UNIVERSITIY OF CALIFORNIA SAN DIEGO

Broad-Host-Range Genetic Tools for Cyanobacterial Engineering.

A thesis submitted in partial satisfaction of the requirements

for the degree of Master of Science

in

Biology

by

Kevin Phan Trieu

Committee in charge:

Professor James W. Golden, Chair
Professor Eric E. Allen
Professor Katherine L. Petrie

2020

The Thesis of Kevin Phan Trieu is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

Chair

University of California San Diego

2020

DEDICATIONS

I would like to dedicate this thesis to my family, for providing their unconditional love and support, encouraging me to take on challenges to pursue my goals, and understanding the decisions I make. Thank you for always having faith in me. You are all the most influential and inspiring people in my life.

I would also like to dedicate this thesis to my closest friends, who have continued to believe in me and supported me through rough times during college. Your unwavering support and encouragement are what helps build my confidence and motivation to continue. Thank you.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

Ap	Ampicillin
Cm	Chloramphenicol
Km	Kanamycin
LB	Lennox Broth
Nm	Neomycin
P	Phosphorus
Phi	Phosphite (PO_3^{2-})
Pi	Phosphate (PO_4^{3-})
Sm	Streptomycin
Sp	Spectinomycin
Tc	Tetracycline

ACKNOWLEDGMENTS

I would like to acknowledge my mentors, Bryan Bishé and Arnaud Taton, and my principal investigator and committee chair, Professor James Golden, for their support and guidance. Thank you Bryan for being one of the best mentors one could hope for. I am grateful for all the work you have done to help me get started, progress, and achieve in the Golden Laboratory. I admire your patience, dedication, and understanding as a well-rounded scientist. Thank you Arnaud for your all of your amazing insights and unwavering support. I really appreciate all the hours of discussion we had and all the work you have put into organizing data and creating figures. Thank you Jim for being one of the most understandable and knowledgeable professors I worked with. Your expertise and leadership have helped guide me through the ups and downs of my research. Thank you for the opportunity to work in your laboratory.

I would also like to thank all members of the Golden Laboratory, both current and past, for their support and understanding throughout my undergraduate and graduate studies. It has been a privilege working with such a welcoming and hard-working group of researchers.

Finally, I would like to thank my committee members, Professor Eric E. Allen and Professor Katherine L. Petrie, for taking time out of their busy schedules to be on my committee. I really appreciate your encouragement during my graduate school experience.

ABSTRACT OF THE THESIS

Broad-Host-Range Genetic Tools for Cyanobacterial Engineering

by

Kevin Phan Trieu

Master of Science in Biology

University of California San Diego, 2020

Professor James W. Golden, Chair

In recent decades, cyanobacteria have garnered interest as promising biological platforms for producing valuable biofuel and feedstock molecules. Genetic tools are instrumental in conducting genetic modifications in cyanobacteria, but tools are generally limited to a few domesticated laboratory strains. Here, we present new genetic tools and techniques using broad-host range, RSF1010-based plasmids to facilitate genetic engineering across diverse cyanobacterial strains. First, we developed a set of orthogonal constitutive promoters constructed in RSF1010-based plasmids. The promoter library generated a wide range of gene expression levels in five diverse

cyanobacterial strains and allowed for strict transcriptional control. Next, we used an RSF1010-based plasmid to conduct biparental conjugal mating experiments with wild cyanobacteria. Serving as a bioprospecting technique, we isolated new, genetically tractable strains with high salt and alkalinity tolerance from a soda lake. Finally, we constructed a plasmid harboring the *ptxD* gene derived from *Pseudomonas stutzeri*, which codes for a phosphite oxidoreductase. We engineered three cyanobacterial strains to express the enzyme, which allows them to convert phosphite to phosphate and grow in media containing phosphite as the sole source of phosphorus. We propose that this plasmid can be used to facilitate research on engineering strains to metabolize phosphite as a strategy for reducing biological contaminants in outdoor cultivation ponds. Overall, these RSF1010-based genetic tools are useful additions to the biotechnological toolkit to enable genetic engineering across diverse cyanobacterial strains.

INTRODUCTION

Cyanobacteria form a diverse phylum of Gram-negative, photoautotrophic bacteria. They perform photosynthesis by harnessing energy from sunlight using light-harvesting chlorophyll a and phycobilin pigments (Garcia-Pichel, 2009). Like green plants, cyanobacteria obtain carbon sources through fixing CO_2 from the atmosphere and releasing oxygen as byproduct. Cyanobacteria have existed for as long as 3.5 billion years and contributed to ancient global oxygenation of Earth's atmosphere (Whitton & Potts, 2012). Research have described cyanobacteria as evolutionary precursors to modern plastids in eukaryotic algae and chloroplasts in plants (Garcia-Pichel, 2009). In current times, cyanobacteria contribute to 25% of the total primary productivity on the planet and are found in diverse ecosystems, including freshwater lakes, oceans, soils, and even deserts (Flombaum *et al.*, 2013; Garcia-Pichel, 2009).

In recent decades, cyanobacteria have garnered interest in being developed as platforms for biofuel and bioplastics production. Certain strains have native metabolic pathways for synthesizing biofuel metabolites that can be exploited for industrial scale production. For example, members of *Cyanothece* and *Nostoc* spp. can convert fatty acid intermediates to alkanes and alkenes (Whitton & Potts, 2012). Moreover, many strains are genetically tractable and can be engineered to express heterologous biosynthesis pathways. Since cyanobacteria perform oxygenic photosynthesis, they serve as renewable and sustainable sources of natural products.

Compared to chemoorganotrophic and eukaryotic organisms, cyanobacteria have several advantages making them more suitable as cellular factories for bio-based products. Unlike cyanobacteria and other photoautotrophs, chemoorganotrophs require organic compounds as carbon and energy sources. Model chemotrophs such as *Escherichia coli*, *Saccharomyces cerevisiae, and Clostridium* would need continuous input of organic compounds in the media to sustain growth on large industrial settings

(Somerville *et al.*, 2010). Food crops would act as feedstock, which could increase the use of water and farmland and reduce overall sustainability efforts (Somerville *et al.*, 2010). In contrast, cyanobacteria grown in large outdoor ponds do not require carbon feedstocks and utilize CO2 for fixation and sunlight for energy. Both resources are unlimited and allow for a sustainable system for production (Gordon & Pfleger, 2018). Moreover, cyanobacteria are more favorable than their eukaryotic counterparts in being developed as cell factories for biofuel and bioplastics production. As prokaryotic organisms, cyanobacteria are easier to genetically manipulate and have faster reproductive cycles, making them better suited organisms to study in laboratory settings (Gordon & Pfleger, 2018). Cyanobacteria require less land for cultivation and potentially can be grown in highly saline, less arable environments (Dismukes, *et al.*, 2008; Gordon & Pfleger, 2018).

Genetic engineering of cyanobacteria is a growing field in biotechnology that is central to developing the organisms as sources of precursor fuel molecules. Liu *et al.* (2011) modulated and exploited the fatty acid synthesis pathway in cyanobacterial laboratory strain *Synechocystis sp.* to produce and secrete fatty acids. Dexter and Fu (2009) inserted ethanol-producing genes from *Zymomonas mobilis* into the *Synechocystis sp.* genome; the genetic alterations yielded bioethanol. To carry out these modifications, researchers utilize genetic tools involving vector plasmids and synthetic regulatory parts such as promoters and transcriptional terminators that regulate gene expression (Wang *et al.* 2012). Because many strains of cyanobacteria take up foreign DNA through natural transformation or conjugation, plasmids can be transferred into cells and used to integrate genes coding for biofuel production pathways into the host chromosome. Alternatively, plasmids that self-replicate can maintain genes of interest in cell populations without genomic integration.

The RSF1010 plasmid is a promising vector with broad host range, allowing for genetic modifications to be done with diverse bacterial strains including cyanobacteria. RSF1010 is a member of the incompatibility group IncQ plasmids and was first identified in *Escherichia coli* by Guerry *et al.* (1974). However, similar IncQ plasmids continued to be found across different species of bacteria, and interest began in developing the plasmid as a vector for a variety of hosts (Meyer, 2009; Nagahari and Sakaguchi, 1978). Like other IncQ plasmids, RSF1010 is not self-transmissible by conjugation. It has been suggested that these plasmids cannot stably house lengthy coding sequences required for conjugation machinery (Rawlings & Tietze, 2001). Instead, RSF1010 propagates by taking advantage of conjugation machinery encoded by cellular conjugal plasmids (Meyer, 2009).

In laboratory settings, conjugation can be performed using *Escherichia coli* (*E. coli*) to transfer RSF1010 into recipient cyanobacteria cells. The donor *E. coli* strain contains a conjugal plasmid, such as pRL443. A variant of the IncP-type RP4 plasmids, pRL443 allows self-mobilization and transfer using the type IV-secretion mechanism (Elhai *et al.*, 1997). Additionally, donor *E. coli* strains can house the optional helper plasmid pRL623, which include genes for methyltransferases. The enzymes protect RSF1010 from restriction activity of *AvaI*, *AvaII*, and *AvaIII* native in the cyanobacterial strain *Anabaena* sp. PCC 7120 (Elhai *et al.*, 1997).

RSF1010 can self-mobilize despite being fairly small (roughly 8.6 kilobase pairs). It harbors mobilization (*mob*) and replication (*rep*) genes, as well as an origin of transfer (*oriT*) (Rawlings & Tietze, 2001). The *oriT* (also referred to as basis of mobility, or *bom*) is an intergenic region of RSF1010 crucial for initiation of DNA replication and plasmid mobilization to the recipient cell (Lanka & Wilkins, 1995). The MobA protein targets the *oriT* site and nicks RSF1010 as a relaxase (Rawlings & Tietze, 2001). MobA covalently links to the 5’ end of the cleaved strand and also functions as a primase, starting DNA

replication in preparation for plasmid transfer (Monzingo *et al.*, 2007; Rawlings & Tietze, 2001). Although essential for self-mobilization, MobA relaxase activity consequently contributes to low yield during RSF1010 plasmid preparations. Nicked, relaxed RSF1010 DNA can denature into single strands during heated or alkaline lysis steps (Taton *et al.*, 2014). Other nicked plasmids will have MobA complexes covalently attached to the strand. Both situations render much of the plasmid yield from standard preparation protocols unusable for following cloning experiments (Taton *et al.*, 2014).

To address this deficiency, a point mutation was made in the *mobA* gene, *mobA*Y25F, to prevent nicking of DNA and increase plasmid yields for further applications (Taton *et al.,* 2014). Furthermore, cloning efficiency was improved by making the RSF1010-based plasmid compatible with the CYANO-VECTOR cloning system for plasmid assembly. Modular devices and functional parts such as selection markers, *E. coli* origins of replication, and reporter genes can be constructed with the RSF1010-based vector using seamless assembly. While cloning efficiency was improved, conjugation efficiency in cyanobacteria reduced by as much as 4 orders of magnitude (Taton *et al.*, 2014).

Bishé *et al.* addressed the concern by performing further modifications to partially restore the conjugal efficiency of the RSF1010-based plasmid while maintaining plasmid yields. The conjugal donor *E. coli* strain contains the conjugal plasmid pRL443, which is a variant of RP4. The closely related RK2 plasmid houses a *bom* site that is recognized by the RK2/RP4 relaxase TraJ (Babic *et al.*, 2008). The RK2-based bom site was added as a second origin of transfer into the RSF1010 *mobAY25F* variant, since RP4-encoded TraJ can recognize the site, nick the plasmid DNA, and increase mobility (Bishé *et al.*, 2019). As a result, conjugation efficiency increased by two orders of magnitude in cyanobacteria laboratory strains (Bishé *et al.*, 2019).

The RSF1010-based vector system serves as a molecular toolkit for a variety of potential applications in cyanobacterial biotechnology. Our investigations aim to improve and apply the available set of tools in the context of research and industrial settings. First, we build on the existing RSF1010 toolkit by developing a promoter library for controlling expression of heterologous genes in diverse strains. Second, we expand the utility of the toolkit by bioprospecting for genetically tractable, halotolerant cyanobacterial strains. Finally, we demonstrate an application of the toolkit by expanding a system for controlling culture contamination in large-scale cyanobacteria ponds.

SECTION I: A PROMOTER LIBRARY FOR CONTROLLING GENE EXPRESSION IN CYANOBACTERIA

1.1 Introduction

Genetic tools are fundamental components of biotechnology research used to manipulate gene expression and metabolic pathways in cyanobacteria. Most available tools focus on studying internal metabolic and physiological features in certain species, especially laboratory strains (Gordon & Pfleger, 2018). Other tools have been developed for engineering cyanobacteria to produce biofuel and other important biomolecules. However, yields for bioproducts, such as chemical feedstocks, produced by cyanobacteria remain lower than those of model heterotrophs like *E. coli* (Nozzi *et al.*, 2013). Tools may also not be compatible across the diverse variety of cyanobacterial strains, especially in wild non-model strains (Gordon & Pfleger, 2018). To improve tools for optimizing production of desired chemicals, scientists have suggested strategies that include minimizing competing native metabolic pathways, improving flux into the targeted pathway, and increasing cell tolerance to high concentrations of the chemicals (Nozzi *et al.*, 2013). Overall, improvements to cyanobacteria genetic toolkits will depend on developing methods for controlling gene expression.

The manipulation of foreign protein expression is crucial for optimizing production of desired products and limiting production of byproducts that may be harmful to the cyanobacterial host. One method for regulating expression is through transcriptional control using specific promoters. Scientists have previously used endogenous promoters, or those that are native to species of cyanobacteria, for expressing heterologous genes (Gordon & Pfleger, 2018). Using RNA-seq data, researchers have discovered native promoters from *Synechococcus* PCC 7002 with potential use as

constitutive and light-sensitive promoter. However, yellow fluorescent protein measurements reveal limited up- or down-regulation of expression, with most promoters studied generating fluorescence around the same level as that of the original promoters (Ruffing *et al.*, 2016). Cyanobacterial native promoters are affected by internal regulatory processes that unintendedly influence heterologous gene expression (Gordon & Pfleger, 2018).

Consequently, scientists have turned to testing heterologous promoters, notably those from *Escherichia coli,* for applications in cyanobacteria biotechnology. For example, the Ptrc promoter has been used for inducible gene expression. However, when tested in *Synechocystis* sp. strain PCC 6803 (hereafter *Synechocystis*), some basal level of expression occurs even in the absence of the inducer, indicating a leaky expression system (Huang *et al.*, 2010). Other *E. coli* promoters include Plac and Ptet; however, both exhibit low detectable activity in the same cyanobacterial strain (Huang *et al.*, 2010). Therefore, it is imperative to look at other strategies, such as promoter libraries, to produce a wider range of expression levels.

A promoter library is a collection of mutated promoter variants generated from a single original promoter, with each member of the library exhibiting a different promoter strength. Libraries are formed by making random point mutations at nucleotide positions along the promoter sequence that significantly influence the strength of the promoter (Jensen & Hammer, 1998). Libraries have been synthesized and characterized in several bacterial model organisms, such as *Escherichia coli* and *Bacillus subtilis* (De Mey *et al.*, 2007; Liu *et al.*, 2018). Moreover, promoter libraries have previously been investigated in cyanobacteria. Markley *et al.* (2015) mutagenized a truncated sequence of the high-performing cpcB promoter from *Synechocystis*, generating a library for testing in *Synechococcus* PCC 7002 (hereafter, PCC 7002). The library of 11 promoters covers

three orders of magnitude in YFP expression levels in PCC 7002. However, the results of this study are only limited to one cyanobacterial strain, making the promoter library unpredictable in other lab strains. The researchers also characterized a set of *E. coli* promoters from the BioBrick promoter family BBa_J23100 to BBa_J23119 (from the iGEM parts collection) in PCC 7002. Analysis revealed a weak correlation between the fluorescence measurements in *E. coli* and those in PCC 7002 (Markley *et al.*, 2015). Variation of expression levels in cyanobacteria as opposed to *E. coli* has been attributed to differences in the promoter sequence, with consensus sequences different between the two organisms (Gordon & Pfleger, 2018). These results suggest that *E. coli* promoter libraries do not necessarily function similarly and must be further developed and studied in cyanobacteria.

To address the need for developing transcriptional control, we sought to develop a new set of tools using the promoter conII. PconII is a derivative of PconI, a synthetic promoter combining -35 and -10 elements of the native trp and lac promoters, respectively (Elledge & Davis, 1989). PconI and PconII both lack repressive operator elements and are constitutively expressive. Both promoters have been optimized using a consensus sequence ideal for expression in *E. coli* (Elledge & Davis, 1989). However, PconII has been used for constitutive expression of foreign protein in cyanobacteria and is a component of the CYANO-VECTOR cloning system (Taton et al., 2014).

The -10 region within a bacterial promoter sequence is the site of recognition by sigma factors that recruit RNA polymerase and drive transcription (Shultzaberger *et al.*, 2007). In *E. coli*, the -10 region is conserved across endogenous promoters, generating the consensus sequence ‘TATAAT’ between positions -12 and -7 relative to the transcriptional start site (Shultzaberger *et al.*, 2007). Similarly, positions -15 and -14 of the promoter are conserved, comprising the extended -10 region. Both conserved regions influence promoter strength, which can be modulated by mutating the nucleotide

bases in the region. In this study, we produced a library of constitutive promoters with varying levels of strength by targeting the -10 region and the -10 extended region of PconII using random mutagenesis. The PconII library allows for strict transcriptional control over a wide range of expression levels across diverse cyanobacterial strains. We demonstrated the characterization of the promoter library in each of three model laboratory strains (*Synechococcus elongatus* PCC 7942, *Synechocystis sp.* PCC 6803, and *Anabaena sp.* PCC 7120) and the robust production strain *Leptolyngbya sp.* BL0902 (Taton *et al.*, 2012).

1.2 Results

Currently, heterologous promoters perform variably across different organisms and are limiting in cyanobacterial biotechnology. Promoter libraries provide a large range of expression levels, but previous studies have not tested for consistency across multiple strains (Gordon & Pfleger, 2018). In this study, we constructed a set of constitutive promoters through random mutagenesis of the conserved -10 regions of PconII using error-prone polymerase chain reaction (PCR). Two strategies were implemented to design the promoter mutants: (1) the three most conserved nucleotides of the -10 region and the two nucleotides of the extended -10 region were randomly mutated, or (2) the three least conserved nucleotides of the -10 region and the two nucleotides of the extended -10 region were randomly mutated (Figure 1.1A). Degree of conservation is based on previous studies that have identified the conservation of these bases across endogenous promoters found in cyanobacterial strains *Synechocystis*, *Anabaena*, and *Synechococcus* (Mitschke, *et al.*, 2011a; Mitschke *et al.*, 2011b; Vijayan *et al.*, 2011). We then tested the promoter mutants across several strains to obtain a diverse set of promoters with a wide range of strengths.

Beginning with *Synechococcus elongatus* PCC 7942 (hereafter *Synechococcus*), we used neutral site 1 vector NS1TC to integrate the PconII-*yfp* reporter construct into neutral site 1 of the *Synechococcus* chromosome through homologous recombination, since *Synechococcus* is not compatible with RSF1010-based plasmids (Figure 1.1B). Neutral sites allow integration of the sequences without known effects on normal cell phenotypes (Clerico *et al.*, 2007).

Preliminary data of the promoter library set designed through random mutagenesis of the most conserved nucleotides of the -10 element and the extended -10 region (strategy #1) produced clones skewed towards higher YFP expression levels and poor intermediate level diversity. For the library designed through random mutagenesis of the least conserved nucleotides of the -10 element and the extended -10 region (strategy #2), measurement of YFP fluorescence levels show high-ranged, evenly distributed expression levels of the PconII variants (Figure 1.2). Hereafter, we used the second set of promoters for successive experiments.

Genomic DNA of *Synechococcus* clones harboring the PconII variants were extracted and PconII-*yfp* fragments were amplified using PCR. Fragments were assembled into the broad-host range RSF1010-based vector using the CYANO-VECTOR cloning system. The plasmid carries the mobA-Y25F mutation, aphI kanamycin/ neomycin antibiotic cassette, and the PconII-*yfp* reporter construct (Figure 1.1C). Plasmids were subsequently transferred into *Synechocystis* sp. PCC 6803 (hereafter *Synechocystis*), *Leptolyngbya* BL0902 (hereafter *Leptolyngbya*), and *Anabaena* sp. PCC 7120 (hereafter *Anabaena*) through conjugation using AM5501 *E. coli* donor strains.

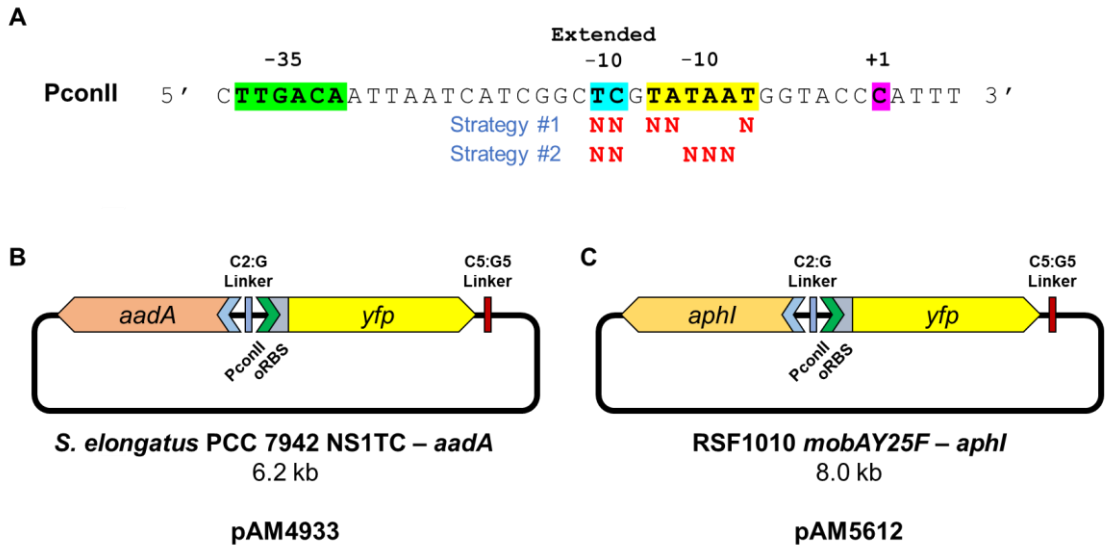


Figure 1.1. Construction of PconII library and assembly of plasmids. (A) PconII sequence with labeled -35 and -10 regions. Transcriptional start site begins at +1 position. Two strategies were used to target specific positions in the promoter. Strategy #1: the extended -10 region (highlighted in blue) and three most conserved nucleotides of the -10 element (yellow). Strategy #2: the extended -10 region (blue) and three least conserved nucleotides of the -10 element (yellow). (B) Diagram of *Synechococcus elongatus* PCC 7942 neutral site 1 plasmid pAM4933 carrying the *aadA* spectinomycin/ streptomycin antibiotic resistance cassette and PconII-YFP reporter gene construct with optimized ribosomal binding site (oRBS). (C) Diagram of RSF1010-based plasmid pAM5612 carrying the *aphI* kanamycin/neomycin antibiotic resistance cassette and P*con*II-YFP reporter gene construct with optimized ribosomal binding site (oRBS).

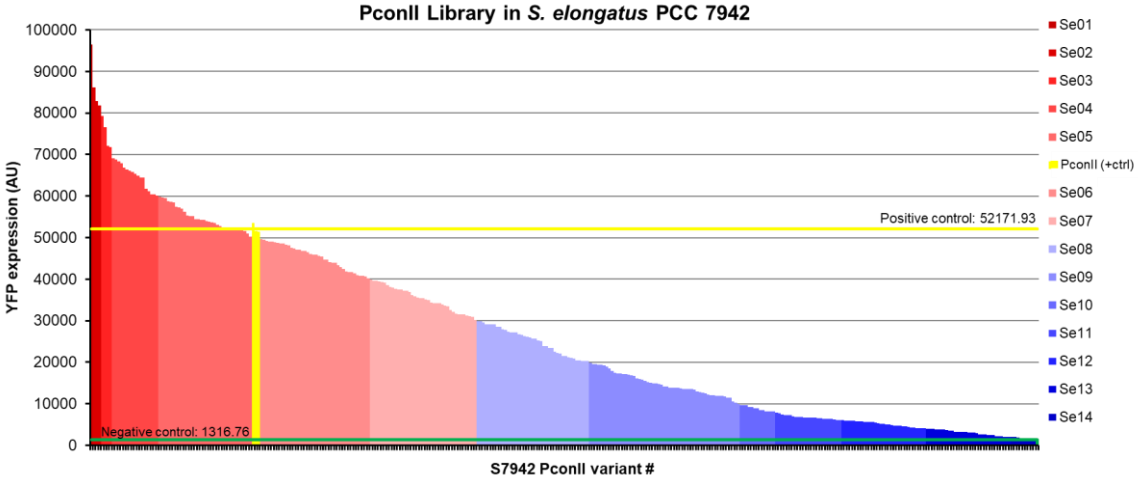


Figure 1.2. Preliminary characterization and grouping of the PconII library in *S. elongatus* PCC 7942. Collection of mutagenized PconII promoters were mass-assembled into a neutral site 1 (NS1) vector and transferred into *S. elongatus* PCC 7942 through natural transformation. For each promoter variant, we measured yellow fluorescent protein (YFP) expression levels of the engineered strain and normalized values to culture chlorophyll *a* measurements. Library clones with similar expression levels were grouped by bar colors, with the darkest red being the group displaying the highest expression (Se01) and the darkest blue displaying the lowest (Se14). Controls: original PconII average YFP expression level, yellow horizontal line; negative control, green line.

To characterize the promoter library in the next strain, *Synechocystis*, as well as additional strains *Leptolyngbya* and *Anabaena*, we used the following parameters to select variants and condense the library. All variants with YFP expression levels below twice that of the negative control were removed. All variants with high expression levels (7x above that of the negative control) were saved. The remaining variants were then grouped based on expression levels of $1x10^4$ a.u. increments. Promoters in each group were pooled and 1 out of every 3.5 of the variants was selected. Grouping ensured that our selected variants maintained the diversity of promoter strengths throughout additional strain characterizations. Additionally, grouping allows us to trace specific promoters back to earlier groups in previous characterized strains. As a result, we obtained 40 variants providing a wide range of expression levels across the four laboratory strains (Supplemental Figure 1). However, we sequenced all 40 library clones to determine duplicate sequences. After removing duplicates, the final set of promoters consisted of 25 variants in addition to the original conII promoter.

To assess the final set of 25 variants of the PconII library in our laboratory strains, we recharacterized each of the four cyanobacterial strains with each of the variants assembled in the RSF1010-based plasmid. Similarly, these plasmids carried the PconII-*yfp* reporter construct for quantifying expression levels. Since wild-type *Synechococcus* is incompatible with RSF1010, we used a mutant strain of *Synechococcus* capable of taking up the plasmid through natural transformation (A. Taton, personal communication). Characterization of the 25 variants, the unaltered promoter, and negative control (empty plasmid) produced a wide range of expression levels in *Synechococcus* (Figure 1.3A). Promoters exhibit levels as high as nearly 125% to as low as 4% of the level exhibited by unaltered PconII, with promoters in between displaying gradual decrements of strength.

We performed biparental conjugation to introduce the 25 PconII variants into *Synechocystis* (Figure 1.3B), *Leptolyngbya* (Figure 1.3C), and *Anabaena* (Figure 1.3D). In *Synechocystis*, the library spans a 150-fold range in expression levels (Figure 1.3B). However, we observed a sharp drop in coverage, where there was a 55% gap between two variants. We also observed that specific promoters exhibiting high expression in *Synechococcus* was much lower in *Synechocystis* relative to the unaltered promoter. For example, variant 'A10.L10.S09.Se03' in *Synechococcus* is the top performing promoter at 123%, but in *Synechocystis* the same variant only exhibits 75%. In contrast, the top performing promoter in *Synechocystis* ('A01.L02.S06.Se05' at 147%) is suppressed in *Synechococcus* (26%). In *Anabaena*, most of the PconII library clones span a 200-fold range in expression levels (Figure 1.3C). Interestingly, we also noticed one outlier displaying over 350 times higher in strength than the unaltered PconII. In *Leptolyngbya*, we observed a 170-fold span of promoter strength levels (Figure 1.3D). Variants 'A01.L02.S06.Se05' and 'A06.L03.S04.Se05' are consistently high-expressing promoters across *Synechocystis*, *Anabaena*, and *Leptolyngbya*, but not in *Synechococcus*.

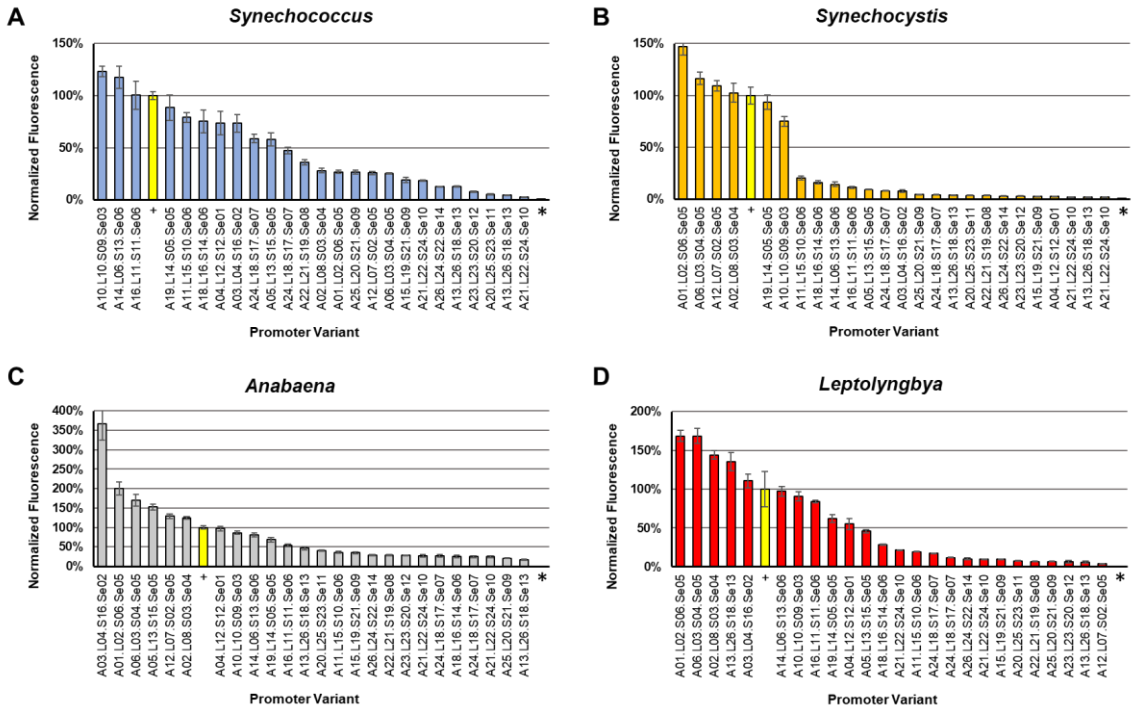


Figure 1.3. YFP expression levels of PconII library clones in cyanobacterial strains *Synechococcus*, *Synechocystis*, *Anabaena*, and *Leptolyngbya*. Cell cultures with each promoter variant were grown for 5 days in multi-well plates, adjusted to an optical density of 0.1 at 750 nm, and measured for yellow fluorescent protein (YFP) expression using a Tecan Infinite® M200 plate reader. Expression levels are normalized to those of the original conII promoter (yellow bar marked with (+)). The negative control (plasmid lacking PconII) is marked by an asterisk (*). Error bars represent standard deviations from biological triplicates.

To examine promoter expression using fluorescence microscopy, we selected three PconII variants demonstrating relatively low, medium, and high YFP expression from each cyanobacterial strain. Micrographs reveal that samples with the relatively weakest PconII variants exhibited the lowest reporter intensity and those with the strongest variants displayed the highest (Figures 1.4A-D). Next, we analyzed the sequences of each of the PconII variants to determine if certain point mutations correspond to patterns of higher or lower relative expression across each individual strain. We used the web-based sequence logo generator WebLogo 2.8.2 (Crooks *et al.*, 2004). Sequence logos represent multiple sequence alignments of the PconII variants (34 nucleotides preceding the transcriptional start site) for each strain characterized

(Figures 1.5A-D). The targeted extended -10 region (positions -15 and -14) appear to have an equal share of the four nucleotide bases in both stronger and weaker promoters. Weaker promoters retained the three less conserved bases of the -10 region (positions -10 to -8) worse than the stronger promoters.

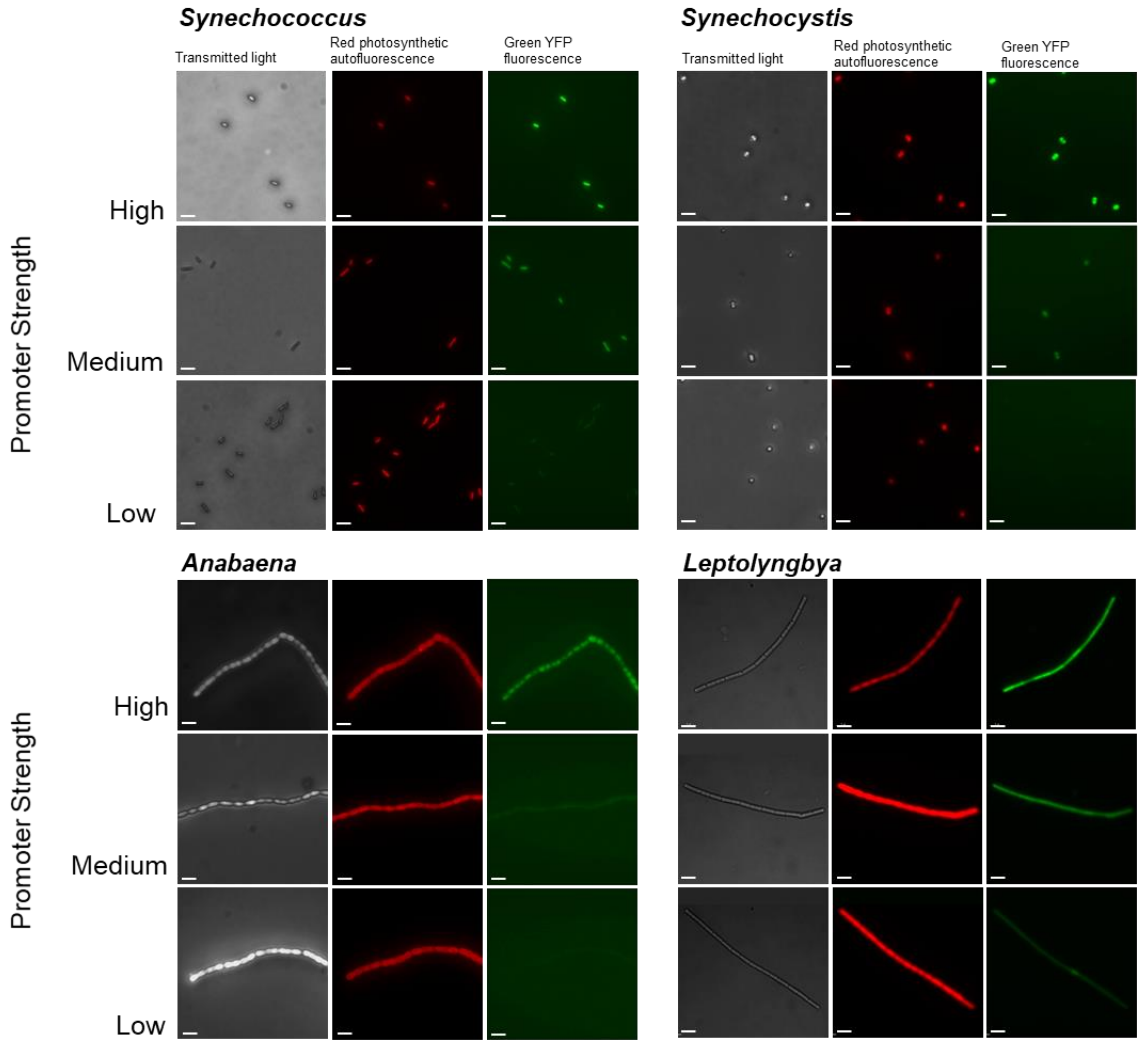


Figure 1.4. Micrographs of cyanobacterial strains harboring low, medium, and high strength PconII variants. Strains (*Synechococcus*, *Synechocystis, Anabaena*, and *Leptolyngbya)* express *yfp* reporter fluorescence driven by library clones of different strengths. For all strains: left panel, transmitted light; middle panel, red photosynthetic (chlorophyll) autofluorescence; right panel, green YFP fluorescence. Scale bars, 5 μm.

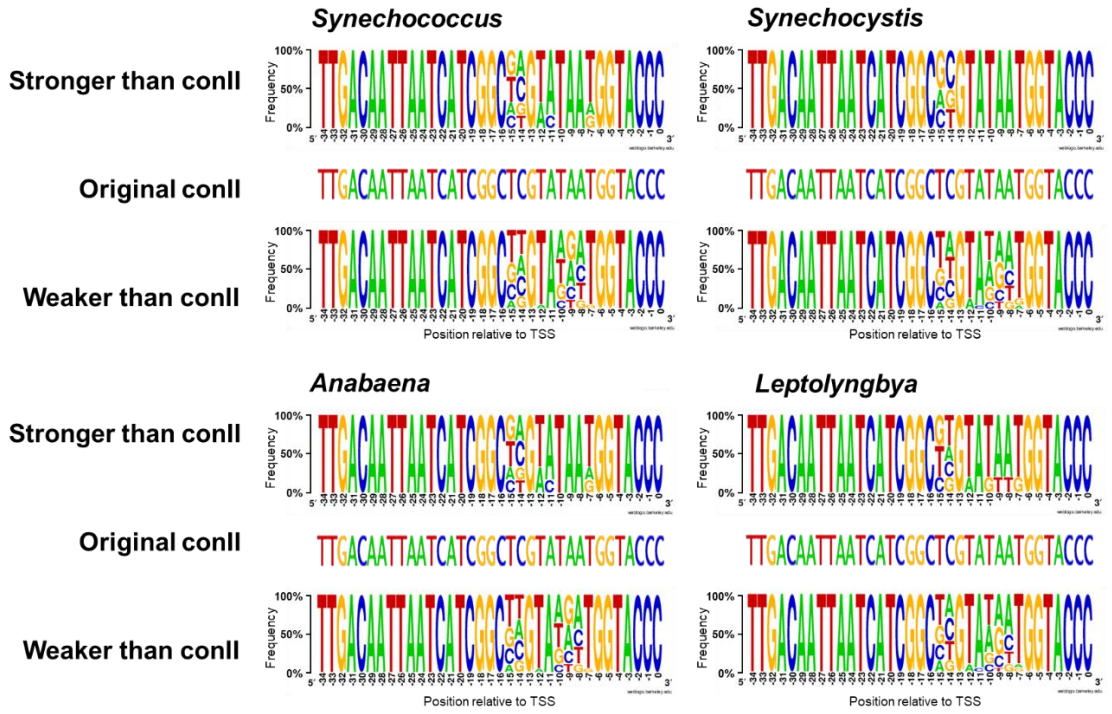


Figure 1.5. Sequence logos depicting multiple sequence alignments of conII library clones with stronger or weaker expression than the original conII. Library clones were characterized in each cyanobacterial strain and categorized based as having either higher ("stronger") or lower ("weaker") YFP expression level than that of the original promoter ("conII") in the same strain. The 34 nucleotides preceding the transcriptional start site (TSS) of the conII promoter were analyzed using WebLogo 2.8.2 (Crooks *et al.*, 2004). The number of promoter variants in each strain were: *Synechococcus* (stronger: n=3, weaker: n=22); *Synechocystis* (stronger: n=4, weaker: n=21); *Anabaena* (stronger: n=6, weaker: n=19); and *Leptolyngbya* (stronger: n=5, weaker: n=20).

1.3 Discussion

Several native and foreign promoters have been developed as genetic tools in cyanobacteria, but are limited to certain strains. For example, the cpcB promoter library has only been tested in *Synechococcus elongatus* PCC 7002. Performance is expected to vary across strains and few papers have addressed this concern (Gordon & Pfleger, 2018). The heterologous conII promoter has been used to study gene regulation across different cyanobacterial strains, including *Synechococcus elongatus* PCC 7942 (hereafter, *Synechococcus*), *Synechocystis* sp. PCC 6803, and *Anabaena* sp. PCC 7120 (Li & Golden, 1993; Taton *et al.*, 2014; Wei *et al.*, 1993). PconII is constitutively

expressive and notably strong. In a previous study, PconII and PpsbAI (positions -54 to +43) were each fused upstream of the bioluminescence reporter luxAB (Michel *et al.*, 2001). PpsbAI is one of the strongest endogenous promoters in *Synechococcus*, and drives expression of the *psbA1* gene necessary for the photosystem II reaction center protein (Nair *et al.*, 2001). In *Synechococcus*, the PconII:*luxAB* construct produced about twice the bioluminescence levels than the PpsbAI:*luxAB* construct (Michel *et al.*, 2001), suggesting that PconII is a strong promoter in *Synechococcus* and other cyanobacteria.

Our promoter library is a collection of 25 conII variants that each cover a certain level of expression. Collectively, they range from low expression (near background levels of the negative control) to high expression (surpassing levels produced by the strong original conII) (Figures 1.3A-D). Developed with RSF1010-based plasmids, the promoter library can be used to conduct genetic experiments in a wide host range. In addition, the promoter library is compatible with the CYANO-VECTOR cloning system and can be combined with antibiotic resistance cassettes and reporter genes during the assembly of new plasmids. Furthermore, plasmids can be transferred into recipient cyanobacteria via conjugation using *E. coli* donor strains.

With precise transcriptional control over heterologous gene pathways, our library has potential applications in cyanobacterial metabolic engineering. Production of biofuel molecules and other natural products not only require optimization of biosynthesis pathways, but also regulation on critical cellular processes for survival. For example, a previous study genetically engineered *Synechococcus* to produce ethylene, which is important in making synthetic chemicals (Takahama *et al.*, 2003). However, ethylene production depleted metabolic intermediates in the tricarboxylic acid (TCA) cycle, which led to a decline in cell growth. In this situation, the conII library can be implemented to maintain metabolic flux and ensure sustainable production. Expression of the ethylene

biosynthesis pathway can be fine-tuned to allow replenishment of TCA cycle intermediates.

Currently, the promoter library has been tested in four distinct laboratory strains (Figures 1.3A-D). Specific variants may not perform similarly from one strain to another, such as variant 'A10.L10.S09.Se03' between *Syenechococcus* (123%), and *Synechocystis* (75%). Despite the differences in expression levels, our library maintains the diversity of promoter strengths so that the same set of promoters can be applied across multiple strains. Nevertheless, the library is still prone to deficiencies in coverage. In *Synechocystis*, we observed a 55% gap in expression levels between two variants (Figure 1.3B). This sharp drop is most likely due to an absence of variants that had been removed during grouping and pooling. While we were able to maintain diversity of promoter strengths overall in *Synechococcus*, *Anabaena*, and *Leptolyngbya*, this deficiency can become apparent in future characterized cyanobacterial strains. However, coverage can be restored by screening for additional clones from the original set of mutagenized conII promoters and selecting those that fill in the gaps in coverage.

We then investigated the sequences of the conII variants to determine if variants stronger than the original conII promoter harbored any notable patterns (Figure 1.5A-D). Sequence logos depict higher frequencies of the consensus 5'-TAA-3' sequence at positions -10 to -8 (the three least conserved nucleotides of the -10 region) for stronger variants. In contrast, weaker variants depict less conservation of the consensus nucleotides at the same positions, where equal proportions of each of the four possible bases exist. This observation is not surprising since these same three positions the least conserved in the -10 region across endogenous promoters in *Synechocystis*, *Anabaena*, and *Synechococcus* (Mitschke, *et al.*, 2011a; Mitschke *et al.*, 2011b; Vijayan *et al.*, 2011). However, the pattern seen in stronger conII clones suggest that the consensus 5'-TAA-3' sequence could lead to increased expression levels, such as in *Synechocystis*

where the sequence is conserved (Figure 1.5B). Changing the nucleotides to the consensus sequence has been suggested to improve expression from orthogonal promoters (Huang & Lindblad, 2013). However, another source has suggested that a fully consensus promoter may be detrimental to activity due to the amount of the times the RNA polymerase interacts with the promoter (Hook-Barnard & Hinton, 2007). This observation may explain why stronger conII promoters tend to have the consensus -10 region sequence, yet the region is not wholly conserved.

Although potentially useful as a new genetic tool, the conII library in RSF1010-based plasmids still has further limitations. Although the plasmid has broad host range, it does not include the wide diversity of cyanobacteria used in research settings and promising new strains that continue to be isolated. Variations between strains may affect overall performance of the promoter library, as strains may differ in genome size and genome content. Although RSF1010 has been characterized as a low-copy number plasmid (around 10-12 copies per cell in *E. coli*), copy numbers in cyanobacteria may vary depending on the strain, number of copies of its chromosome, and even culturing conditions (Frey *et al.*, 1992; Huang *et al.*, 2010). Some strains will not be compatible with the RSF1010 system, such as wild-type *Synechococcus elongatus* PCC 7942. Moreover, our library has currently only been tested with the expression of fluorescent reporter proteins in cyanobacteria. It will be necessary to test specific genes of interest (biomolecule synthesis pathways, for example) to determine any effects on promoter performance.

SECTION II: BIOPROSPECTING FOR HALOTOLERANT, GENETICALLY TRACTABLE STRAINS

2.1 Introduction

Bioprospecting is a technique for discovering and deriving chemical products from biotic sources with potential industrial, research, or medical applications. Bioprospecting experiments using cyanobacteria have revealed that these microorganisms can be reservoirs of new natural products and chemical compounds. For example, extracts from wild specimens collected from the Aegean Sea showed antibacterial, antifungal, anti-inflammatory, and anticancer effects (Montalvão *et al.*, 2016). However, because countless branches within the cyanobacteria taxon have yet to be explored, there is great potential to discover new specialized metabolites that could be used in future pharmaceuticals, bioplastics, or sources of renewable energy (Lennen & Pfleger, 2013; Niedermeyer, 2015).

In our laboratory, bioprospecting also serves as a method for discovering and domesticating new strains of cyanobacteria for biotechnological applications. With growing focus on developing renewable sources for industrial feedstock, engineered cyanobacteria serve as platforms for production of biofuel and bioplastics. Scientists continue to search for new strains with advantageous traits for industrial use, including genetic tractability, robust growth, and tolerance to grow in non-potable water or other harsh conditions. Growth in non-potable brackish water allows sustainable large-scale production without competing for freshwater resources allocated to food production and human consumption. Algae cultivation has been proposed as a strategy for removing excess nutrients from agricultural wastewater in which case the environmental pollutants are converted into valuable cyanobacterial biomass and natural products (Rawat *et al.*, 2011). Wastewater can contain elevated levels of dissolved carbonate salts, leading to high cyanobacterial growth rates (Markou & Georgakakis, 2011). Additionally, harsh

growth conditions discourage contamination from undesired strains and predators in production pools (Taton *et al.*, 2012).

Several cyanobacterial species have demonstrated growth in extreme conditions. *Arthrospira* species (common name Spirulina) is a genus of edible cyanobacteria that is grown in large-scale, highly alkaline pools (Furmaniak *et al.*, 2017). *Arthrospira* is recognized as a historical food source and continues to be commercially grown for producing dietary supplements. However, current cultivated species of *Arthrospira* are unfavorable for genetic modifications, with very limited success for transformation (Furmaniak *et al.*, 2017). Other strains with potential for large-scale growth, such as *Leptolyngbya* BL0902 and *Synechococcus elongatus* PCC 11901, are more genetically tractable. The potential production strain *Leptolyngbya* BL0902 (hereafter, *Leptolyngbya*) was discovered after screening for strains with desired traits from an environmental sample (Taton *et al.*, 2012). *Leptolyngbya* grows robustly, can be genetically manipulated using RSF1010-based plasmids, and demonstrates tolerance to high salt, highly alkaline, and light-intensive environments. Moreover, *Leptolyngbya* accumulates more fatty acids than two strains of *Arthrospira* spp., suggesting it may be a better model organism for producing biofuel (Taton *et al.*, 2012). One recent study showed that expression of DesB and DesD acyl-lipid desaturases enabled high level production of stearidonic acid, a type of fatty acid, in *Leptolyngbya* (Poole et al., 2020). Similarly, researchers recently isolated the fast-growing and genetically tractable strain *Synechococcus elongatus* PCC 11901 from a sample of seawater (Włodarczyk *et al.*, 2020). Engineering the strain to produce free fatty acids showed higher productivity than that of model strain *Synechococccus* PCC 7002 (Włodarczyk *et al.*, 2020). Although new strains continue to be identified, techniques for isolating genetically tractable cyanobacteria have been underdeveloped. Researchers isolate new strains of cyanobacteria by obtaining samples from wild sources and

enriching for the growth of culturable strains. These isolated strains may not be necessarily be genetically tractable, and further laborious screening must be done to assess for genetic tractability. Only recently has there been a promising method of bioprospecting specifically for genetically tractable cyanobacteria.

Plasmids based on the broad-host-range RSF1010 plasmid serve both as a tool for conducting genetic modifications in cyanobacteria and as a practical method for isolating genetically tractable strains through bioprospecting. Bishé *et al.* (2019), using the modified RSF1010-based plasmid pAM5409, demonstrated the isolation of genetically tractable strains by transferring the plasmid into wild mixed cultures collected from freshwater lakes using biparental conjugation. Exconjugant cultures expressed antibiotic resistance and the fluorescence marker carried on the pAM5409 plasmid. This bioprospecting technique allows for the selection of tractable strains even before the isolation of colonies and eliminates the need for extensive screening of individual isolated clones. Exconjugant clones were shown to maintain the pAM5409 plasmid stably over time even when grown under nonselective conditions for several months. However, by removing strains from selection, strains can ultimately be cured of pAM5409 and used for subsequent genetic experiments (Bishé *et al.*, 2019). Since RSF1010-based plasmids have broad host range, genetic modifications can be performed across diverse new strains as well as model laboratory strains.

In this study, we expand the utility of the pAM5409 plasmid for cyanobacterial bioprospecting by isolating genetically tractable strains from a soda lake. Unlike freshwater lakes, soda lakes contain higher concentrations of dissolved salts and are highly alkaline (Budinoff & Hollibaugh, 2007). Native cyanobacteria adapted to these extreme conditions are ideal for cultivation in brackish water. Compared with freshwater, using brackish water avoids future resource conflicts and is more sustainable. Moreover,

strains exhibit a faster growth rate due to the high availability of dissolved carbon in the growth medium.

We selected Mono Lake, an alkaline lake in eastern California, as our source for bioprospecting tractable, halotolerant cyanobacteria. In a previous study, researchers isolated a picocyanobacterium from Mono Lake (Budinoff & Hollibaugh, 2007). The strain, MLCB, exhibits tolerance to high salinity, but it is unknown if the cyanobacterium is genetically tractable. Applying the bioprospecting technique as described by Bishé *et al.* (2019) for Mono Lake, we hoped that we might isolate a tractable strain of *Arthrospira*. Because *Arthrospira* is already an established commercial genus of cyanobacteria, *Arthrospira* would make an ideal production strain cultivated in brackish water. We conjugated the RSF1010-based plasmid pAM5409 into a mixed wild culture derived from a Mono Lake water sample. Isolated exconjugant cyanobacteria were selected and confirmed by the antibiotic resistance and fluorescence markers carried by pAM5409. We selected one of the isolated strains, ML3B, to determine its high carbonate growth tolerance and for the characterization of the conII promoter library.

2.2 Results

With rising interest in using cyanobacteria for producing biofuel and biomass feedstocks, bioprospecting serves as a method for discovering new strains suitable for industrial scale cultivation. Bioprospecting methods have been underdeveloped, and only recently has a technique been described for efficiently isolating genetically tractable strains. The methods described by Bishé *et al.* (2019) screened for genetically tractable cyanobacteria before isolating colonies and eliminated the need for extensive screening as employed in previous studies.

Our objective was to isolate a new cyanobacterial strain that was genetically tractable and halotolerant. Due to the appeal of using waste or brackish water for

sustaining large scale cultures, we hoped to derive a genetically tractable strain of halophilic *Arthrospira*. We collected water samples from Mono Lake, a highly alkaline (pH >9.0), hypersaline lake in eastern California. Due to its limited drainage, the lake is subject to high rates of evaporation and accumulation of salts. Mono Lake has a salinity of 8.5% (about 2.5 times as salty as the ocean) and carbonate levels at 400 mmol L^{-1} (Budinoff & Hollibaugh, 2007). The lake's environmental conditions make it a potential source for deriving genetically amenable cyanobacteria with halophilic and alkaliphilic properties.

Samples were collected from the shore of Mono Lake and inoculated in liquid BG-11 medium. Mixed cultures were grown up under constant medium light at 30°C. Biparental mating was performed using the AM5501 conjugal *E. coli* donor strain, which enabled the transfer of pAM5409 into recipient cyanobacteria. The mixture of conjugated cells was grown on plates with antibiotic selection until isolated colonies were formed. Colonies were picked, grown up in individual liquid cultures, and then observed by microscopy to distinguish between cyanobacteria and eukaryotic green algae based on cell size and morphology.

We obtained a total of four clonal strains of cyanobacteria demonstrating potential for genetic tractability. All strains were able to grow on spectinomycin/ streptomycin selective media, suggesting that they expressed the *aadA* antibiotic resistance cassette. In addition, all 4 strains expressed YFP reporter fluorescence, which confirmed the presence of the pAM5409 plasmid (Figure 2.1). Using microscopy, we found that all 4 strains grew as chains of cells in filaments. Three of the strains formed straight filaments but strain ML2C1 formed helical filaments (Figure 2.1). All strains grew as distinct colonies on agar plates, and three of the four routinely formed clumps when grown in liquid BG-11. ML3B was the only strain that grew as dispersed filaments in liquid medium.

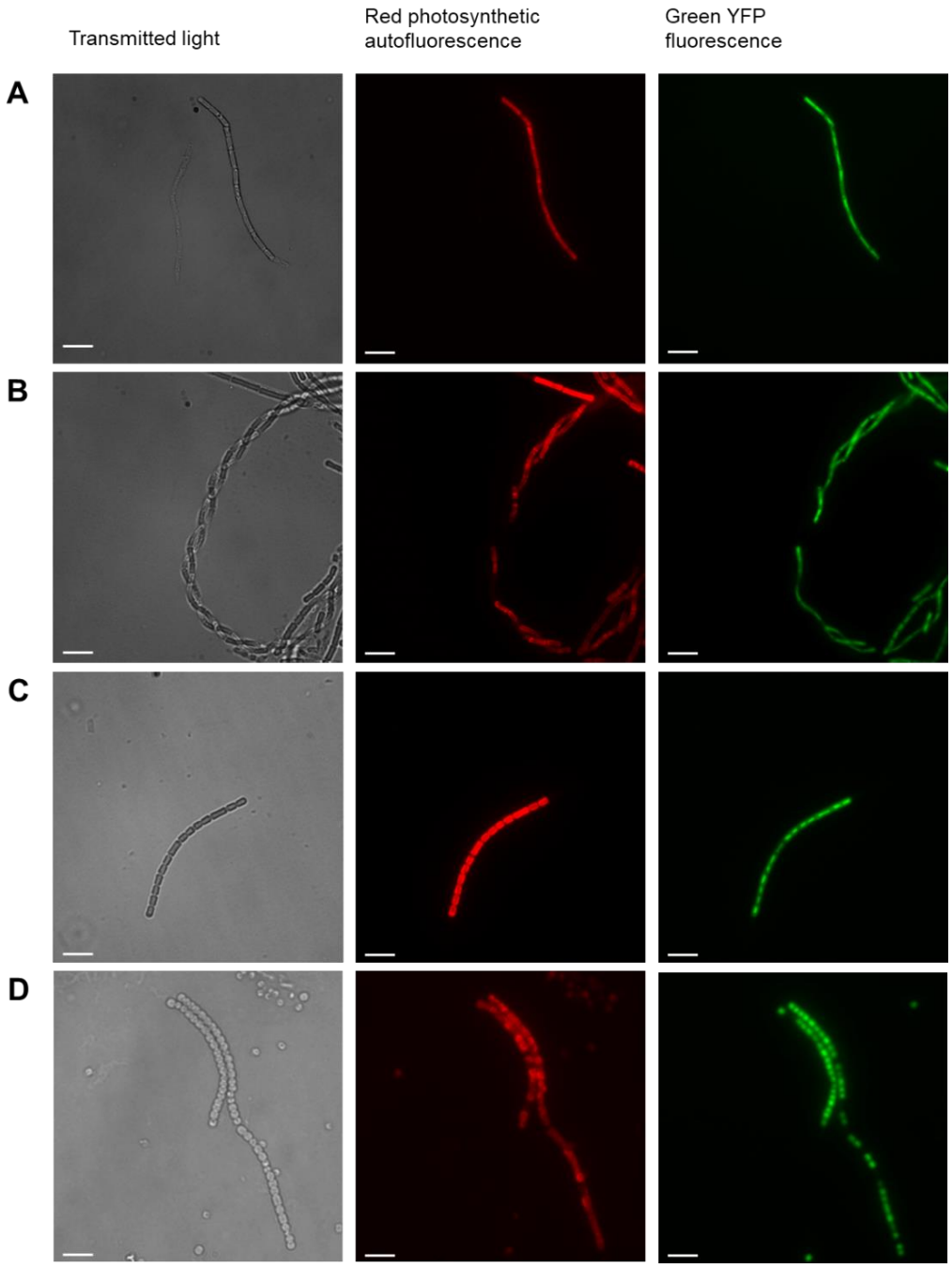


Figure 2.1. Micrographs of bioprospected exconjugant strains. Isolated cyanobacterial strains from Mono Lake expressing spectinomycin/streptomycin resistance and YFP fluorescence after conjugation with pAM5409. Left column, transmitted light; middle column, red autofluorescence of photosynthetic pigments; right column, green YFP fluorescence. Bioprospected strains: (A) ML2A, (B) ML2C1, (C) ML2C2, and (D) ML3B. Scale bars, 5μm.

To determine the phylogeny of each of the strains, we performed colony polymerase chain reactions (PCRs) to amplify the 16S ribosomal RNA genes using cyanobacterial-specific primers (Taton *et al.*, 2003). We sequenced the PCR products and queried the sequences using BLASTN in the National Center for Biotechnological Information database. The top 5 BLASTN hits for each bioprospected strain are shown in Table 2.1. All strains had hits with at least 97% identity within the *Nodosilinea* and *Leptolyngbya* genera, along with additional unclassified strains. Queries for strains ML2-C1 and ML2-C2 returned the same highest match with "Uncultured bacterium clone GBII-15" from The Netherlands. We then used the multiple sequence alignment tool Clustal Omega to determine the sequence similarities of the four 16S ribosomal RNA genes. The alignment showed all 4 strains to be at least 94% identical with one another.

Table 2.1. Top 5 BLASTN hits for each bioprospected strain obtained by querying the 16S rRNA gene sequences in the GenBank database.

Description*	Accession	Query Coverage	Score	E-value	Identity	Origin of strain
ML2A						
Uncultured bacterium clone GBII-87	GQ441350.1	100%	2348	0	98.07%	The Netherlands
Uncultured bacterium clone T-05_10	KP793940.1	100%	2337	0	97.92%	El Tatio, Chile
Leptolyngbya sp. 0BB30S02	AJ639892.1	100%	2324	0	97.77%	-
Leptolyngbya sp. LEGE 07080	HM217085.1	100%	2302	0	97.47%	Porto, Portugal
Leptolyngbya antarctica ANT.LAC.1	AY493588.1	99%	2298	0	97.54%	Antarctica

*Redundant, closely related clones obtained from the same source were removed.

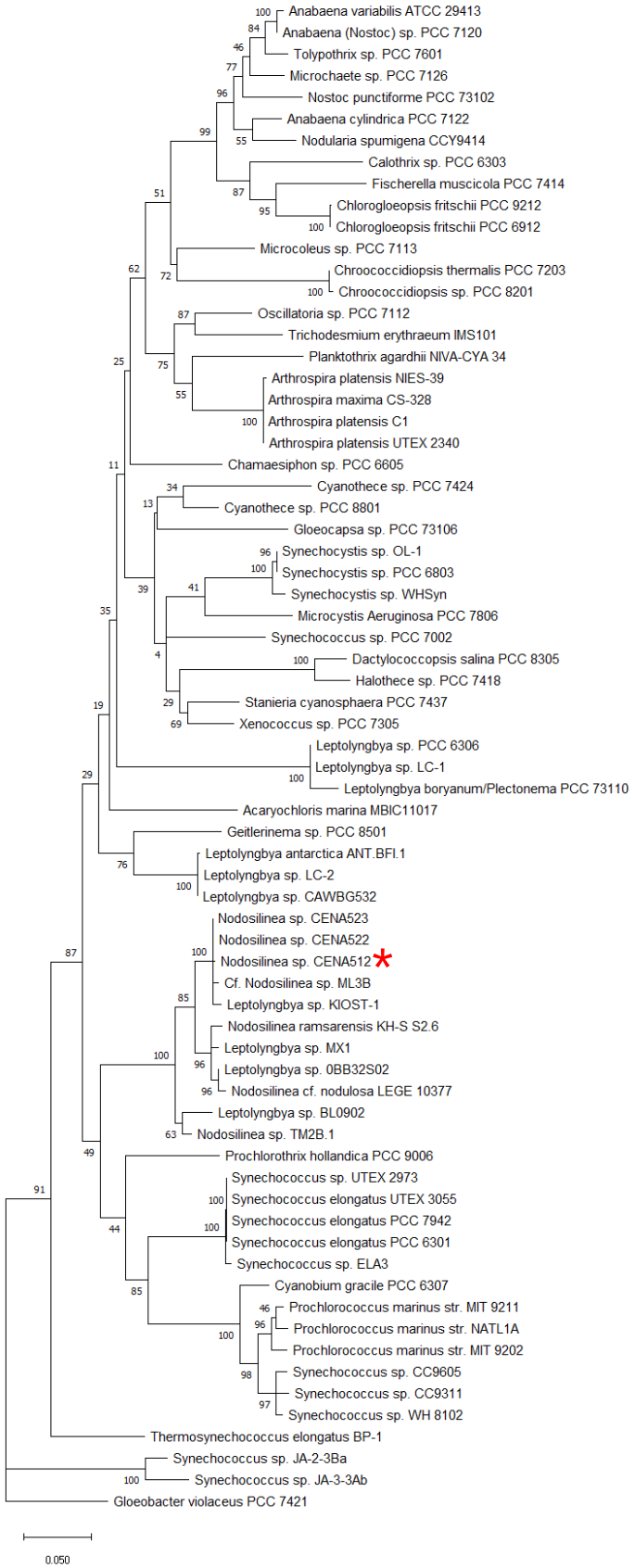
Table 2.1 (continued). Top 5 BLASTN hits for each bioprospected strain obtained by querying the 16S rRNA gene sequences in the GenBank database.

Description*	Accession	Query Coverage	Score	E-value	Identity	Origin of strain
ML2-C1						
Uncultured bacterium clone GBII-15	GQ441289.1	100%	2206	0	98.33%	The Netherlands
Uncultured cyanobacterium clone T-05_10	KP793940.1	100%	2200	0	98.25%	El Tatio, Chile
Leptolyngbya antarctica ANT.LAC.1	AY493588.1	99%	2165	0	97.85%	Antarctica
Nodosilinea nodulosa LEGE 06104	KU569325.1	100%	2161	0	97.69%	-
Nodosilinea sp. ACSSI 330	MT425944.1	100%	2159	0	97.69%	-
ML2-C2						
Uncultured bacterium clone GBII-15	GQ441289.1	100%	2222	0	99.59%	The Netherlands
Uncultured cyanobacterium clone T-05_10	KP793940.1	100%	2217	0	99.51%	El Tatio, Chile
Uncultured cyanobacterium clone R8-R56	DQ181691.1	99%	2182	0	99.09%	Antarctica
Leptolyngbya antarctica ANT.LAC.1	AY493588.1	99%	2180	0	99.09%	Antarctica
Leptolyngbya sp. B-CY-NM2	KX086280.1	96%	2150	0	99.74%	Turkey
ML3B						
Nodosilinea sp. CENA523	KF246490.1	99%	2475	0	99.13%	Pantanal wetlands, Brazil
Leptolyngbya sp. KIOST-1	JX401929.1	100%	2438	0	99.48%	Ansan, South Korea
Leptolyngbya sp. 0BB32S02	AJ639894.1	100%	2351	0	97.32%	-
Nodosilinea cf. nodulosa LEGE 10377	JQ927349.1	99%	2337	0	97.24%	Coastal Portugal
Uncultured bacterium clone BJGMM-3s-409	JQ801062.1	100%	2329	0	97.04%	Yellow River, China

*Redundant, closely related clones obtained from the same source were removed.

We selected the strain ML3B for additional characterization and genetic experiments. Because ML3B grows as dispersed filaments in liquid medium, the strain is preferable for taking accurate culture density and fluorescence measurements. Non-clumping strains also facilitate maintaining axenic cultures and the removal of contaminants when necessary. To obtain the ML3B strain cured of the pAM5409 plasmid used for bioprospecting, we grew the exconjugant ML3B strain in liquid growth medium without antibiotic selection and passaged the culture to fresh media every 5-6 days for a month. The resultant strain was susceptible to antibiotic selection, did not express the fluorescence YFP marker, and was further confirmed by PCR to have lost the plasmid. Phylogenetic analysis identified ML3B to be within the *Nodosilinea* genus, closely related to the strains CENA512, CENA522, and CENA523 (Figure 2.2). Additionally, ML3B is closely related to *Leptolyngbya* sp. KIOST-1. Morphologically, ML3B is a filamentous strain with cells measuring 1.17 ± 0.12 µm long and 1.16 ± 0.13 µm wide. ML3B is naturally resistant to kanamycin, but susceptible to neomycin, spectinomycin, and streptomycin. The strain was unable to grow on growth medium lacking a source of fixed nitrogen and failed to form heterocysts.

Figure 2.2. Phylogenetic tree deduced 16S rRNA gene sequences by maximum likelihood for identification of strain ML3B. The evolutionary history was inferred by using the Maximum Likelihood method and General Time Reversible model. The tree with the highest log likelihood (-14053.35) is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the Maximum Composite Likelihood (MCL) approach, and then selecting the topology with superior log likelihood value. A discrete Gamma distribution was used to model evolutionary rate differences among sites (5 categories (+G, parameter = 0.6326)). The rate variation model allowed for some sites to be evolutionarily invariable ([+I], 56.68% sites). The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. This analysis involved 70 nucleotide sequences. There were a total of 1362 positions in the final dataset. Evolutionary analyses were conducted in MEGA X. Strain ML3B is emphasized with an asterisk (*).


Anabaena variabilis ATCC 29413
Anabaena (Nostoc) sp. PCC 7120
Tolypothrix sp. PCC 7601
Microchaete sp. PCC 7126
Nostoc punctiforme PCC 73102
Anabaena cylindrica PCC 7122
Nodularia spumigena CCY9414
Calothrix sp. PCC 6303
Fischerella muscicola PCC 7414
Chlorogloeopsis fritschii PCC 9212
Chlorogloeopsis fritschii PCC 6912
Microcoleus sp. PCC 7113
Chroococcidiopsis thermalis PCC 7203
Chroococcidiopsis sp. PCC 8201
Oscillatoria sp. PCC 7112
Trichodesmium erythraeum IMS101
Planktothrix agardhii NIVA-CYA 34
Arthrospira platensis NIES-39
Arthrospira maxima CS-328
Arthrospira platensis C1
Arthrospira platensis UTEX 2340
Chamaesiphon sp. PCC 6605
Cyanothece sp. PCC 7424
Cyanothece sp. PCC 8801
Gloeocapsa sp. PCC 73106
Synechocystis sp. OL-1
Synechocystis sp. PCC 6803
Synechocystis sp. WHSyn
Microcystis Aeruginosa PCC 7806
Synechococcus sp. PCC 7002
Dactylococcopsis salina PCC 8305
Halothece sp. PCC 7418
Stanieria cyanosphaera PCC 7437
Xenococcus sp. PCC 7305
Leptolyngbya sp. PCC 6306
Leptolyngbya sp. LC-1
Leptolyngbya boryanum/Plectonema PCC 73110
Acaryochloris marina MBIC11017
Geitlerinema sp. PCC 8501
Leptolyngbya antarctica ANT.BFI.1
Leptolyngbya sp. LC-2
Leptolyngbya sp. CAWBG532
Nodosilinea sp. CENA523
Nodosilinea sp. CENA522
Nodosilinea sp. CENA512
Cf. Nodosilinea sp. ML3B
Leptolyngbya sp. KIOST-1
Nodosilinea ramsarensis KH-S S2.6
Leptolyngbya sp. MX1
Leptolyngbya sp. 0BB32S02
Nodosilinea cf. nodulosa LEGE 10377
Leptolyngbya sp. BL0902
Nodosilinea sp. TM2B.1
Prochlorothrix hollandica PCC 9006
Synechococcus sp. UTEX 2973
Synechococcus elongatus UTEX 3055
Synechococcus elongatus PCC 7942
Synechococcus elongatus PCC 6301
Synechococcus sp. ELA3
Cyanobium gracile PCC 6307
Prochlorococcus marinus str. MIT 9211
Prochlorococcus marinus str. NATL1A
Prochlorococcus marinus str. MIT 9202
Synechococcus sp. CC9605
Synechococcus sp. CC9311
Synechococcus sp. WH 8102
Thermosynechococcus elongatus BP-1
Synechococcus sp. JA-2-3Ba
Synechococcus sp. JA-3-3Ab
Gloeobacter violaceus PCC 7421
0.050

We determined the growth characteristics of ML3B in media supplemented high sodium bicarbonate compared to our common laboratory strains (Figure 2.3). Wild-type cultures of *Synechococcus*, *Synechocystis*, *Anabaena*, *Leptolyngbya*, and ML3B were grown in liquid BG-11 medium supplemented with different concentrations of sodium bicarbonate, ranging from 0 to 2 M. All strains grown in the standard medium (0 M sodium bicarbonate) exhibited high growth rates. Growth of *Anabaena* was suppressed at 0.1 M sodium bicarbonate, and *Synechococcus* and *Synechocystis* were suppressed at 0.25 M. At 0.25 M, *Leptolyngbya* grew normally for the first 4 days but the OD dropped on day 5. Strain ML3B maintained normal growth in concentrations of sodium bicarbonate up to 0.5 M. At 1 M, ML3B cultures did not grow for the first 3 days, but then showed a small amount of growth on days 4 and 5.

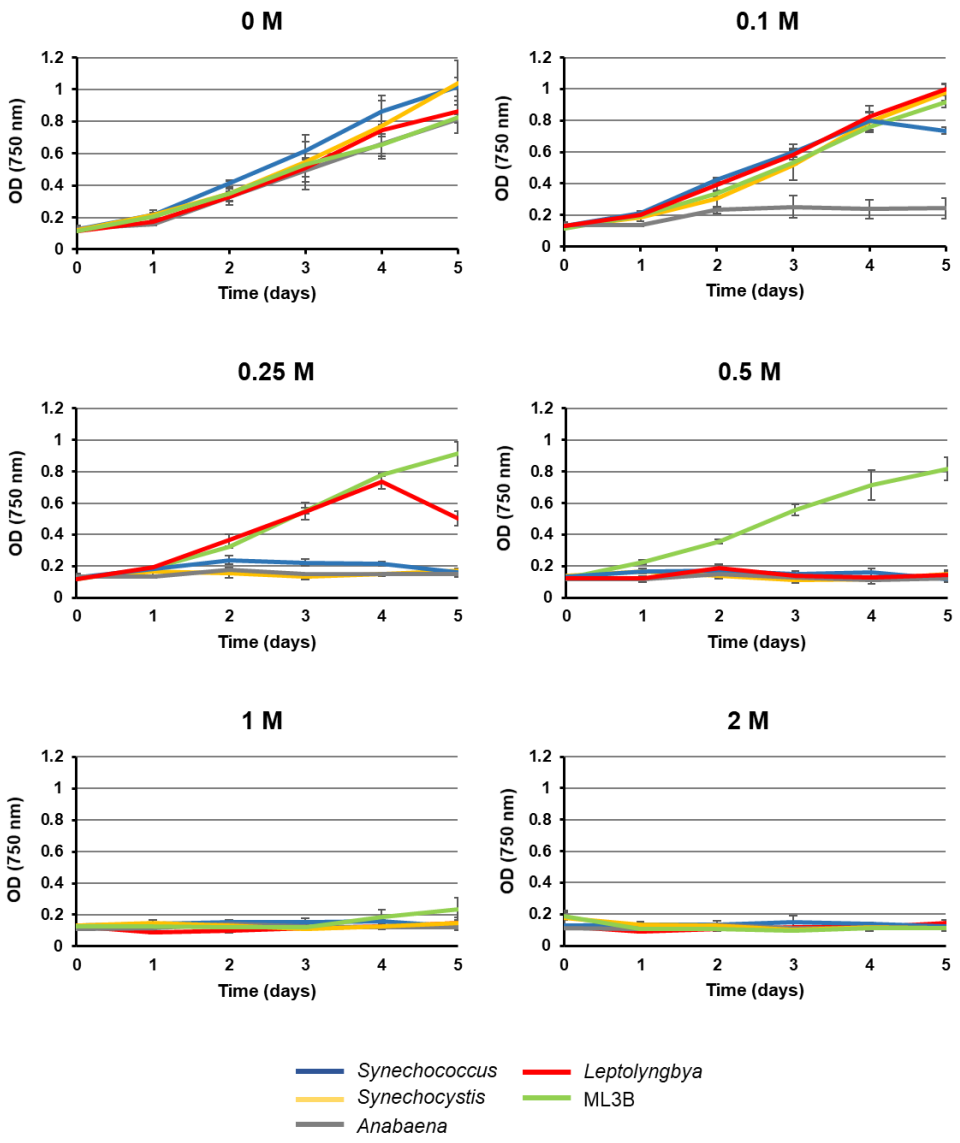
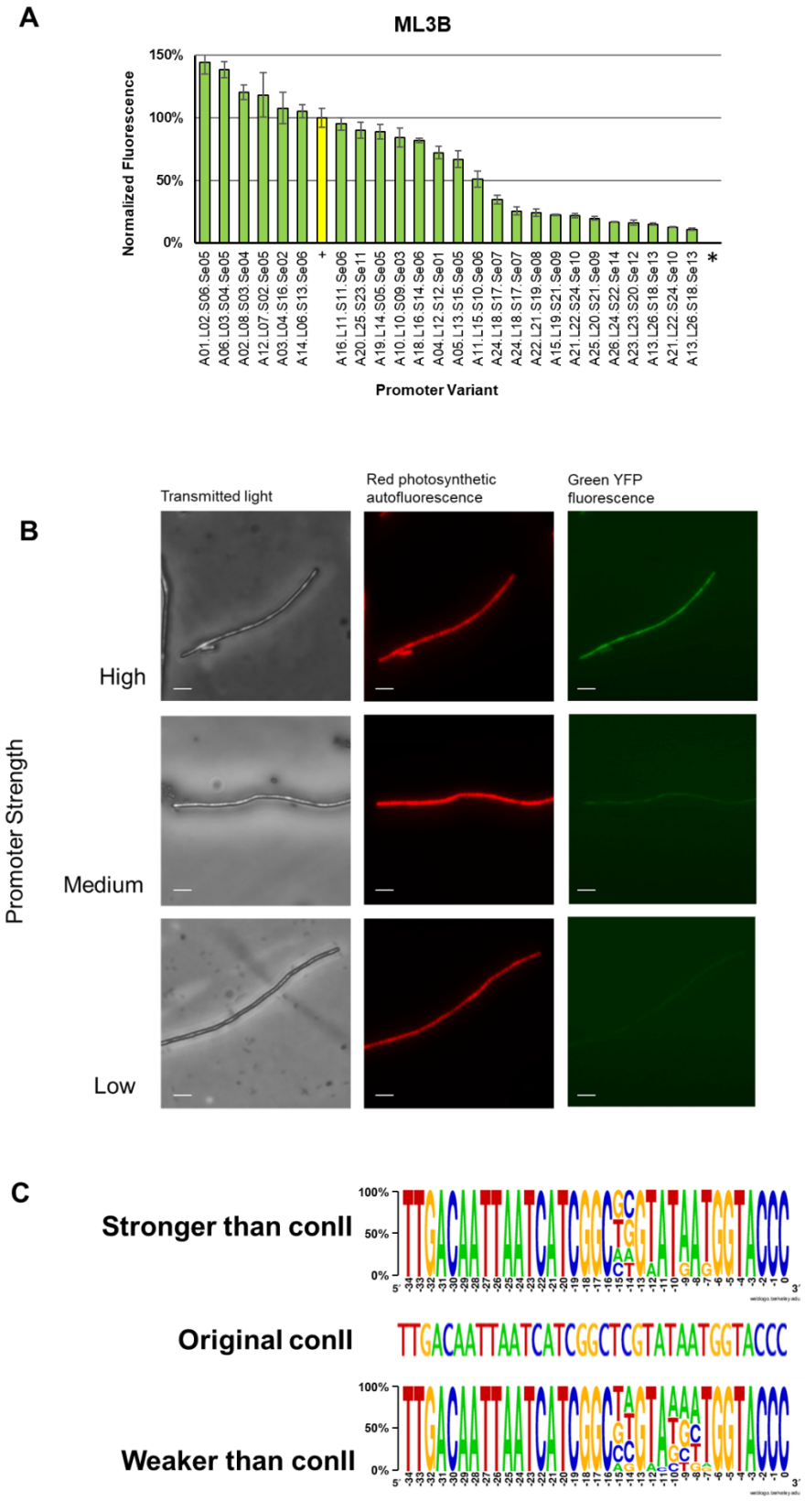


Figure 2.3. Growth tolerance of strains *Synechococcus*, *Synechocystis*, *Anabaena*, *Leptolyngbya* and ML3B in high sodium bicarbonate media. Strains were grown in BG-11 medium supplemented with sodium bicarbonate ($NaHCO_3$) at concentrations ranging from 0 to 2 M. Cultures were grown for 5 days and the optical density (OD) was measured at 750 nm each day. Error bars represent standard deviations from biological triplicates.

We characterized the PconII promoter library in the bioprospected ML3B strain to demonstrate its genetic tractability. Using *E. coli* conjugal donor strain AM5501, we conjugated the 25 variants of PconII, each cloned upstream of the YFP fluorescence marker in pAM5612, into strain ML3B. The 25 PconII variants, the original PconII promoter, and the negative control produced a 150-fold range of expression levels in ML3B (Figure 2.4A). Figure 2.4B shows YFP reporter fluorescence in 3 strains of ML3B carrying low, medium, and high strength PconII variants. As expected, the strain carrying the weakest promoter variant exhibited the lowest reporter intensity and the sample with the strongest variant displayed the highest (Figures 2.4B). We used the sequence logo generator WebLogo 2.8.2 to analyze the 34 nucleotides preceding the transcriptional start site of PconII relative to YFP expression levels in ML3B (Figure 2.4C) (Crooks *et al.*, 2004). Promoter library clones characterized in ML3B were categorized based on either higher ("stronger") or lower ("weaker") YFP expression levels than that of the original promoter in the same strain. The extended -10 region (positions -15 and -14) within the region targeted for mutagenesis displayed nearly equal frequencies of each nucleotide base across both stronger and weaker promoters. For the targeted -10 region (positions -10 to -8), stronger promoters retained the canonical 'TAA' sequence more than did the weaker promoters (Figure 2.4C).

Figure 2.4. Characterization of conII promoter library in bioprospected strain ML3B. (A) Fluorescence levels of conII promoter library clones in ML3B. Cell cultures with each promoter variant were grown for 5 days in multi-well plates, adjusted to an optical density of 0.1 at 750 nm, and measured for yellow fluorescent protein (YFP) expression using a Tecan Infinite® M200 plate reader. Expression levels are normalized to those of the original conII promoter (yellow bar marked with (+)). The negative control (plasmid lacking PconII) is marked by an asterisk (*). Error bars represent standard deviations from biological triplicates. (B) Fluorescence micrographs of ML3B strains carrying low, medium, and high strength PconII variants. For all strains: left column, transmitted light; middle column, red photosynthetic autofluorescence; right column, green YFP fluorescence. Scale bars, 5 μm. (C) Sequence logos depicting multiple sequence alignments of conII library clones with stronger or weaker expression than the original conII. Library clones were characterized in ML3B and categorized based on either higher ("stronger") or lower ("weaker") YFP expression levels than that of the original promoter in the same strain. The 34 nucleotides preceding the transcriptional start site of the conII promoter were analyzed using WebLogo 2.8.2 (Crooks et al., 2004). For number of samples: 'stronger than conII': n=6; 'weaker than conII': n=19.


A
ML3B
Normalized Fluorescence
150%
100%
50%
0%
A01.L02.S06.Se05
A06.L03.S04.Se05
A02.L08.S03.Se04
A12.L07.S02.Se05
A03.L04.S16.Se02
A14.L06.S13.Se06
+
A16.L11.S11.Se06
A20.L25.S23.Se11
A19.L14.S05.Se05
A10.L10.S09.Se03
A18.L16.S14.Se06
A04.L12.S12.Se01
A05.L13.S15.Se05
A11.L15.S10.Se06
A24.L18.S17.Se07
A24.L18.S17.Se07
A22.L21.S19.Se08
A15.L19.S21.Se09
A21.L22.S24.Se10
A25.L20.S21.Se09
A26.L24.S22.Se14
A23.L23.S20.Se12
A13.L26.S18.Se13
A21.L22.S24.Se10
A13.L26.S18.Se13
*
Promoter Variant
B
Transmitted light
Red photosynthetic autofluorescence
Green YFP fluorescence
Promoter Strength
High
Medium
Low
C
Stronger than conII
Original conII
Weaker than conII
100%
50%
0%

2.3 Discussion

The establishment of a bioprospecting method specifically for identifying genetically tractable cyanobacteria will enable the identification of new strains for biotechnological applications and industrial large-scale growth. In this study, we isolated four cyanobacterial strains that replicate the shuttle plasmid pAM5409 and express the spectinomycin/streptomycin resistance and YFP fluorescence genes carried on the plasmid. We demonstrated that the bioprospecting methods developed by Bishé *et al.* are broadly applicable for isolating cyanobacteria from a new environmental source with harsh growth conditions. These newly isolated strains are adapted to the high alkalinity and dissolved carbonate levels of Mono Lake, making them promising candidates for cultivation in brackish water. All four strains are closely related and share at least 95% sequence similarity for the 16S rRNA gene. Although these strains lack phylogenetic diversity, future studies that obtain samples from different locations within Mono Lake or from other sources, and that use different laboratory culturing methods may reveal more diverse species of cyanobacteria.

One of our strains, ML3B, was selected for subsequent characterization and genetic experiments. Unlike the other three strains, ML3B grows as uniformly dispersed filaments in liquid medium, which is preferable for taking culture measurements during cultivation. Phylogenetic analysis revealed ML3B to be genetically closest to a group of cyanobacteria isolated from the Pantanal Wetlands in western Brazil, most notably strain *Nodosilinea* sp. CENA 523 (Andreote *et al.*, 2014) (Figure 2.2). Not surprisingly, the group of *Nodosilinea* cyanobacteria was extracted from hypersaline, alkaline lakes. ML3B also shares similar morphology to *Nodosilinea* sp. CENA 522, growing as long, thin filaments (Andreote *et al.*, 2014). The next most closely related strain to ML3B is *Leptolyngbya* sp. KIOST-1, which was isolated from a culture pond of *Arthrospira* in South Korea (Kim *et al.*, 2015) (Figure 2.2). Researchers characterized the strain KIOST-

1 as a potential biomass producer comparable to *Arthrospira*, with protein content comprising 50% of cellular components. Additionally, strain KIOST-1 is tolerant to saline conditions and exhibits no cytotoxic effects on vero cell cultures (kidney epithelial cells) (Kim *et al.*, 2015). Based on these characteristics from its genetic neighbors, strain ML3B may possess similar qualities that make it suitable for large-scale cultivation. We demonstrated that ML3B can tolerate high carbonate levels, with growth surpassing laboratory strains in carbonate concentrations as high as 0.5 M (Figure 2.3).

We characterized the PconII promoter library in ML3B to assess its compatibility with components of the CYANO-VECTOR genetic toolkit that were used to construct the library (Taton et al., 2014). Similar to the previously characterized strains *Synechococcus*, *Synechocystis*, *Anabaena*, and *Leptolyngbya* sp. BL0902, the promoter library in ML3B also produced a wide range of YFP expression levels (Figure 2.4A). As evident in our previous experiments, the expression strengths of specific PconII variants do not correlate consistently between ML3B and the other 4 cyanobacterial strains. However, some variants that were consistently high-expressing in *Synechocystis*, *Anabaena,* and *Leptolyngbya* behaved similarly in ML3B. Variants A01.L02.S06.Se05, A06.L03.S04.Se05, and A02.L08.S03.Se04 exhibited YFP expression levels higher than that of the original PconII across all four strains. Predictably, the lowest-expressing variant in ML3B, A13.L26.S18.Se13, also performed poorly across all strains, including *Synechococcus*.

Our bioprospected strain has potential applications for biotechnological research as it is genetically tractable and works with our promoter library and presumably other genetic engineering tools in the CYANO-VECTOR toolkit. ML3B can serve as a model strain for studying the enzymes and metabolic capabilities that allow cyanobacteria to survive in hypersaline and alkaline conditions. Currently, researchers have speculated that cyanobacteria can be used to couple both biomass production and removal of

excess nutrients and chemical pollutants from eutrophic water sources (Markou & Georgakakis, 2011). ML3B should be useful for the development of cyanobacterial cultivation techniques in brackish and wastewater. Furthermore, ML3B can be developed as a production strain, facilitated by its genetic tractability. Considering that ML3B is genetically close to the protein-rich strain *Leptolyngbya* sp. KIOST-1, ML3B may be favorable for protein production. The ability to make genetic modifications would allow metabolic engineering to enhance production pathways and boost protein yield. ML3B can be cultivated in growth medium rich in nitrogen (such as wastewater) because increased nitrogen levels in growth media appears to correlate with increased protein content in cyanobacteria (Markou & Georgakakis, 2011; Wu & Pond, 1981). Moreover, ML3B may contain endogenous biosynthesis pathways for natural products yet to be discovered, such as new pharmaceuticals and biofuel molecules.

In conclusion, we have demonstrated efficient bioprospecting methods and genetic tools for identifying and testing genetically tractable cyanobacteria for biotechnological applications. We showed that the bioprospecting approach established by Bishé *et al.* is broadly applicable by isolating strain ML3B from a soda lake, which represents a new harsh environmental source. We demonstrated that the newly isolated ML3B exconjugant strain can be cured of the original broad-host range plasmid used for selection, which allows ML3B to be used as a new platform strain for conjugation with other plasmids for further genetic studies. Furthermore, strain ML3B has the potential to become a model strain suitable for the sustainable production of biofuel and other biomolecules.

SECTION III: CHARACTERIZATION OF THE PHOSPHITE OXIDOREDUCTASE PTXD IN CYANOBACTERIA

3.1 Introduction

Although there is great interest in using cyanobacteria as cellular factories for renewable production of bioproducts, cultivation techniques are underdeveloped. Currently, there are two distinct methods of cultivation: open-culture systems (such as outdoor raceway ponds) and closed-culture systems (photo-bioreactors, or PBRs). Open-culture systems include natural lakes and artificial ponds that are open to the surrounding environment and allow direct exchange of gases between the culture and the atmosphere (Mata *et al.*, 2010). In contrast, closed-culture PBRs strictly control growth conditions, such that the cultures have limited exposure to the environment. PBRs can regulate growth factors such as CO_2 concentration, pH, temperature, and light intensity. Control over these factors allow for the optimization of culture growth better than open-culture systems (Mata *et al.*, 2010). However, open-culture systems are the more promising method for the future of large-scale cultivation. Large raceway ponds are currently the most popular method for cultivating cyanobacteria, accounting for 90% of all algae biomass production worldwide (Zhang *et al.*, 2014). Open-culture ponds are less expensive to construct and maintain than PBRs or other closed-culture systems (Mata *et al.*, 2010). Ponds have higher capacity for production and are believed to be more sustainable to operate (Mata *et al.*, 2010; Smith *et al.*, 2010).

Despite the promise of open-culture raceway ponds for cyanobacteria cultivation, one of their biggest limitations is the risk of contamination. Ponds are exposed to the external environment, making them considerately prone to contamination from undesired and predatory microbes, including other algae, other bacteria, fungi, and protozoa (Mata et al., 2010; Zhang et al., 2014). Zooplankton, including ciliates and rotifers, graze on

cyanobacteria cells and can cause outdoor cultures to collapse within a few days (Wang *et al.*, 2013). Other species of microalgae may outcompete for nutrient resources and ultimately take the place of the original production strain in the culture (Wang *et al.*, 2013). Currently, the most appealing method of controlling for contamination is to subject the culture to extreme growth conditions, such as high acidity, high alkalinity, and high salt. One study discovered that lowering the pH of cyanobacterial cultures to 3.0 killed invading ciliates (Liu & Lu, 1990). In contrast, cultivated strains like *Spirulina* and *Oscillatoria* are grown in high pH and salt conditions, which are advantageous for lowering contamination risk (Pulz & Gross, 2004). However, most cyanobacteria are unable to grow under these extreme conditions, limiting the scope of potential strains for biomolecule and biomass production. When shifting growth conditions to extreme measures is not desirable, other methods should be considered to address contamination.

Currently, most research in cyanobacterial biotechnology is centered around engineering cyanobacteria genetically and metabolically to produce biofuel molecules and other natural products (Loera-Quezada *et al.*, 2016). However, there are few studies recognizing the possibilities of genetic tools for other applications, such as cultivation. Genetic engineering is a promising method to design strains with favorable characteristics to outcompete against undesirable organisms when grown in open ponds. One major strategy is to engineer strains with the ability to convert essential nutrients from a nonmetabolizable form to a metabolizable one. This ability would give a competitive advantage for desired strains to utilize sources of nutrients that are not directly metabolizable in a selective medium.

Phosphorus (hereafter, P) is an essential nutrient required by all organisms to sustain cellular metabolism and synthesis of biomolecules. In earth's biosphere, nearly all phosphorus is found as inorganic phosphate (+5 oxidation state) (Metcalf & Wolfe,

1998). Phosphate forms phosphodiester bonds with pentose sugar to produce the backbone of DNA and RNA. Phosphate is also a major component of phospholipids, which forms the protective plasma membrane. However, reduced compounds of phosphorus exist within the biological realm, including phosphite (+3 oxidation state) and hypophosphite (+1). Studies have found a few microorganisms that were capable of oxidizing the reduced forms when phosphate (Pi) was absent as the source of P, including *Escherichia coli*, *Saccharomyces cerevisiae*, and some species of *Pseudomonas* and *Rhizobium* (Metcalf & Wolfe, 1998).

The metabolic pathway for phosphite (Phi) oxidation was first genetically characterized in *Pseudomonas stutzeri* WM88 (Metcalf & Wolfe, 1998). This Gram-negative bacterium harbors the *ptxABCD* gene operon, which is responsible for the oxidation of phosphite to the usable phosphate form. The *ptxABC* cluster encodes for an ATP-binding cassette (ABC)-type transporter protein for phosphite (Metcalf & Wolfe, 1998). The *ptxD* gene encodes for a NAD-dependent phosphite oxidoreductase enzyme, sharing homology with a family of D-hydroxyacid dehydrogenases (Grant, 1989; Metcalf & Wolfe, 1998). Homologous gene clusters highly similar to *ptxABCD* were identified in cyanobacteria including *Cyanothece* sp. ATCC 51142, *Trichodesmium erythraeum* IMS101, and *Marinobacter aquaeolei* YT8 (Martinez *et al.*, 2012). Researchers also determined that the marine cyanobacterium *Prochlorococcus* MIT9301, which carries the *ptxABCD* operon, can use Phi as the sole source of P (Martinez *et al.*, 2012).

The metabolic capabilities of the phosphite oxidoreductase PtxD have raised interest within algal biotechnology. Engineering cyanobacteria to express the phosphite oxidation pathway can allow production strains to grow in selective media containing Phi as the only source of P. This growth strategy would reduce overall risk of contamination of open-culture ponds by discouraging the growth of microbes that cannot metabolize phosphite. One previous study engineered the green algae *Chlamydomonas reinhardtii*

to express the PtxD enzyme, which generated clones that could oxidize phosphite to the usable phosphate form (Loera-Quezada *et al.*, 2016). The study demonstrated that engineered *C. reinhardtii* dominated mixed cultures with *Scenedesmus obliquus* (green algae) when grown in media containing only Phi as the sole P source (Loera-Quezada *et al.*, 2016). Likewise, researchers recently used neutral site 1 (NS1) homologous recombination to integrate the *ptxD* gene into the cyanobacterial *Synechococcus elongatus* PCC 7942 genome (González-Morales *et al.*, 2020). The engineered strain, grown in media containing only Phi as the P source, was able to grow robustly in 100-mL nonsterile bioreactors. In contrast, a culture of wild-type *S. elongatus* grown in normal Pi-supplemented media collapsed within 4 days due to contamination (González-Morales *et al.*, 2020). Overall, the study confirms the potential of engineering cyanobacteria to metabolize phosphite as a method for controlling contaminants. However, the study's results are limited to *S. elongatus* and further investigations in other strains of cyanobacteria will require a different experimental approach.

To facilitate the expression of the phosphite oxidoreductase PtxD across multiple strains, we assembled the *ptxD* gene from *P. stutzeri* into a broad-host range RSF1010-based plasmid. We demonstrated the transfer of the plasmid into three cyanobacterial strains: *Synechocystis* sp. PCC 7803, *Leptolyngbya* sp. BL0902, and *Anabaena* sp. PCC 7120. The engineered strains were able to metabolize Phi as the sole source of P.

3.2 Results

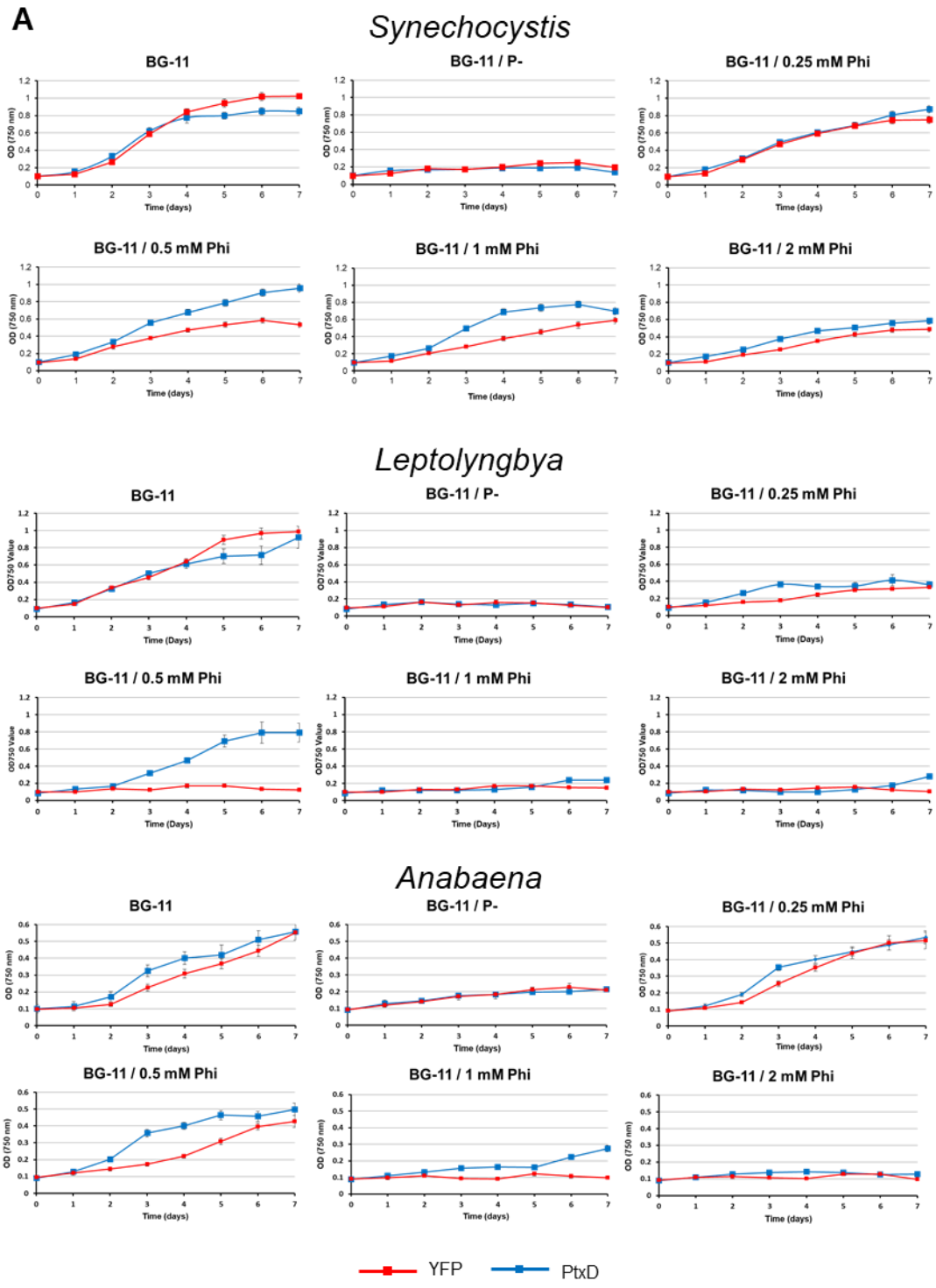
Currently, the development of outdoor pond cultivation is undermined by the risks of biological contaminants outcompeting cyanobacterial cultures. Previous literature has demonstrated that engineering green algae and cyanobacteria to metabolize phosphite established an effective strategy in controlling contamination. However, further research into expressing the *ptxD* gene in cyanobacteria has been negligible due to the lack of

available genetic tools. To enable the expression of the phosphite oxidoreductase PtxD across a wider range of cyanobacterial strains, we constructed the broad-host range RSF1010-based plasmid pAM5684, which contains the *ptxD* gene driven by the original conII promoter. pAM5684 harbors the *mobAY25F* mutation, the additional RK2*bom* site, and a spectinomycin/streptomycin *aadA* antibiotic resistance cassette used for selecting exconjugant colonies. We performed biparental mating using AM5501 conjugal *E. coli* cells to transfer pAM5684 into three cyanobacterial strains: *Synechocystis* sp. PCC 7803 (hereafter, *Synechocystis*), *Leptolyngbya* sp. BL0902 (hereafter, *Leptolyngbya*), and *Anabaena* sp. PCC 7120 (hereafter, *Anabaena*). We obtained spectinomycin/ streptomycin-resistant colonies from conjugation experiments for each strain. Likewise, we conjugated pAM5409 (RSF1010-*mobA*Y25F-RK2*bom-aadA-yfp*) into the three strains and obtained exconjugant colonies. Exconjugant strains harboring pAM5409 were used as controls for subsequent experiments.

We tested the engineered strains expressing either PtxD or YFP for their ability to grow in media containing only Phi as the sole P source. We first cultured the strains in BG-11 liquid media containing phosphate (Pi) for 5 days. To deplete cellular reserves of Pi, we replaced the media of all the cultures with liquid BG-11 lacking any P source (P-) and grew the cultures for another 3 days. Strains were then grown in multi-well plates containing three types of media: (1) BG-11 with Pi as the P source, (2) BG-11 with no P source (P-), and (3) BG-11 with Phi as the sole P source (at concentrations ranging from 0.25 to 2 mM). After 7 days of cultivation, we observed that the cloned strains harboring *ptxD* could grow in media containing Phi as the sole P source (Figure 3.1). Most of the *ptxD* clones grown in media supplemented with 0.25 and 0.5 mM Phi displayed robust growth rates similar to those of cultures growing with Pi. The main exception was the cultures of *Leptolyngbya* in 0.25 mM Phi, which had diminished growth curves compared to counterparts in Pi (Figure 3.1A).

Surprisingly, we found that some YFP clones displayed growth rates matching those of their ptxD counterparts across all three strains, especially in media supplemented with the lowest concentration of Phi (Figure 3.1A). Cultures of the *Synechocystis*-YFP clone appeared to grow steadily throughout the experiment despite the media containing only Phi as the P source. However, all the cultures containing YFP clones grown in Phi physically appeared greenish-yellow in color after 7 days of cultivation, indicating stress (Figure 3.1B). In contrast, cultures of ptxD clones grown in Phi were greener. Interestingly, we observed some cultures of ptxD clones grown in higher Phi concentrations that appeared colorless during the first 5-6 days of cultivation and began growing up afterward. For example, cultures of *Leptolyngbya*-ptxD in 2 mM Phi was colorless for 6 days before appearing light green on day 7 (Figure 3.1B).

Figure 3.1. Growth of engineered strains in media containing phosphite as the sole phosphorus source. Each strain (*Synechocystis*, *Leptolyngbya*, and *Anabaena*) harbored either pAM5409 (RSF1010-aadA-PconII-YFP) or pAM5684 (RSF1010-aadA-PconII-ptxD). Strains were grown in three types of media: (1) BG-11 with normal Pi concentrations, (2) BG-11 with no P source (P-), and (3) BG-11 with Phi as the sole P source (at concentrations ranging from 0.25 to 2 mM). (A) Growth curves of cultures grown for 7 days. Measurements for optical density at 750 nm (OD750 nm) were taken each day. Error bars represent standard deviations from biological triplicates. (B) Photographs of cyanobacterial 6-well cultures for each strain after 7 days of cultivation. First column, *Synechocystis*; middle column, *Leptolyngbya*; right column, *Anabaena*. Upper row: strains harboring the *ptxD* gene; lower row, strains harboring the *yfp* gene.


A
Synechocystis
BG-11
BG-11 / P-
BG-11 / 0.25 mM Phi
BG-11 / 0.5 mM Phi
BG-11 / 1 mM Phi
BG-11 / 2 mM Phi
OD (750 nm)
Time (days)
Leptolyngbya
BG-11
BG-11 / P-
BG-11 / 0.25 mM Phi
BG-11 / 0.5 mM Phi
BG-11 / 1 mM Phi
BG-11 / 2 mM Phi
OD750 Value
Time (Days)
Anabaena
BG-11
BG-11 / P-
BG-11 / 0.25 mM Phi
BG-11 / 0.5 mM Phi
BG-11 / 1 mM Phi
BG-11 / 2 mM Phi
OD (750 nm)
Time (days)
YFP
PtxD

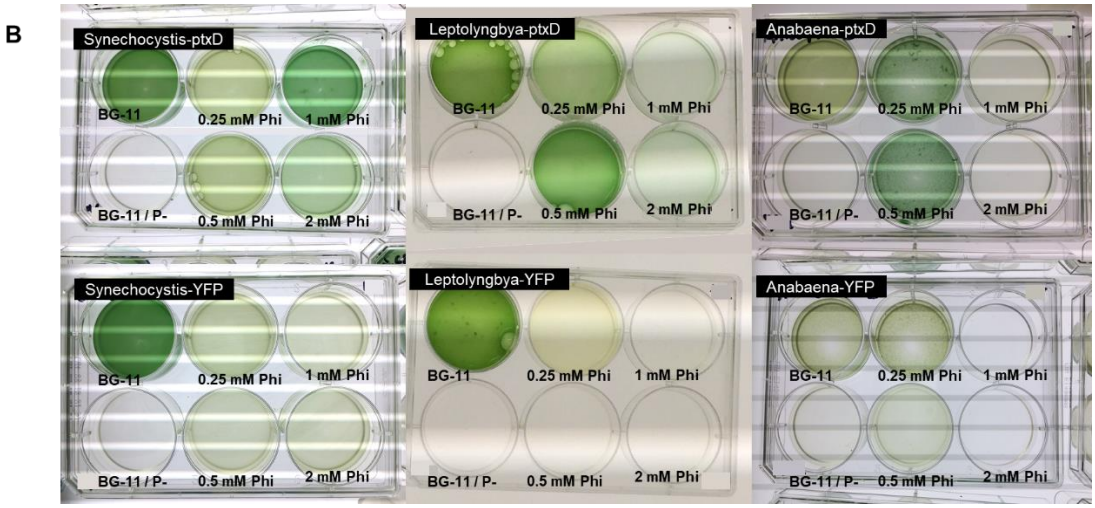


Figure 3.1 (continued). Growth of engineered strains in media containing phosphite as the sole phosphorus source.

3.3 Discussion

Biological contamination has long been a major challenge of open-culture raceway ponds for cyanobacteria cultivation. Ponds are easier and less expensive to maintain than sterile bioreactors but are often invaded by undesired and predatory microbes, such as other algae and bacteria (Mata *et al.*, 2010). One promising solution is to replace the necessary nutrient phosphate (Pi) with its reduced, nonmetabolizable form phosphite (Phi) in the growth media. Although phosphorus (P) is necessary for all organisms, most can solely take up P in the form of Pi. Only organisms that possess metabolic pathways for Phi oxidation can convert Phi to the usable Pi form and grow in the altered media. Previous literature demonstrated that *Synechococcus elongatus* PCC 7942 engineered to express the phosphite oxidoreductase PtxD could metabolize Phi and outcompete contaminating microorganisms not able to take up Phi as a P source (González-Morales *et al.*, 2020). However, the study's experimental approach is strain-specific to *S. elongatus*, which hinders the ability to test PtxD expression in other cyanobacterial strains.

To allow for expanded availability of genetic tools to test expression of PtxD across different strains of cyanobacteria, we constructed a RSF1010-based plasmid harboring the ptxD gene derived from *P. stutzeri*. We performed conjugal mating of the assembled plasmid (pAM5684) into three laboratory strains and obtained engineered cells that could metabolize Phi as the sole P source. In most cases, cultures of the PtxD clones significantly outgrew those of the YFP clones (Figure 3.1). Surprisingly, few cultures of the YFP clones displayed similar growth rates as PtxD equivalents, leading us to suspect that the strains may harbor endogenous genes for a phosphite oxidoreductase that allows them to metabolize Phi. We tested the growth of wild-type strains in agar media containing Phi as the only P source. Both *Synechocystis* and *Leptolyngbya* failed to grow in Phi media. However, *Anabaena* did grow, albeit not as well as when grown in Pi media. Previous literature has identified a genetic homolog of the *ptxD* gene in *Anabaena*, confirming our suspicions that the strain harbors an innate metabolic pathway for Phi oxidation (Martinez *et al.*, 2012). However, the physical appearances of the YFP clone cultures signify that cells are still under stress compared to PtxD clones (Figure 3.1B). The addition of the heterologous *ptxD* gene improved culture growth in *Anabaena* grown in Phi.

Although the expression of PtxD alone appears to promote Phi metabolization in engineered cyanobacteria, there have been no studies investigating the phosphite transporter protein. The ABC-type transporter, encoded by *ptxA*, *ptxB*, and *ptxC*, allows *P. stutzeri* to take up Phi molecules from the environment (Metcalf & Wolfe, 1998). We hypothesize that cloning the transporter protein gene cluster (*ptxABC*) with *ptxD* in cyanobacteria will allow greater uptake of Phi from the media and support faster cellular growth. Further investigations aiming to further test PtxD expression in cyanobacteria should also consider the transporter protein.

Here, we demonstrated that the RSF1010-based plasmid can be used to express the phosphite oxidoreductase enzyme PtxD across different strains of cyanobacteria, especially potential production strains. Production strains, such as *Leptolyngbya*, will benefit from this strategy as they are developed for outdoor cultivation in nonsterile conditions. With continuing discoveries for new production strains, our broad host range plasmid can be used to test expression of PtxD in emerging cyanobacteria of interest. Moreover, we demonstrated that the RSF1010-based cloning system has wider applications than solely biofuel or biomolecule production. Genetic engineering can be used to design strains with advantageous metabolic capabilities for cultivation.

CONCLUSION

Currently, the world is experiencing major challenges to develop new energy resources and combat environmental pollution. Cyanobacteria offer many opportunities to develop sustainable and renewable production of biofuel compounds, biomass, and nutritional products using simple inputs of carbon dioxide, water, and sunlight. Cyanobacteria have fast growth rates, are tolerant to saline and alkaline conditions, and require less land to cultivate than eukaryotic plants. In the past few decades, advances in cyanobacterial biotechnology have facilitated the development of gene regulatory tools to engineer cyanobacteria to produce biochemicals. However, these genetic tools are often developed in *Escherichia coli* or limited to certain strains of cyanobacteria, producing tools that tend to behave inconsistently across different strains.

Our projects focused on demonstrating potential applications for the broad-host range plasmid RSF1010 and the CYANO-VECTOR cloning system, which allows for genetic engineering to be performed in diverse strains of cyanobacteria. My first project produced a set of constitutive promoters with a wide range of expression strengths that can be used across multiple strains. The promoter library, which is constructed in RSF1010-based plasmids, can be used to study the expression of biofuel production pathways in cyanobacteria with engineered transcriptional control. My second project demonstrated a useful application for using the broad-host-range plasmids for bioprospecting. We isolated a new, genetically tractable cyanobacterial strain with high salt tolerance and showed that this bioprospecting technique can be efficient for discovering potentially new production strains. Finally, my third project focused on expanding a promising technique to reduce the risks of contamination for open-culture cultivation of cyanobacteria. We demonstrated that a RSF1010-based plasmid harboring the *ptxD* gene for the phosphite oxidoreductase can be used to engineer diverse strains to metabolize phosphite. When grown in conditions selective for phosphite-metabolizing

organisms, engineered cyanobacteria will have a competitive advantage over biological contaminants.

The work presented here describes the several applications and possibilities for RSF1010-based plasmids in cyanobacteria biotechnology. We believe that these genetic tools are valuable additions to the continually expanding engineering toolkit available to cyanobacterial researchers to help facilitate further investigations and development of cyanobacteria as cellular factories.

MATERIALS AND METHODS

Strain and Growth Conditions

All strains and plasmids used in these studies are listed in Table S1. Cyanobacterial strains were grown in BG-11 at 30°C under continuous medium light (50-100 µmol photons m^{-2} s^{-1}), unless otherwise noted. Light measurements were made with a QSL2100 PAR Scalar Irradiance sensor (Biospherical Instruments, San Diego, CA). Liquid cultures grown in shaking flasks and agar plates were supplemented with appropriate antibiotics. *E. coli* strains were grown in Lennox broth (LB) liquid medium or as LB agar plates at 37°C with appropriate antibiotics.

Section I

Plasmid construction

Plasmid pAM4933 is a shuttle vector carrying the *Synechococcus elongatus* PCC 7942 (hereafter, *Synechococcus*) neutral site 1 (NS1) for chromosomal integration by homologous recombination. pAM4933 has a tetracycline resistance marker, an *aadA* cassette for spectinomycin/streptomycin resistance, and a PconII promoter driving expression of yellow fluorescent protein (YFP). The sequence for the *yfp* gene harbors an optimized ribosomal binding sequence (oRBS). Construction of pAM4933 is composed of CYANO-VECTOR devices, including the chromosomal integration site NS1 for *Synechococcus* elongatus PCC 7942 from pCVD020, the *aadA* cassette for spectinomycin and streptomycin resistance from pCVD002, and the promoter-reporter PconII-oRBS-*yfp* construct from pCVD044 (Taton et al., 2014).

Plasmid pAM5612 is a broad-host range plasmid that contains RSF1010 mobilization (*mob*) and replication (*rep*) genes. The *mobA* gene contains a Y25F mutation that decreases plasmid mobilization and improves yields during plasmid preparations (Taton *et al.*, 2014). pAM5612 also contains an *aphI* cassette for

kanamycin/neomycin resistance, and a PconII promoter driving expression of yellow fluorescent protein (YFP). The sequence for the *yfp* gene harbors an optimized ribosomal binding sequence (oRBS). To construct pAM5612, the PconII-*yfp* sequence was amplified from pAM4933 using PCR. The *aphI* cassette was digested from pCVD003 using restriction enzymes (New England BioLabs). The promoter-reporter gene construct and *aphI* cassette were assembled with a RSF1010*mobA*Y25F backbone.

For plasmid construction, QIAprep Spin Miniprep Kit (Qiagen) was used for plasmid preparations. DNA purification and concentration were performed using DNA Clean and Concentrator^{TM}-5 (Zymo Research) and subsequent DNA concentrations were measured using UV-Vis spectrophotometer NanoDrop 2000c. Assembly of shuttle plasmids were carried out using a Quick Ligation^{TM} Kit (New England BioLabs) or GeneArt Seamless Cloning and Assembly Kit (Life Technologies) following manufacturer's instructions.

Design of Promoters

The conII promoter is a constitutive, synthetic promoter combining the -35 and -10 elements of the native trp and lac promoters, respectively, from *E. coli* (Elledge & Davis, 1989). To generate variants of the PconII, two strategies were used to target specific positions in the promoter. The first strategy consists of targeting the three most conserved nucleotides of the -10 element and the two nucleotides of the extended -10 region. The second strategy consists of targeting the three least conserved nucleotides of the -10 element and the two nucleotides of the extended -10 region. Degree of nucleotide conservation is based on consensus of the -10 element sequence from cyanobacterial native promoters found in *Synechococcus*, *Synechocystis* sp. PCC 6803 (hereafter, *Synechocystis*), and *Anabaena* sp. PCC 7120 (hereafter, *Anabaena*)

(Mitschke *et al.*, 2011a; Mitschke *et al.*, 2011b; Vijayan *et al.*, 2011). Error-prone, non-specific primers were used to amplify the PconII region, targeting nucleotides that were either randomly replaced by different nucleotides or remained the same. Throughout all experiments, the positive control consists of the unmutated, original PconII sequence and the negative control lacks the PconII promoter upstream of the *yfp* gene.

Characterization, grouping, and selection of PconII promoters

Syenechococcus: Preliminary screening of the PconII library was conducted in *Synechococcus* using pAM4933 to integrate the PconII variants and *yfp* reporter into the cyanobacterial genome. Fluorescence measurements indicated that variants developed using Strategy #1 produced YFP expression levels skewed extremely high and low. The set of variants developed using Strategy #2 produced a more even distribution of expression levels, which we used to continue for our study.

To organize and condense the number of variants, we set up a grouping system based on the sample's YFP fluorescence to chlorophyll fluorescence ratios (YFP/CHL) normalized to a $1x10^5$ a.u. maximum. All variants measuring below twice the background level (negative control YFP/CHL value) were removed, whereas all variants measuring above 7x the background level were retained entirely. All remaining variants were issued into 15 groups based on similar expression levels. Within the $1x10^4$ to $1x10^5$ a.u. range, 10 groups of $1x10^4$ a.u.-increments were formed. For example, the first group consisted of variants with expression levels above $1x10^5$ a.u., and the second group consisted of those with levels between 9x104 and $1x10^5$ a.u. For all variants measuring below $1x10^4$ a.u., 5 groups of $2x10^3$ increments were formed. To condense the library and retain diversity in expression levels, all members were pooled into their respective groups and randomly selected at a frequency of 1 to 3.5 variants for characterization in future strains.

Groups are labeled numerically, starting from ‘Se01’ for the group of the strongest variants, and ending with ‘Se15’ for the group of the weakest.

Synechocystis: To characterize the PconII variants in *Synechocystis*, we used the broad-host range plasmid pAM5612. We pooled samples of the PconII library in *Synechococcus* into their respective groups and performed colony PCRs to amplify the PconII-*yfp* constructs. We assembled the constructs into the RSF1010-*mobA*Y25F backbone with *aphI* antibiotic cassette to form pAM5612 and performed biparental conjugation to transfer the plasmids into *Synechocystis*.

To select variants for subsequent strain characterization, all variants measuring above 10x the background levels were retained, whereas those measuring below 10x the background were subject to the same grouping system employed in *Synechococcus*. Groups 2-4 were retained without discrimination and the remaining groups underwent selection at a frequency of 1 in 5 variants. However, unlike the selection process in *Synechococcus*, the system in *Synechocystis* considers the groups already formed from the previous strain characterization. Groups formed during characterization in the latter strain are classified as subgroups, such that their performance in the current strain can be tracked with information about their performance in previously characterized strains. As a result, 24 subgroups are formed and labeled numerically from ‘S01’ to ‘S24’, attached to the associated group labels from *Synechococcus*.

Leptolyngbya: To continue characterization in the next strain, we extracted the pAM5612 plasmids from *Synechocystis* and transferred them into *Leptolyngbya* as their corresponding *Synechocystis* subgroups using biparental conjugation. Using the similar subgrouping strategy employed in the characterization of previous strain, all variants measuring above 10x the background levels were retained whereas those measuring below 10x the background were grouped accordingly into another 26 subgroups. We

selected these grouped variants at a frequency of 1 in 5 variants. Subgroups are labeled numerically from ‘L01’ to ‘L26’.

Anabaena: Continuing our characterization of the PconII variants in the final strain, we extracted the pAM5612 plasmids from *Leptolyngbya* and transferred them into *Anabaena* as their corresponding *Leptolyngbya* subgroups using biparental conjugation. Since *Anabaena* was the final cyanobacterial strain used to characterize the PconII variants, we chose not to form additional subgroups. Instead, we selected 40 individual variants from the following criteria: one variant demonstrating the highest expression level (YFP/CHL ratio) from each of the 26 subgroups formed from previous characterization in *Leptolyngbya* and retained in *Anabaena*, and another 14 variants demonstrating moderate expression in their respective groups. Additionally, we ensured that at least two variants from each of the original 14 main groups created in Synechoccocus was included to obtain a wide distribution of expression levels across all characterized strains.

Transformation into cyanobacteria

Plasmid pAM4933 (NS1TC-aadA-PconII-yfp) was introduced into an actively dividing culture of *Synechococcus* using natural transformation, based on established protocols (Golden *et al.*, 1987). The aadA antibiotic resistance cassette and PconII-*yfp* promoter-reporter construct was integrated into the *Synechococcus* genome at the neutral site 1 (NS1) by homologous recombination. Transformed cells were plated onto BG-11 agar plates supplemented with kanamycin and incubated at 30°C in high light (100-150 µmol photons m^{-2} s^{-1}).

Bi-parental conjugal mating

We performed biparental conjugation using *E. coli* conjugal donor strains to transfer pAM5612 (RSF1010-mobAY25F-aphI-PconII-*yfp*), pAM5612 lacking the *yfp* gene sequence, and all 25 variants of pAM5612 harboring clones of the PconII library into cyanobacterial recipient strains. Day 0: Electrocompetent DH10B *E. coli* cells carrying the conjugal plasmid pRL443 and helper plasmid pRL623 (Elhai, *et al.*, 1997) were electroporated with 10 ng of plasmid DNA and incubated in SOC medium at 37°C for 1 hour. Cells were then plated on LB plates with chloramphenicol, ampicillin, and kanamycin. The plates were incubated at 37°C overnight. Day 1: *E. coli* colonies were selected and grown in LB liquid media with kanamycin at 37°C overnight. In preparation for conjugation, filamentous cyanobacteria strains (*Leptolyngbya* and *Anabaena*) were fragmented to shorter filaments using probe sonication (20% amplitude for 50 sec., 5 sec. on/off). Cultures were pelleted by centrifugation and allowed to recover in fresh liquid BG-11 overnight. Day 2: 1.8 mL of each *E. coli* culture were sedimented, washed twice with LB to remove antibiotics, and resuspended in 200 µL LB. 30 mL of actively dividing cyanobacterial cultures were pelleted by centrifugation, washed once with BG-11, and resuspended in 6 mL of BG-11. 40 µL of each *E. coli* culture was mixed with 200 µL of each cyanobacterial culture. Each mixture of *E. coli* and cyanobacterial cells was then pelleted by centrifugation at 4000 g for 5 minutes at room temperature and resuspended in 200 µL BG-11. Mixtures were then plated onto Petri plates containing 40 mL BG-11+5% LB solidified with 1.5% Bacto agar. Plates were incubated at 30°C overnight in low light (10-30 µmol photons $m^{-2}s^{-1}$. Day 3: Plates were underlaid with the antibiotic kanamycin. Since *Anabaena* is naturally resistant to kanamycin, neomycin was instead used to underlay plates containing *Anabaena* exconjugants. All plates were transferred to medium light at 30°C. Day 4-6: Colonies were selected and streaked onto BG-11 agar plates supplemented with appropriate antibiotics (kanamycin or neomycin).

Fluorescence measurements

Liquid cultures were grown in triplicate for each strain. Cultures were adjusted to an optical density of 0.1 at 750 nm (OD750) at the start of the culture, grown for three days, and adjusted again prior to measurement. Using the Tecan Infinite® M200 plate reader (TECAN), optical densities and fluorescence intensities were measured from 200µL of culture in black-walled, clear-bottom 96-well plates (Greiner). Excitation and emission wavelengths were set to 490 and 535 nm, respectively, for YFP. Excitation and emission wavelengths were set to 425 and 680 nm, respectively, for chlorophyll a (Chla).

Microscopy

Micrographs were taken using a Zeiss Axioscope using a 40x objective lens. Images exhibiting cyanobacteria autofluorescence was acquired using the Texas Red filter (596-nm excitation and 615-nm emission). Images exhibiting yellow fluorescence protein (YFP) expression was acquired using the YFP filter (500/20-nm excitation, 535/30-nm emission). Micrographs were taken using a Retiga R3 camera controlled by Micromanager software (Edelstein *et al.*, 2014). Images were processed using ImageJ (National Institute of Health). Figures were assembled in Microsoft Office PowerPoint.

Section II

Strain Isolation from Environmental Samples

Bioprospected strains ML2A, ML2C1, ML2C2, and ML3B were derived from cultures sampled from Mono Lake, California. Samples were collected from the lakeshore, comprised of water and sediment. Biparental mating using mixed cultures was performed as described in Section I above, but with the following changes: 200 µL *E. coli* (instead of 40 µL) and 1 mL cyanobacterial mixed culture (instead of 200 µL) were combined. Environmental samples were sonicated prior to biparental mating to break up

cell clumps. 100 µL of the conjugation mix were plated on BG-11+5% LB+0.5 M $NaHCO_3$ plates solidified with 1% Gelrite (gellan gum). Plates were incubated at low light for 24 hours. Potentially exconjugant cyanobacteria were streaked on BG-11 plates with spectinomycin and streptomycin to isolate antibiotic-resistant colonies. To obtain an axenic culture and sustain growth of exconjugant cyanobacteria, colonies were re-streaked alternately between BG-11+0.5 M $NaHCO_3$ and BG-11+spectinomycin/ streptomycin to ensure that high carbonate levels and antibiotic selection are maintained.

Axenic cultures of each cyanobacterial strain were grown up in 50 mL of BG-11 in 125 mL flasks under antibiotic selection. Cultures were grown at 30°C in medium light. Colony PCRs were performed to amplify the 16S rRNA gene from each strain using a pair of cyanobacteria-specific primers: 16S27F and 23S30R (Taton *et al.*, 2003). Sequencing of the 16S rRNA region was performed by GENEWIZ (South Plainfield, NJ) using primers 16S378F and 16S784R (Taton *et al.*, 2003).

To cure strain ML3B of RSF1010-based plasmid pAM5409, a culture of engineered ML3B was grown in BG-11+0.5 M $NaHCO_3$ at 30°C in medium light without antibiotics. Over the course of a month, cultures were passaged regularly to new media when they reached high density (OD750 > 0.7). After 30 days, the culture was plated on BG-11 agar media supplemented with spectinomycin/ streptomycin, which failed to grow after 7 days at 30°C. Plasmid loss was also confirmed by fluorescence microscopy, which the cured strain displayed no detectable YFP expression.

Fluorescence measurements

Liquid cultures were grown in triplicate for each strain. Cultures were adjusted to an optical density of 0.1 at 750 nm (OD750) at the start of the culture, grown for three days, and adjusted again prior to measurement. Using the Tecan Infinite® M200 plate reader (TECAN), optical densities and fluorescence intensities were measured from

200µL of culture in black-walled, clear-bottom 96-well plates (Greiner). Excitation and emission wavelengths were set to 490 and 535 nm, respectively, for YFP. Excitation and emission wavelengths were set to 425 and 680 nm, respectively, for chlorophyll a (Chla).

Microscopy

Micrographs were taken using a Delta Vision (Applied Precision, Inc.) Olympus IX71 inverted microscope using 100x oil immersion objective. Images exhibiting cyanobacteria autofluorescence was acquired using the TRITC filter (555/28-nm excitation and 617/73-nm emission). Images exhibiting yellow fluorescence protein (YFP) expression was acquired using the YFP filter (500/20-nm excitation, 535/30-nm emission). Micrographs were taken using Resolve3D softWoRx-Acquire v. 4.0.0 and processed using ImageJ (National Institute of Health). Figures were assembled in Microsoft Office PowerPoint.

Strain Cultivation in High Salt

Five cyanobacterial strains were tested: *Synechococcus elongatus* PCC 7942, *Synechocystis* sp. PCC 6803, *Anabaena* sp. PCC 7120, *Leptolyngbya* sp. BL0902, and Cf. Nodosilinea sp. ML3B. All cultures were grown in regular BG-11 liquid medium for 4-5 days prior to the experiment. BG-11 buffered with 5mM MOPS was prepared with varying concentrations of supplemented sodium bicarbonate ($NaHCO_3$): 0M (normal salt concentrations), 0.1 M, 0.25 M, 0.5 M, 1 M, and 2 M. Each strain was added to 2-mL cultures of BG-11 at each concentration of $NaHCO_3$, adjusted to an OD750 of 0.1. All culture assays were performed with biological triplicates using clear 6-well plates (Greiner). Cultures were grown up for 5 days, shaking at 30°C under continuous light (60-80 µmol photons $m^{-2} s^{-1}$). OD750 readings for cell density were measured daily using the Tecan Infinite® M200 plate reader (TECAN).

Section III

Plasmid Construction

To construct pAM5684, pAM5409 was digested using SbfI and AflII restriction enzymes (New England BioLabs) to remove the *yfp* reporter gene immediately downstream of the conII promoter. The *ptxD* gene sequence was amplified using PCR from pWM239 (Metcalf & Wolfe, 1998); PCR fragments contained sticky ends with pAM5409 sequence homology. Assembly of pAM5684 was carried out using a GeneArt Seamless Cloning and Assembly Kit (Life Technologies) following manufacturer's instructions. QIAprep Spin Miniprep Kit (Qiagen) was used for plasmid preparations. DNA purification and concentration were performed using DNA Clean and Concentrator^{TM}-5 (Zymo Research) and subsequent DNA concentrations were measured using UV-Vis spectrophotometer NanoDrop 2000c.

Bi-parental conjugal mating

We performed biparental conjugation using *E. coli* conjugal donor strains to transfer pAM5684 (RSF1010-*mobA*Y25F-*aadA*-PconII-*ptxD*) and pAM5409 (RSF1010-*mobA*Y25F-*aadA*-PconII-yfp) into cyanobacterial recipient strains. Biparental conjugal mating was performed as described in Section I above. For antibiotic selection, spectinomycin/streptomycin was used instead of kanamycin or neomycin.

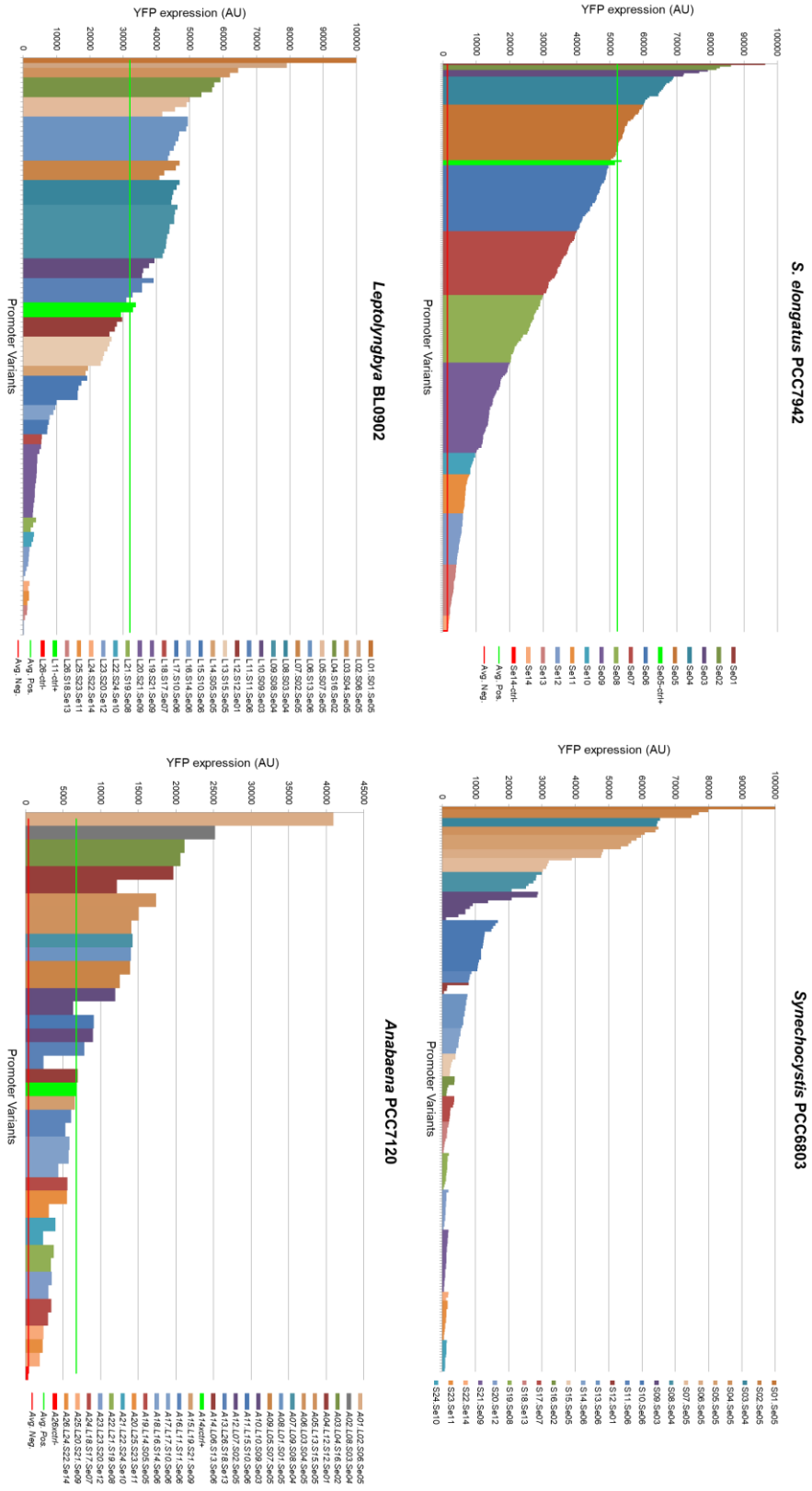
Strain Cultivation with Phosphite as the Sole Source of Phosphorus

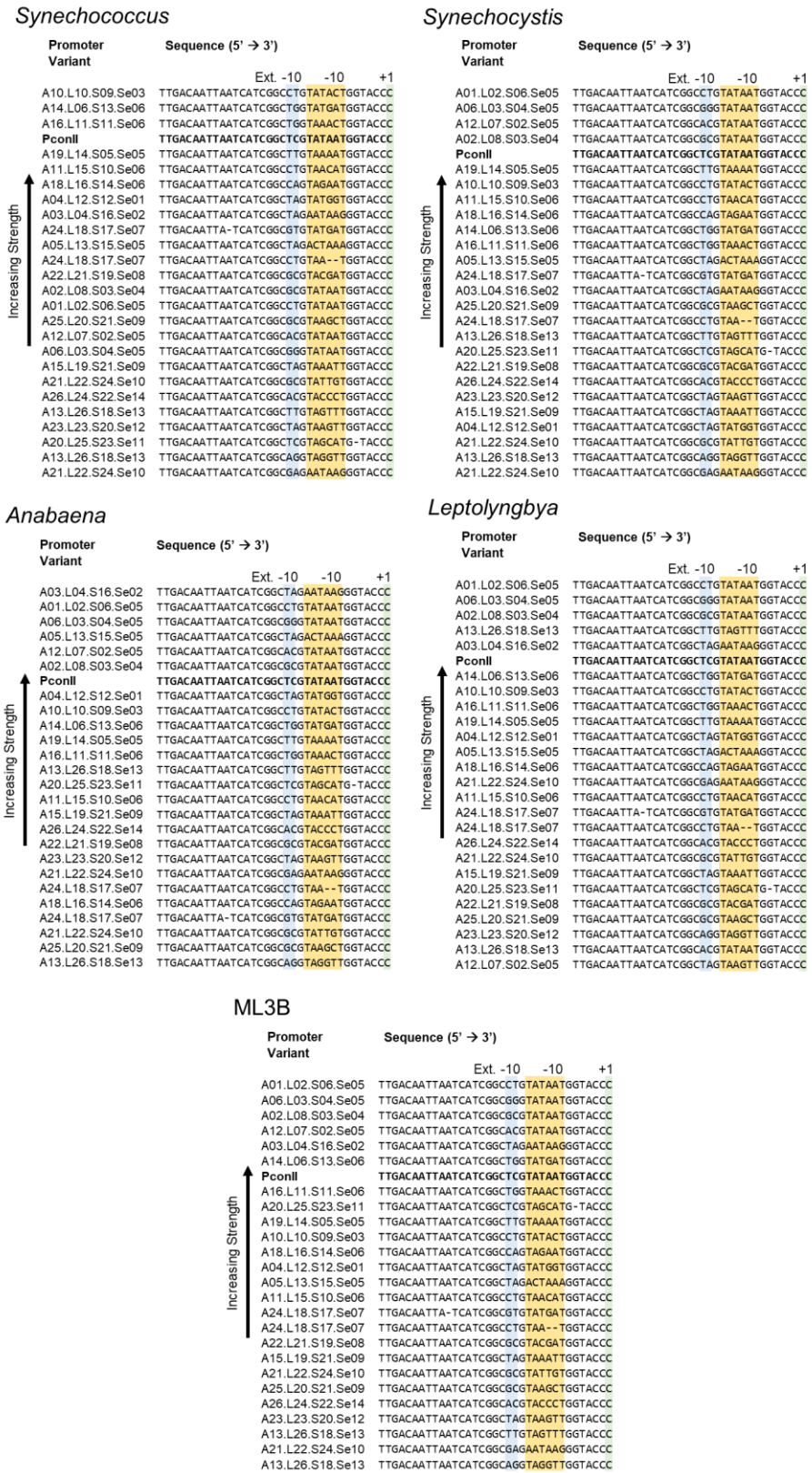
Engineered strains harboring either pAM5684 or pAM5409 were grown in regular BG-11 with spectinomycin/streptomycin for 5 days at 30°C under continuous light (60-80 µmol photons m^{-2} s^{-1}). To deplete cellular reserves of Pi, media from all cultures were replaced with liquid BG-11 lacking any P source (P-) and grown for another 3 days under antibiotic selection and similar growth conditions. Three different types of liquid BG-11

were then prepared: (1) regular BG-11 with 0.5 mM phosphate (Pi) as the positive control, (2) BG-11 with no P source (P-) as the negative control, and (3) BG-11 with Phi, at concentrations 0.25 mM, 0.5 mM, 1 mM, and 2 mM. Each strain was inoculated in each 2-mL cultures. Each strain was added to 2-mL cultures adjusted to an OD750 of 0.1. All culture assays were performed with biological triplicates using clear 6-well plates (Greiner). Cultures were grown up for 7 days, shaking at 30°C under continuous light (60-80 µmol photons $m^{-2} s^{-1}$). OD750 readings for cell density were measured daily using the Tecan Infinite® M200 plate reader (TECAN).

APPENDIX

Supplementary Figure 1. Preliminary characterization and grouping of PconII library clones in *Synechococcus* (*S. elongatus*), *Synechocystis*, *Leptolyngbya*, and *Anabaena*. Selected library clones were extracted from *Synechococcus* after preliminary characterization (Figure 1.2) and assembled into pAM5612. Plasmids were transferred into cyanobacterial strains by conjugation from donor *E. coli* strains. Isolated exconjugant colonies were grown in liquid culture for 3 days, adjusted to an optical density of 0.1 at 750 nm (OD750), and measured for yellow fluorescent protein (YFP). Expression levels were normalized to culture chlorophyll *a* measurements. After *Synchococcus*, a subset of library clones was characterized in *Synechocystis*, then *Leptolyngbya*, and finally *Anabaena*. Sub-groups were formed based on similar expression levels, with each bar color representing a different group. Controls: unaltered PconII, green horizontal line; no promoter control, red line. Control data are not available for the *Synechocystis* experiments.


S. elongatus PCC7942
YFP expression (AU)
Promoter Variants
Leptolyngbya BL0902
YFP expression (AU)
Promoter Variants
Synechocystis PCC6803
YFP expression (AU)
Promoter Variants
Anabaena PCC7120
YFP expression (AU)
Promoter Variants



Supplementary Figure 2. PconII library promoter sequences ordered by strength in each characterized strain.

Supplementary Table 1. List of strains and plasmids.

Strain	Description	Reference
***E. coli* strains**		
DH10B	Cloning strain	Gibco BRL
AM1359	Conjugal donor strain, Ap^r, Tc^r, Cm^r	Yoon & Golden, 1998
AM5501	DH10B strain harboring helper plasmid pAM5505 and conjugal plasmid pRL443, Ap^r, Tc^r, Cm^r	Bishé *et al.*, 2019
One Shot TOP10	Cloning host	Invitrogen
Cyanobacterial strains		
Synechococcus elongatus PCC 7942	Wild type	Laboratory collection
Synechococcus elongatus PCC 7942	Argonaute (AGO) knockout	Laboratory collection
Synechocystis sp. PCC 6803	Wild type	Laboratory collection
Anabaena sp. PCC 7120	Wild type	Laboratory collection
Leptolyngbya BL0902	Wild type	Taton *et al.*, 2012
ML2-A	Cyanobacterial strain isolated from Mono Lake, CA	This study
ML2-C1	Cyanobacterial strain isolated from Mono Lake, CA	This study
ML2-C2	Cyanobacterial strain isolated from Mono Lake, CA	This study
ML3-B	Cyanobacterial strain isolated from Mono Lake, CA	This study
Plasmids		
pCVD002	CYANO-VECTOR plasmid containing *aadA* (Sp/Sm resistance), Sp^r, Sm^r	Taton et al., 2014
pCVD003	CYANO-VECTOR plasmid containing *aphI* (Km/Nm resistance), Km^r, Nm^r	Taton et al., 2014
pCVD020	CYANO-VECTOR plasmid containing chromosomal integration site NS1 for *Synechococcus elongatus* PCC 7942, Ap^r, Tc^r	Taton et al., 2014
pCVD044	CYANO-VECTOR plasmid containing promoter-reporter device (PconII driving *yfp*), Ap^r	Taton et al., 2014
pAM4933	NS1TC-*aadA*-PconII-oRBS-*yfp* with Sp^r, Sm^r,	This study
pAM5404	RSF1010-mobAY25F-RK2*bom*, Ap^r	Bishé *et al.*, 2019
pAM5409	RSF1010-*mob*AY25F-RK2*bom*-*aadA*-PconII-oRBS-*yfp*, Sp^r, Sm^r (assembled from	Bishé *et al.*, 2019
pAM5612	RSF1010-*mobA*Y25F-*aphI*-Pconii-oRBS-*yfp*, Km^r, Nm^r	This study
pAM5684	RSF1010-m*obA*Y25F-RK2*bom*-*aadA*-PconII-*ptxD*, Sp^r, Sm^r	This study
pWM239	Source plasmid containing gene cluster *ptxABCD*	Metcalf & Wolfe, 1998
pRL443	Conjugal plasmid, derivative of RP4, Ap^r	Elhai *et al.*, 1997
pRL623	Helper plasmid carrying Mob, ColK and methylase genes M.AvaI, M.Eco47II, M.EcoT22I, Cm^r	Elhai *et al.*, 1997

Supplementary Table 2. List of primers

Primer Name	Primer Sequence	Reference
conII-n5-F_b	CTTGACAATTAATCATCGGCNNGTANNNTGGTACCC ATTT	This work
conII-n5-R_b	AAATGGGTACCANNNTACNNGCCGATGATTAATTGTC AAGACGTC	This work
aphi_seq_FW	ATGGCAAGATCCTGGTATCG	This work
aphi_seq_RV	TTATGCCTCTTCCGACCATC	This work
aadA_seq_FWD	TACTGCGCTGTACCAAATGC	This work
aadA_seq_REV	TGATTTGCTGGTTACGGTGA	This work
YFP_RSF_sq_FW	CCTGAAGTTCATCTGCACCA	This work
YFP_RSF_sq_RV	GAACTCCAGCAGGACCATGT	This work
pconii_seq_fw	GATGGTCGGAAGAGGCATAA	This work
pconii_seq_rv	GAACTTCAGGGTCAGCTTGC	This work
16S27F	AGAGTTTGATCCTGGCTCAG	Taton et al., 2003
23S30R	CTTCGCCTCTGTGTGCCTAGGT	Taton et al., 2003
16S378F	GGGGAATTTTCCGCAATGGG	Taton et al., 2003
16S784R	GGA CTA CWG GGG TAT CTA ATC CC	Taton et al., 2003
RSF_hom_ptxD_F	GTACCCTATTTAAATAAAGGAGGTCatgctgccgaaactcgtt at	This work
RSF_hom_ptxD_R	GCCATCCGTCAGGATGGCCTTCTCCtgctcttcaagcttctgc aa	This work
ptxD_Seqprimer_R	GCGAAGAGTTCGCTGCAC	This work
ptxD_Seqprimer_F	CTGGCAACCACAGTTCTACG	This work

REFERENCES

Andreote, A. P. D., Vaz, M. G. M. V., Genuário, D. B., Barbiero, L., Rezende-Filho, A. T., & Fiore, M. F. (2014). Nonheterocytous cyanobacteria from Brazilian saline-alkaline lakes. *Journal of Phycology*, *50*(4), 675–684. https://doi.org/10.1111/jpy.12192

Babic, A., Guérout, A.-M., & Mazel, D. (2008). Construction of an improved RP4 (RK2)-based conjugative system. *Research in Microbiology*, *159*(7), 545–549. https://doi.org/https://doi.org/10.1016/j.resmic.2008.06.004

Bishé, B., Taton, A., & Golden, J. W. (2019). Modification of RSF1010-Based Broad-Host-Range Plasmids for Improved Conjugation and Cyanobacterial Bioprospecting. *IScience*, *20*, 216–228. https://doi.org/10.1016/j.isci.2019.09.002

Budinoff, C. R., & Hollibaugh, J. T. (2007). Ecophysiology of a Mono Lake picocyanobacterium. *Limnology and Oceanography*, *52*(6), 2484–2496. https://doi.org/10.4319/lo.2007.52.6.2484

Clerico, E. M., Ditty, J. L., & Golden, S. S. (2007). Specialized Techniques for Site-Directed Mutagenesis in Cyanobacteria. In *Circadian Rhythms* (pp. 155–177). https://doi.org/10.1007/978-1-59745-257-1_27

Crooks, G., Hon, G., Chandonia, J., & Brenner, S. (2004). NCBI GenBank FTP Site. WebLogo: a sequence logo generator. *Genome Res*, *14*, 1188–1190. https://doi.org/10.1101/gr.849004.1

De Mey, M., Maertens, J., Lequeux, G. J., Soetaert, W. K., & Vandamme, E. J. (2007). Construction and model-based analysis of a promoter library for E. coli: An indispensable tool for metabolic engineering. *BMC Biotechnology*, *7*, 1–14. https://doi.org/10.1186/1472-6750-7-34

Dismukes, G. C., Carrieri, D., Bennette, N., Ananyev, G. M., & Posewitz, M. C. (2008). Aquatic phototrophs: efficient alternatives to land-based crops for biofuels. *Current Opinion in Biotechnology*, *19*(3), 235–240. https://doi.org/10.1016/j.copbio.2008.05.007

Edelstein, A. D., Tsuchida, M. A., Amodaj, N., Pinkard, H., Vale, R. D., & Stuurman, N. (2014). Advanced methods of microscope control using μManager software. *Journal of Biological Methods*, *1*(2), 10. https://doi.org/10.14440/jbm.2014.36

Elhai, J., Vepritskiy, A., Muro-Pastor, A. M., Flores, E., & Wolk, C. P. (1997). Reduction of conjugal transfer efficiency by three restriction activities of Anabaena sp. strain PCC 7120. *Journal of Bacteriology*, *179*(6), 1998–2005. https://doi.org/10.1128/jb.179.6.1998-2005.1997

Elledge, S. J., & Davis, R. W. (1989). Position and density effects on repression by stationary and mobile DNA-binding proteins. *Genes & Development*, *3*(2), 185–197. https://doi.org/10.1101/gad.3.2.185

Flombaum, P., Gallegos, J. L., Gordillo, R. A., Rincón, J., Zabala, L. L., Jiao, N., … Martiny, A. C. (2013). Present and future global distributions of the marine Cyanobacteria Prochlorococcus and Synechococcus. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(24), 9824–9829. https://doi.org/10.1073/pnas.1307701110

Frey, J., Bagdasarian, M. M., & Bagdasarian, M. (1992). Replication and copy number control of the broad-host-range plasmid RSF1010. *Gene*, *113*(1), 101–106. https://doi.org/10.1016/0378-1119(92)90675-F

Furmaniak, M. A., Misztak, A. E., Franczuk, M. D., Wilmotte, A., Waleron, M., & Waleron, K. F. (2017). Edible cyanobacterial genus Arthrospira: Actual state of the art in cultivation methods, genetics, and application in medicine. *Frontiers in Microbiology*, *8*(DEC), 1–21. https://doi.org/10.3389/fmicb.2017.02541

Garcia-Pichel, F. (2009). Cyanobacteria. In M. B. T.-E. of M. (Third E. Schaechter (Ed.), *Encyclopedia of Microbiology* (Third, pp. 107–124). https://doi.org/https://doi.org/10.1016/B978-012373944-5.00250-9

Golden, S. S., Brusslan, J., & Haselkorn, R. (1987). Genetic Engineering of the Cyanobacterial Chromosome. *Methods in Enzymology*, *153*(C), 215–231. https://doi.org/10.1016/0076-6879(87)53055-5

González-Morales, S. I., Pacheco-Gutiérrez, N. B., Ramirez-Rodríguez, C. A., Brito-Bello, A. A., Esrella-Hernández, P., Herrera-Estrella, L., & López-Arredondo, D. L. (2020). Biotechnology for Biofuels Metabolic engineering of phosphite metabolism in Synechococcus elongatus PCC 7942 as an effective measure to control biological contaminants in outdoor raceway ponds. *Biotechnology for Biofuels*, *13*(119), 1–19. https://doi.org/10.1186/s13068-020-01759-z

Gordon, G. C., & Pfleger, B. F. (2018). Synthetic Biology of Cyanobacteria. In W. Zhang & X. Song (Eds.), *Advances in Experimental Medicine and Biology* (Vol. 1080, pp. 281–315). https://doi.org/10.1007/978-981-13-0854-3_12

Grant, G. A. (1989). A New Family of 2-Hydroxyacid Dehydrogenases. *Biochemical and Biophysical Research Communications*, *165*(3), 1371–1374.

Hook-Barnard, I. G., & Hinton, D. M. (2007). Transcription Initiation by Mix and Match Elements: Flexibility for Polymerase Binding to Bacterial Promoters. *Gene Regulation and Systems Biology*, *1*, 275–293. Retrieved from papers2://publication/uuid/2AAF1F9B-FEFC-4DAB-9119-3880456EC171

Huang, H. H., Camsund, D., Lindblad, P., & Heidorn, T. (2010). Design and characterization of molecular tools for a synthetic biology approach towards developing cyanobacterial biotechnology. *Nucleic Acids Research*, *38*(8), 2577–2593. https://doi.org/10.1093/nar/gkq164

Huang, H. H., & Lindblad, P. (2013). Wide-dynamic-range promoters engineered for cyanobacteria. *Journal of Biological Engineering*, *7*(1), 1–11. https://doi.org/10.1186/1754-1611-7-10

Jensen, P. R., & Hammer, K. (1998). The sequence of spacers between the consensus sequences modulates the strength of prokaryotic promoters. *Applied and Environmental Microbiology*, *64*(1), 82–87. https://doi.org/10.1128/aem.64.1.82-87.1998

Kim, J. H., Choi, W., Jeon, S., Kim, T., Park, A., Kim, J., … Kang, D.-H. (2015). *Isolation and characterization of Leptolyngbya sp. KIOST-1, a basophilic and euryhaline filamentous cyanobacterium from an open paddle-wheel raceway Arthrospira culture pond in Korea*. *119*, 1597–1612. https://doi.org/10.1111/jam.12961

Lanka, E., & Wilkins, B. M. (1995). DNA Processing Reactions in Bacterial Conjugation. In *Annual Review of Biochemistry* (Vol. 64, pp. 141–169).

Lennen, R. M., & Pfleger, B. (2013). Microbial production of fatty acid-derived fuels and chemicals. *Current Opinion in Biotechnology*, *24*(6), 1–17. https://doi.org/10.1038/jid.2014.371

Li, R., & Golden, S. S. (1993). Enhancer activity of light-responsive regulatory elements in the untranslated leader regions of cyanobacterial psbA genes. *Proceedings of the National Academy of Sciences of the United States of America*, *90*(24), 11678–11682. https://doi.org/10.1073/pnas.90.24.11678

Liu, D., Mao, Z., Guo, J., Wei, L., Ma, H., Tang, Y., … Zhao, X. (2018). Construction, Model-Based Analysis, and Characterization of a Promoter Library for Fine-Tuned Gene Expression in Bacillus subtilis. *ACS Synthetic Biology*, *7*(7), 1785–1797. https://doi.org/10.1021/acssynbio.8b00115

Liu, Z. G., & Lu, G. L. (1990). The sterilizing studies of flagellate and ciliate in marine unicellular algae liquid. *Zhanjiang Aquacult. Coll*, *2*, 36–41.

Loera-Quezada, M. M., Leyva-González, M. A., Velázquez-Juárez, G., Sanchez-Calderón, L., Do Nascimento, M., López-Arredondo, D., & Herrera-Estrella, L. (2016). A novel genetic engineering platform for the effective management of biological contaminants for the production of microalgae. *Plant Biotechnology Journal*, *14*(10), 2066–2076. https://doi.org/10.1111/pbi.12564

Markley, A. L., Begemann, M. B., Clarke, R. E., Gordon, G. C., & Pfleger, B. F. (2015). Synthetic Biology Toolbox for Controlling Gene Expression in the Cyanobacterium Synechococcus sp. strain PCC 7002. *ACS Synthetic Biology*, *4*(5), 595–603. https://doi.org/10.1021/sb500260k

Markou, G., & Georgakakis, D. (2011). Cultivation of filamentous cyanobacteria (blue-green algae) in agro-industrial wastes and wastewaters: A review. *Applied Energy*, *88*(10), 3389–3401. https://doi.org/10.1016/j.apenergy.2010.12.042

Martinez, A., Osburne, M. S., Sharma, A. K., Delong, E. F., & Chisholm, S. W. (2012). Phosphite utilization by the marine picocyanobacterium Prochlorococcus MIT9301. *Environmental Microbiology*, *14*(6), 1363–1377. https://doi.org/10.1111/j.1462-2920.2011.02612.x

Mata, T. M., Martins, A. A., & Caetano, N. S. (2010). Microalgae for biodiesel production and other applications: A review. *Renewable and Sustainable Energy Reviews*, *14*(1), 217–232. https://doi.org/10.1016/j.rser.2009.07.020

Metcalf, W. W., & Wolfe, R. S. (1998). Molecular genetic analysis of phosphite and hypophosphite oxidation by Pseudomonas stutzeri WM88. *Journal of Bacteriology*, *180*(21), 5547–5558.

Michel, K., Pistorius, E. K., & Golden, S. S. (2001). *Unusual Regulatory Elements for Iron Deficiency Induction of the idiA Gene of Synechococcus elongatus PCC 7942*. *183*(17), 5015–5024. https://doi.org/10.1128/JB.183.17.5015

Mitschke, J., Georg, J., Scholz, I., Sharma, C. M., Dienst, D., Bantscheff, J., … Hess, W. R. (2011). An experimentally anchored map of transcriptional start sites in the model cyanobacterium Synechocystis sp. PCC6803. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(5), 2124–2129. https://doi.org/10.1073/pnas.1015154108

Mitschke, J., Vioque, A., Haas, F., Hess, W. R., & Muro-Pastor, A. M. (2011). Dynamics of transcriptional start site selection during nitrogen stress-induced cell differentiation in Anabaena sp. PCC7120. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(50), 20130–20135. https://doi.org/10.1073/pnas.1112724108

Montalvão, S., Demirel, Z., Devi, P., Lombardi, V., Hongisto, V., Perälä, M., … Tammela, P. (2016). Large-scale bioprospecting of cyanobacteria, micro- and macroalgae from the Aegean Sea. *New Biotechnology*, *33*(3), 399–406. https://doi.org/10.1016/j.nbt.2016.02.002

Monzingo, A. F., Ozburn, A., Xia, S., Meyer, R. J., & Robertus, J. D. (2007). The Structure of the Minimal Relaxase Domain of MobA at 2.1 Å Resolution. *Journal of Molecular Biology*, *366*(1), 165–178. https://doi.org/10.1016/j.jmb.2006.11.031

Nair, U., Thomas, C., & Golden, S. S. (2001). *Functional Elements of the Strong psbAI Promoter of Synechococcus elongatus PCC 7942*. *183*(5), 1740–1747. https://doi.org/10.1128/JB.183.5.1740

Niedermeyer, T. H. ors. J. (2015). Anti-infective Natural Products from Cyanobacteria. *Planta Medica*, *81*(15), 1309–1325. https://doi.org/10.1055/s-0035-1546055

Nozzi, N. E., Oliver, J. W. K., & Atsumi, S. (2013). *Cyanobacteria as a platform for biofuel production*. *1*(September), 1–6. https://doi.org/10.3389/fbioe.2013.00007

Poole, L. B., Parsonage, D., Sergeant, S., Miller, L. R., Lee, J., Furdui, C. M., & Chilton, F. H. (2020). Acyl-lipid desaturases and Vipp1 cooperate in cyanobacteria to produce novel omega-3 PUFA-containing glycolipids. *Biotechnology for Biofuels*, *13*(1), 1–15. https://doi.org/10.1186/s13068-020-01719-7

Pulz, O., & Gross, W. (2004). Valuable products from biotechnology of microalgae. *Applied Microbiology and Biotechnology*, *65*(6), 635–648. https://doi.org/10.1007/s00253-004-1647-x

Rawat, I., Ranjith Kumar, R., Mutanda, T., & Bux, F. (2011). Dual role of microalgae: Phycoremediation of domestic wastewater and biomass production for sustainable biofuels production. *Applied Energy*, *88*(10), 3411–3424. https://doi.org/10.1016/j.apenergy.2010.11.025

Rawlings, D. E., & Tietze, E. (2001). Comparative Biology of IncQ and IncQ-Like Plasmids. *Microbiology and Molecular Biology Reviews*, *65*(4), 481–496. https://doi.org/10.1128/mmbr.65.4.481-496.2001

Ruffing, A. M., Jensen, T. J., & Strickland, L. M. (2016). Genetic tools for advancement of Synechococcus sp. PCC 7002 as a cyanobacterial chassis. *Microbial Cell Factories*, *15*(1), 1–14. https://doi.org/10.1186/s12934-016-0584-6

Shultzaberger, R. K., Chen, Z., Lewis, K. A., & Schneider, T. D. (2007). Anatomy of Escherichia coli σ70 promoters. *Nucleic Acids Research*, *35*(3), 771–788. https://doi.org/10.1093/nar/gkl956

Smith, V. H., Sturm, B. S. M., deNoyelles, F. J., & Billings, S. A. (2010). The ecology of algal biodiesel production. *Trends in Ecology and Evolution*, *25*(5), 301–309. https://doi.org/10.1016/j.tree.2009.11.007

Somerville, C., Youngs, H., Taylor, C., Davis, S. C., & Long, S. P. (2010). Feedstocks for Lignocellulosic Biofuels. *Science*, *329*(5993), 790 LP – 792. https://doi.org/10.1126/science.1189268

Takahama, K., Matsuoka, M., Nagahama, K., & Ogawa, T. (2003). Construction and analysis of a recombinant cyanobacterium expressing a chromosomally inserted gene for an ethylene-forming enzyme at the psbAI locus. *Journal of Bioscience and Bioengineering*, *95*(3), 302–305. https://doi.org/https://doi.org/10.1016/S1389-1723(03)80034-8

Taton, A., Grubisic, S., Brambilla, E., Wit, R. De, & Wilmotte, A. (2003). Cyanobacterial Diversity in Natural and Artificial Microbial Mats of Lake Fryxell (McMurdo Dry Valleys, Antarctica): a Morphological and Molecular Approach. *Applied and Environmental Microbiology*, *69*(9), 5157–5169. https://doi.org/10.1128/AEM.69.9.5157

Taton, A., Lis, E., Adin, D. M., Dong, G., Cookson, S., Kay, S. A., … Golden, J. W. (2012). Gene transfer in Leptolyngbya sp. strain BL0902, a cyanobacterium suitable for production of biomass and bioproducts. *PLoS ONE*, *7*(1). https://doi.org/10.1371/journal.pone.0030901

Taton, A., Unglaub, F., Wright, N. E., Zeng, W. Y., Paz-Yepes, J., Brahamsha, B., … Golden, J. W. (2014). Broad-host-range vector system for synthetic biology and biotechnology in cyanobacteria. *Nucleic Acids Research*, *42*(17), e136–e136. https://doi.org/10.1093/nar/gku673

Vijayan, V., Jain, I. H., & O'Shea, E. K. (2011). A high resolution map of a cyanobacterial transcriptome. *Genome Biology*, *12*(5), R47. https://doi.org/10.1186/gb-2011-12-5-r47

Wang, H., Zhang, W., Chen, L., Wang, J., & Liu, T. (2013). The contamination and control of biological pollutants in mass cultivation of microalgae. *Bioresource Technology*, *128*, 745–750. https://doi.org/10.1016/j.biortech.2012.10.158

Wei, T., Ramasubramanian, T. S., Pu, F., & Golden, A. J. W. (1993). *Anabaena sp . Strain PCC 7120 bifA Gene Encoding a Sequence-Specific DNA-Binding Protein Cloned by In Vivo Transcriptional Interference Selection*. *175*(13), 4025–4035.

Whitton, B. A., & Potts, M. (2012). Introduction to the Cyanobacteria. In B. A. Whitton (Ed.), *Ecology of Cyanobacteria II: Their Diversity in Space and Time* (pp. 1–13). https://doi.org/10.1007/978-94-007-3855-3_1

Włodarczyk, A., Selão, T. T., Norling, B., & Nixon, P. J. (2020). Newly discovered Synechococcus sp. PCC 11901 is a robust cyanobacterial strain for high biomass production. *Communications Biology*, *3*(1). https://doi.org/10.1038/s42003-020-0910-8

Wu, J. F., & Pond, W. G. (1981). Amino Acid Composition and Microbial Contamination of Spirulina maxima, a Blue-Green Alga, Grown on the Effluent of Different Fermented Animal Wastes. *Bulletin of Environmental Contamination and Toxicology*, *27*, 151–159.

Yoon, H. S., & Golden, J. W. (1998). Heterocyst pattern formation controlled by a diffusible peptide. *Science (New York, N.Y.)*, *282*(5390), 935–938. https://doi.org/10.1126/science.282.5390.935

Zhang, X., Rong, J., Chen, H., He, C., & Wang, Q. (2014). Current status and outlook in the application of microalgae in biodiesel production and environmental protection. *Frontiers in Energy Research*, *2*(AUG), 1–15. https://doi.org/10.3389/fenrg.2014.00032