). Transformation in cyanobacteria. *CRC Critical reviews in microbiology, 13*(2), 111-132.

Racharaks, R., Arnold, W., & Peccia, J. (2021). Development of CRISPR-Cas9 knock-in tools for free fatty acid production using the fast-growing cyanobacterial strain Synechococcus elongatus UTEX 2973. *Journal of Microbiological Methods*, 106315.

Rodriguez Ayala, F., Bartolini, M., & Grau, R. (2020). The Stress-Responsive Alternative Sigma Factor SigB of Bacillus subtilis and Its Relatives: An Old Friend With New Functions. *Front Microbiol, 11*, 1761. doi:10.3389/fmicb.2020.01761

Rumyantsev, K. A., Shcherbakova, D. M., Zakharova, N. I., Emelyanov, A. V., Turoverov, K. K., & Verkhusha, V. V. (2015). Minimal domain of bacterial phytochrome required for chromophore binding and fluorescence. *Sci Rep, 5*, 18348. doi:10.1038/srep18348

Shi, T., & Falkowski, P. G. (2008). Genome evolution in cyanobacteria: the stable core and the variable shell. *Proc Natl Acad Sci U S A, 105*(7), 2510-2515. doi:10.1073/pnas.0711165105

Shi, Y. (2009). Serine/threonine phosphatases: mechanism through structure. *Cell, 139*(3), 468-484. doi:10.1016/j.cell.2009.10.006

Shih, P. M., Wu, D., Latifi, A., Axen, S. D., Fewer, D. P., Talla, E., . . . Kerfeld, C. A. (2013). Improving the coverage of the cyanobacterial phylum using diversity-driven genome sequencing. *Proc Natl Acad Sci U S A, 110*(3), 1053-1058. doi:10.1073/pnas.1217107110

Springstein, B. L., Nies, F., & Dagan, T. (2020). Natural competence in Chlorogloeopsis fritschii PCC 6912 and other ramified cyanobacteria. *bioRxiv*.

Su, J., Schlicker, C., & Forchhammer, K. (2011). A third metal is required for catalytic activity of the signal-transducing protein phosphatase M tPphA. *J Biol Chem, 286*(15), 13481-13488. doi:10.1074/jbc.M109.036467

Tan, X., Hou, S., Song, K., Georg, J., Klahn, S., Lu, X., & Hess, W. R. (2018). The primary transcriptome of the fast-growing cyanobacterium Synechococcus elongatus UTEX 2973. *Biotechnol Biofuels, 11*, 218. doi:10.1186/s13068-018-1215-8

Taton, A., Erikson, C., Yang, Y., Rubin, B. E., Rifkin, S. A., Golden, J. W., & Golden, S. S. (2020). The circadian clock and darkness control natural competence in cyanobacteria. *Nature Communications, 11*(1), 1-11.

Toyoshima, M., Tokumaru, Y., Matsuda, F., & Shimizu, H. (2020). Assessment of Protein Content and Phosphorylation Level in Synechocystis sp. PCC 6803 under Various Growth Conditions Using Quantitative Phosphoproteomic Analysis. *Molecules, 25*(16). doi:10.3390/molecules25163582

Trehan, K., & Sinah, U. (1981). Genetic transfer in a nitrogen-fixing filamentous cyanobacterium. *Microbiology, 124*(2), 349-352.

Ungerer, J., Lin, P.-C., Chen, H.-Y., & Pakrasi, H. B. (2018). Adjustments to photosystem stoichiometry and electron transfer proteins are key to the remarkably fast growth of the cyanobacterium Synechococcus elongatus UTEX 2973. *MBio, 9*(1), e02327-02317.

Ungerer, J., Lin, P. C., Chen, H. Y., & Pakrasi, H. B. (2018). Adjustments to Photosystem Stoichiometry and Electron Transfer Proteins Are Key to the Remarkably Fast Growth of the Cyanobacterium Synechococcus elongatus UTEX 2973. *MBio, 9*(1). doi:10.1128/mBio.02327-17

Ungerer, J., & Pakrasi, H. B. (2016). Cpf1 is a versatile tool for CRISPR genome editing across diverse species of cyanobacteria. *Scientific Reports, 6*, 39681.

Ungerer, J., & Pakrasi, H. B. (2016). Cpf1 Is A Versatile Tool for CRISPR Genome Editing Across Diverse Species of Cyanobacteria. *Scientific Reports, 6*. doi:ARTN 39681

10.1038/srep39681

Ungerer, J., Wendt, K. E., Hendry, J. I., Maranas, C. D., & Pakrasi, H. B. (2018). Comparative genomics reveals the molecular determinants of rapid growth of the cyanobacterium Synechococcus elongatus UTEX 2973. *Proceedings of the National Academy of Sciences, 115*(50), E11761-E11770.

Ungerer, J., Wendt, K. E., Hendry, J. I., Maranas, C. D., & Pakrasi, H. B. (2018). Comparative genomics reveals the molecular determinants of rapid growth of the cyanobacterium

Synechococcus elongatus UTEX 2973. *Proc Natl Acad Sci U S A*. doi:10.1073/pnas.1814912115

Wendt, K. E., Walker, P., Sengupta, A., Ungerer, J., & Pakrasi, H. B. (2021). Engineering natural competence into the fast-growing cyanobacterium Synechococcus elongatus UTEX 2973. *Appl Environ Microbiol*, AEM0188221. doi:10.1128/AEM.01882-21

Will, S. E., Henke, P., Boedeker, C., Huang, S., Brinkmann, H., Rohde, M., . . . Petersen, J. (2019). Day and Night: Metabolic Profiles and Evolutionary Relationships of Six Axenic Non-Marine Cyanobacteria. *Genome Biol Evol, 11*(1), 270-294. doi:10.1093/gbe/evy275

Wu, X., Lee, D. W., Mella, R. A., & Golden, J. W. (2007). The Anabaena sp. strain PCC 7120 asr1734 gene encodes a negative regulator of heterocyst development. *Mol Microbiol, 64*(3), 782-794. doi:10.1111/j.1365-2958.2007.05698.x

Yang, X., Kang, C. M., Brody, M. S., & Price, C. W. (1996). Opposing pairs of serine protein kinases and phosphatases transmit signals of environmental stress to activate a bacterial transcription factor. *Genes Dev, 10*(18), 2265-2275. doi:10.1101/gad.10.18.2265

Yoshihara, S., Suzuki, F., Fujita, H., Geng, X. X., & Ikeuchi, M. (2000). Novel putative photoreceptor and regulatory genes Required for the positive phototactic movement of the unicellular motile cyanobacterium Synechocystis sp. PCC 6803. *Plant Cell Physiol, 41*(12), 1299-1304. doi:10.1093/pcp/pce010

Yu, J., Liberton, M., Cliften, P. F., Head, R. D., Jacobs, J. M., Smith, R. D., . . . Pakrasi, H. B. (2015). Synechococcus elongatus UTEX 2973, a fast growing cyanobacterial chassis for biosynthesis using light and CO2. *Scientific Reports, 5*.

Yu, J., Liberton, M., Cliften, P. F., Head, R. D., Jacobs, J. M., Smith, R. D., . . . Pakrasi, H. B. (2015). Synechococcus elongatus UTEX 2973, a fast growing cyanobacterial chassis for biosynthesis using light and CO(2). *Sci Rep, 5*, 8132. doi:10.1038/srep08132

Zhang, W., & Shi, L. (2004). Evolution of the PPM-family protein phosphatases in Streptomyces: duplication of catalytic domain and lateral recruitment of additional sensory domains. *Microbiology (Reading), 150*(Pt 12), 4189-4197. doi:10.1099/mic.0.27480-0

Chapter 4:
Conclusions, future directions, and additional data

Chapter Contributions:
All data in this Chapter were collected by PLW

4.1 Summary

The research presented in this thesis provides new information about components and mechanisms for high-light (HL) tolerance in cyanobacteria. In my thesis work, I sought to identify the unique genetic factors in *Synechococcus elongatus* UTEX 2973, as compared to *Synechococcus elongatus* PCC 7942, that allow this strain to thrive under HL conditions. To accomplish this, I investigated the effects of individual protein-coding SNPs and indels on HL tolerance and compared the global transcriptomes of both strains during HL acclimation. As I show in this thesis, I have identified transcriptional changes and important genes in HL response and bioinformatically characterized a novel sensory protein necessary for tolerance to HL in *S.* 2973. The conclusions, future directions, and additional data not included in Chapters 2 and 3 will be presented in the following sections.

Overall, these studies expand our understanding of the HL adaptive processes of *S.* 2973 and could help lead to the development of ways to improve light resilience in other phototrophic organisms. Parts of this work may also apply to advancing tolerance to other environmental stresses in photosynthetic organisms, as the stress response pathways of light, heat, and oxidative stress are highly intertwined, as discussed in Chapter 1. The two *Synechococcus* genotypes adopted different transcriptional strategies to cope with light stress. Understanding adaptive mechanisms down to the cellular response may help in advancing technological solutions to the industrial use of cyanobacteria.

4.2 Conclusions and Future directions: Chapter 2

4.2.1 Conclusions

The goal of Chapter 2 was to identify critical regulation and transcriptional differences between two *Synechococcus* genotypes during high light (HL) acclimation. To achieve this objective, I first identified important light stress conditions and time points indicated by growth and photosynthetic activity. RNA-seq was then used to generate high-quality transcriptome reads of both *Synechococcus* strains 7942 and 2973 grown in normal light and after exposure to 30 minutes and 4 hours of HL.

In my first objective, I identified key light intensities and timepoints that reflect the divergence in HL stress endured of *S.* 2973 and *S.* 7942 as measured by photosynthetic efficiency. I found that within the first hour of HL exposure *S.* 7942 underwent more damage to PSII than *S.* 2973, and after 4 hours *S.* 2973 was able to recover photosynthetic efficiency to pre-stress levels, unlike *S.* 7942 which maintained a photosynthetic efficiency of 80% of the pre-stress conditions. These data indicated that the HL tolerance of *S.* 2973 may rely on the prompt response capacity to recover photosynthetic activity. Avoiding low light stress response for the control condition was important, and thus appropriate light intensities had to be selected. Typically, a light intensity of ~ 400 μmol photons m^{-2} s^{-1} is used as a normal growth condition, although we found that even at this intensity there was a divergence between the strains in photosynthetic activity and damage to PSII. Only at a lower light condition of 200 μmol photons m^{-2} s^{-1} was there minimal enough damage to the photosynthetic apparatus that even without repair to PSII the photosynthetic efficiency did not drop after an hour of light exposure.

Exposure to HL above photosynthetic capacity induces adaptive responses that allow organisms to continue normal metabolic and growth functions. As discussed in Chapter 1, at the

cellular level adaptive responses involve transcriptional and translational changes and require several interconnected global transcription factors that mediate various processes in the cell, particularly photosynthesis and metabolism gene regulation. In Chapter 2, I outlined the transcriptional changes observed in response to HL including downregulation of photosynthetic pigments, and upregulation of antioxidants and chaperone proteins to enable cell survival. Genes involved in nitrogen uptake and assimilation were distinctly regulated between the two genotypes, resulting in, or because of, their different tolerances to light stress.

Gene expression analysis allows for simultaneous evaluation of the expression patterns of hundreds of genes, providing cellular insights into genes that are involved in light response. The differential expression, clustering, and pathway analysis revealed considerably different gene expression characteristics between the strains during high-light acclimation. Many known light responsive genes experienced different levels of transcriptional regulation in *S.* 2973 and *S.* 7942, most of which underwent larger expression changes in *S.* 2973. Three gene expression clusters that were found to contain strong expression profile divergence between the strains were predominantly regulated by RpaA and NtcA. The opposing expression profile of Cluster 4, enriched for photosynthetic and NtcA regulated genes, reflected the difference in carbon/nitrogen balance during HL acclimation of the strains.

We found considerable variation in transcriptomic changes in response to HL between the two strains, in particular the expression of many circadian genes central to photosynthesis and carbon transport. The *rpaA* allele was differentially expressed likely due to the 7-bp difference in promoter sequences. This difference in RpaA resulted in global metabolic differences between the two strains. The RpaA allele in *S.* 2973 results in its downregulation and subsequent alterations to regulation of metabolism and downstream interconnected regulators

including SigD3, and SigF. *S.* 2973 highly upregulated carbon uptake and metabolism, which created a sink for excess absorbed light energy, leading to reduced generation of reactive oxygen species and allowing for increased growth. We determined that the difference in circadian gene regulation that resulted in changes to various interconnected regulators is a key mechanism for *S.* 2973 to properly adapt to HL.

This study is the first transcriptome comparison of these strains as well as the first time-course transcriptome study in *S.* 2973. Overall, these findings provide crucial data for models aiming to predict the light adaptive capacity of different cyanobacteria and to identify genes that may be useful for future studies on oxidative stress and light tolerance.

4.2.2 Future directions

In addition to the 31 SNPs, one indel, and one inversion of protein coding genomic regions, there are also 20 non-coding genetic differences between *S.* 2973 and *S.* 7942 (Yu et al., 2015b). Adaptation through genetic variation acts as a driving force for environmental acclimatization in new or changing environments. By studying the small genetic variations found in different ecotypes we can associate genes with phenotypes such as HL tolerance. However, non-coding genetic variation is also important for the regulation of mRNAs and cellular physiology. Therefore, integration of regulation from non-coding regulatory RNAs into our current knowledge of HL response could help clarify the gene regulatory network of HL acclimation in *S.* 2973. For example, the HL-induced small regulatory RNA PsrR1 together with RpaB controls PSI gene expression in *Synechocystis* 6803 (Kadowaki et al., 2016). Furthermore, post-translational modifications (PTMs) are already known to play an important role in HL tolerance, such as the PTMs by Hik33 and Trx, which modify RpaB activity (Ibrahim et al., 2022, Kato et al., 2022). RpaA and HltA both undergo PTM, which are not captured by

transcriptome studies. Proteomic experiments and non-coding RNA studies could provide more insight into the PTM occurrences and elucidate additional protein-protein interactions that give rise to HL tolerance in *Synechococcus elongatus*.

S. 7942 exhibits circadian oscillations in the abundance of mRNA transcripts (Gutu, 2013), but this has not been studied in *S.* 2973. The circadian oscillator is composed of KaiA, KaiB, and KiaC proteins that transmit circadian timekeeping information through RpaA. In Chapter 2 we found that the expression of *rpaA* and the core clock genes *sasA* and *kaiC* were differentially regulated solely in *S.* 2973. These data suggest that the RpaA allele may not only affect the light response mechanisms, but could alter the circadian oscillator in *S.* 2973. Further, analysis of circadian gene oscillations in *S.* 2973 should reveal the scope of the RpaA allelic variation.

Eight of the top 33 differentially regulated genes are annotated as hypothetical proteins, 4 of which were DE in both strains and 4 that were found to be DE only in *S.* 2973 (Table 1.1). While this study clustered gene expression, further analysis into gene co-expression can identify putative proteins that may be important not only for high-light tolerance but also for heat and oxidative stress response.

The ratios of photosynthetic electron transport chain components have a large effect on the energy and reaction oxygen species production. Our data indicated that the turnover of the PSII reaction center D1 protein was very similar between genotypes, with an interesting difference in the D1 gene copies *psbA2* and *psbA3*, which showed sustained upregulation to a slightly greater extent in *S.* 7942 than *S.* 2973. However, our data showed a rapid decrease in expression of PSII, PBS, and cytochrome c-oxidase types aa_3 and cbb_3 genes exclusively in *S.* 2973. In addition, PSI genes were downregulated to a greater extent in *S.* 2973 than *S.* 7942.

Previous comparisons found that *S.* 2973 contained higher amounts of PSI protein (encoded by *psaA*) and similar levels of D1 protein, although this study was carried out with strains grown at different light intensities (*S.* 2973 at 900 µmol photons m^{-2} s^{-1} light and *S.* 7942 at 400 µmol photons m^{-2} s^{-1} light) (Ungerer, et al., 2018a). Preliminary absorption spectra support the premise of decreased PBS levels in *S.* 2973 after transfer to high light intensity (Fig. 4.1a and 4.1b). Additional analyses to determine the PSI to PSII ratio of both strains at high light intensities (1500 µmol photons m^{-2} s^{-1}) can further investigate the divergent changes that occur in the strains.

4.2.2 Additional data

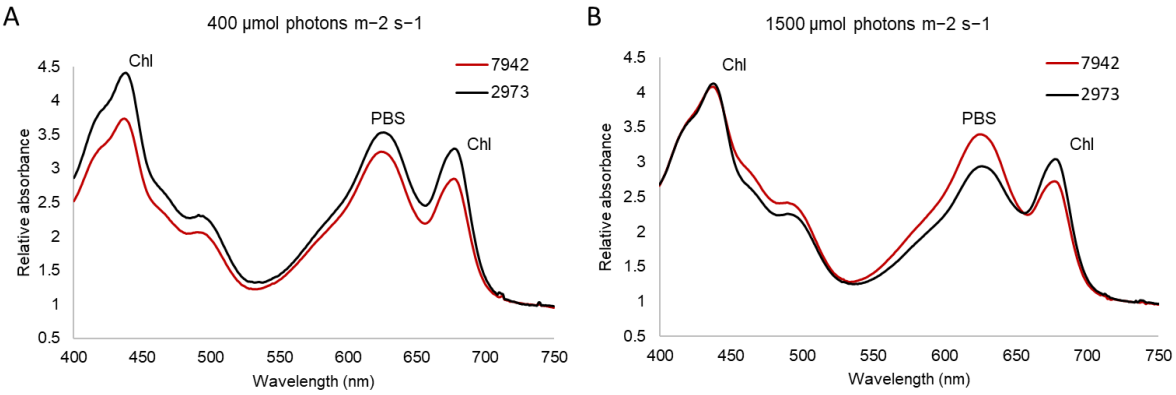

Figure 4.1: Whole-cell absorption spectra of *S.* 7942 and *S.* 2973 at (a) growth light, 400 µmol photons m^{-2} s^{-1} light and (b) high light, 1500 µmol photons m^{-2} s^{-1}. Peaks of photosynthetic pigments are indicated: chlorophyll *a* (680 nm) and phycobilisome (625 nm). Spectra were normalized to the absorbance at 750 nm.

4.3 Conclusions, additional data, and future directions: Chapter 3

4.3.1 Conclusions

High light tolerance is most likely a quantitative trait that results from the interactions of many genes. Therefore, by changing any one component of light tolerance the overall trait would

likely be affected. In Chapter 3, I examined CRISPR strains with individual allelic mutations and identified polymorphic loci that resulted in a loss of high-light tolerance in *S.* 2973.

To characterize light tolerance, I assessed strains under high-light conditions and measured both growth and physiological parameters indicative of light tolerance, which included photosynthetic efficiency, oxygen evolution, and pigment content. I focused on genes not known to be stress-relevant, which led to the focus Chapter 3, the characterization of a hypothetical protein now known as HltA for high-light tolerant A. Results demonstrated that HltA is a conserved cyanobacterial phosphatase, essential for high-light tolerance in *S.* 2973, and is not involved in a general stress response including heat and nutrient starvation. Additionally, this work provided evidence that PP2Cs did not evolve in eukaryotic organisms but rather originated in prokaryotes. Sensor domain PP2Cs have not been found in eukaryotic organisms and have been quite limited to mostly cyanobacteria and streptomyces (or at least until more prokaryotic organism genomes are sequenced). HltA may represent a key light-response adaptive mechanism likely involved in light signaling (acting as a light adaptive sensor).

4.3.2 Future directions

Screening of CRISPR mutants led to the identification of several SNPs that contribute to light tolerance, a CTP synthetase, and the sigma factor interacting protein SinA. In Chapter 2 we found that the expression of SinA, a protein only recently characterized in 2020 (Hasegawa et al., 2020), was different in the strains during light acclimation. Further exploration of SinA could be promising in our understanding of light response regulation. Many sigma factors are already connected to or believed to be involved in light response, such as those regulated by RpaA and RpaB; SinA could further our understanding of why the regulon of certain sigma factors changes in response to different environmental stresses.

There were no apparent alterations in circadian growth due to *hltA* mutations in light/dark cycles compared to wild-type *S.* 2973. While *S.* 7942 has emerged as the leading model system to study the cyanobacterial circadian clock, *S.* 2973 has never been explicitly tested for rhythmic gene expression. We performed preliminary experiments in which cultures were entrained with light/dark cycles for three days before being transferred to free-running continuous light conditions. Once transferred to free-running conditions, *S.* 2973 wild type and mutants continued to grow in light with no signs of circadian entrainment as determined by growth in diurnal light cycles. For this experiment to provide more meaningful results, it should first be established that *S.* 2973 displays rhythmic growth in response to light/dark entrainment. This work only utilized *S.* 2973 for studying mutations in *hltA.* However, between *Synechococcus elongatus* strains, the differences in RpaA (the transcriptional regulator of hltA) may lead to the divergent expression of *hltA*, which may affect the quantities of HltA protein levels. Additionally, the N-terminal SNP of HltA may cause an alteration in the interaction with the activator molecule (such as a bilin). Therefore, it may be the case that mutations to *hltA* in *S.* 7942 would result in a different light tolerance phenotype.

4.3.3 Additional data

HltA protein proximity labelling

Identification of the proteins that interact with HltA will help to elucidate its function in the cell and possibly identify a novel light response pathway. A proximity-based protein labelling experiment can be used to identify protein partners and targets of HltA. Spatiotemporal protein labeling followed by mass spectrometry of modified proteins can be used to determine the network of proteins located in the same region as the protein of interest. By fusing the protein

of interest with the APEX2 enzyme, an ascorbate peroxidase that was recently developed for use in *Synechococcus* (Dahlgren et al, 2021), we can detect the regional proteome around HltA.

I designed and constructed a plasmid containing an APEX2-tagged *hltA* that targets the native *hltA* sequence and transformed *S.* with this construct (Fig. 4.2a). After continuous re-streaking on BG11 plates supplemented with gentamycin for several rounds, PCR confirmed that the mutant strains have been fully segregated (Fig. 4.2b). By using the native location, the APEX tagged HltA will be under the control of its native promoter and should be expressed in the natural subcellular location. I expect these factors to be important for deciphering its function in a physiologically relevant setting.

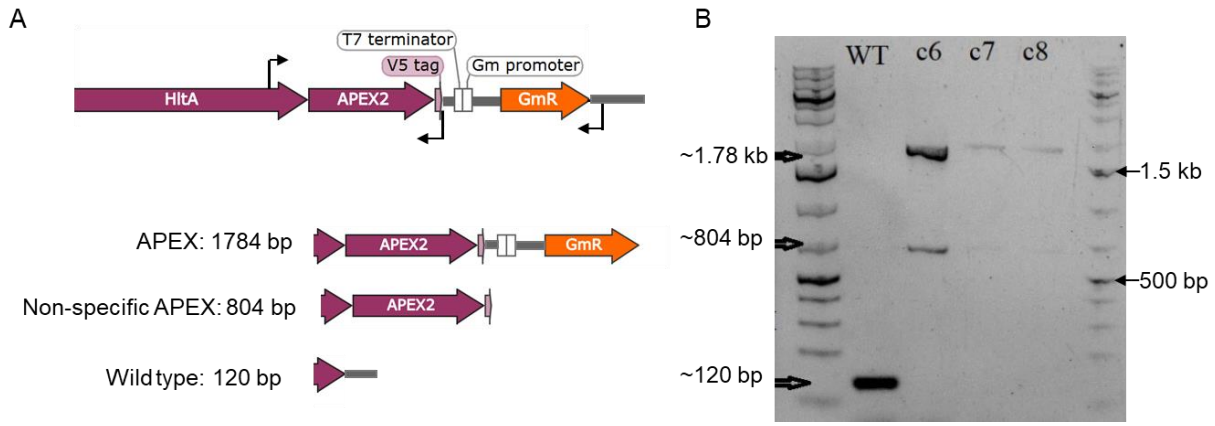

Figure 4.2: Analysis of APEX2 tagged chromosomal *hltA* in *Synechococcus.* (a) Arrangement of sequence and various expected targets with nucleotide lengths from primers. (b) Agarose gel electrophoresis results with three bands confirming APEX2 tag of *hltA* in mutants c6-c8 and absent in WT. Primers shown in black arrow, forward primer sequence (CGACCTATTCAACCGAGTCC), reverse primer sequence (tctacagcagatcgctcacc). Reverse primer binds to two locations in the sequence shown in (a).

Due to the presence of the GAF domain, it is reasonable to believe that HltA is activated in response to increased light intensity. As with other PP2Cs containing sensory domains, binding to the sensor causes a conformational switch that allows for phosphatase activity. Therefore, for the identification of the HltA protein network, the best time to label would be

about 5-20 minutes following a light upshift (Fig. 4.3). Both a change from dark to low light, to mimic dawn, as well as the shift from low light to high light which mimics a sun fleck, could be used. I hypothesize that at these times following an upshift in light the HltA protein senses the change and utilizing a partner-switching mechanism in which it dephosphorylates its target protein, leading to a signal cascade which ultimately regulates the transcription initiation activity of a light-responsive sigma factor.

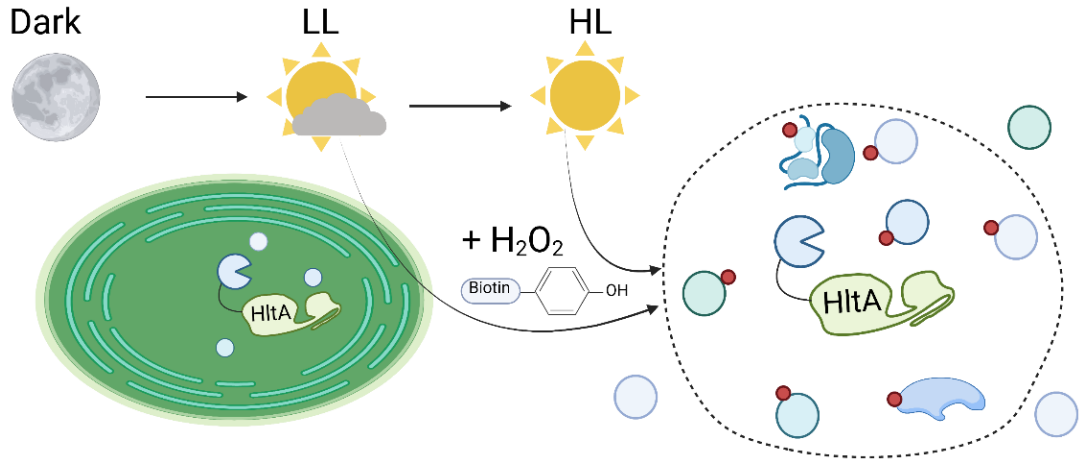

Figure 4.3: Proposed experimental set up for APEX2-HltA protein proximity labeling. Cultures with APEX-2 tagged HltA (and controls) are exposed to an upshifted light, from dark to low light (LL) and LL to high light (HL). Following the upshift cultures are induced with H_2O_2, hydrogen peroxide and biotin phenol. Within minutes the engineered APEX2 enzyme will then biotinylate proteins within a distance of the HltA-APEX2 location, illustrated by red dots, determined by the concentrations of biotin added.

HltA phosphatase activity

To assay HltA for phosphatase activity I used an IPTG inducible promoter to overexpress histidine-tagged *hltA* in E. coli. I purified his-tagged HltA and analyzed the purification fractions (Fig. 4.4) and established a phosphatase protocol for HltA that resulted in low levels of metal-dependent activity (Fig. 4.5).

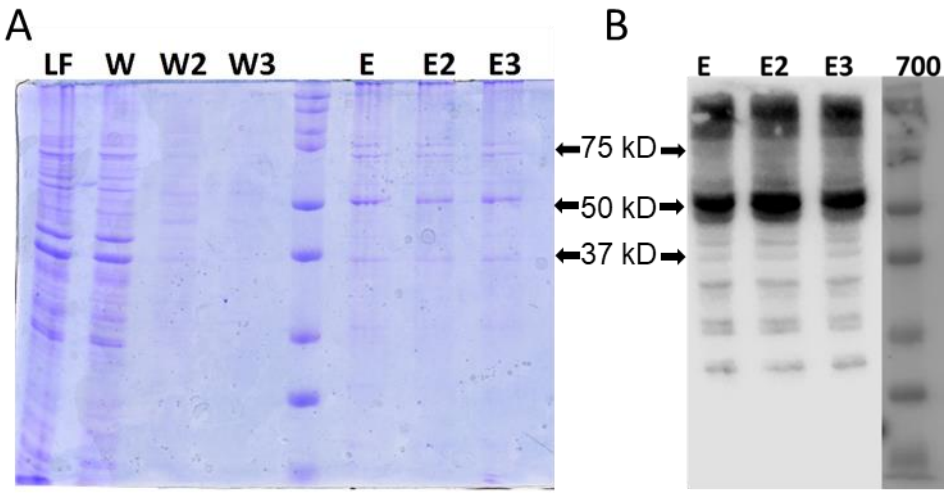

Figure 4.4: Protein purification fraction analysis of recombinant HltA expression. (a) SDS-PAGE gel stained with Coomassie blue; each lane was loaded with 5 µl of the following fractions. (b) Anti-his tag western blot of elution fractions. The expected size of HltA is 51 kDa. LF, loading fraction; W, wash; W2, second wash; W3, third wash; E, first elution; E2, second elution; E3, third elution.

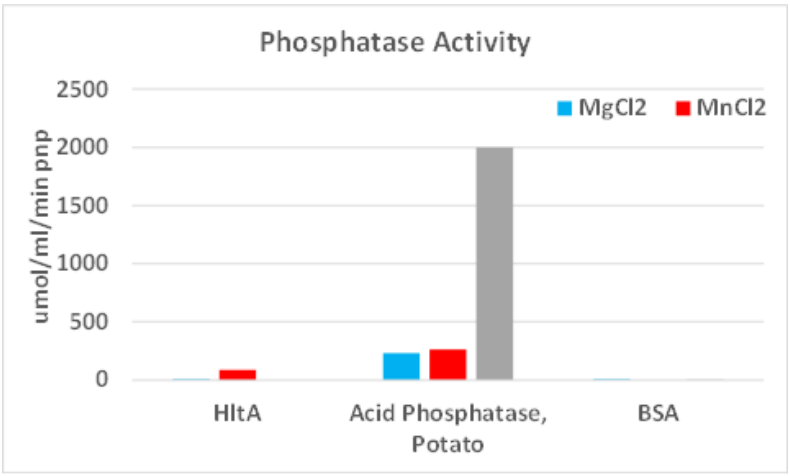

Figure 4.5: Phosphatase activity rates of enzymes on pNPP with different metals. pNP absorption measured at 405 nm, normalized by concentration of protein added per minute. Reactions in MES buffer, pH 5.5. Data shown are averaged from two replicates. Acid phosphatase from potato was used as the positive control, BSA from bovine was used as the negative control.

To establish the phosphatase assay, varying concentrations of purified HltA were assayed in numerous phosphatase activity experiments. HltA protein phosphatase activity was determined (as described in methods) to be active in at very low detectable levels towards p-nitrophenylphosphate (pNPP) substrate with the required addition of at least 10 mM $MnCl_2$.

Phosphatase activity: future directions

Although the sequence of the GAF sensory domain was not highly homologous to any characterized proteins with known ligands, the predicted secondary structure matches the topology of the GAF domain ligand binding pocket, which has a five antiparallel β-sheets connected by surrounding α-helices. At the N-terminus of the GAF domain is a region that resembles the tongue-like protrusion connected to the GAF domain on phytochromes, which stabilizes ligand binding and conformational change to the active state (Anders et al., 2013). The HltA structure also contains the PP2C-α0 that may act as a ligand switch mechanism to coordinate phosphatase activity (Brody et al., 2009, Bradshaw et al., 2017). Therefore, it is likely that the ligand that the HltA GAF domain binds is required for high activity. Although the ligand has been undetermined, there are a few likely candidates, including tetrapyrrole chromophores and secondary molecules such as c-di-GMP and c-di-AMP. I demonstrated in the previous section that HltA can dephosphorylate the generic substrate pNPP. The HltA activating ligand can be determined by including an extra incubation step with different ligands proceeding the phosphatase activity assay outlined above.

4.3.4 Material and methods

E. coli overexpression cells were grown in LB media and induced by the addition of IPTG. Cells were harvested, washed, and resuspended in lysis buffer. Cells were broken by sonication; cell debris was removed by centrifugation at 4 °C. The affinity purification of His-tagged HltA was performed with nickel column purification by incubating cell extract with Ni-NTA agarose beads. This was followed by a series of wash and imidazole elution steps. The

protein content was determined using bicinchoninic acid (BCA) protein assay reagent (Thermo Scientific).

To visualize products from the purification 5 µl of fractions were loaded into an SDS-PAGE gel. Proteins were transferred to a polyvinylidene difluoride (PVDF) membrane (Millipore) and incubated in 5% bovine serum albumin (BSA) blocking reagent for 3 hours at room temperature, and then separately incubated with the primary anti-6x-his polyclonal antibody raised in rabbit overnight at 4°C, followed by washing and incubation with HRP-conjugated secondary anti-rabbit antibody. Bands were visualized using chemiluminescence reagents (EMD Millipore, Billerica, MA, USA) with an ImageQuant LAS-4000 imager (GE Healthcare).

HltA phosphatase activity was measured using pNPP substrate. To achieve HltA phosphatase activity many elements to the pNPP phosphatase assay were tested at different concentrations, temperatures, and for different periods of time. Briefly, protein purified from lysates of E. coli expressing HltA-6xhis were incubated at room temperature or 38°C in a volume of various metal concentrations and buffers, and 50-500 mM P-labeled phosphoprotein substrate (p-nitrophenylphosphate, pNPP). Production of pNP was measured as a function of enzyme concentrations (1-10 µg), pH (4.0-7.0), $MgCl_2$ concentration (0-10mM), $MnCl_2$ concentration (0-10mM), $ZnCl_2$ concentration (0-10mM) at room temperature or 38°C, in 50 mM Tris, 0.5M MES, or 100mM Sodium Acetate buffer. The reaction was terminated every 10 minutes for an hour, by the addition of NaOH and EDTA, and measured spectroscopically.

4.4 References

Anders, K., Daminelli-Widany, G., Mroginski, M. A., von Stetten, D., & Essen, L. O. (2013). Structure of the cyanobacterial phytochrome 2 photosensor implies a tryptophan switch

for phytochrome signaling. *J Biol Chem, 288*(50), 35714-35725. doi:10.1074/jbc.M113.510461

Bradshaw, N., Levdikov, V. M., Zimanyi, C. M., Gaudet, R., Wilkinson, A. J., & Losick, R. (2017). A widespread family of serine/threonine protein phosphatases shares a common regulatory switch with proteasomal proteases. *Elife, 6*. doi:10.7554/eLife.26111

Dahlgren, K. K., Gates, C., Lee, T., & Cameron, J. C. (2021). Proximity-based proteomics reveals the thylakoid lumen proteome in the cyanobacterium *Synechococcus* sp. PCC 7002. *Photosynth Res, 147*(2), 177-195. doi:10.1007/s11120-020-00806-y

Hasegawa, H., Tsurumaki, T., Kobayashi, I., Imamura, S., and Tanaka, K. (2020). Identification and analysis of a principal sigma factor interacting protein SinA, essential for growth at high temperatures in a cyanobacterium *Synechococcus elongatus* PCC 7942. J Gen Appl Microbiol *66*, 66-72.

Appendix

A.1 Restoration of *hltA* full length sequence in PCC 6803

A.1.1 Background

The ortholog of *hltA* in the model cyanobacterium *Synechocystis* sp. PCC 6803 (hereafter; *PCC 6803)* is encoded by *slr2031*. This locus is commonly known as it is used as a neutral site for the introduction of sequences. When digging deeper into this unexpected finding, I also uncovered a few reports of conflicting data around the function of *slr2031*. A 1998 study disrupted *slr2031* with a kanamycin resistance cassette and found that this mutant showed improved growth when streaked onto BG11 plates and grown under 350 μmol photons m^{-2} s^{-1}. This result led to the conclusion that the mutant was defective in high-light response (Kamei, 1998). Two years later another study also produced a *slr2031* mutant in PCC 6803, which was not able to recover after long-term N- and S- starvation. They also reported that in wild-type cells mRNA levels of *slr203*1 did not change significantly after applying salt, heat, and high-light stress. They concluded that *slr2031* was not involved in high light or salt stress response but might be involved in nutrient deficiency and subsequent degradation of the phycobilisome antenna (Huckauf, Nomura et al. 2000). The background strain utilized for both studies was a glucose tolerant non-motile strain of PCC 6803 missing a 154 bp sequence located in the intergenic and n-terminal region of *slr2031*. Whole-genome resequencing found this deletion in all but one glucose tolerant (GT) / non-motile strain (Ding et al., 2015), and I confirmed the 154 bp deletion in our own lab's GT strain as well (Fig 5.1).

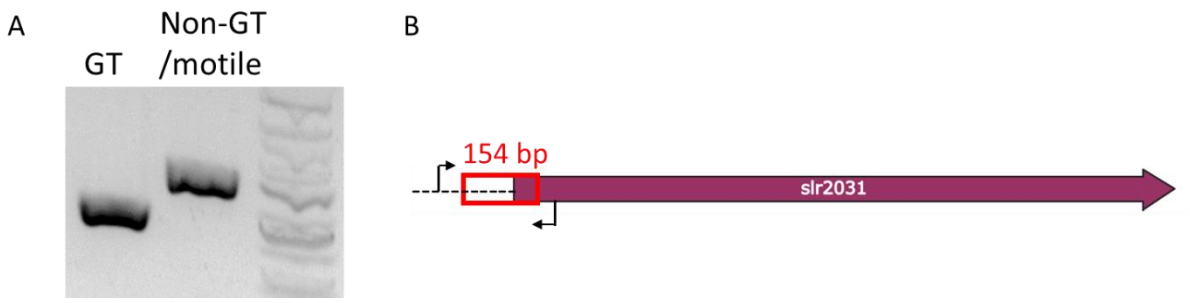

Figure 5.1: Difference in slr2031 sequence length of GT and non-GT strains. (a) PCR products from sequencing the DNA region outside of the 154 bp area of GT and non-GT/motile 6803 strains. (b) Diagram of location of the154 bp deletion, represented by the red box, in GT strains, and location of primers, indicated by black arrows.

In the bioinformatic analysis of more than 400 cyanobacterial *hltA* orthologs, this truncation was not observed in any other sequence. The truncation likely occurred as an unintended consequence during the lab-directed evolution process to obtain a GT PCC 6803 strain. The strong conservation for *hltA* in cyanobacteria and few genetic polymorphisms between the GT and non-GT strains suggest that *hltA* should still have a function in PCC 6803. I showed in Chapter 3, that in addition to *slr2031*, PCC 6803 also encodes four additional Group II PP2Cs and three Group I PP2Cs. I hypothesize that the perplexing and non-essential nature of *slr2031* is due to an overlap in function of the several PP2Cs in PCC 6803, which can compensate for the truncated *slr2031* function. To clarify the role of *slr2031* in light response the full-length sequence needs to first be restored.

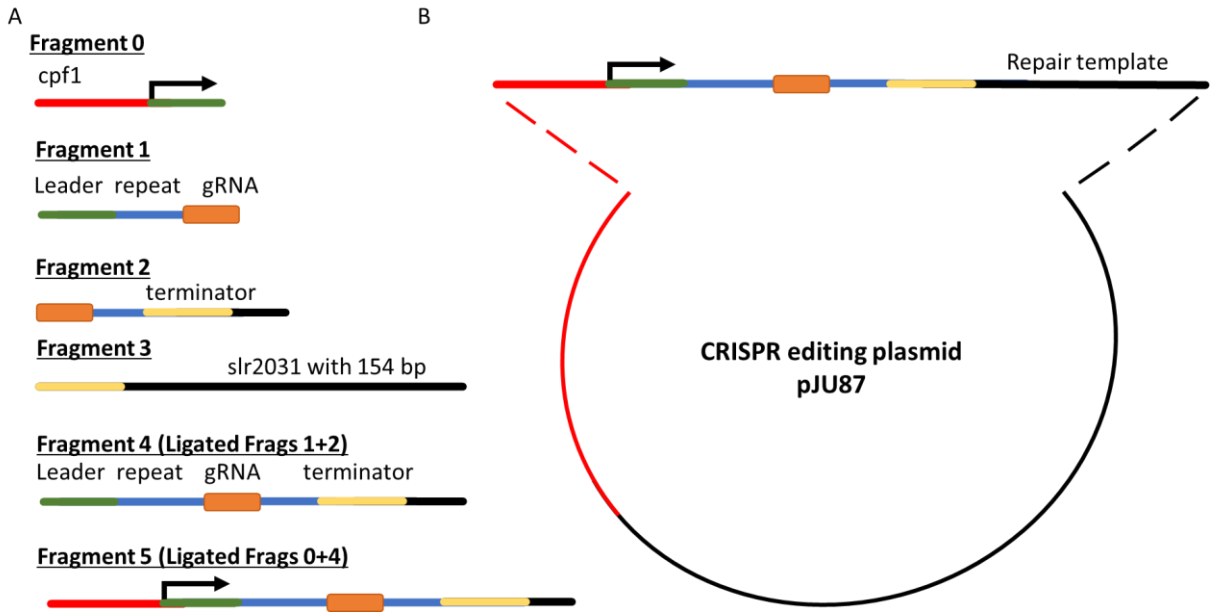

Figure 5.2: CRISPR construct to reintroduce the 154 bp sequence into GT PCC 6803. (a) Fragments to assemble using Gibson assembly in order of fragment number. (b) Overview of the final plasmid design including assembled fragments and pJU87 CRISPR backbone.

A.1.2 Results

The introduction of the 154 bp region into the GT strain (Fig. 5.1a) was performed using a CRISPR construct (Fig. 5.2). After reintroducing the 154 bp intergenic and n-terminal sequence region of *slr2031* (Fig. 5.3), the strain was not fully segregated (Fig. 5.3b). An initial growth assay with non-segregated mutants for changes in high-light response (Fig. 5.4a and 5.4b) did not yield any differences. The mutant strains did not appear to grow much differently from the wild type (Fig. 5.4), although when checked for the segregation of the 154 bp region after high-light growth no bands corresponding to the 154 bp insertion were observed (Fig. 5.4c).

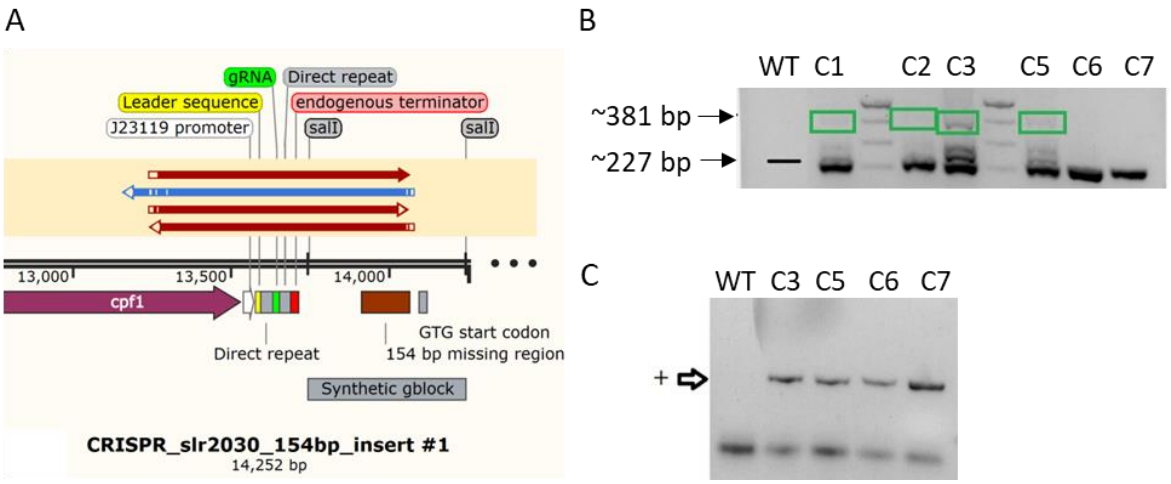

Figure 5.3: **Insertion of 154 bp region into "WT" GT PCC 6803.** (a) Positive sequence alignment (red and blue arrows) of ex-transformant colonies for the 154 bp region introduced from a synthetic gblock sequence. (b) PCR bands showing non-segregated colonies (C1-C3, C5-C7) with corresponding non-154 bp insertion and 154 bp insertion, highlighted by green boxes. (c) PCR results using primers targeting inside the 154 bp region to validate the presence of the insertion in ex-conjugate colonies (C3, C5-C7).

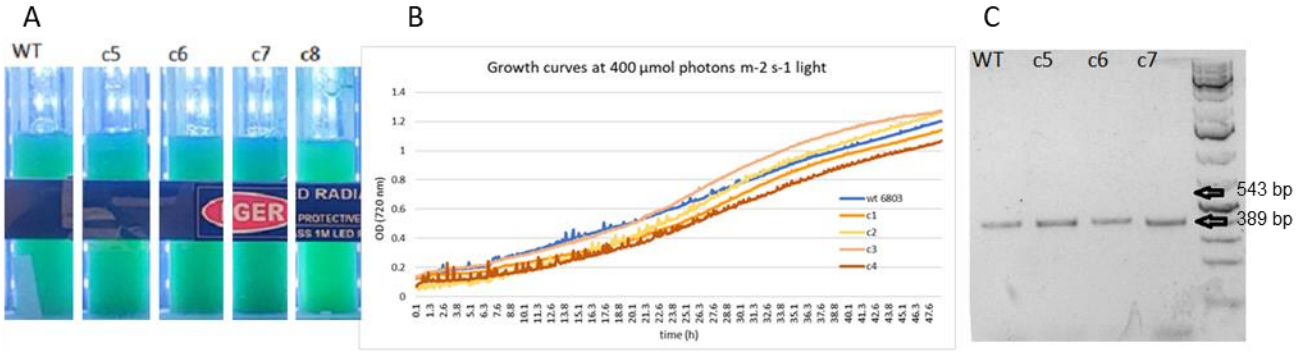

Figure 5.4: Growth of non-segregated 154bp insertion mutants (a, b) in high light intensity and (c) check for presence of insertion after high light growth. Primers: slr2030_inter_seq_fwd (TCCCGCCATTGCTAAAAAACCC), inter_seq_R (AAAGCACTGCCATCGGCATC)

A.1.3 Conclusions and future direction

Similar to the issues in creating mutations in *Synechococcus*, PCC 6803 was also resistant to mutating *hltA*. To encourage the segregations of the full-length *hltA* sequence, the mutant strain needs to be grown with higher kanamycin antibiotic concentrations and continuously streaked out onto new BG11 plates. After the strain is fully segregated, it can be assayed using a dilution series spot plates grown under increasing light intensity. The advantage of the dilution

series spot plates is that it allows for more sensitive selection to identify small differences in high-light growth, and plate assays do not have the same issues of shading as seen in the bulk liquid cultures. The obstacle to shading in bulk liquid growth assays could be overcome with a chemostatic culturing bioreactor although unfortunately, such equipment with the correct light parameters are not widespread.

A.1.4 Methods

Strains and cultivation

Wild-type and non-segregated mutant strains were pre-cultured in a multicultivator with 1% bubbling CO_2 at 30C and 50 μmol photons m^{-2} s^{-1}, mutant cultures were supplemented with Kan_{10}. For high light treatment, a light intensity of 400 μmol photons m^{-2} s^{-1} was used on 50 mL cultures, and grown under high light for several days. These cultures were harvested at the end of growth to check for the presence of the 154bp insertion.

The plasmid for the CRISPR-154bp was constructed using a gRNA targeting the intergenic *slr2030-slr2031* sequence region, and pJU87 as the backbone CRISPR editing template (Fig. 5.2). Using a gBlock gene fragment synthesized by IDT (idtdna.com) that contained the 154 bp sequence of interest and flanking sequence regions as the repair template. Fragments were assembled using Gibson assembly and transferred into PCC 6803 using conjugation. Ex-conjugant colonies were grown on Kan for selection of putative insertion mutants. A dozen mutant colonies were selected and tested using PCR for the inserted sequence. To promote segregation of the inserted sequence colonies were grown on BG11 plates supplemented with kanamycin.

Table 5.1: Primers used create and check for the 154 bp slr2031 insertion.

				Expected length (bp)	
Forward primer		Reverse primer		mutant	WT
slr2030_inter_seq_f	TCCCGCCATTGCTAAAAAACCC	slr2030_inter_seq_r	AAAGCACTGCCATCGGCATC	543	389
slr2031_154bp_f	ATGGACGAGCCGATTTTCACC	slr2031_154bp_r	TCCTTGAGAATGACCACCGGAG	224	70
154bp_cPCR_f	GCTTCATCGCTCCAACTAAC	154bp_cPCR_r	AATTGCCGGGGGGAATAAC	93	0
learder_F2	Cgctgatttaggcaaaaacgggtctaagaactttaaataatttctac	Leader_R1	ccgtttttgcctaaatcagcGC		
gRNA_fwd_slr2030	AGCCCATGGCTGTCTCCCGCGtctaagaactttaaataatttctac	gRNA_rev_slr2030	GCGGGAGACAGCCATGGGCTAtctacaacagtagaaattatttaaag		
slr2030insert_terminator_fwd	cggtagttGGTACCGCGTTAATTTTTCTTACCCATGG	slr2030_Term_rev	TAAGAAAAATTAACGCGGTACCaactaccgcattaaagc		

A.2
Engineering Natural Competence into the Fast-Growing Cyanobacterium *Synechococcus elongatus* UTEX 2973

Wendt KE, Walker P, Sengupta A, Ungerer J, Pakrasi HB. 2022. Engineering natural competence into the fast-growing cyanobacterium *Synechococcus elongatus* UTEX 2973. Appl Environ Microbiol 88:e0188221.

Author Contribution: PLW generated Figure 5.5

A.2.1 Summary

Natural transformation is the process by which bacteria actively take up and integrate extracellular DNA into their genomes. In cyanobacteria, natural transformation has only been experimentally demonstrated in a handful of species. Although cyanobacteria are important model systems for studying photosynthesis and circadian cycling, natural transformation in cyanobacteria has not been characterized to the degree that the process has been studied in other gram-negative bacteria. Two cyanobacterial species that are 99.8% genetically identical provide a unique opportunity to better understand the nuances of natural transformation in cyanobacteria: *Synechococcus elongatus* PCC 7942 and *Synechococcus elongatus* UTEX 2973 (hereafter *S.* 7942 and *S.* 2973 respectively). *S.* 7942 is a naturally transformable model system, while *S.* 2973 is a recently discovered species that is not naturally competent. Taking only 1.5 hours to replicate, *S.* 2973 is the fastest growing cyanobacterial species known, and thus is a strong candidate for serving as a model organism. However, the organism's inability to undergo natural transformation has prevented it from becoming a widely used model system. By substituting polymorphic alleles from *S.* 7942 for native *S.* 2973 alleles, natural transformation was introduced into *S.* 2973. Two genetic loci were found to be involved in differential natural competence between the two organisms: transformation pilus component pilN and circadian transcriptional master regulator rpaA. By using targeting genome editing and enrichment outgrowth, a strain that was both naturally transformable and fast-growing was created. This new *S.* 2973-T strain will serve as a valuable resource to the cyanobacterial research community.

A.2.2 Introduction

Synechococcus elongatus UTEX 2973 has the fastest doubling time of any cyanobacterium known to date and is quickly rising as a preferred model organism in the cyanobacteria research community (Yu et al., 2015a). In addition to having a doubling time of ~1.5 hours under optimal conditions, this organism is also high-light tolerant and accumulates biomass at a remarkable rate (Ungerer et al., 2018a). However, even though *S.* 2973 is a promising system for biological investigation, there is one feature that prevents it from reaching its full potential as a widely used model cyanobacterium in the laboratory: natural competence.

Natural competence is the biological capacity of a bacterium to take up and maintain naked DNA. During natural transformation, cells are incubated with donor DNA, and the cell uses its native machinery to take up and integrate this extracellular DNA into its genome. Because natural transformation requires no special equipment or donor *E. coli* strains, it is a preferred method for generating targeted mutants in many laboratories. With conjugal transformation, donor DNA must be cloned into a plasmid and transformed into *E. coli* prior to transforming the cyanobacteria. Unfortunately, some cyanobacterial genomic sequences are toxic in *E. coli*, and therefore certain mutant lines cannot be made with conjugal transformation (Golden et al., 1986, Nagarajan et al., 2011). Additionally, natural transformation allows both linear fragments and circularized DNA to be used as donor DNA. The linear fragments can be generated in the lab via PCR or ordered from commercial synthesis services, an approach that is becoming increasingly popular with the decreasing cost of these services. Natural competence is relatively uncommon across bacteria, and among the hundreds of species of cyanobacteria that have been cultivated, only a few have been naturally transformed experimentally (Nies et al.,

2020, Onai et al., 2004, Porter, 1986, Springstein et al., 2020, Trehan and Sinah, 1981, Jaiswal et al., 2018).

The unique genetic nature of *S.* 2973 has made it the topic of several scientific investigations. *S.* 2973 is >99.8% genetically identical to the model cyanobacterium *Synechococcus elongatus* PCC 7942; however, these organisms are phenotypically distinct (Yu et al., 2015a). *S.* 7942 is naturally competent and exhibits and optimal growth rate of 4.9 hours, while *S.* 2973 is not capable of undergoing natural transformation and divides in 1.5 hours under optimal conditions (Ungerer et al., 2018d, Yu et al., 2015a). There are only a handful of genetic polymorphisms that differ between the two genomes (Yu et al., 2015a). These polymorphisms include 55 SNPs, one large inversion, and one small deletion (Yu et al., 2015a). The genetic similarities between these two organisms provided a unique opportunity to investigate the genetic basis for accelerated growth.

In our previous work we used comparative genomics to identify three polymorphic loci that were responsible for the difference in growth rate between *S.* 7942 and *S.* 2973 (Ungerer et al., 2018d). Using a CRISPR/Cas12a genome editing system, individual polymorphic alleles in *S.* 2973 were substituted out for their counterpart sequences in the naturally competent sister strain in *S.* 7942. From this analysis, three genetic loci were found to cause the difference in growth rate between the two organisms: *atpA, nadK,* and *rpaA.* At the conclusion of this study, we presented a *S.* 7942 background strain with substituted *S.* 2973 alleles for these three loci that grew at a rate on par with wild type *S.* 2973. However, while this strain had gained accelerated growth, it lost one of the key features that has made *S.* 7942 a desirable model system: natural transformation.

In this study we assess which of the accelerated growth polymorphisms also affects natural transformation. Additionally, we identify a secondary polymorphic locus that contributes to differential natural transformability between *S.* 7942 and *S.* 2973. Using this comparative genomics strategy and an accelerated growth rate enrichment, we were ultimately able to engineer a version of the *S.* 2973 strain that has the potential to fill the niche for a fast-growing, naturally competent model system in cyanobacterial research.

A.2.3 Results

Past reports have shown that the structure and methylation state of the donor DNA have an effect on cyanobacterial transformation efficiency; therefore, preliminary experiments were performed to determine the optimal form of donor DNA (Beauchamp et al., 2017). It was important to determine which form of donor DNA would be best to use in future experiments to determine the quantitative restoration of natural transformation. Therefore, wild-type *S.* 7942 was assayed for natural transformation efficiency with donor DNA that had structural and methylation state variations.

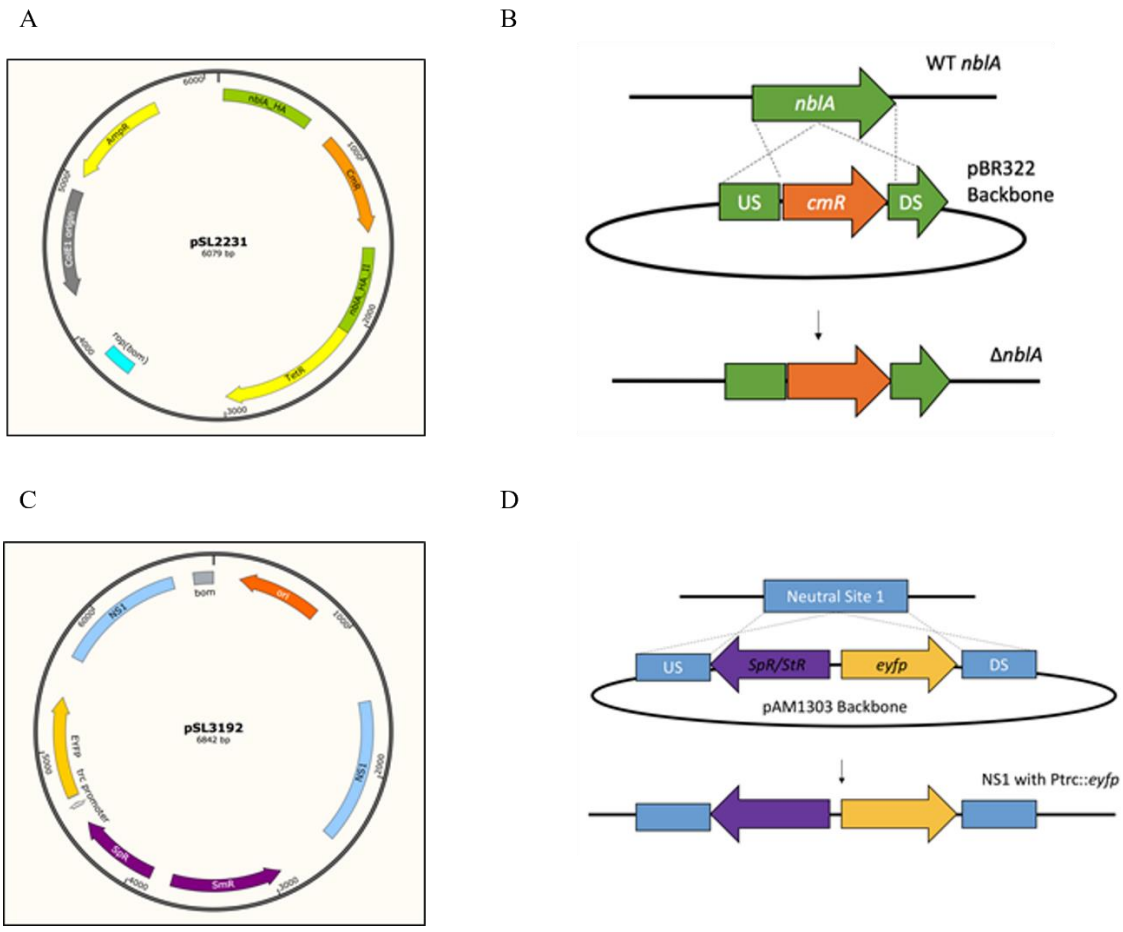

Figure A1: The pSL2231 construct was used to assay natural transformation efficiency, and the pSL2231 and pSL3192 constructs were used to assay restored natural transformation capacity. The pSL2231 construct is designed to integrate a chloramphenicol resistance cassette to disrupt the *nblA* locus of *S.* 7942/2973. This construct has a pBR322 backbone and is 6079 bp in size. (A) Plasmid map of pSL2231. Orange arrow indicates the chloramphenicol resistance gene and green blocks represent homology arms. (B) pSL2231 diagram of integration of Cm^R gene into the *nblA* locus of the *S.* 7942/2973 genome. The pSL3192 construct integrates a spectinomycin cassette and PTrc::eYFP cassette into Neutral Site I of *S.* 2973 mutant lines. The plasmid has a pAM1303 backbone and is 6842 bp in size. (C) Plasmid map of pSL3192. (D) pSL3192 diagram of integration into NS1 of the *S.* 7942/2973 genome.

Structurally different forms of the pSL2231 construct sequence with varying lengths of plasmid backbone flanking the homology arms were used as the donor DNA, as shown in Fig A1. This plasmid integrates a chloramphenicol resistance cassette into the *nblA* gene of *S.* 7942 (Synpcc7942_2127). The pSL2231 plasmid has a pBR332 backbone, a suicide backbone lacking

a cyanobacterial origin of replication (Bolivar et al., 1977). NblA is a cyanobacterial protein that is involved in the disassembly of the phycobilisome photosynthetic antenna under nitrogen deprivation conditions (Collier and Grossman, 1994). When nitrogen resources are lacking, cells disassemble the large phycobilisomes, resulting in a yellowing appearance of the culture. Disrupting *nblA* prevents cells from bleaching in nitrogen replete conditions (Collier and Grossman, 1994). Thus, the pSL2231 construct yielded a visual phenotype that was used to confirm natural transformants in addition to PCR analysis. Four different donor DNA structures were tested in *S.* 7942: intact pSL2231 plasmid, linearized pSL2231 plasmid, linearized pSL2231 plasmid with one half of the plasmid backbone truncated, and only the homology region and integrating cassette of pSL2231, as shown in Fig A2. Additionally, two different methylation states were tested for Construct 2, Construct 3, and Construct 4 (Fig. A2). While one version of the donor DNA was made with plasmid that was isolated from *E. coli* and then digested with appropriate restriction endonucleases, the second version of the DNA was amplified with PCR to generate an unmethylated fragment. It has not previously been tested whether methylation decorations from *E. coli* have a demonstrable effect in protecting donor DNA during natural transformation in *S.* 7942. Table A1 shows the results of natural

transformation experiments with various forms of donor DNA.

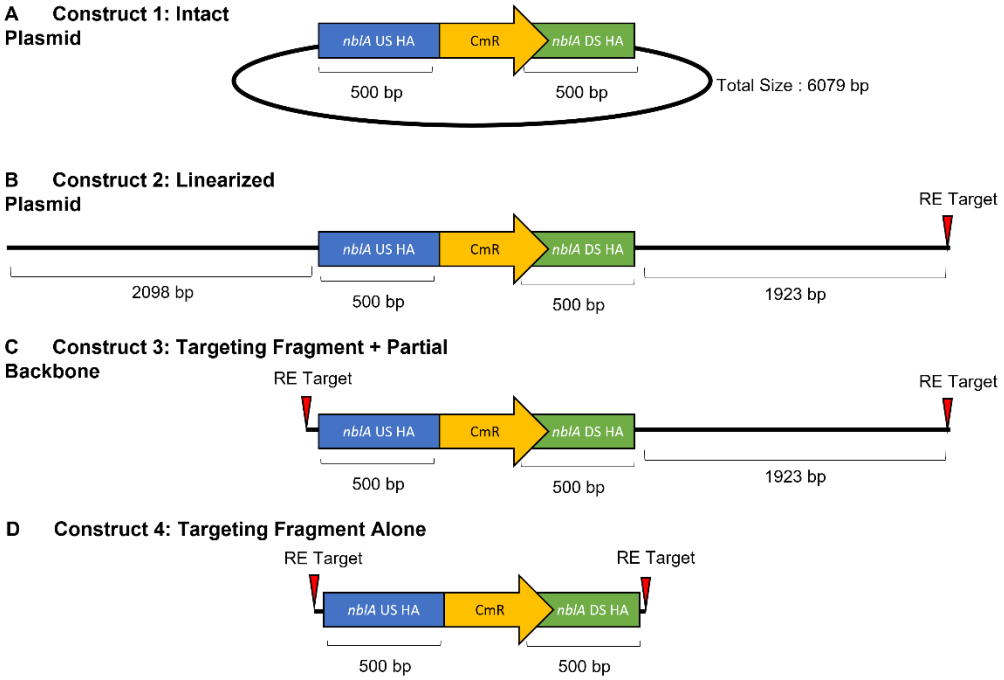

Figure A2: Various lengths of flanking backbone used to assay natural transformation efficiency in *S.* 7942. Blue denotes the upstream homology arm, yellow denotes the chloramphenicol resistance gene, and green denotes the downstream homology arm. The thick black line is used to show the pBR322 backbone. Red triangles denote loci that were targeted by restriction enzymes for linearization. (A) Intact plasmid was isolated from DH5- E. coli and used directly for transformation. For B, C, and D, the pSL2231 plasmid was isolated from DH5-α E. coli and linearized by restriction digestion for one set of experiments and PCR for a parallel set of experiments. (B) The backbone DNA flanking the homology arms was not truncated during linearization. (C) The backbone DNA flanking the homology arms was truncated on one side of the recombinant region. (D) The backbone DNA flanking the homology arms was truncated on both sides of the homologous regions.

Overall, it was found that intact plasmid transformed with the highest efficiency and thus this form of donor DNA was used to assess the restoration of natural transformation for subsequent experiments. After creating a baseline for natural transformation efficiency in wild-type *S.* 7942, the next step was to attempt to introduce natural transformation into *S.* 2973. To introduce natural competence into *S.* 2973, alleles from the naturally transformable sister strain *S.* 7942 were substituted for alleles in the *S.* 2973 strain. A CRISPR/Cas12a approach was

implemented to avoid integrating selective markers into the genome (Ungerer and Pakrasi, 2016b). With this strategy, mutant lines with multiple allele substitutions were generated without integrating antibiotic resistance cassettes into the genome. Various markerless mutant lines were therefore created and assayed for natural transformation efficiency.

Two intact plasmids, pSL2231 and pSL3192, were used to assess the restoration of natural transformation to mutant lines. These constructs were selected because successful transformation results in phenotypes that can be visually assessed. As previously discussed, the integrated pSL2231 construct generates a non-bleaching phenotype under nitrogen deprivation conditions. The pSL3192 construct integrates an eYFP cassette under the constitutive Ptrc promoter into Synpcc7942_2498 locus (widely known as Neutral Site 1) of *S.* 7942 and *S.* 2973, as shown in Fig. A1. Cells that successfully transform with the pSL3129 construct exhibit YFP fluorescence, which can be measured on a fluorescence plate reader to confirm successful transformation. The results of natural transformation experiments with these cultures on various mutant lines are displayed in Table A2.

Table A1. Methylation and length of flanking DNA regions affect natural transformation efficiency in wild-type *S.* 7942. Donor pSL2231 DNA with the structures displayed in Figure A2 were used to transform wild-type *S.* 7942 cells. All molar ratios of DNA molecules to cells were kept constant in the various transformation trials. Intact plasmid and digestion linearized donor DNA were isolated from DH5-α *E. coli* cells. A minimum of three replicates were performed for all experimental data presented.

	Transformation Efficiency (Transformants/Viable Cell)		
Donor DNA Structure	Intact Plasmid	Linearization by Digestion	Linearization by PCR
Construct 1: Intact Plasmid	4.6E-5 ± 5.2E-6	-	-
Construct 2: Linearized Plasmid	-	3.0E-5 ± 3.4E-6	2.1E-6 ± 6.6E-7
Construct 3:	-	4.7E-6 ± 4.4E-7	5.1E-7 ± 1.1E-7

Targeting Fragment with Partial Backbone			
Construct 4: Targeting Fragment Alone	-	1.1E-7 ± 1.3E-8	4.4E-8 ± 1.2E-8

Initially, we set out to identify which of the three accelerated growth loci described in our previous manuscript was also responsible for the lack of natural transformation in the engineered fast-growing *S.* 7942 *atpA*$_{2973}$ *nadK*$_{2973}$ *rpaA*$_{2973}$ strain (Ungerer et al., 2018d). The first strain that we tested lacked the substituted *rpaA* allele (*S.* 7942 *atpA*$_{2973}$ *nadK*$_{2973}$); this strain was tested first because we already had it on hand from the accelerated growth study. Experiments showed that the *S.* 7942 version of the *rpaA* allele was necessary for natural transformation, but not sufficient to engineer natural transformation into *S.* 2973 on its own, as seen in Table A2. Additionally, when the *rpaA* promoter and coding mutations in *S.* 7942 were separately substituted for their *S.* 2973 counterparts, neither strain exhibited a detectable level of natural transformation.

Table A2. Transformation efficiencies in *S.* 2973/7942 substituted allele strains using pSL2231 and pSL3192 donor DNA. Error values represent the standard error. Samples without values did not exhibit a detectable level of natural transformation.

	Transformation Efficiency (Transformants/Viable Cell)	
Strain	pSL2231	pSL3192
Synechococcus 7942	4.6E-5 ± 5.2E-6	4.9E-7 ± 1.7E-7
Synechococcus 2973	-	-
Synechococcus 7942 *atpA*$_{2973}$ *nadK*$_{2973}$ *rpaA*$_{2973}$*	-	-
Synechococcus 7942 *atpA*$_{2973}$ *nadK*$_{2972}$*	5.4E-5 ± 1.8E-5	1.6E-5 ± 5.6E-6
Synechococcus 2973 promoter *rpaA*$_{7942}$	-	-
Synechococcus 7942 promoter *rpaA*$_{2973}$	-	-
Synechococcus 2973 coding *rpaA*$_{7942}$	-	-
Synechococcus 7942 coding *rpaA*$_{2973}$	-	-
Synechococcus 2973 promoter + coding *rpaA*$_{7942}$	-	-
Synechococcus 2973 *pilN*$_{7942}$	-	-
Synechococcus 7942 *pilN*$_{2973}$	-	-

Synechococcus 2973 $pilN_{7942}$ $rpaA_{7942}$ (Clumpy)	1.7E-6 ± 5.0E-7	5.7E-7 ± 1.2E-7
Synechococcus 2973 $pilN_{7942}$ $rpaA_{7942}$ promoter $pilMNO_{7942}$ (Clumpy)	2.3E-6 ± 4.9E-7	8.6E-7 ± 3.5E-7
Synechococcus 2973-T	1.1E-7 ± 5.0E-8	6.7E-8 ± 2.0E-8

*Strains were generated as described in (Ungerer et al., 2018d)

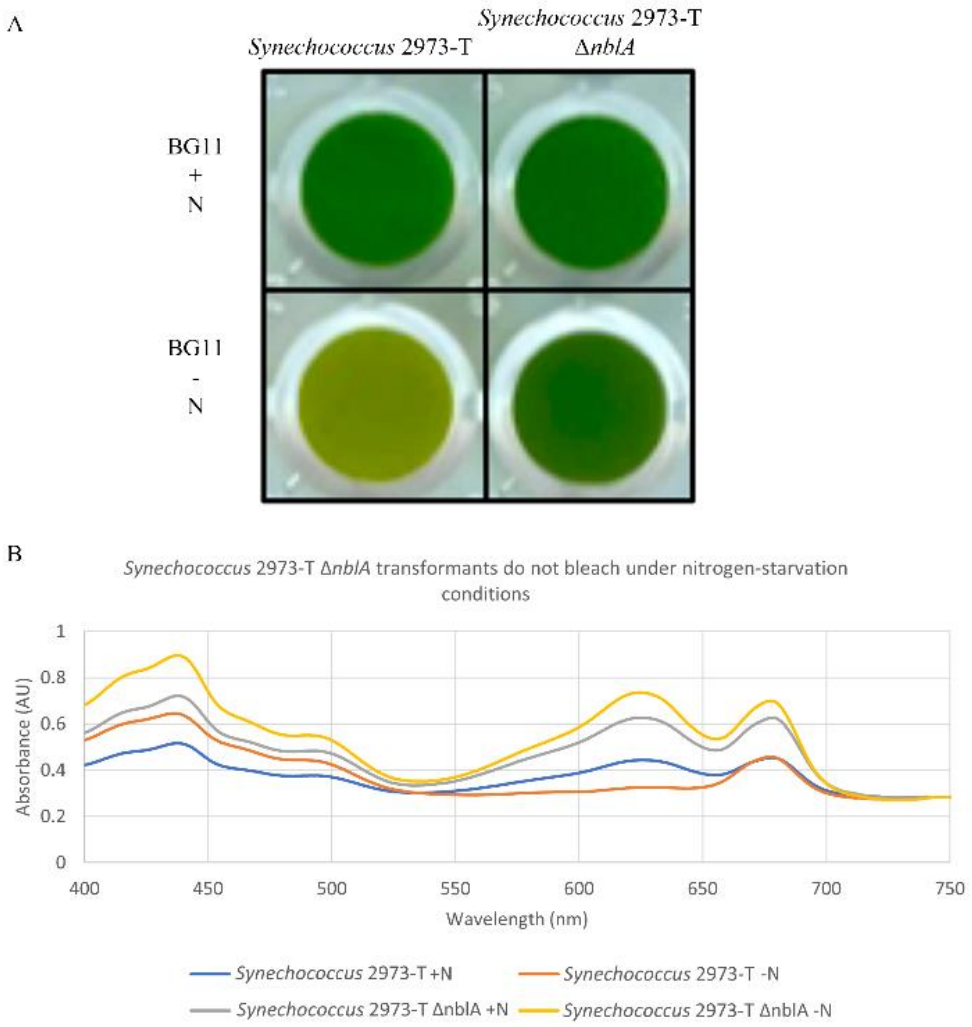

Figure A3: Transformants from the pSL2231 exhibit the expected nonbleaching phenotype in addition to chloramphenicol resistance. (A) The untransformed *S.* 2973-T strain bleaches under nitrogen deprivation conditions, however, a representative colony from the natural transformation of *S.* 2973-T plasmid with pSL2231 loses this capacity under nitrogen deprivation conditions. (B) An example *S.* 2973-T nblA interruption transformant shows no change in absorbance under nitrogen deprivation conditions. PC, phycocyanin and Chl, chlorophyll.

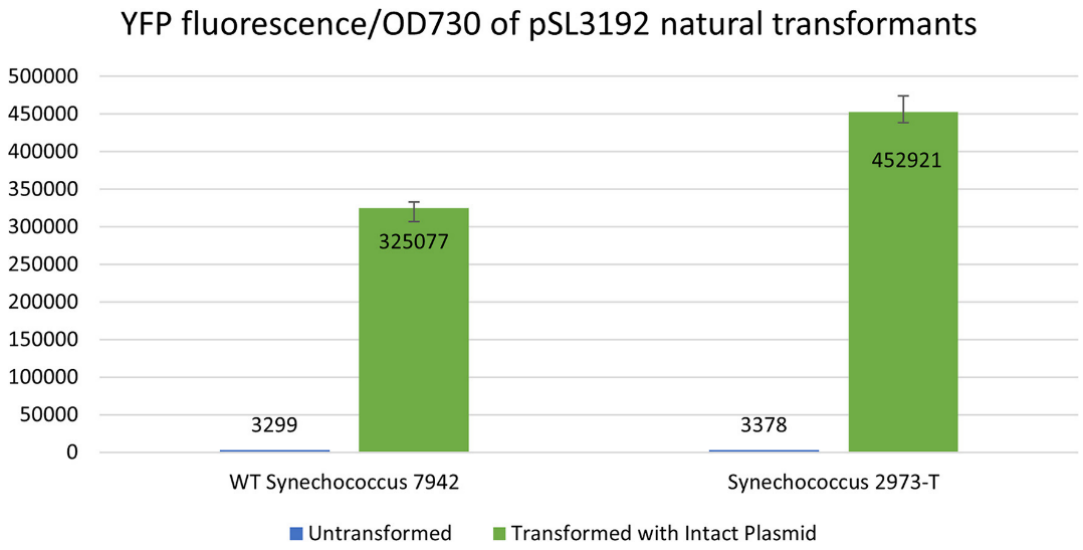

Figure A4: Quantitative fluorescence measurements show pSL3192 natural transformants fluorescence more than untransformed cells. Plate reader fluorescence data are normalized for optical density (OD) based on absorbance at 730 nm. Error bars show standard error.

These data indicated that at least one additional polymorphic locus would need to be converted to restore natural transformation to *S.* 2973. Therefore, we cross-referenced previously published literature on natural transformation in cyanobacteria with the annotated list of polymorphic SNPs. This is how the *pilN* allele was identified as a candidate locus for causing the difference in natural transformation between *S.* 7942 and *S.* 2973 (Bhaya et al., 2000). When the *S.* 2973 *rpaA* and *pilN* alleles are replaced with the *S.* 7942 versions of these alleles, natural transformation is restored, as seen in Table A2. Additionally, it was found that substituting the *S.* 7942 *pilN* allele for the *S.* 2973 version was enough to disable natural transformation in the resulting strain, but substitution of the *S.* 2973 *pilN* allele for the *S.* 7942 allele was not sufficient to restore natural transformation.

A third polymorphism was also engineered into this mutant line in an effort to further increase natural transformation efficiency. The *pilMNO* operon has a 243 base pair insertion directly upstream of *pilM* in *S.* 2973 that is not present in *S.* 7942, and it was previously shown

that expressing an intact version if the gene resulted in natural transformability (Li et al., 2018). With this in mind, the *S.* 7942 version of this region was also substituted into the genome of the *S.* 2973 *pilN*$_{7942}$ *rpaA*$_{7942}$ mutant. However, the new line did not exhibit a significant increase in natural transformation efficiency, so no further analysis was performed on this mutant line (Table A2).

As we generated strains with allele substitutions to introduce natural transformation into *S.* 2973, we qualitatively observed that the mutants that we were generating were growing more slowly than wild type *S.* 2973. Having identified *rpaA* as a locus associated with growth rate in our previous publication, this slower growth was not wholly unexpected; however when we performed quantitative doubling time analysis, we realized that the *S.* 2973 *pilN*$_{7942}$ *rpaA*$_{7942}$ strain was growing at a rate slower than that of both single mutants and wild-type *S.* 7942, as seen in Table A3.

Table A3: *S.* 2973-T has a doubling time comparable to WT *S.* 2973. Cells were grown in a PSI multicultivator at 38°C bubbling 1% CO_2. Light intensities used: 400 µmol·m^{-2}·s^{-1} and 900 µmol·m^{-2}·s^{-1} (Ungerer et al., 2018a). Strains with the "Clumpy" indicator exhibit morphology of cells grown in liquid culture.

	Doubling Time (hr)	
Strain	400 µmol · m^{-2} · s^{-1}	900 µmol · m^{-2} · s^{-1}
Synechococcus 7942	4.31 ± 0.49	3.12 ± 0.19
Synechococcus 2973	2.64 ± 0.28	2.16 ± 0.47
Synechococcus 2973 *pilN*$_{7942}$	3.17 ± 0.69	2.39 ± 0.70
Synechococcus 2973 *rpaA*$_{7942}$	3.09 ± 0.56	3.15 ± 0.77
Synechococcus 2973 *pilN*$_{7942}$ *rpaA*$_{7942}$ (Clumpy)	4.96 ± 0.44	3.98 ± 0.72
Synechococcus 2973 -T	2.26 ± 0.22	2.22 ± 0.40

It was observed that the cells in the engineered transformable strain were clumping together in liquid culture; this physiological change was inferred to be responsible for the growth rate increase. With this in mind, a directed evolution strategy was implemented to reduce the clumping in the *S.* 2973 $pilN_{7942}$ $rpaA_{7942}$ line. Cells were grown under fast growth conditions, 900 $\mu mol \cdot m^{-2} \cdot s^{-1}$ light and 1% CO_2 bubbled through the culture in a PSI multicultivator. Once the culture was fairly dense, the tube was removed from the chamber and left undisturbed on the benchtop for two hours. After the two-hour period, most of the cells had settled down to the bottom of the tube, which resulted in a dense green layer at the bottom of the tube covered by clear liquid. A volume of 500 µL was then pipetted from the clear liquid at the top of the tube and inoculated into a fresh tube of BG11. The cells were returned to fast-growth conditions in the multicultivator. This strategy was continued for several weeks with dilutions around every 60 hours. At the end of this enrichment, cells were diluted to a low density and spread on BG11 plates to form single colonies. One of these colonies was grown in liquid medium to ensure that the clumpy phenotype had been eradicated.

The newly isolated strain was assayed for natural transformation capacity and accelerated growth rate. The strain was found to be naturally transformable and was named *S.* 2973-T for its Transformable phenotype. Although the strain was found to be naturally transformable, it exhibited a transformation efficiency lower than that of the clumpy *S.* 2973 $pilN_{7942}$ $rpaA_{7942}$ strain (Table A2). Transformants of the *S.* 2973-T strain were initially confirmed via PCR. Subsequently, successful transformation was confirmed with phenotypic assays as shown in Fig. A3 and Fig. A4. The new strain also exhibited a growth rate on par with *S.* 2973 (Table A3) and did not clump when grown in liquid culture.

Additionally, we wanted to determine the relative rate of expression of genes that had been engineered to have the substituted *S.* 7942 alleles. Quantitative PCR was used to compare the expression of genes in the *pilMNO* operon and *rpaA* in *Synechococcus* 2973-T to their expression in WT *S.* 2973. These data showed that all of the engineered loci were expressed at a higher rate in *S.* 2973-T, as seen in Fig. A5.

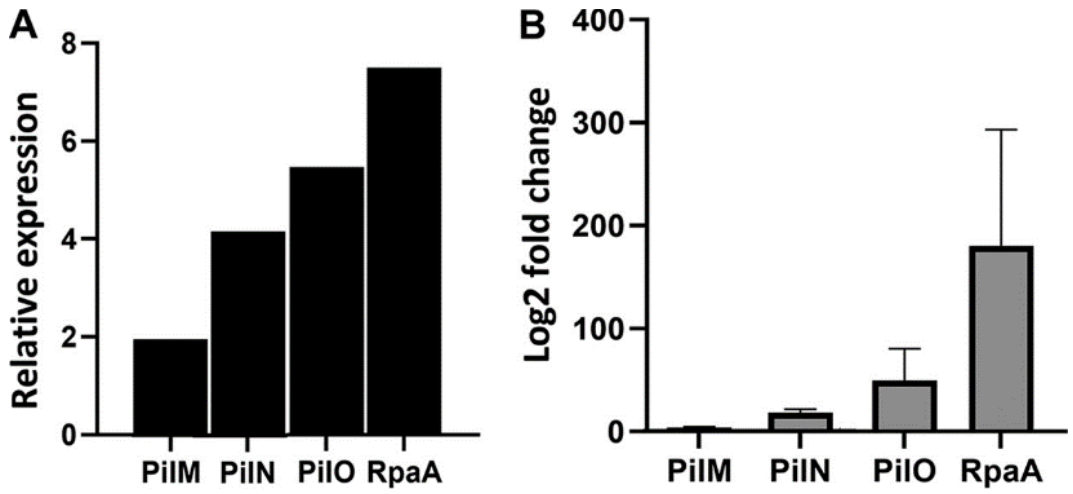

Figure A5: Genes in the pilMNO operon and rpaA are expressed at a higher level in *S.* 2973-T than in WT *S.* 2973. (A) Relative expression (ΔCt) values of pilMNO and rpaA genes of *S.* 2973T relative to those of *S.* 2973, shown in absolute values. (B) Relative fold change (2−ΔΔCt) in pilMNO and rpaA gene expression of *S.* 2973T in relation to WT *S.* 2973. Bars and whiskers represent the mean and standard deviation of 2−ΔΔCt values (n = 3 or 4). Total RNA was extracted from cultures that underwent the same treatment as cells being prepared for experimental natural transformation.

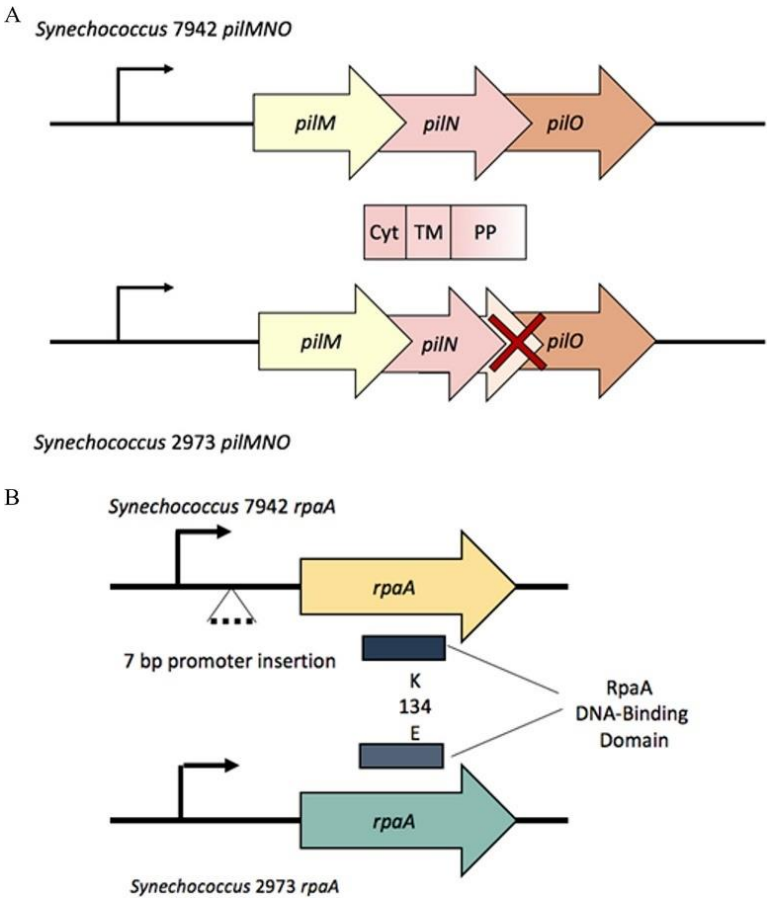

Figure A6: Modification of pilN and rpaA loci to generate the naturally transformable *S.* 2973-T strain. (A) Modifications to *pilN* included replacement of the truncated version of the gene in *S.* 2973 with the intact *S.* 7942 allele. Cyt is the cytoplasmic domain of PilN, TM is the transmembrane domain of PilN, and PP is the periplasmic domain of PilN. (B) Changes to the rpaA locus included a 7-bp promoter insertion and the replacement of the *S.* 2973 *rpaA* DNA-binding domain SNP with the *S.* 7942 sequence.

Figure A7: Clumping morphology of *Synechococcus* strains with restored natural transformation capacity. Left, *S.* 2973 $pilN_{7942}$ $rpaA_{7942}$ strain prior to enrichment. Right, *S.* 2973-T strain.

In order to determine how the enrichment outgrowth had genetically altered the fast-growing, naturally transformable organism, we used whole genome sequencing. Sequencing revealed that there were seven loci with genetic differences in *S.* 2973-T that could be causing the clumping phenotype, as seen in Table A4. Genome sequencing also revealed that there were several unintended mutations that were present in both strains. Potentially, these mutations may have come from off-target editing from the genome modification system that was used to generate the allele substitution strains. These shared mutations were ruled out as candidates for causing the difference in clumping morphology because they were present in both strains , as displayed in Table SA1. The remaining differences between the two genomes are displayed in Table A4. Additionally, the *pilN* and *rpaA* loci were sequenced via Sanger sequencing to ensure that no mutations had occurred at these loci.

TABLE A4: Whole genome sequencing comparison of clumping and non-clumping strains identifies several polymorphic loci[a]

Gene tag(s)	Gene function	Nucleotide change	Protein sequence change(s)
M744_03955 and M744_03965	ABC transporter (coding)/substrate-binding protein	Δ 1 bp	Frameshift
M744_06035	PilQ-like (coding)	+G, A → G, A → G	Frameshift, I344V, T345A
M744_13325	HAD family hydrolase (coding)	T → C	I92T
M744_11110 and M744_11115	Serine-threonine protein kinase/CheY-like protein (intergenic)	$(G)_{9 \rightarrow 10}$	Frameshift
M744_11615	PilJ (coding)	A → C	Q779P

[a]Whole genome Illumina sequencing analysis of non-clumping *S.* 2973-T and clumping *S.* 2973 promoter $pilMNO_{7942}$ $pilN_{7942}$ $rpaA_{7942}$ strains was performed. Loci in the table are those that were polymorphic between the two strains, suggesting that one or more of the listed locus/loci are responsible for the clumping phenotype.

A.2.4 Discussion

The initial experiments to determine the best form of donor DNA to use for efficient natural transformation revealed that intact plasmid transformed with the highest efficiency into *S.* 7942. Previous studies in other bacteria showed that DNA enters the cell in a linearized form during natural transformation, and it was thus hypothesized that linearizing the DNA prior to uptake would increase natural transformation efficiency (Hamilton and Dillard, 2006). Random shearing or shearing caused by extracellular restriction enzymes could result in linearization at inopportune locations such as within the homology arms or the selective marker. However, it is possible that pre-linearizing the donor DNA made the material more susceptible to degradation via extracellular exonucleases by increasing the amount of time that the DNA was in a linear form prior to DNA uptake.

Additionally, using *E. coli*-methylated donor DNA, as opposed to unmethylated donor DNA generated with PCR amplification, had a significant effect on natural transformation efficiency (Table A1). *E. coli*-methylated donor DNA transformed with a higher efficiency than PCR-amplified DNA for all of the constructs tested. The efficiency difference between methylated and unmethylated DNA is around one order of magnitude for both the linearized plasmid and partially truncated backbone samples. This suggests that the native *E. coli* Restriction Modification system is decorating the donor DNA in a manner that helps protect it from cyanobacterial nucleases. Comparison of the *E. coli* methylation system to the cyanobacterial nuclease system suggests that the Dcm methylase in the DH5-α *E. coli* methylates donor DNA in a manner that protects against the Type II endonuclease coded by the Synpcc7942_2459 locus.

The most interesting finding of the experiment with the various forms of donor DNA was that the presence of the plasmid backbone had a remarkable effect on natural transformation

efficiency. In theory, because the pBR322 backbone is non-homologous to the *S.* 7942 genome, there should not be a notable difference in transformation efficiency between Construct 2, Construct 3, and Construct 4. However, truncating the backbone resulted in a significant decrease in natural transformation efficiency. For the *E. coli* methylated donor DNA, the loss of one half of the flanking backbone resulted in a six-fold decrease in transformation efficiency and the loss of all of the backbone resulted in a 270-fold decrease in natural transformation efficiency. One explanation for these data are that the backbone is acting as a buffer region that protects the homology sequence from degradation by exonucleases. Exonucleases eat away at the end of a linearized fragment and the presence of the backbone DNA buffers the homology arms from being degraded, thus increasing overall recombination and transformation efficiency. Removing the backbone that flanks the upstream homology arm decreases efficiency less than removing both flanking backbone regions because one homology arm is still protected from exonuclease activity. Previous publications on the activity of exonuclease RecJ in cyanobacteria suggest that this protein may be responsible for the phenomenon (Kufryk et al., 2002, Racharaks et al., 2021). Additionally, it is possible that the donor DNA is linearized at a random site during uptake, and the larger backbone reduces the chance that linearization occurs within the region of the donor DNA that is intended to undergo recombination (Chen and Dubnau, 2004).

Overall, these experiments provided insight on ways to optimize natural transformation efficiency. Particularly, if a fragment is difficult to transform into cells, it might be possible to achieve transformation by isolating the donor DNA from *E. coli* so that it has a protective methylation signature or by adding buffer sequences to flank the recombining fragment. In this project, experiments determined that the highest transformation efficiency was yielded from intact

plasmid donor DNA. Therefore, this was the form of donor DNA that was used to assess the restoration of natural competence in subsequent experiments on *S.* 2973 mutant lines.

Differential natural transformation between *S.* 7942 and *S.* 2973 was found to be affected by the *pilN* and *rpaA* loci. Functionally, PilN aids in aligning the outer membrane pore with the transformation pilus to facilitate DNA uptake. When DNA is brought into the cell via natural transformation, the Type IV-like transformation pilus actively attaches to extracellular DNA and transports it into the cell via a ratcheting motion. Electron cryotomography has been used to characterize the structure of the Type IVa pilus in the gram-negative bacterium *Myxococcus xanthus*, and these structures show that PilN is anchored in the inner membrane and extends into the periplasm (Chang et al., 2016). The sequence of the *M. xanthus pilN* is similar to that of *pilN* in *S.* 7942, and studies of PilN in the cyanobacterium *Synechocystis* sp. PCC 6803 suggest that the protein has a similar functionality in cyanobacteria. In *S.* 2973, there is a point mutation in *pilN*, which replaces a glutamine codon with an early stop codon, as shown in Fig. A6a. This results in a protein with a truncated periplasmic domain, which would render it structurally inoperable.

Furthermore, *pilN* lies in the *pilMNO* operon, which contains other proteins that are in the transformation pilus. Because of polar effects in bacterial polycistronic gene expression, it seemed likely that the mutation in the *S.* 2973 *pilN* would lead to reduced *pilO* expression. Analysis of RNA-sequencing data presented in (Ungerer et al., 2018d) confirmed this phenomenon, as shown in Table A5. While wild-type *S.* 2973 exhibited a one to two-fold reduced expression of *pilM* and *pilN*, the *pilO* expression was decreased 58-fold compared to wild-type *S.* 7942.

TABLE A5: RNA-sequencing analysis shows differential expression between *Synechococcus* 7942 and *Synechococcus* 2973 *pilMNO* genes[a]

Gene	TPM fold change[b]
pilM	−1.40
pilN	−2.07

Gene	TPM fold change[b]
pilO	−58.90

[a] RNA-sequencing data were collected for wild-type (WT) *Synechococcus* 7942 and WT *Synechococcus* 2973. [b] TPM (Transcripts Per Million) is a relative expression value normalized for read depth and gene length (**11**).

The *rpaA* locus also affects natural transformability between *S.* 7942 and *S.* 2973. The *S.* 2973 version of *rpaA* has multiple differences from the *S.* 7942 allele; there is a seven base pair promoter insertion and a codon-changing point mutation in the DNA-binding domain of the protein (Fig. A6b). The mutation in the DNA binding region converts residue 134 from a lysine in *S.* 7942 to a glutamic acid in *S.* 2973. Because RpaA is a transcription factor, these polymorphisms result in a change in both the dosage of the transcription factor and the affinity to which it binds to its targets. There has been a large body of work performed on cyanobacterial RpaA because of its role in facilitating circadian transcriptional cycling in these photosynthetic organisms, however, the link between RpaA and natural transformation was only published recently (Taton et al., 2020).

The physiological differences in *S.* 2973 and *S.* 7942 strains with different native versions of RpaA were thoroughly analyzed in our previous publication discussing the polymorphisms that give rise to accelerated growth (Ungerer et al., 2018d). The same *rpaA* polymorphisms that were found to give rise to accelerated growth in *S.* 2973 are those that also cause a loss of natural competence in the organism (Ungerer et al., 2018d). For *S.* 2973 to be a desirable model system, its accelerated growth needs to be maintained when natural competence is restored. Thus, a strain that was capable of both accelerated growth and natural transformation was created to fill this niche in cyanobacterial research.

Recombinant strains in which the *pilN* allele and *rpaA* allele are independently substituted do not exhibit detectable levels of natural transformation. A previous publication reported that placing a copy of the *S.* 7942 *pilN* in trans under a constitutive promoter gave rise to low efficiency natural transformation (Li et al., 2018). However, even though this approach would have generated plenty of intact PilN, it would not have mitigated the polar effects reflected in *pilO* transcription levels, which is likely part of why the natural transformation efficiency was low for this transgenic line (4.00E-8) (14). The quantitative PCR data displayed in Fig. A5 underscore the reduction in *pilO* expression that results from the truncated *pilN* allele. Replacing only the *S.* 2973 *pilN* with the *S.* 7942 allele did not restore natural transformation even nominally, as seen in Table A2.

Overall, it was observed that the pSL2231 donor DNA transformed at a higher rate than pSL3129 (Table A2). One explanation is that the region of the plasmid undergoing recombination is smaller in pSL2231; it is thought that smaller insertions are more efficiently integrated into the genome because homologous regions of the recombining fragments line up more readily.

Genome editing was used to replace two alleles in *S.* 2973, and the *S.* 2973-T strain was later generated by growing this strain under conditions that enriched for the non-clumpy phenotype. When cultured in liquid, cells of the *S.* 2973 $pilN_{7942}$ $rpaA_{7942}$ strain would nucleate and form particulate rafts (Fig. A7, left). This phenotype was also observed in the *S.* 2973 $pilN_{7942}$ single mutant line (data not shown). It seemed likely that the clumping morphology limits growth rate by limiting the cells' ability to access light. Cells on the interior of the clump do not get as much light as cells at the exterior and therefore the growth rate is reduced across the culture.

The accelerated growth enrichment was performed in hopes that random mutations would accumulate and result in a non-clumpy strain. However, with this approach it was unknown which specific loci underwent mutations that gave rise to the conversion from clumping to planktonic morphology. Therefore, the *S.* 2973-T strain and *S.* 2973 $pilN_{7942}$ $rpaA_{7942}$ lines were sequenced for comparison.

Several of the polymorphisms detected during sequencing stand out as good causative candidates for the clumping morphology. PilJ, for instance, is a strong candidate because it has been classified in the literature as a minor pilin for other species of naturally competent bacteria (Piepenbrink et al., 2015). Since the reduction in clumping morphology led to a decrease in natural transformation efficiency, this locus seems like a good candidate. The amino acid change to PilJ occurs in the methyl-accepting chemotaxis domain in the protein tail, which is used as a localization signal to ensure transport of pilins to the membrane. Prior to being integrated as subunits into the transformation pilus, the signal sequence is cleaved. The biochemical change from an amino acid with a polar side chain to an amino acid with a non-polar side chain could change the structure of the signaling sequence tail and could alter cleavage of the signal sequence. Furthermore, a study in *Neisseria gonorrhoeae* reported type IV pilus-medidiated cellular aggregation in *pilJ* mutant lines, a phenotype that was not exhibited by the deletion *pilJ* strains (Carbonnelle et al., 2006).

Additionally, the PilQ-like protein is a good candidate for causing the clumping morphology and difference in natural transformation efficiency. PilQ monomers compose the outer membrane pore through which the transformation pilus protrudes from the cell (Pakrasi and Wendt, 2019). Because there is a large frameshift mutation in the M744_06035 protein sequence, it is difficult to make predictions about how this protein could be affecting clumping

and natural transformation. However, the protein's similarity to PilQ makes it a promising candidate for causing the phenotypic differences between the clumpy and un-clumpy strains.

Although there are some promising candidate loci in this set, any of the individual changes or any combination of the differences may be responsible for the decrease in natural transformation efficiency, increase in growth rate, and abolishment of clumping morphology. Even though the clumping morphology was not the primary focus of this study, and it was in fact intentionally abolished, it could be a desirable attribute for researchers seeking a cost-effective way to harvest cells for biosynthetic production with cyanobacteria. In any case, identifying the limited number of genomic differences between the clumpy and un-clumpy strains will enable future research endeavors on this topic.

Overall, this work used two sister cyanobacteria strains to combine two desirable attributes: accelerated growth and natural competence. By substituting the alleles from one sister organism for the alleles of the other, two polymorphic loci were found to play a role in differential natural competence between *S.* 7942 and *S.* 2973. One locus had a well-established role in natural transformation (*pilN*), but the other locus has only recently been associated with natural transformation in cyanobacteria (*rpaA*). However, just combining these two alleles was not enough to yield an organism with the desired properties. In order to combine both accelerated growth and natural competence into a single strain, natural selection was implemented to remove the clumping morphology and restore accelerated growth to the organism. The resulting *S.* 2973-T strain will fill a niche in cyanobacterial research for a fast-growing strain that is capable of undergoing natural transformation.

The link between circadian control and natural transformation was initially suggested by data from experiments that were performed decades ago (Golden and Sherman, 1984). Natural

transformation was found to happen more efficiently in cells that underwent an overnight dark incubation followed by an incubation of several hours in the light. With new data in hand on RpaA, it is now possible to say that this treatment was synchronizing the circadian rhythm of cells that were grown in constant light. A recent study performed in *S.* 7942 screened a randomly barcoded transposon library for mutants with natural transformation fitness biases (Taton et al., 2020). This study identified *rpaA* as essential for natural transformation. Although the study identified the genes that are needed for natural transformation to occur, the manner through which RpaA controls this process is unclear.

The unique form of the *S.* 2973 *rpaA* allele, which behaves differently than null *rpaA*, provides an opportunity to better understand the mechanism by which natural transformation is linked to circadian cycling (Markson et al., 2013b). While the null mutant is locked in a constant dawn-like transcriptional state, the *S.* 2973 allele maintains cycling, albeit in a dampened manner compared to *S.* 7942. Identifying the specific *rpaA* targets that are enabling temporal natural competence will be a new research avenue in this field of study. Furthermore, experiments to understand the prevalence of circadian regulation of natural transformation across cyanobacterial species will be key in better understanding this process from an evolutionary perspective.

In conclusion, this study produced a fast-growing strain that is capable of undergoing natural transformation and will be an asset to the cyanobacteria research community. However, perhaps the more important finding of this work is in the greater context of understanding cyanobacterial physiology; the *S.* 7942/ *S.* 2973 natural transformation dichotomy has provided key insights into cyanobacterial natural transformation and will enable future studies to better understand this process.

A.2.5 Materials and methods

Generating mutant *S.* 2973 lines

S. 2973 mutant lines were generated using CRISPR/Cpf1 system (Ungerer and Pakrasi, 2016b). Individual SNPs were substituted with the sequence from the sister strain without the integration of a selective marker. The repair template was designed so that guide RNA target sequences were mutated to result in silent protein mutations post homologous recombination. The plasmid pSL2680 was used as the backbone for all of the editing plasmids that were constructed. The 3975R, 3975ML, 3975L and 3975MR primers were used to amplify the *pilN* repair template from *Synechococcus* 7942 wild-type genomic DNA and the 3975gRNAL and 3975gRNAR oligos were inserted into the construct to serve as the gRNA. The *pilMNO* promoter repair template was generated with the 03965L_V2 and 03965R_V2 primers and the 03965gRNAR and 03965gRNAL oligos were used for the gRNA on this construct. The *rpaA* promoter repair template was generated with the 131058R, 131058ML, 131058L, and 131058MR primers and the gRNA was made from the 131058gRNAL and 131058gRNAR oligos. The *rpaA* coding region repair template was generated with the 2R, 2605ML, 2605L, and 2605MR primers and the 2605gRNAL and 2605gRNAR oligos served as the gRNA. Sequences for oligonucleotides are displayed in Table SA2.

Natural transformation

Natural transformation was performed in the same way for all of the data presented in this manuscript. Cells were initially streaked onto BG11 medium and grown overnight at 38°C and ~40 $\mu mol \cdot m^{-2} \cdot s^{-1}$ light with ambient CO_2 in a Percival tissue culture chamber. The plate was then transferred to a Caron plant growth chamber where it was grown at 38°C and 75 $\mu mol \cdot m^{-2} \cdot s^{-1}$ light with 0.6% CO_2. Cells were then resuspended in BG11 medium and transferred to 125 mL

Erlenmeyer flasks to be grown overnight under the same conditions in the Caron chamber. Subsequently, cells were inoculated back to a particular OD_{730} (based on growth rate) and grown overnight in liquid culture under the same conditions to allow strains with different growth rates to be harvested at roughly the same OD_{730}. Cells were harvested at OD_{730} = 0.32-0.37 on the Biotek µQuant micro plate reader. Cells were pelleted for each experimental sample at a low speed (2000 x g) to avoid breaking off transformation pili, in a fixed angle rotor of a Sorvall benchtop centrifuge. Cells were washed twice and then resuspended in sterile BG11 and transferred to microcentrifuge tubes where donor DNA was added to each experimental sample. For most experiments, 125 ng of donor DNA was added to each sample; however, for the truncated constructs used for some of the experiments in Table 1, the amount of DNA was adjusted to ensure the molar ratio of donor DNA to cells was maintained. Additionally, negative control samples were run for each line to calculate cell viability and overall transformation efficiency.

After samples received donor DNA, tubes were kept in the dark at room temperature overnight, and then transferred to a 38°C chamber under 40 $\mu mol \cdot m^{-2} \cdot s^{-1}$ light. Cells remained in the light for four hours prior to plating on selective medium (or non-selective in the case of the viability test samples). For pSL2231 samples, BG11 with 10 µg/mL chloramphenicol was used for selection and for pSL3192, BG11 with 10 µg/mL kanamycin for selection. For the viability test, the cells were diluted 1E-3 prior to plating on BG11 in order to ensure that the colonies were countable. Plates were incubated at 30°C until colonies formed. Natural transformation efficiency was calculated by dividing the number of transformants in a sample by the estimated number of viable cells in the sample calculated from the viability assay data. This ensured that

cells that doubled during the transformation or cells that died during the transformation procedure were taken into account.

A representative subset of natural transformants were confirmed via PCR analysis for each experimental iteration. The KW001 and KW005 primers were used to confirm natural transformants with pSL2231 donor DNA and Spec_seq_F and NS1_Chr_Seq_R_V2 primers were used to confirm natural transformants with pSL3192 donor DNA. Additionally, a subset of transformants were phenotypically analyzed as later described.

Preparation of donor DNA for wild-type *Synechococcus* 7942 natural transformation

Donor DNA was generated by miniprep isolation for intact plasmids, miniprep isolation followed by digestion for methylated linearized fragments, and PCR amplification for unmethylated linearized fragments. For Fragment 2 (Fig. A2) the restriction enzyme NdeI was used for linearization. Fragment 3 was linearized with the NdeI and BamHI restriction enzymes, and Fragment 4 was linearized with BamHI and HindIII restriction enzymes. The JY11_whole_BB_F and JY11_whole_BB_R PCR primers were used for PCR linearization of Fragment 2, JY11_whole_BB_F and JY11_DS_HA_R primers for Fragment 3, and JY11_US_HA_F and JY11_DS_HA_R primers for Fragment 4. After linearization, samples were column purified with a Thermo Fisher GeneJet PCR purification kit for both the PCR amplified and restriction enzyme treated fragments. A NanoDrop Spectrophotometer was then used to quantify the isolated DNA prior to using it for natural transformation.

Preparation of pSL2193 donor DNA for transformation assays

The pSL2193 isolated from *E. coli* via miniprep prior to transformation. A NanoDrop Spectrophotometer was then used to quantify the isolated DNA prior to using it for natural transformation.

Generating growth curves and calculating doubling time

Growth curves were generated using a PSI multicultivator (Czech Republic). Cells were grown on a plate in a Caron plant growth chamber at 38°C and 75 µmol · m^{-2} · s^{-1} with 0.6% CO_2. Cells were then inoculated into liquid and grown overnight in the Caron plant growth chamber under the same conditions. The culture was diluted back with BG11 and cultured overnight in the PSI multicultivator under either 400 or 900 µmol·m^{-2}·s^{-1} light with 1% CO_2 prior to generating the growth curve to allow for acclimation. Once cells reached exponential phase, they were diluted back again, and grown to linear phase to generate a growth curve. The growth curve was then used to calculate doubling time as described in (Ungerer et al., 2018d).

NblA phenotype assays

The *nblA* deletion phenotype was first confirmed by growing transformants to linear phase in BG11 medium. Cells were then pelleted and washed in BG11-Nitrogen two times. After pelleting a third time, cells were grown in BG11-Nitrogen medium for four days in twelve well plates at 38°C. Cells were then analyzed for bleaching visually and quantitatively by measuring the absorbance spectrum on an Olis DW2000 spectrometer.

YFP phenotype assays

Integration of the Ptrc::eYFP cassette was confirmed phenotypically using a BioTek fluorescence plate reader. Cells were grown to linear phase in BG11 medium and quantitatively analyzed for YFP fluorescence. For the fluorescence measurements, an excitation wavelength of 488 and emission wavelength of 528 were used. Relative fluorescence levels were then normalized by OD_{730} for each sample.

RNA Isolation and cDNA synthesis preparation for quantitative PCR

Cells were subjected to the same treatment used in experimental natural transformation prior to RNA isolation. Cells were spun down and total RNA was isolated using RNAwiz reagent (Ambion) according to manufacturer's instructions and treated with RNase-free DNase I to eliminate contaminating chromosomal DNA. RNA quality and quantity were evaluated by using electrophoresis in 2% agarose gel and a NanoDrop ND-1000 instrument (Nanodrop Technologies). cDNA was synthesized using oligo-dT primers containing a 3'-adapter and 14-16 cycles of PCR amplification.

Quantitative real-time PCR

RT-qPCR was performed in a CFX96 Real-time PCR system (Bio-Rad), according to the manufacturer's instructions. Reactions were performed in a 20 µl volume. Reaction mixtures contained 10 µl PowerTrack SYBR Green Master Mix (Thermo Fisher), 200 nM of each primer, and 100 ng cDNA. The following primer sets were used for each transcript: Synpcc7942_1615_F and Synpcc7942_1615_R were used for *rnaP,* Synpcc7942_2113_F and Synpcc7942_2113_R were used for *prs,* Synpcc7942_2453_F and Synpcc7942_2453_R were used for *pilM,* Synpcc7942_2452_F and Synpcc7942_2452_R were used for *pilN,* Synpcc7942_2451_F and Synpcc7942_2451_R were used for *pilO,* and Synpcc7942_0095_F and Synpcc7942_0095_R were used for *rpaA.* Reactions were loaded into 96-well PCR plates. No-template reactions were used as negative controls. The RT-qPCR temperature profile included an initial denaturation step at 95 °C for 3 minutes, followed by 45 cycles of 15 s at 95 °C, 15 seconds at 55 °C, and 15 seconds at 72 °C. A melting curve was performed at the end of the RT-qPCR run by stepwise (+1 °C/10 seconds) ramping from 65 °C to 95 °C. A minimum of three replicates were performed for each of the cDNA samples.

Quantification of gene expression

For the determination of Cq (quantification cycle) values, the same baseline threshold was set for all experiments in the single threshold mode of the CFX Manager Software (Bio-Rad). ΔCt was calculated by subtracting the Ct of the endogenous reference genes (*rnpA* and *prs*) from the Ct of the target genes (*pilNMO* and *rpaA*). Two-way ANOVA with sidak's multiple comparison test was performed to compare calculated ΔCt values with significance reported for $p < 0.05$. The relative change in gene expression of the target genes was analyzed using the $2^{-\Delta\Delta CT}$ method. To calculate ΔΔCt the WT *S.* 2973 ΔCt was subtracted from the ΔCt of *S.* 2973-T. The fold change of the target gene of *S.* 2973-T vs. the control WT *S.* 2973 was calculated according to the formula: $2^{-\Delta\Delta Ct}$ (Livak and Schmittgen, 2001).The results of the relative expression and log2 fold change were graphed using Graph Pad Prism (GraphPad, La Jolla, CA, USA).

Whole genome sequencing

Whole genome sequencing was performed by the Microbial Genome Sequencing Center (MiGS Center). Cells were grown overnight in a Caron plant growth chamber and genomic DNA was isolated with a Qiagen DNeasy Blood and Tissue Kit. The resulting isolated DNA was then concentrated in a SpeedVac prior to sending it for sequencing. Illumina technology was used to perform the sequencing run. Both sequenced genomes yielded around 2,300,000 reads with average read lengths of 123 bases, giving complete 104X coverage of both of the 2.7 Mb genomes. Using the BreSeq software package to assess coverage, data was fitted to a Poisson curve and yielded a fit mean value of 191.6 and fit dispersion value of 11.4 to the *S.* 2973 main chromosome (Deatherage and Barrick, 2014). The fit mean value for the CP006472 plasmid was 64.5 and fit dispersion value was 5.6. The BreSeq software package was used for variant calling (Deatherage and Barrick, 2014).

A2.6 Supporting data

TABLE SA1: Shared mutations in the *Synechococcus* 2973 pilN7942 rpaA7942 strain and *Synechococcus* 2973-T strain that differ from the native *Synechococcus elongatus* 2973 sequence.

Gene Tag(s)	**Gene Function**	**Nucleotide Change**	**Change(s) in Protein Sequence**
M744_06570	hydrolase	C →G	R242G
M744_068675	hypothetical protein	+A	Frameshift
M744_9300	phosphate acyltransferase	C→T	D14N
M744_10890	lipoprotein	+T	Frameshift

TABLE SA2: Oligonucleotides used in this study. Capital letters represent homology with the sequence undergoing amplification and lowercase letters represent newly introduced sequences that do not have genomic homology.

Name	Sequence
3975R	cgctgcccggattacagatcctctagagtcgacGGTACCGGACTGCAGCTGCAAG CCAG
3975ML	CACGTCCTTAATaCAGGTCTTCCCGCAGGTCAAATC
3975L	ggtcatttttttgtctagctttaatgcggtagttGGTACCCCAGTACGGATTGCAAGA AGC
3975MR	GATTTGACCTGCGGGAAGACCTGtATTAAGGACGTG
3975gRNAL	agatAGGTCTTCCCGCAGGTCAAA
3975gRNAR	agacTTTGACCTGCGGGAAGACCT
03965L_V2	tttgtctagctttaatgcggtagttGGTACGCTCCACGGAAACAAAATCCTGAA GC
03965R_V2	ggattacagatcctctagagtcgaCGGTACTCATGCGGGACTGGAAGTGCGA
03965gRNAR	agatAAATTGTTGCCGTAGTTCTT
03965gRNAL	agaCAAGAACTACGGCAACAATTT
131058R	cgctgcccggattacagatcctctagagtcgacGGTACCgcgcagatgacggatgtgg
131058ML	ctcaccagaacagagccACTATCATttttaagaaaacataaatagatccgctg
131058L	ggtcatttttttgtctagctttaatgcggtagttGGTACCgttgtcagcaaggattccgagg
131058MR	cagcggatctatttatgttttcttaaaaATGATAGTggctctgttctggtgag
131058gRNAL	agatatccgctgtaagccgctgcc
131058gRNAR	agacggcagcggcttacagcggat
2605R	cgctgcccggattacagatcctctagagtcgacGGTACC GAGTCCTGAGCTGCTACTGCC
2605ML	GCTGCAGCGCACCGATCGCATaCCCCATGCAGCCCGCCATAGC
2605L	ggtcatttttttgtctagctttaatgcggtagttGGTACC GCCCTGCGTCTCAATTAATCTCC
2605MR	GCTATGGCGGGCTGCATGGGGtATGCGATCGGTGCGCTGCAGC
2605gRNAL	agat CCCATGCAGCCCGCCATAGC
2605gRNAR	agac GCTATGGCGGGCTGCATGGG
Spec_seq_F	ACACCGTGGAAACGGATGAA
NS1_Chr_Seq_R_V2	TTGCGCTGCTGATTGTTCTAGGC
KW001	GTCGGCCCCAGTCCATCG
KW005	CCTGTACAAAGGCCGCAATCGA
JY11_whole_BB_F	CCGCACAGATGCGTAAGGAGAAAATAC
JY11_whole_BB_R	ACTCTCAGTACAATCTGCTCTGATGCC
JY11_DS_HA_R	ATGCTGCTGGAGTTCTACGCC
JY11_US_HA_F	GCGCAGATAGCCTGACTGTTCC
Synpcc7942_1615_F	CTCGTTGCCATCGACTGCG
Synpcc7942_1615_R	CGACTTTCGTTGACGGACTCTATTT
Synpcc7942_2113_F	TCTTGCCCTACTACGGTTACGC
Synpcc7942_2113_R	TCGCTCCAGCCTGAGTGATTAG

A.2.7 References

Beauchamp, J. M., Leveque, R. M., Dawid, S., & DiRita, V. J. (2017). Methylation-dependent DNA discrimination in natural transformation of Campylobacter jejuni. *Proceedings of the National Academy of Sciences, 114*(38), E8053-E8061.

Bhaya, D., Bianco, N. R., Bryant, D., & Grossman, A. (2000). Type IV pilus biogenesis and motility in the cyanobacterium Synechocystis sp. PCC6803. *Molecular microbiology, 37*(4), 941-951.

Bolivar, F., Rodriguez, R. L., Greene, P. J., Betlach, M. C., Heyneker, H. L., Boyer, H. W., . . . Falkow, S. (1977). Construction and characterization of new cloning vehicle. II. A multipurpose cloning system. *Gene, 2*(2), 95-113.

Carbonnelle, E., Helaine, S., Nassif, X., & Pelicic, V. (2006). A systematic genetic analysis in Neisseria meningitidis defines the Pil proteins required for assembly, functionality, stabilization and export of type IV pili. *Molecular microbiology, 61*(6), 1510-1522.

Chang, Y.-W., Rettberg, L. A., Treuner-Lange, A., Iwasa, J., Søgaard-Andersen, L., & Jensen, G. J. (2016). Architecture of the type IVa pilus machine. *Science, 351*(6278), aad2001.

Chen, I., & Dubnau, D. (2004). DNA uptake during bacterial transformation. *Nature Reviews Microbiology, 2*(3), 241-249.

Collier, J. L., & Grossman, A. (1994). A small polypeptide triggers complete degradation of light-harvesting phycobiliproteins in nutrient-deprived cyanobacteria. *The EMBO journal, 13*(5), 1039-1047.

Deatherage, D. E., & Barrick, J. E. (2014). Identification of mutations in laboratory-evolved microbes from next-generation sequencing data using breseq. In *Engineering and analyzing multicellular systems* (pp. 165-188): Springer.

Golden, S. S., Brusslan, J., & Haselkorn, R. (1986). Expression of a family of psbA genes encoding a photosystem II polypeptide in the cyanobacterium Anacystis nidulans R2. *The EMBO journal, 5*(11), 2789.

Golden, S. S., & Sherman, L. A. (1984). Optimal conditions for genetic transformation of the cyanobacterium Anacystis nidulans R2. *Journal of bacteriology, 158*(1), 36-42.

Hamilton, H. L., & Dillard, J. P. (2006). Natural transformation of Neisseria gonorrhoeae: from DNA donation to homologous recombination. *Molecular microbiology, 59*(2), 376-385.

Jaiswal, D., Sengupta, A., Sohoni, S., Sengupta, S., Phadnavis, A. G., Pakrasi, H. B., & Wangikar, P. P. (2018). Genome Features and Biochemical Characteristics of a Robust, Fast Growing and Naturally Transformable Cyanobacterium Synechococcus elongatus PCC 11801 Isolated from India. *Scientific Reports, 8*(1), 16632.

Kufryk, G. I., Sachet, M., Schmetterer, G., & Vermaas, W. F. (2002). Transformation of the cyanobacterium Synechocystis sp. PCC 6803 as a tool for genetic mapping: optimization of efficiency. *FEMS Microbiology letters, 206*(2), 215-219.

Li, S., Sun, T., Xu, C., Chen, L., & Zhang, W. (2018). Development and optimization of genetic toolboxes for a fast-growing cyanobacterium Synechococcus elongatus UTEX 2973. *Metabolic engineering*.

Livak, K. J., & Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2− ΔΔCT method. *methods*, *25*(4), 402-408.

Markson, J. S., Piechura, J. R., Puszynska, A. M., & O'Shea, E. K. (2013). Circadian control of global gene expression by the cyanobacterial master regulator RpaA. *Cell*, *155*(6), 1396-1408.

Nagarajan, A., Winter, R., Eaton-Rye, J., & Burnap, R. (2011). A synthetic DNA and fusion PCR approach to the ectopic expression of high levels of the D1 protein of photosystem II in Synechocystis sp. PCC 6803. *Journal of Photochemistry and Photobiology B: Biology*, *104*(1), 212-219.

Nies, F., Mielke, M., Pochert, J., & Lamparter, T. (2020). Natural transformation of the filamentous cyanobacterium Phormidium lacuna. *PloS one*, *15*(6), e0234440.

Onai, K., Morishita, M., Kaneko, T., Tabata, S., & Ishiura, M. (2004). Natural transformation of the thermophilic cyanobacterium Thermosynechococcus elongatus BP-1: a simple and efficient method for gene transfer. *Molecular Genetics and Genomics*, *271*(1), 50-59.

Pakrasi, H. B., & Wendt, K. E. (2019). Genomics approaches to deciphering natural transformation in cyanobacteria. *Frontiers in microbiology*, *10*, 1259.

Piepenbrink, K. H., Maldarelli, G. A., De La Peña, C. F. M., Dingle, T. C., Mulvey, G. L., Lee, A., . . . Sundberg, E. J. (2015). Structural and evolutionary analyses show unique stabilization strategies in the type IV pili of Clostridium difficile. *Structure*, *23*(2), 385-396.

Porter, R. D. (1986). Transformation in cyanobacteria. *CRC Critical reviews in microbiology*, *13*(2), 111-132.

Racharaks, R., Arnold, W., & Peccia, J. (2021). Development of CRISPR-Cas9 knock-in tools for free fatty acid production using the fast-growing cyanobacterial strain Synechococcus elongatus UTEX 2973. *Journal of Microbiological Methods*, 106315.

Springstein, B. L., Nies, F., & Dagan, T. (2020). Natural competence in Chlorogloeopsis fritschii PCC 6912 and other ramified cyanobacteria. *bioRxiv*.

Taton, A., Erikson, C., Yang, Y., Rubin, B. E., Rifkin, S. A., Golden, J. W., & Golden, S. S. (2020). The circadian clock and darkness control natural competence in cyanobacteria. *Nature Communications*, *11*(1), 1-11.

Trehan, K., & Sinah, U. (1981). Genetic transfer in a nitrogen-fixing filamentous cyanobacterium. *Microbiology*, *124*(2), 349-352.

Ungerer, J., Lin, P.-C., Chen, H.-Y., & Pakrasi, H. B. (2018). Adjustments to photosystem stoichiometry and electron transfer proteins are key to the remarkably fast growth of the cyanobacterium Synechococcus elongatus UTEX 2973. *MBio*, *9*(1), e02327-02317.

Ungerer, J., & Pakrasi, H. B. (2016). Cpf1 is a versatile tool for CRISPR genome editing across diverse species of cyanobacteria. *Scientific Reports*, *6*, 39681.

Ungerer, J., Wendt, K. E., Hendry, J. I., Maranas, C. D., & Pakrasi, H. B. (2018). Comparative genomics reveals the molecular determinants of rapid growth of the cyanobacterium Synechococcus elongatus UTEX 2973. *Proceedings of the National Academy of Sciences, 115*(50), E11761-E11770.

Yu, J., Liberton, M., Cliften, P. F., Head, R. D., Jacobs, J. M., Smith, R. D., . . . Pakrasi, H. B. (2015). Synechococcus elongatus UTEX 2973, a fast growing cyanobacterial chassis for biosynthesis using light and CO2. *Scientific Reports, 5*.