SAND2021-11486

SANDIA REPORT
SAND20XX-XXXX
Printed Click to enter a date

Sandia National Laboratories

CERES: CRISPR Engineering for the Rapid Enhancement of Strains

SNL: Anne M. Ruffing, Raga Krishnakumar, Chuck R. Smallwood, Joshua Podlevsky, Tessa Dallo, Stephanie D. Kolker, and Xavier M. Torres

Purdue University, Davidson School of Chemical Engineering: John A. Morgan, Nathaphon Yu Hing King, and Melissa Marsing

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico
87185 and Livermore,
California 94550

Issued by Sandia National Laboratories, operated for the United States Department of Energy by National Technology & Engineering Solutions of Sandia, LLC.

Available to DOE and DOE contractors from
U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831

Telephone: (865) 576-8401
Facsimile: (865) 576-5728
E-Mail: reports@osti.gov
Online ordering: http://www.osti.gov/scitech

Available to the public from
U.S. Department of Commerce
National Technical Information Service
5301 Shawnee Rd
Alexandria, VA 22312

Telephone: (800) 553-6847
Facsimile: (703) 605-6900
E-Mail: orders@ntis.gov
Online order: https://classic.ntis.gov/help/order-methods/

ABSTRACT

Previous strain development efforts for cyanobacteria have failed to achieve the necessary productivities needed to support economic biofuel production. We proposed to develop CRISPR Engineering for Rapid Enhancement of Strains (CERES). We developed genetic and computational tools to enable future high-throughput screening of CRISPR interference (CRISPRi) libraries in the cyanobacterium *Synechococcus* sp. PCC 7002, including: (1) Operon-SEQer: an ensemble of algorithms for predicting operon pairs using RNA-seq data, (2) experimental characterization and machine learning prediction of gRNA design rules for CRISPRi, and (3) a shuttle vector for gene expression. These tools lay the foundation for CRISPR library screening to develop cyanobacterial strains that are optimized for growth or metabolite production under a wide range of environmental conditions. The optimization of cyanobacterial strains will directly advance U.S. energy and climate security by enabling domestic biofuel production while simultaneously mitigating atmospheric greenhouse gases through photoautotrophic fixation of carbon dioxide.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACRONYMS AND DEFINITIONS

Abbreviation	Definition
13-C MFA	Carbon-13 metabolic flux analysis
7002	*Synechococcus* sp. PCC 7002
aTc	anhydrotetracycline
CERES	CRISPR Engineering for the Rapid Enhancement of Strains
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
CRISPRa	CRISPR activation
CRISPRi	CRISPR interference
dCas9	dead Cas9
FACS	Fluorescence activated cell sorting
gRNA	guide RNA
INST-MFA	Isotopically non-stationary flux analysis
ML	Machine Learning
PAM	Protospacer adjacent motif
PCR	Polymerase Chain Reaction
qPCR	Quantitative polymerase chain reaction
RNAP	RNA polymerase
SNL	Sandia National Laboratories
TSS	Transcription start site

This page left blank.

1. INTRODUCTION

Contributing Author: Anne M. Ruffing

1.1. Cyanobacteria for biofuel and biochemical production

Cyanobacteria are prokaryotes with the ability to fix CO_2 via photosynthesis, making them attractive chassis for renewable biofuel and biochemical production. Research efforts have successfully engineered cyanobacteria to produce a wide range of biofuel precursors, biofuels, and biochemicals, including ethanol,[1-3] 2,3-butanediol,[4-9] alkanes,[10,11] free fatty acids,[12-16] isoprenoids,[17-25] sucrose,[26-30] and polyhydroxybutyrate.[31-36] While a few of these cyanobacterial systems have achieved production titers of greater than 1 g/L,[30] most strains produce titers that are too low for economical, large-scale production. Therefore, new genetic tools are needed for cyanobacterial strain optimization.

Only a few cyanobacterial species have been developed as chassis organisms for both fundamental genetics research and potential biofuel and biochemical production systems. The most well-studied cyanobacteria include two freshwater species (*Synechococcus elongatus* PCC 7942 and *Synechocystis* sp. PCC 6803), a marine species (*Synechococcus* sp. PCC 7002), and a nitrogen-fixing species (*Anabaena* sp. PCC 7120). Recent efforts have also focused on engineering newly isolated, fast-growing, freshwater species including *S. elongatus* UTEX 2973 and *S. elongatus* PCC 11801 and 11802. Due to limited freshwater resources, a marine strain is desirable for large-scale biofuel production. Therefore, we selected *Synechococcus* sp. PCC 7002 (hereafter abbreviated 7002) as our chassis cyanobacterium. Additionally, 7002 has a fast growth rate and high light tolerance, beneficial traits for a production strain.

1.2. CRISPR engineering for strain development

Within the last decade, CRISPR technology has emerged as a new tool for strain development. While most initial applications have focused on gene editing in eukaryotic systems, CRISPR technologies have also proven to be useful for prokaryotic strain development. In addition to rapid and high-throughput gene editing, CRISPR interference (CRISPRi) and CRISPR activation (CRISPRa) have enabled down- and up-regulation of targeted genes, along with high-throughput CRISPRi/a screens.[37-43] Most of these bacterial CRISPR applications are in *E. coli*, yet CRISPR gene editing and CRISPRi have been demonstrated in a number of bacterial organisms, including cyanobacteria.[44-52] High-throughput CRISPRi/a screening using guide RNA (gRNA) libraries enables genome-wide metabolic engineering to help identify non-intuitive, unannotated, or combinatorial tuning of native genes to optimize growth or production of a metabolic product that can be readily screened.

The use of CRISPR for metabolic engineering of cyanobacteria was recently reviewed in 2018.[44] To date, the only published report of CRISPR technology in 7002 is the development of inducible CRISPRi using dead Cas9 (dCas9) driven by the *tet* promoter from Brian Pfleger's group at the University of Wisconsin.[45] High-throughput CRISPRi gRNA library screening was recently demonstrated in a cyanobacterium in 2020, approximately a year after the start of the CERES project. In this work, a gRNA library was developed and transformed into *Synechocystis* sp. PCC 6803 using genome integration, and the resulting CRISPRi population was screened for gRNAs leading to enhanced growth, increased resistance to L-lactate stress, and L-lactate production.[52]

1.3. CERES project objectives and report summary

The long-term objective of the CERES project is to apply high-throughput gRNA library screenings of CRISPRi/a for improved growth of 7002 under outdoor light and temperature conditions to improve biomass productivity for biofuel and biochemical production in outdoor systems. A schematic of the long-term CERES objective is shown in Figure 1-1.

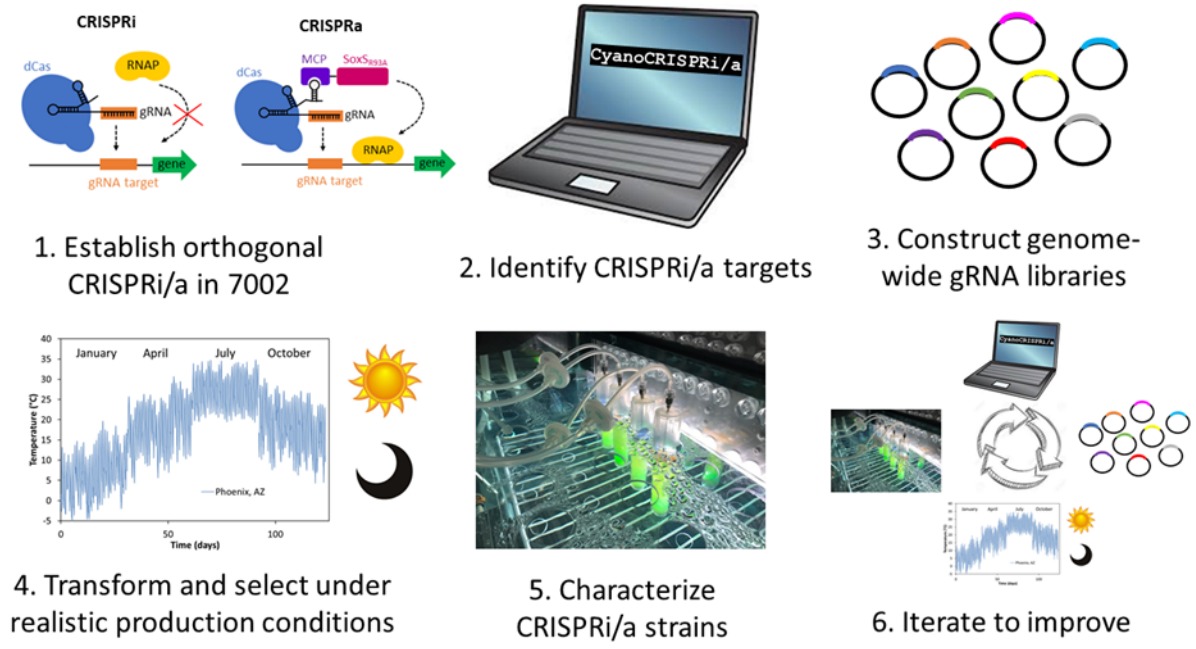


Figure 1-1. Schematic for high-throughput CRISPRi/a screening in 7002

In order to develop a high-throughput CRISPRi/a screening platform for 7002, several additional technologies must be established in 7002:

First, operon prediction in 7002 should be improved to facilitate gRNA design, particularly for gene activation which requires targeting of the promoter region. Section 2 of this report describes the development of Operon-SEQer, an ensemble of algorithms to improve operon prediction using short-read RNA-seq data.

Second, we need a better understanding of gRNA design rules for both CRISPRi and CRISPRa in 7002. The prior demonstration of CRISPRi in 7002 used an inducible promoter to tune gene expression with gRNAs targeting the non-template strand near the start codon of the target gene;[45] however, when performing genome-wide CRISPRi gRNA library screens, an improved understanding of gRNA design rules will help to design more effective and robust gRNA libraries. In Section 3 of this report, we use high-density gRNA tiling of reporter genes along with correlation analysis and machine learning to identify gRNA design rules for CRISPRi in 7002 and to predict gRNA performance.

Third, development of a shuttle vector for gRNA expression is more desirable than relying on genome integration of the gRNA construct. As 7002 and many cyanobacteria are polyploid, complete, homozygous genome integration requires multiple rounds of growth on solid selection media with PCR verification of genome integration.[53] This is challenging with gRNA library

integration, and therefore, genome integration may lead to biased results based on the number of gRNA copies integrated into the multiple genome copies. Additionally, the polyploidy nature of 7002 will result in multiple copies of the target gene requiring dCas9-gRNA binding for interference or activation. Thus, higher gRNA copy number from plasmid-based expression may also help to saturate CRISPRi/a binding of all copies of the target gene. Section 4 of this report summarizes the development of a shuttle vector for gene expression in 7002.

Fourth, CRISPRa must be established in 7002. While CRISPRi has been reported for many bacteria, including at least 5 strains of cyanobacteria,[44,54] CRISPRa has only been reported in a few model heterotrophic bacteria.[37,38,55-57] CRISPRa is inherently more challenging than CRISPRi, as it requires a transcriptional activator to recruit the native RNA polymerase (RNAP) and precise positioning of the gRNA upstream of the promoter region to enable gene activation. In Section 5 of this report, we describe our attempts to establish CRISPRa in 7002.

Lastly, the CERES project established an Academic Alliance LDRD collaboration with Dr. John Morgan's laboratory at Purdue University. Through this collaboration, the Morgan laboratory performed 13C metabolic flux analysis (13C-MFA) to understand how carbon fluxes change in 7002 under diurnal light conditions and to investigate changes in carbon fluxes of a CRISPRi strain of 7002 showing an altered growth phenotype. The results from this collaboration are detailed in Section 6.

Section 7 of this report summarizes our conclusions from the CERES project and discusses ideas for future work.

This page left blank.

2. OPERON-SEQER

Contributing Authors: Raga Krishnakumar and Anne M. Ruffing

2.1. Executive Summary

Prokaryotes often organize their genes into transcriptional units known as operons. This involves transcription from a single promoter, with downstream processing to yield separate gene products.[58-63] This allows for functional grouping of gene expression to optimize the production of specific proteins as they are needed.[58,62] Traditional operon prediction utilizes genomic sequence and functional conservation information to determine whether a gene pair is within the same operon.[64-66] However, with the advent of next-generation sequencing, we are now able to directly measure total transcription and use this information to determine operon status of gene pairs species.[67,68] There have been a number of studies that have used RNA sequencing data to determine operon structure in prokaryotes.[69-74] The limitations of these studies are that they are often based on a number of well-studied organisms, the rules for which may not be broadly applicable. Another limitation is the lack of flexibility that could account for a certain amount of bias in the algorithms behind the prediction. We sought to address both of these issues by designing an operon prediction algorithm that (1) is trained across a broad range of organisms, (2) uses a statistical method to circumvent species- or experiment-specific bias and (3) employs a flexible voting system allowing the user to decide the stringency with which operon determinations should be made. To this end, we created Operon-SEQer, which is a flexible set of machine learning (ML) algorithms for determine operon membership. The first step uses a non-parametric Kruskal Wallis test to determine the likelihood that the expression pattern across two genes come from the same underlying distribution, and this likelihood and the confidence in the call (along with distance between genes) become features for the ML algorithms. We tuned the algorithms using Bayesian Optimization, and we implement a multi-threshold voting system to increase both robustness and context-based flexibility of the software. We demonstrate that Operon-SEQer can identify a high percentage of known operons in organisms with a range of GC content and size. In addition, Operon-SEQer can identify previously unknown operons, as confirmed by published long-read sequencing (i.e. RNA sequencing with read lengths that can encompass multiple genes and intergenic sequences, thus providing empirical proof for the existence of an operon).[67]

The impact of this work is to advance the state-of-the-art in identification of prokaryotic operons, which is a critical step not only for understanding the biology of these organisms, but to determine the methods by which to best engineer them. For example, design of CRISPR activation and CRISPR interference gRNAs requires knowledge of promoter regions versus transcribed regions.[75-77] In addition, knowing operon structure can help with functional annotation of genes of unknown function in organisms that are less well-studied.[78-80] Understanding prokaryotic biology and how to engineer prokaryotic organisms has broad applicability in many fields, including agriculture, manufacturing and healthcare.

This page left blank.

3. GUIDE RNA DESIGN RULES FOR CRISPR INTERFERENCE

Contributing Authors: Tessa Dallo, Raga Krishnakumar, Stephanie D. Kolker, and Anne M. Ruffing

3.1. Executive Summary

Guide RNA (gRNA) design rules for CRISPRi are primarily based on studies performed in eukaryotic systems, particularly mammalian cell lines.[81-83] Furthermore, prior studies on bacterial systems indicate that gRNAs should target the non-template strand of DNA near the transcription start site (TSS) for effective gene repression by CRISPRi.[84,85] However, experimental results show some gRNAs are ineffective, suggesting other factors influence gRNA performance. To account for these unknown factors, most researchers design two to five gRNAs for each CRISPR gene target to achieve successful gene knockdown.[45-48,51] When designing a gRNA library to target every gene in the genome of a microorganism, this high percentage of ineffective gRNAs greatly reduces the robustness and potential of the high-throughput screening application.

To establish CRISPRi in 7002, Gordon *et al.* tested six different gRNAs targeting a yellow fluorescent protein reporter gene with two gRNAs targeting the template strand and four gRNAs targeting the non-template strand.[45] Moreover, five of the six gRNAs targeted the reporter gene near the TSS (within 220 bp). The results agreed with prior studies, showing greater repression closer to the TSS and with targeting of the non-template strand. Many regions of the reporter operon remain untested, however, including the upstream, promoter, and terminator regions as well as the template strand away from the TSS. To gain a better understanding of gRNA design rules for CRISPRi in 7002 for gRNA library design, we constructed 76 strains of 7002 with gRNAs targeting three reporter gene operons. These strains tested gRNA effectiveness with regard to position relative to the start codon, strand preference, protospacer adjacent motif (PAM) preference, GC content of the CRISPR RNA (crRNA), minimum free energy to evaluate secondary structure of the gRNA, and the number of predicted off-target sites in the 7002 genome. We also used this data to evaluate machine learning methods for predicting gRNA effectiveness of CRISPRi gene repression in 7002.

A detailed discussion of the results and methods may be found in our manuscript: High Density Guide RNA Tiling and Machine Learning for Predicting CRISPR Interference in *Synechococcus* sp. PCC 7002, to be submitted for publication in ACS Synthetic Biology. This study determined the following gRNA design rules for CRISPRi in 7002: (1) CRISPRi is capable of activating gene expression when the gRNA targets upstream of the promoter region. We hypothesize that the slight, but statically significant, activation is due to steric effects of the dCas9, promoting binding of RNA polymerase (RNAP). (2) CRISPRi with gRNAs targeting the promoter region results in gene repression with targeting of either the template or non-template DNA strand. (3) gRNAs targeting the coding region demonstrate increased gene knockdown when targeting the non-template strand, and the strength of gene repression generally decreases with increasing distance from the start codon. (4) CRISPRi with the gRNA targeting the terminator region is effective for gene repression. (5) Higher GC content in the crRNA leads to enhanced interference, presumably through stronger binding to the target sequence. (6) Guanine and cytosine PAM sites (GGG and CGG) are correlated with greater repression. (7) More negative minimum free energy (i.e. more stable gRNA structures) result in more effective interference. (8) The number of off-target gRNAs in the genome is correlated with reduced CRISPRi gene knockdown. In addition to these design rules, this work also shows that these CRISPRi datasets can predict gRNA effectiveness in 7002 using machine learning, with decision tree and random forest algorithms demonstrating lower error than linear regression or k-nearest neighbors algorithms. These gRNA design rules and ML-based gRNA performance

predictions will inform future CRISPRi studies in 7002 with specific gene targets and also facilitate the design of more effective gRNA libraries for CRISPRi in 7002.

4. SHUTTLE VECTOR FOR GENE EXPRESSION IN 7002

Contributing Authors: Joshua Podlevsky, Chuck R. Smallwood, Tessa Dallo, Xavier Torres, and Anne M. Ruffing

4.1. Development of a shuttle vector for recombinant gene expression in *Synechococcus* sp. PCC 7002 (Executive Summary)

The standard method for genetic modification of 7002 requires genome integration of the recombinant DNA construct via homologous recombination.[86] Due to the fact that 7002 is polyploid,[87] multiple rounds of selection are required to segregate homozygous transformants. This is undesirable for high-throughput gRNA library screening, as gRNA copy number in the population may be variable based on the number of genomes containing the integrated gRNA construct. Variability in gRNA copy number will lead to biased results in subsequent screenings. Therefore, we developed an *E. coli*-7002 shuttle vector to facilitate gRNA library construction and evaluation. A detailed discussion of this work is presented in our manuscript: Development of a Shuttle Vector for Recombinant Gene Expression in *Synechococcus* sp. PCC 7002, to be submitted for publication in *Frontiers in Bioengineering and Biotechnology*.

To construct the *E. coli*-7002 shuttle vector, we first identified the origin of replication for pAQ1, the smallest native plasmid in 7002. The fragment containing the origin of replication was integrated into two backbones for cloning in *E. coli*: pBISNL2, a plasmid containing a broad-host-range origin with low copy number in *E. coli* to facilitate expression in multiple bacterial hosts, and pTOPO, a plasmid with high copy number in *E. coli* to facilitate cloning efforts requiring high amounts of DNA. Both shuttle vectors containing the pAQ1 origin of replication were maintained in 7002 transformants for up to seven re-streaks on solid medium containing antibiotic selection. To confirm stable maintenance of the shuttle vector in 7002, plasmid was isolated from these re-streaked transformants and transformed back into *E. coli* for verification of plasmid size and plasmid sequencing. A small percentage of 7002 transformants maintained the intact shuttle vector across the seven sub-cultivations, while most of the transformants contained both the intact shuttle vector and the shuttle vector recombined with pAQ1. This confirmed successful maintenance of the shuttle vector in 7002 with antibiotic selection.

To demonstrate recombinant gene expression from the *E. coli*-7002 shuttle vector, we integrated two operons expressing a yellow fluorescent protein reporter (YPET), one using the strong A2579 promoter to drive YPET expression and one using the moderate *rbc* promoter to drive YPET expression.[86] After transformation of both reporter plasmids into 7002, YPET expression is being evaluated with and without antibiotic selection over five sequential sub-cultivations with 7 day growth periods. We are also comparing YPET expression from the shuttle vector to YPET expressed from genome-integrated constructs containing the same operons. As expected, preliminary measurements show increased YPET expression from the shuttle vector strains compared to the genome-integrated strains. Stability of recombinant gene expression from the shuttle vector is measured from YPET fluorescence, and plasmid stability is evaluated by quantitative PCR (qPCR). This experiment is on-going at the time of writing this report, and the results will be included in the aforementioned manuscript.

The *E. coli*-7002 shuttle vector developed in the CERES project may be used in future applications to construct a gRNA library for high-throughput CRISPRi/a screenings in 7002. Additionally, this shuttle vector will be useful for other strain development efforts in 7002. First, gene expression from the shuttle vector will likely be higher than a genome-integrated construct due to higher copy number, and second, additional sub-cultivation on selection media to obtain homozygous transformants will not be required with a plasmid-based expression system, thereby reducing the time required for generating 7002 strains.

4.2. Attempts to integrate endogenous 7002 plasmid into host chromosome

4.2.1. Introduction

Synechococcus sp. PCC 7002 lacks self-replicating shuttle vectors that are essential for the rapid development and genetic engineering of cyanobacterial strains.[88,89] To facilitate this, we explored the possibility of coopting an endogenous 7002 plasmid[90] and fusing the minimal cyanobacterial replication element with an *Escherichia coli* replicating plasmid. 7002 naturally harbors six endogenous plasmids, pAQ1 (4809 bp), pAQ3 (16,103 bp), pAQ4 (31,972 bp), pAQ5 (38,515 bp), pAQ6 (124,030 bp), and pAQ7 (186,459 bp).[91] We choose the smallest plasmid, pAQ1, that has four open reading frames: ORF943, ORF64, ORF71, and ORF93[90] and has the highest copy number of genetic elements in the cell. The pAQ1 contains essential genes,[92] thus the plasmid could not be cured, as reported for a shuttle plasmid system recently developed for *Synechococcus* sp. PCC 7942[93]. To overcome the inability to directly eliminate the pAQ1 plasmid and to use the minimal cyanobacterial replication element for a shuttle vector, we attempted to integrate the four pAQ1 open reading frames into the host chromosome using the previously described neutral site 2 (NS2)[94]. Integrating pAQ1 essential genes into the host chromosome would then free the pAQ1 replication element for our desired shuttle vector. Despite our efforts, the pAQ1 genes appears to recombine with the native pAQ1 plasmid and not integrate into the host chromosome. This was likely the results of the numerous recombination sites previously reported for pAQ1[95] and persisted despite site directed mutagenesis of these eight recombination hot spots. Herein, we report the efforts to integrate the pAQ1 genes into the host chromosome.

4.2.2. Results and Discussion

4.2.2.1. Construction of pAQ1 open reading frame integration system

A putative origin of replication for pAQ1 had been proposed for a region overlapping the 3'-end of ORF943, a 367 bp region at position 2,835 to 3,201 bp.[95] To avoid integrating the pAQ1 origin of replication into the host chromosome, we designed primers to amplify a 4,625 bp fragment from 3,200 to 3,015 bp. This region would retain approximately 181 bp of the putative origin of replication yet was necessary for the 3'-end of ORF943. This PCR product (Table 4-1) was cloned into a plasmid harboring two approximately 500 bp regions homologous to NS2 in the host chromosome[94] to generate the pSaq1 plasmid. Sequencing indicated a single point mutation at C3179 in the native PCR amplified product. This C3179T transition mutation introduced a nonsense mutation, Q460*, that would abort translation approximately halfway through the ORF943 sequence. Multiple cloning attempts resulted in nonsense mutations in similar locations. Commercial vendor (i.e., GenScript) attempts to rectify the mutation were similarly unsuccessful; however, they were successful at mutating the eight recombination hot spots[95] within the pSaq1 plasmid (Table 4-2) to generate pSaq1(mut).

4.2.2.2. Transformation of 7002 with pSaq1(mut)

The pSaq1(mut) plasmid was linearized and transformed into 7002 cells and screened for integration into the host chromosome (

Table 4-3). The initial screening was performed using primers that annealed to the NS2 homology regions, NS2-f1 and NS2-r1. These primers generated two distinct bands of 351 bp and 5,835 bp for the non-integrated native host chromosome sequence and the integration of the linearized pSaq1(mut) sequence (Figure 4-1). We performed multiple re-streaks of these colonies to try to eliminate the lower molecular weight product and have homogeneous integration of pSaq1(mut) at NS2. However, after 3 re-streaks, there was no change in the ratio of the higher and lower molecular weight bands. To ensure these bands were due to integration into the host chromosome and not either pAQ1 or a non-specific location in the chromosome, we redesigned the NS2 screening primers to anneal outside of the NS2 homology regions, NS2-f2 and NS2-r2 (

Table 4-3). These redesigned primers were expected to generate a 934 bp larger product at bands of 1,285 bp and 6,769 bp for the non-integrated native host chromosome sequence and the integration of the linearized pSaq1(mut) sequence. PCR screening generated exclusively the expected lower molecular weight band and not the higher molecular weight band (Figure 4-2). This additional screening revealed that the pSaq1(mut) sequence was not integrating into NS2 in the host chromosome. It is likely that there remained recombination between pAQ1 and pSaq1(mut) despite mutagenesis of the eight palindromic sequences. Additionally, it is possible that there was non-specific recombination elsewhere in the 7002 genome.

4.2.3. Conclusions

This study attempted to integrate the essential open reading frames from pAQ1 into the host chromosome NS2 region. Despite rigorous attempts to correct a point mutation in the pSaq1 plasmid, a nonsense mutation remained that interrupted ORF943 approximately halfway through the protein sequence. This would indicate there was toxicity associated with either this genetic element alone or in combination with the other three open reading frames in pSaq1. Additionally, mutagenesis of the eight palindromic sequences to eliminate recombination between pAQ1 and pSaq1 failed to prevent recombination. While the sequence for pSaq1(mut) could be detected, it was revealed this was not within the NS2 region of the host chromosome and likely from recombination with pAQ1. Overall, this reveals that there are additional factors for pAQ1 recombination outside of the eight reported palindromic sequences.

4.2.4. Methods

4.2.4.1. Construction of pSaq1(mut) plasmid

The 4,625 bp fragment spanning position 3,200 to 3,015 from the native pAQ1 plasmid, exuding most of the origin of replication. The pAQ1 fragment was amplified as 3 separate regions, appended with compatible restriction sites taking advantage of these existing in the pAQ1 fragment sequence for reassembly. For the PCR reaction, 1 µl of 7002 cells was used as template in a 25 µL reaction with 0.5 µM DNA primers (

Table 4-3), 1x Q5 Reaction buffer (25 mM TAPS- HCl at pH 9.3, 50 mM KCl, 2 mM MgCl2, and 1 mM β-mercaptoe- thanol), 0.2 mM each dNTP, and 0.5 U of Q5 DNA Polymerase (NEB). The PCR products were size verified on a 1% agarose gel and purified with Wizard SV Gel and PCR Clean-Up System (Promega). The pSrbcYpet plasmid and purified PCR products were digested with AatII and XhoI to insert the first fragment. The second fragment was inserted with NheI and XhoI and the final fragment was inserted with ApaI and XhoI. For each digestion, 1.5 µg of plasmid DNA and 1.0 µg of purified PCR product was digested in a 15 µl reaction with 1x CutSmart buffer (50 mM KOAc, 20 mM Tris-acetate, 10 mM MgOAc, and 100 µg/ml BSA) with 10.0 U of appropriate

restriction enzyme (NEB). The purified restriction digested products were ligated in a 10 µL reaction with 1x T4 DNA Ligase buffer (50 mM Tris-HCl, 10 mM MgCl2, 1 mM ATP, 10 mM DTT), 200 U of T4 DNA Ligase (NEB). The ligated products were electroporated using a Gene Pulser Xcell electroporator (BioRad) into *E. coli* E. cloni 10G (Lucigen) and plated onto LB medium with appropriate antibiotic. Single isolated colonies were screened for presence of plasmid and verified by sequencing.

4.2.4.2. 7002 transformation and screening

Four micrograms of plasmid were linearized in a 25 µl reaction with 1x CutSmart buffer (50 mM KOAc, 20 mM Tris-acetate, 10 mM MgOAc, and 100 µg/ml BSA) with 10.0 U of SpeI (NEB) and heat inactivated at 80°C. The linearized plasmid was added to 1 ml of OD_{730} 1.0 *Synechococcus* sp. PCC 7002 and incubated for 24 hours at 30°C with 150 rpm and 60 µE light. Cells were plated onto A+ media with 50 µg/ml kanamycin. For screening, transformants were diluted into 100 µl water and 1 µl was used in a 20 µl reaction with 0.32 µM DNA primers (

Table 4-3), 1x LongAmp Taq Reaction Buffer (60 mM Tris-SO_4, 20 mM $(NH4)_2SO_4$, 2 mM $MgSO_4$, 3% Glycerol, 0.06% IGEPAL CA-630, and 0.05% Tween 20) 0.24 mM each dNTP, and 2.0 U of LongAmp Taq DNA Polymerase (NEB).

Table 4-1. The pAQ1 open reading frame fragment

Name	Sequence	Notes
pAQ1 ORF PCR fragment	TCGCCCAAGAGATTCAGCGACCGGGGCGATCGCCCTTGGTAATTCTC TCAGGCGCTCAACTTTTTTAGGCGCTGATCGAGCTCTTGGATTTGCT GTTGGATGTCAGCAAGCTCTGGAATTTCCCGATTCTCTGATGGGAGA TCCAAAAATTCTCGCAGTCCCTCAATCACGATATCGGTCTTGGATCG CCCTGTAGCTTCCGACAACTGCTCAATTTTTTCGAGCATCTCTACCG GGCATCGGAATGAAATTAACGGTGTTTTAGCCATGTGTTATACAGTG TTTACAACTTGACTAACAAATACCTGCTAGTGTATACATATTGTATT GCAATGTATACGCTATTTTCACTGCTGTCTTTAATGGGGATTATCGC AAGCAAGTAAAAAAGCCTGAAAACCCCAATAGGTAAGGGATTCCGAG CTTACTCGATAATTATCACCTTTGAGCGCCCCTAGGAGGAGGCGAAA AGCTATGTCTGACAAGGGGTTTGACCCCTGAAGTCGTTGCGCGAGCA TTAAGGTCTGCGGATAGCCCATAACATACTTTTGTTGAACTTGTGCG CTTTTATCAACCCCTTAAGGGCTTGGGAGCGTTTTATACGAGTGCGG GGAAATTTCGGGGGCGATCGCCCCTATATCGCAAAAAGGAGTTACCC CATCAGAGCTATAGTCGAGAAGAAAACCATCATTCACTCAACAAGGC TATGTCAGAAGAGAAACTAGACCGGATCGAAGCAGCCCTAGAGCAAT TGGATAAGGATGTGCAAACGCTCCAAACAGAGCTTCAGCAATCCCAA AAATGGCAGGACAGGACATGGGATGTTGTGAAGTGGGTAGGCGGAAT CTCAGCGGGCCTAGCGGTGAGCGCTTCCATTGCCCTGTTCGGGTTGG TCTTTAGATTTTCTGTTTCCCTGCCATAAAAGCACATTCTTATAAGT CATACTTGTTTACATCAAGGAACAAAAACGGCATTGTGCCTTGCAAG GCACAATGTCTTTCTCTTATGCACAGATGGGGACTGGAAACCACACG CACAATTCCCTTAAAAAGCAACCGCAAAAAATAACCATCAAAATAAA ACTGGACAAATTCTCATGTGTTCTTCTCAATTTCCACACTGTTTATC CACAGGAAATTAAGGGGCTGTAGCGTTGGTGCTACAGAATAAATGTA GGGATCGCCCATAGCTTTATTGCTAGCCACAGTGCTATGGGGAAAAG GAAAAGAAAAAATACCACCATGAATGGGGGTGTCAAATCTTTTGGAT ACTGTAAAATGATAGAGACTTTCTTAGGCGATCCCATGACGACTAGA CCGAATAAGAATTTAAAATCAGCGAGCGCGGTTCGTTTTCCCCCACT CATGTACAGCGCGGCTACTAAAAAAGCCAATGAGCAAGGCTTAAATT TCAGCGACTATATCCGGGAGCTTGTTTTACGAGATTTGCTCGAAGTC TATAACAATGATGAGGCGGATCAAGATGCTGCCTAAACACGAGAAAA CCCCCAAAAGTTTGGCGACCGGCGGGGGCTTCTCTAAACCAAATAAA	Fragment contains the four open reading frames. Location of the single point mutation denoted.

Name	Sequence	Notes
	AGCAAACAGGGACGGCAAAACCGAGCCCTAATTGATCATGATCGTCA AGGCGAAGGAGGCCTTAACTCACTGTGATCGTATCAGATAATTTTAC ATCTAGCAGTAATACCGGGGCAGATTTAGCGAGTAGCGTCAGTGAGA ACCGTTCTCATTTTATCGCAGAATCTCACTATAGGGAATGGGTAGAG GGTTCCGGGGTCGCCCCTGAGATTGCCCGGTTAAATGTTAGATCGTT GTCAGGGATGACGCCCTATGAATACCTGCTCTACAGTGATGAGCCAG CATTACGCCGCAATGATGGCAGGCTTCGGGATACTTGGCTTAAGCGA TATGCCTTTGTCGAGCATGGCGGTTGGTGGTGCTCAGGCATCGATAT TAAGACCGGGAAAGATTCTCTTTGGGGTTGCTTCAAAGGCGATCGCC CCCGTAAAGATCGCGAAGATAAGAAACCGATCAAGTATGAGCACCCG CCACGGGTAGCCACCGAAATTTTCACTCTCAAGGTAGACCGGGGCAC CTGGCGCAAGATTGCCAAGCGCCACAAAGTCGAGCTACCAGAAACCG ATCAAGGCTTCTGGGAATGGGTACTAGCCCATCCTGAGCTACCGATC ATCATCACTGAGGGCGCGAAGAAGGCAGGGGCCCTCCTAACCGCTGG TTATTGCGCCATCGGTCTACCGGGGATTTACAACGGCTACAGAACGC CAAAAAATGACCATGGCGAGCCAATGCGACAGCTACGGCACCTCATT CCAGAGCTTGACCTGTTGGCGAAAAATAACCGGGCGATCGCCTTCTG TTTTGACCAAGACAAGAAACCCAAGACGATCAAGGCAGTGAACGGGG CGATCCAAACTACCGGGGCACTCCTAGAGAAGGCCGGGGCGAAAGTA TCGGTGATCACCTGGCACCAGGACGCGAAAGGTGTTGATGATCTGAT CGTCGAGCACGGAGCGAAAGCACTCCATAACCGCTACAAGCACCGCA AGCCCTTAGCAGTCTGGGAGATGGATAATCTCACCGATATCACCACG CAAGTCGATCTAACGGTCGATCAGCGCTATCTCGACATCGATCCGCG TGCTATCCCCAAGGATGCTCAGATTATTTTCATTAAATCTGCCAAGG GCACCGGGAAAACAGAATGGTTAGGGAAAATCGTTAAGCTCGCCCAA GATGATTGCGCTCGCGTACTGGTTTTGACTCACCGCATCCAATTAGC CAAGGAGCTCGCCCGCCGTCTCGATATCGATCACATTAGCGAGCTCG ACAGTAGCCCGACCGGGGGCGCTCTAGGGATGGCGATGTGTATCGAT AGCCTACATCCCGATAGCCAAGCTCATTTTAATTTCATGGAATGGCA CGGCGCTCACATTGTCCTAGACGAAATCGAGCAAGTTTTAGGGCACG CTTTGGGTAGCTCGACCTGTACCCAAGACCGGGCGAAAATCCTTGAA ACGTTCTACAACCTAATCCTTTATGCCCTAAGGACAGGCGGAAAACT CTACTGCTCTGATGCTGATTTATCTCCCATCTCCTATGAGCTAATCA AGTACATTCTTGACGGTTGTGAGTTCAAACCATTCACCATTCTGAAT ACCTATAAGCCCTGTTTAGAG**C**AACAAAGAGACCTGTTTTTCTATGA AGGGAATGACCCTAGAGACCTGTTAACCAATCTCAGACAAGCGATTG AGAACGGTGAGAAAACACTTGTCTTTACCGCTGCCCAAAAGACCGCA TCGACCTACAGCACGCAGAACTTAGAATCACTCTTTAGGGAGAAATA TCCCGATAAGAGAATCTTGAGAATCGATGCTGAATCGGTCGCTGAAC CGGGGCACCCTGCCTATGGTTGTATTGACTCGCTGAATGCAATTTTG CCCCTGTACGATATTGTGCTCTGCTCGCCCGCTGTCGAGACCGGGGT GAGTATCGATATCAAGGATCATTTTGATTCTGTTTGGGGCATGGGTT CAGGGGTTCAGACCGTTAACGGTTTCTGCCAAGGGCTAGAGCGGTTA CGGGATAACGTCCCTCGCCATGTTTGGATACCGAAATTTTCCCCACA CTCGAACCGGATCGCCAATGGGGGCTACACCGCTAAGGCGATCGCCC GTGACCAGCACCGCTATGCAGAGCTCACCCACAAATTAATCGGTGAG CACGCCGCTGAATGCAGTGGGTTAGAAGATTCTTTAAAACCATTCCT TTGGGCCTATTGTCGCTATGCGGCGCTTGCTAACCGTGGCTTTGGCA GTTACCGGGAAGCGATTTTAAATAAGCTGCTTTCTGAGGGCTATGTA CAGAAAGATTTGAGCGAAATCGATCCAGCATTGGCTAAGGATTATCG AGACGAATTAAAAGCGGTGAAAGACCATAATTATCTACAGGAAAGGG TTGCGATTTCTAAAGTAGAAAATCCTGACGATCGCCAGTACGAAAAA CTGAAGCGTCAGCGGGCGAAATCTGAGACGGAACGGCACCAAGAACG CCACGGGAAACTTTCTCGCTCCTATGGGTTAACTGTGACCCCTGAGC TTGTCGAGAAAGATGATGATGGGTGGTACTCTCAGCTCCAGCTCGAA	

Name	Sequence	Notes
	TACTACTTAACCGTTGGGAAAGCATTCTGCTCTGCCCGCGACCGGGC GAAATATGACCAGCTCCAACATGAGGGCTTTGTATTTAAGCCGGATA TCAACAGGCGATCGCTCTCACCAAAGATTCACCTGTTAGAGCTACTC AACATCCATCAGTTCTTAAAACCAGGGGTGACATTCACCGGGGCGAG CCTTGAAGGGTTCAAGGAAAATTGTTTGCGGTATGCCAAGCCGATCA AGTGGATTCTTGGCAGAACGATCACCGACAAAATGAGCCCGCTCGAA ATTGCTCAGGCGCTCCTAGGCAAGCTTGACCGGAAATTGGAATACAA GGGGCGCTTTGGATCGCGGGATAACCGTCAGCGGGTCTATGAGGCGA TCGCCCCTAACGATCAGCGCGAAAAGGTCTTTGCTCATTGGTTACAG CGTGACCAAGCAAAATTAGGGGCCGTGTCCAACCCCTGTATAAATAG ATTTATTCAGGAGGCTTAG	

Table 4-2. Mutations to disrupt the 8 palindromic sequences for recombination in pAQ1

Name	Native sequence	Mutated sequence
PS1	GGCGACCGGC	GGCGATCGCC
PS2	GGCTATAGCC	GGCGATCGCC
PS3	GGCTATAGCC	GGCGATCGCC
PS4	GACGATAGGC	GACGATCGCC
PS5	GGAGGTCGCT	GGCGATCGCT
PS6	GGCTATAGCC	GGCGATCGCC
PS7	GGCGATTGCC	GGCGATCGCC
PS8	GGCGATTGCC	GGCGATCGCC

Table 4-3. Primers for screening pSaq1(mut) integration and for PCR amplifying the pAQ1 open reading frame fragment

Name	Sequence	Length
NS2-f1	ACTCAATTAGAAATATAGTGTCGTTGC	27 bp
NS2-r1	TTGCTGAGATTATCTTTTTCCTAAAGG	27 bp
NS2-f2	ATCAAGGGATTCTTCTAGGGTTTGG	25 bp
NS2-r2	TTCGGGAAAGAGATAACCTGC	21 bp
pAQ1-f1	ACTGGACGTCGCCCAAGAGATTCAGCGACC	30 bp
pAQ1-r1	TGACCTCGAGATCTGAGGGCCCTATCCAGCTAGCAATAAAGCTATGGGCG	50 bp
pAQ1-f2	TCCAGCTAGCCACAGTGCTATGGGGA	26 bp
pAQ1-r2	TGACCTCGAGATCTGAGGGCCCCTGCCTTCTTCGCG	36 bp
pAQ1-f3	CTGAGGGCCCTCCTAACCGCTGGTTATTG	29 bp
pAQ1-r3	TGACCTCGAGCTAAGCCTCCTGAATAAATCTATTTATAC	39 bp

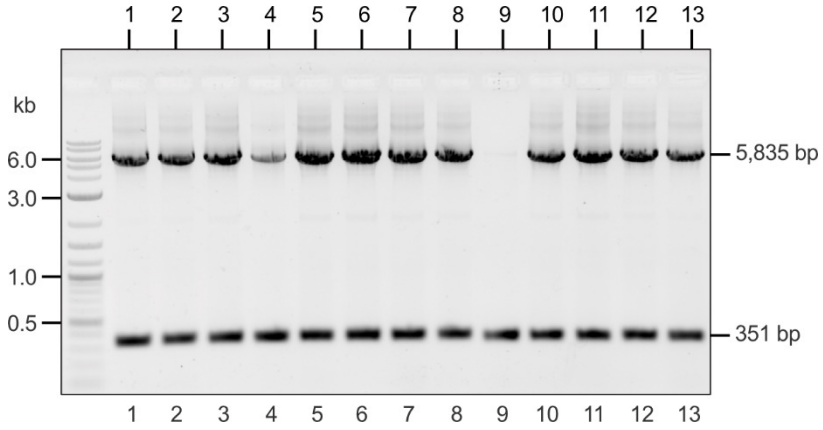


Figure 4-1. Screening for pSaq1(mut). PCR analysis for genome integration into NS2 in the host chromosome using primers NS2-f1 and NS2-r1. The lower molecular weight band represents the unmodified chromosome and the higher molecular weight product is due to the integration of pSaq1(mut).

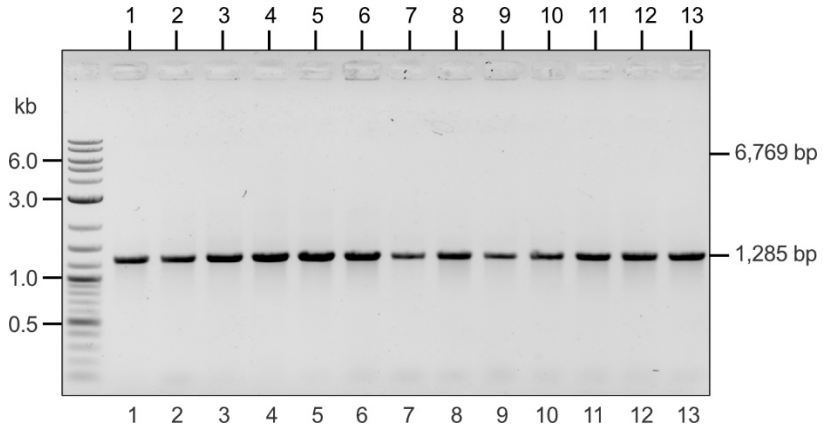


Figure 4-2. Further screening for pSaq1(mut). PCR analysis for genome integration into NS2 in the host chromosome using primers NS2-f2 and NS2-r2. The lower molecular weight band represents the unmodified chromosome and the higher molecular weight product not present would have been pSaq1(mut) integration.

This page left blank.

5. ATTEMPTS TO ESTABLISH CRISPR ACTIVATION IN 7002

Contributing Authors: Anne M. Ruffing and Joshua Podlevsky

5.1. Introduction

CRISPR activation (CRISPRa) tools have been established in several bacteria, including *E. coli,* to enable targeted up-regulation of gene expression using dCas9, a transcriptional activator, and a gRNA.[37,38,55-57,96] CRISPRa has enabled 50-fold up-regulation of gene expression in *E. coli*.[55] The development of CRISPRa tools is inherently more challenging than CRISPRi, as many physical factors affect CRISPRa function, including the distance at which the gRNA binds upstream of the target gene, the binding of the gRNA within the helical turn of the DNA, possible 3-dimensional folding of the DNA upstream of the target gene, and effectiveness of the transcriptional activator within the host microorganism.[37] With these many challenges, CRISPRa has yet to be demonstrated in any cyanobacterium.

Two main designs have successfully demonstrated CRISPRa in bacteria: (1) direct fusion of the transcriptional activator to dCas9[38,96] and (2) a two-hybrid approach, which adds an MS2 phage hairpin to the gRNA to enable binding to MS2 coat protein (MCP) fused to the transcriptional activator.[37,55,57] As the latter design was shown to be more effective in *E. coli*, we used this two-hybrid approach to attempt to establish CRISPRa in 7002 (Figure 5-1). Because RNA polymerase (RNAP) in cyanobacteria differs structurally from the RNAPof *E. coli* and other bacteria,[97-99] it's likely that non-native bacterial transcriptional activators will not be effective at recruiting RNAP in 7002. Therefore, in this study, we investigated both the best performing transcriptional activator in *E. coli*, $SoxS_{R93A}$, along with native, putative transcriptional activators identified from the genome of 7002. To test for transcriptional activation, these CRISPRa designs were genome integrated into a strain of 7002 previously engineered to express a yellow fluorescent protein (YPET) along with a gRNA targeted the promoter region of the *ypet* operon.[86] The fluorescence produced by YPET production was monitored to test for gene activation.

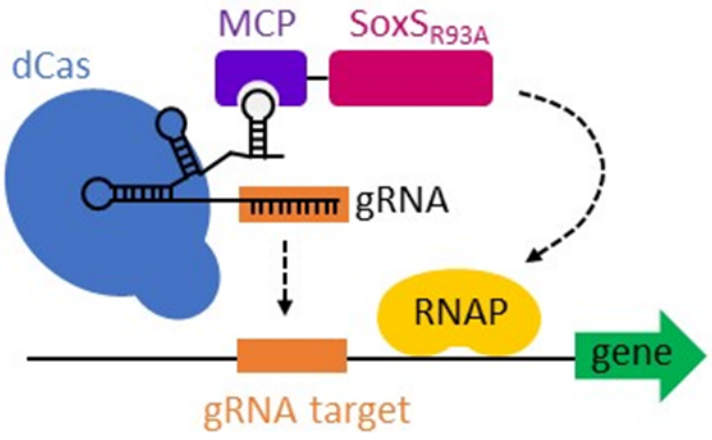
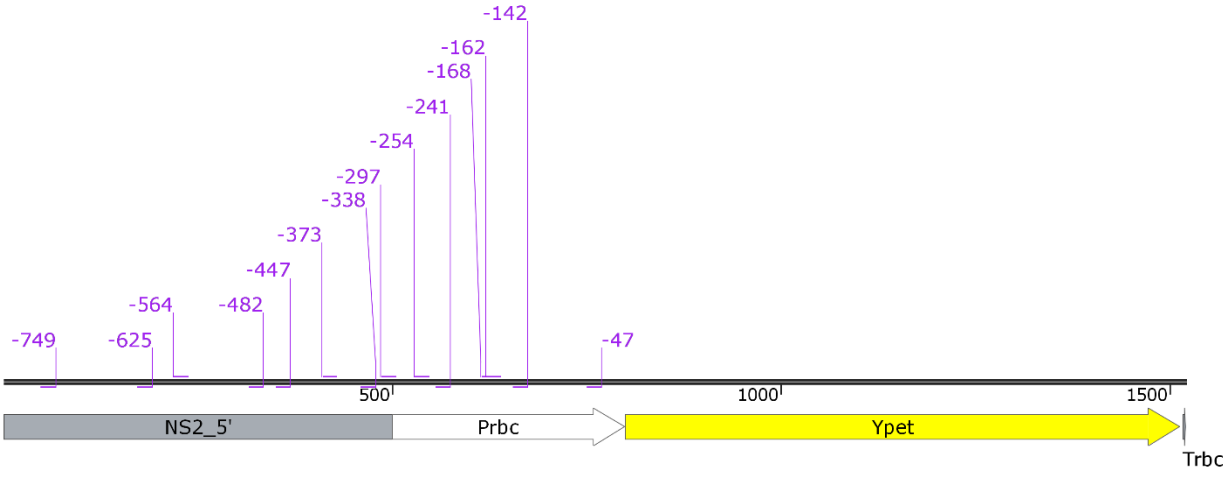


Figure 5-1: Schematic of two-hybrid CRISPRa system.

5.2. Results and Discussion

5.2.1. CRISPRa with E. coli transcriptional activator, $SoxS_{R93A}$

The two-hybrid CRISPRa system previously established in *E. coli* (dCas9-gRNA'-MCP-$SoxS_{R93A}$) was constructed in 7002 by modifying the previously constructed CRISPRi system (Section 0). Since the two-hybrid CRISPRa uses the same dCas9 as the CRISPRi system, the constructs for CRISPRi provided the foundation for building the CRISPRa system. No modification was needed for the dCas9 integration construct (pCas2C).[45] The gRNA integration construct was modified by fusing the MS2 RNA hairpin loop to the gRNA scaffold and inserting the fragment for MCP-$SoxS_{R93A}$ expressed from the aTc-inducible *tet* promoter, along with a second copy of *tetR*. With this design, all three components, dCas9, gRNA, and MCP-$SoxS_{R93A}$ are genome integrated and induced by addition of aTc. The dCas9 construct was integrated into 7002 expressing YPET from the *rbc* promoter (7002-P_{rbc}-*ypet*), which was previously shown to have moderate expression of YPET.[86] As the *rbc* promoter is uncharacterized and binding to the helical turn of the DNA was previously shown to influence effectiveness of CRISPRa in *E. coli*,[37] we designed 14 CRISPR RNA (crRNA) targeting sequences, spanning from -47 to -749 nucleotides upstream of the *ypet* start codon with varying position on the helical turn of the DNA and targeting both the sense (+) and antisense (-) DNA strands (Figure 5-2). All crRNA DNA sequences were inserted into the pNS1D-sgRNA9A01 construct and genome integrated into 7002-P_{rbc}-*ypet*-P_{tet}-dCas9. These potential CRISPRa strains were then cultured with and without inducer for 48 hours, and YPET fluorescence and cell growth were measured to test for CRISPR activation of *ypet* expression. Unfortunately, no evidence of activation was detected for any of the 14 crRNA targets relative to a random control gRNA (Figure 5-2).



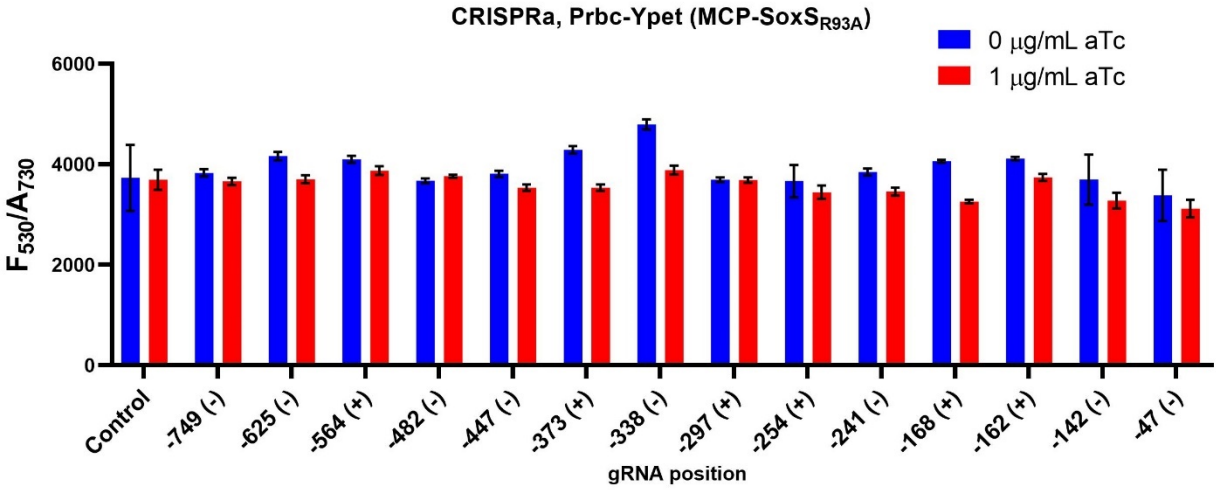


Figure 5-2: CRISPRa testing with SoxS$_{R93A}$ in 7002-P$_{rbc}$-*ypet*. Top: crRNA targets in gRNAs. Bottom: Normalized fluorescence (fluorescence at 530 nm / absorbance at 730 nm) in CRISPRa strains of 7002-P$_{rbc}$-*ypet* with (red) and without (blue) aTc inducer after 48 hours of cultivation under standard growth conditions. The control is a randomized crRNA sequence with no predicted binding site in 7002. Data are averages of at least three biological replicates (PCR-positive transformants) with the error bars indicating the standard deviation.

5.2.2. CRISPRa with 7002 transcriptional activators

Since no activation was detected using the SoxS$_{R93A}$ transcriptional activator from *E. coli*, we selected 10 putative native transcriptional activators (TAs) from 7002 to test in the CRISPRa system (Table 5-1). Testing of all 14 crRNA targets for P$_{rbc}$-*ypet* along with a control crRNA would require construction, verification, and characterization of 160 different CRISPRa strains. To reduce this to a more reasonable number for experimental verification, we selected four crRNA positions based on the activation detected with the CRISPRi system when targeting these regions of the NS2-P$_{rbc}$ region: -241, -297, -338, and -373. MCP-TA fusions were successfully constructed for the following TAs: RpoZ, RpoD, SigC, SigB, SigE, SigF, A1970, A0641, A1229, and A1232. Only one CRISPRa system with native TA was able to be tested under this effort: dCas9-gRNA-MCP-SigC. No evidence of *ypet* activation was detected for the 4 crRNAs targeted by the SigC CRISPRa system (Figure 5-3).

Table 5-1: List of putative transcriptional activators from 7002 for CRISPRa.

Name	Locus	Location	Description
SigE	A0270	Chromosome	SigE RNAP sigma factor
SigC	A0364	Chromosome	SigC, Group II sigma-70 type sigma factor (RpoD family)
RpoZ	A0614	Chromosome	RpoZ omega subunit
A0641	A0641	Chromosome	Transcriptional regulator protein, LuxR family
SigB	A1202	Chromosome	SigB RNAP sigma factor

Name	Locus	Location	Description
A1229	A1229	Chromosome	Transcriptional regulator, TetR family domain protein
A1232	A1232	Chromosome	Transcriptional regulator, Crp/Fnr family
SigF	A1924	Chromosome	SigF RNAP sigma factor
SigG	A1970	Chromosome	SigG RNAP sigma factor
RpoD	A2014	Chromosome	RpoD sigma factor
G0007	G0007	pAQ7	Transcriptional regulator, AraC family
G0096	G0096	pAQ7	Transcriptional regulator, AraC family, closest homolog to SoxS
E0011	E0011	pAQ5	MarR family transcriptional regulator
E0020	E0020	pAQ5	Fis family transcriptional regulator

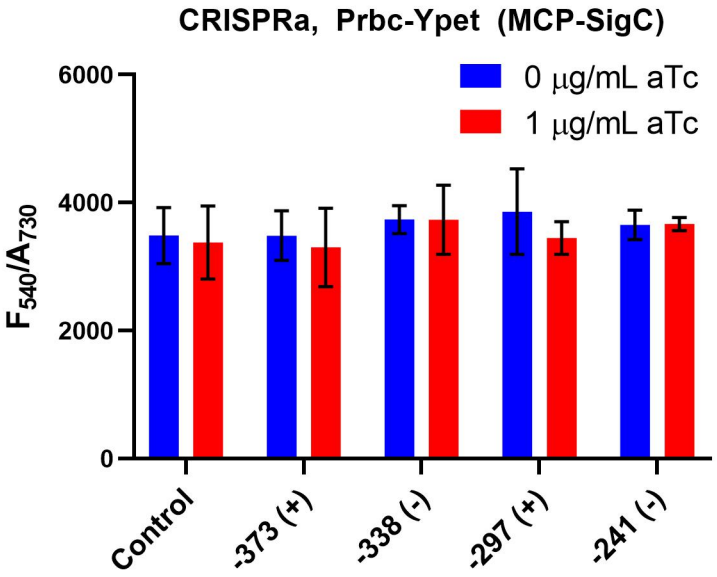


Figure 5-3: CRISPRa testing with SigC in 7002-P$_{rbc}$-*ypet*. Normalized fluorescence (fluorescence at 530 nm / absorbance at 730 nm) in CRISPRa strains of 7002-P$_{rbc}$-*ypet* with (red) and without (blue) aTc inducer after 48 hours of cultivation under standard growth conditions. The control is a randomized crRNA sequence with no predicted binding site in 7002. Data are averages of at least three biological replicates (PCR-positive transformants) with the error bars indicating the standard deviation.

5.3. Conclusions

While a 2-hybrid CRISPRa system was successful at activating gene expression in *E. coli*, no activation of gene expression in 7002 was detected using a similarly designed 2-hybrid CRISPRa system in this study. This study tested both the $SoxS_{R96A}$ transcriptional activator from *E. coli* and the native SigC transcriptional activator from 7002 fused to MCP, with 14 and 4 different crRNA targets for the $SoxS_{R96A}$ and SigC systems, respectively. We have identified 9 additional putative transcriptional activators from 7002 which remain to be tested for activity with the CRISPRa 2-hybrid system. However, there may be other issues preventing successful demonstration of CRISPRa in 7002. Insufficient binding of the dCas9-sgRNA complex to the MCP-TA may explain the absence of gene activation in our 2-hybrid CRISPRa systems. As mentioned previously, crRNA targeting is very important for gene activation, with both the position relative to the promoter and the 3D position along the helical turn of the DNA being important factors.[37] While this study attempted to address these issues by designing multiple crRNA targeting sequences, it is possible that none of these sequences are optimal for activation of gene expression from the *rbc* promoter. Additionally, expression levels for the three main components of the CRISPRa system may not be sufficient. Since we used the same dCas9 construct and a similar gRNA expression system for demonstration of the CRISPRi system (Section 0), we do not anticipate that expression is a major issue, but this remains to be verified. Lastly, expression levels of *ypet* from the *rbc* promoter may be high enough that additional activation is challenging. Prior work in the Ruffing laboratory demonstrated that much higher levels of *ypet* expression can be achieved from the strong A2579 promoter in 7002,[86] so again we do not anticipate that this is a major issue. However, we could also test for activation with even weaker promoters.

Due to the large number of transcriptional activator and crRNA combinations to optimize for demonstration of CRISPRa, we recommend the application of high-throughput screening methods for future efforts. Libraries can be generated for both crRNAs and transcriptional activators and transformed into 7002 expressing *ypet*.[57,100] Fluorescence activated cell sorting (FACS) may then be applied to screen and sort for transformants with increased Ypet fluorescence in order to identify the best transcriptional activator and crRNA.

5.4. Methods

5.4.1. Materials

All DNA oligonucleotides were ordered from Integrated DNA Technologies (IDT). All restriction enzymes, DNA polymerases, ligases, Gibson assembly master mix, and ATP were ordered from New England Biolabs (NEB). Plasmid and PCR purifications were performed using the plasmid mini-prep and PCR purification kits from Qiagen. Chemicals were purchased from Sigma-Aldrich ($Na_2EDTA{\cdot}2H_2O$, $CaCl_2{\cdot}2H_2O$, KH_2PO_4, vitamin B_{12}, $ZnCl_2$, spectinomycin sulfate, and acrylic acid), MP Biomedicals (DTT, $FeCl_3{\cdot}6H_2O$, $CuSO_4{\cdot}5H_2O$, and $CoCl_2{\cdot}6H_2O$), Acros ($Na_2MoO_4{\cdot}2H_2O$), Amresco (NaOAc, 3M, pH 5.2), Fermentas (Tango buffer), BD Difco (granulated agar), Fisher Chemical ($MnCl_2{\cdot}4H_2O$ and $Na_2S_2O_3$), and Fisher BioReagents (LB broth, agarose, NaCl, $MgSO_4{\cdot}7H_2O$, KCl, $NaNO_3$, Tris base, H_3BO_3, TAE 50x buffer, and kanamycin monosulfate).

5.4.2. Strain construction

The plasmids and strains used and constructed in this study are listed in

Table 5-2 and Table 5-3, respectively.

To develop a construct for the 2-hybrid CRISPRa system using the $SoxS_{R96A}$ transcriptional activator (pNS1D-sgRNA9A01), the sgRNA integration plasmid previously constructed for CRISPRi (pNS1D-sgRNA9-SpR) was modified by adding the MS2 RNA stem loop to the RNA scaffolding sequence and inserting the operon for MCP-$SoxS_{R96A}$ from pJF104B. Oligos for the MS2 RNA stem loop were designed based on the sequence of pJF121 (oligo sequences listed in Table A-1.[55] The plasmid backbone containing NS1D homology arms and gRNA scaffold sequence was PCR amplified from pNS1D-sgRNA9-SpR using Q5 DNA polymerase (primers listed in Table A-1) and purified using the Qiagen PCR purification kit. The MS2 RNA stem loop oligos were annealed and ligated to this purified backbone using Gibson assembly master mix to produce pNS1D-sgRNA9A. A fragment containing the MCP-$SoxS_{R96A}$ and TetR operons were amplified from pJF104B using Q5 DNA polymerase (primers listed in Table A-1). Following purification of the PCR product using the Qiagen PCR purification kit, the fragment containing MCP-$SoxS_{R96A}$ and TetR was inserted into pNS1S-sgRNA9A using SphI digestion and ligation with the Quick Ligation Kit (New England Biolabs). The resulting plasmid, pNS1D-sgRNA9A01, served as the $SoxS_{R96A}$-based CRISPRa backbone for inserting crRNA sequences targeting P_{rbc}-*ypet*.

To develop a construct for the 2-hybrid CRISPRa system using the native SigC transcriptional activator (pNS1D-sgRNA9A02), the $SoxS_{R96A}$-based CRISPRa backbone (pNS1D-sgRNA9A01) was modified by digestion with BbvCI and NotI to replace SoxS with the PCR amplified open reading frame for SigC, resulting in pNS1D-sgRNA9A-04r_A0364_SigC ligated with Quick Ligation Kit (New England Biolabs). The resulting plasmid, pNS1D-sgRNA9A01, served as the SigC-based CRISPRa backbone for inserting crRNA sequences targeting P_{rbc}-*ypet*.

The crRNA sequences were inserted into pNS1D-sgRNA9A01 and pNS1D-sgRNA9A02 using BbsI digestion and a Golden Gate assembly method, developed previously by Ran *et. al.*[101] The oligos for each crRNA sequence are listed in Table A-1.

All plasmid modifications, including insertion of crRNA oligos, were verified using Sanger sequencing from Eurofins Genomics (sequencing primers listed in Table A-2).

To construct the parental strain for CRISPRa testing (7002-P_{rbc}-*ypet*-P*tet*-dCas9), pCas2C was digested and transformed into 7002 expressing Ypet from the *rbc* promoter (7002-P_{rbc}-*ypet*). Transformants were selected on medium A+ plates with 100 μM acrylic acid. The 2-hybrid CRISPRa strains were constructed by digesting and transforming pNS1S-sgRNA9A01-X into the parental strain 7002-P_{rbc}-*ypet*-P*tet*-dCas9 (where X designates the crRNA sequence). Transformants were selected on medium A+ plates with 40 μg/mL of spectinomycin hydrochloride. The 7002 transformation protocol previously described in Ruffing *et. al.* 2016 was used for all 7002 transformations with slight modification.[86] Due to decreased transformation efficiencies, the transformation protocol was modified by growing the 7002 parental strain under standard growth conditions until an OD_{730} of 0.5 was reached. Two mL of this culture was centrifuged for 3 minutes at 5,000 x g, and the resulting cell pellet was resuspended in 1 mL of supernatant to concentrate the culture to an approximate OD_{730} of 1.0. One to five μg of linearized and purified integration constructs were added to the concentrated culture and incubated for 24 hours under standard growth conditions. After 24 hours, the transformation cultures were centrifuged for 3 minutes at 5,000 x g, resuspended in 100 μL of supernatant and spread on medium A+ agar plates with selection. For each transformation, 25 to 50 colonies were restreaked onto fresh plates two to three

times to segregated homozygous transformants, and colony PCR screening was used to verify complete integration at the target site in the genome (primers listed in Table A-2). For each strain, three PCR-positive transformants were tested for CRISPRa characterization.

Table 5-2: List of plasmids used and constructed in this study.

Plasmid Name	Description	Source
pCas2C	7002 genome integration plasmid for dCas9 for CRISPRi: dCas9 is expressed from a *tet* promoter with optimized RBS and genome integrated at the *ascA* site in 7002.	Gordon 2016[45]
pNS1D-sgRNA9-SpR	7002 genome integration plasmid for sgRNA for CRISPRi: sgRNA is expressed from a *tet* promoter and genome integrated at the NS1D site in 7002.	Section 0
pJF104B	*E. coli* expression construct for 2-hybrid CRISPRa system, includes MCP-$SoxS_{R96A}$ expressed from the *tet* promoter.	Dong 2018[55]
pNS1D-sgRNA9A01	7002 genome integration plasmid for sgRNA scaffold and MCP-$SoxS_{R96A}$ for CRISPRa: sgRNA and MCP-$SoxS_{R96A}$ are expressed from a *tet* promoter and genome-integrated at the NS1D site in 7002.	This study
pNS1D-sgRNA9A02	7002 genome integration plasmid for sgRNA scaffold and MCP-SigC for CRISPRa: sgRNA and MCP-SigC are expressed from a *tet* promoter and genome-integrated at the NS1D site in 7002.	This study
pNS1D-sgRNA9A01-cont	7002 genome integration plasmid for randomized crRNA (control) and MCP-$SoxS_{R96A}$ for CRISPRa: sgRNA and MCP-$SoxS_{R96A}$ are expressed from a *tet* promoter and genome-integrated at the NS1D site in 7002.	This study
pNS1D-sgRNA9A01-rbc749	7002 genome integration plasmid for crRNA targeting P_{rbc}-*ypet* 749 bp upstream of the *ypet* start codon and MCP-$SoxS_{R96A}$ for CRISPRa: sgRNA and MCP-$SoxS_{R96A}$ are expressed from a *tet* promoter and genome-integrated at the NS1D site in 7002.	This study
pNS1D-sgRNA9A01-rbc625	7002 genome integration plasmid for crRNA targeting P_{rbc}-*ypet* 625 bp upstream of the *ypet* start codon and MCP-$SoxS_{R96A}$ for CRISPRa: sgRNA and MCP-$SoxS_{R96A}$ are expressed from a *tet* promoter and genome-integrated at the NS1D site in 7002.	This study

Plasmid Name	Description	Source
pNS1D-sgRNA9A01-rbc564	7002 genome integration plasmid for crRNA targeting P_{rbc}-*ypet* 564 bp upstream of the *ypet* start codon and MCP-$SoxS_{R96A}$ for CRISPRa: sgRNA and MCP-$SoxS_{R96A}$ are expressed from a *tet* promoter and genome-integrated at the NS1D site in 7002.	This study
pNS1D-sgRNA9A01-rbc482	7002 genome integration plasmid for crRNA targeting P_{rbc}-*ypet* 482 bp upstream of the *ypet* start codon and MCP-$SoxS_{R96A}$ for CRISPRa: sgRNA and MCP-$SoxS_{R96A}$ are expressed from a *tet* promoter and genome-integrated at the NS1D site in 7002.	This study
pNS1D-sgRNA9A01-rbc447	7002 genome integration plasmid for crRNA targeting P_{rbc}-*ypet* 447 bp upstream of the *ypet* start codon and MCP-$SoxS_{R96A}$ for CRISPRa: sgRNA and MCP-$SoxS_{R96A}$ are expressed from a *tet* promoter and genome-integrated at the NS1D site in 7002.	This study
pNS1D-sgRNA9A01-rbc373	7002 genome integration plasmid for crRNA targeting P_{rbc}-*ypet* 373 bp upstream of the *ypet* start codon and MCP-$SoxS_{R96A}$ for CRISPRa: sgRNA and MCP-$SoxS_{R96A}$ are expressed from a *tet* promoter and genome-integrated at the NS1D site in 7002.	This study
pNS1D-sgRNA9A01-rbc338	7002 genome integration plasmid for crRNA targeting P_{rbc}-*ypet* 338 bp upstream of the *ypet* start codon and MCP-$SoxS_{R96A}$ for CRISPRa: sgRNA and MCP-$SoxS_{R96A}$ are expressed from a *tet* promoter and genome-integrated at the NS1D site in 7002.	This study
pNS1D-sgRNA9A01-rbc297	7002 genome integration plasmid for crRNA targeting P_{rbc}-*ypet* 297 bp upstream of the *ypet* start codon and MCP-$SoxS_{R96A}$ for CRISPRa: sgRNA and MCP-$SoxS_{R96A}$ are expressed from a *tet* promoter and genome-integrated at the NS1D site in 7002.	This study
pNS1D-sgRNA9A01-rbc254	7002 genome integration plasmid for crRNA targeting P_{rbc}-*ypet* 254 bp upstream of the *ypet* start codon and MCP-$SoxS_{R96A}$ for CRISPRa: sgRNA and MCP-$SoxS_{R96A}$ are expressed from a *tet* promoter and genome-integrated at the NS1D site in 7002.	This study
pNS1D-sgRNA9A01-rbc241	7002 genome integration plasmid for crRNA targeting P_{rbc}-*ypet* 241 bp upstream of the *ypet* start codon and MCP-$SoxS_{R96A}$ for CRISPRa: sgRNA and MCP-$SoxS_{R96A}$ are	This study

Plasmid Name	Description	Source
	expressed from a *tet* promoter and genome-integrated at the NS1D site in 7002.	
pNS1D-sgRNA9A01-rbc168	7002 genome integration plasmid for crRNA targeting P_{rbc}-*ypet* 168 bp upstream of the *ypet* start codon and MCP-$SoxS_{R96A}$ for CRISPRa: sgRNA and MCP-$SoxS_{R96A}$ are expressed from a *tet* promoter and genome-integrated at the NS1D site in 7002.	This study
pNS1D-sgRNA9A01-rbc162	7002 genome integration plasmid for crRNA targeting P_{rbc}-*ypet* 162 bp upstream of the *ypet* start codon and MCP-$SoxS_{R96A}$ for CRISPRa: sgRNA and MCP-$SoxS_{R96A}$ are expressed from a *tet* promoter and genome-integrated at the NS1D site in 7002.	This study
pNS1D-sgRNA9A01-rbc142	7002 genome integration plasmid for crRNA targeting P_{rbc}-*ypet* 142 bp upstream of the *ypet* start codon and MCP-$SoxS_{R96A}$ for CRISPRa: sgRNA and MCP-$SoxS_{R96A}$ are expressed from a *tet* promoter and genome-integrated at the NS1D site in 7002.	This study
pNS1D-sgRNA9A01-rbc47	7002 genome integration plasmid for crRNA targeting P_{rbc}-*ypet* 47 bp upstream of the *ypet* start codon and MCP-$SoxS_{R96A}$ for CRISPRa: sgRNA and MCP-$SoxS_{R96A}$ are expressed from a *tet* promoter and genome-integrated at the NS1D site in 7002.	This study
pNS1D-sgRNA9A02-cont	7002 genome integration plasmid for randomized crRNA (control) and MCP-SigC for CRISPRa: sgRNA and MCP-SigC are expressed from a *tet* promoter and genome-integrated at the NS1D site in 7002.	This study
pNS1D-sgRNA9A02-rbc373	7002 genome integration plasmid for crRNA targeting P_{rbc}-*ypet* 373 bp upstream of the *ypet* start codon and MCP-SigC for CRISPRa: sgRNA and MCP-SigC are expressed from a *tet* promoter and genome-integrated at the NS1D site in 7002.	This study
pNS1D-sgRNA9A02-rbc338	7002 genome integration plasmid for crRNA targeting P_{rbc}-*ypet* 338 bp upstream of the *ypet* start codon and MCP-SigC for CRISPRa: sgRNA and MCP-SigC are expressed from a *tet* promoter and genome-integrated at the NS1D site in 7002.	This study
pNS1D-sgRNA9A02-rbc297	7002 genome integration plasmid for crRNA targeting P_{rbc}-*ypet* 297 bp upstream of the *ypet* start codon and MCP-SigC	This study

Plasmid Name	Description	Source
	for CRISPRa: sgRNA and MCP-SigC are expressed from a *tet* promoter and genome-integrated at the NS1D site in 7002.	
pNS1D-sgRNA9A02-rbc241	7002 genome integration plasmid for crRNA targeting P_{rbc}-*ypet* 241 bp upstream of the *ypet* start codon and MCP-SigC for CRISPRa: sgRNA and MCP-SigC are expressed from a *tet* promoter and genome-integrated at the NS1D site in 7002.	This study

Table 5-3: List of strains used and constructed in this study.

Strain Name	Description	Source
7002-P_{rbc}-*ypet*	*Synechococcus* sp. PCC 7002 with P_{rbc}-*ypet*-T_{rbc} genome integrated at neutral site 2 (NS2)	Ruffing 2016[86]
7002-P_{rbc}-*ypet*-P*tet*-dCas9	*Synechococcus* sp. PCC 7002 with P_{rbc}-*ypet*-T_{rbc} genome integrated at neutral site 2 (NS2) and P_{tet}-dCas9 integrated at *ascA* (pCas2C).	This study
7002-A01-r-cont	7002-P_{rbc}-*ypet*-P*tet*-dCas9 with P_{tet}-MCP-$SoxS_{R96A}$ and P_{tet}-sgRNA9A-cont integrated at NS1D.	This study
7002-A01-r-749	7002-P_{rbc}-*ypet*-P*tet*-dCas9 with P_{tet}-MCP-$SoxS_{R96A}$ and P_{tet}-sgRNA9A-rbc749 integrated at NS1D.	This study
7002-A01-r-625	7002-P_{rbc}-*ypet*-P*tet*-dCas9 with P_{tet}-MCP-$SoxS_{R96A}$ and P_{tet}-sgRNA9A-rbc625 integrated at NS1D.	This study
7002-A01-r-564	7002-P_{rbc}-*ypet*-P*tet*-dCas9 with P_{tet}-MCP-$SoxS_{R96A}$ and P_{tet}-sgRNA9A-rbc564 integrated at NS1D.	This study
7002-A01-r-482	7002-P_{rbc}-*ypet*-P*tet*-dCas9 with P_{tet}-MCP-$SoxS_{R96A}$ and P_{tet}-sgRNA9A-rbc482 integrated at NS1D.	This study
7002-A01-r-447	7002-P_{rbc}-*ypet*-P*tet*-dCas9 with P_{tet}-MCP-$SoxS_{R96A}$ and P_{tet}-sgRNA9A-rbc447 integrated at NS1D.	This study
7002-A01-r-373	7002-P_{rbc}-*ypet*-P*tet*-dCas9 with P_{tet}-MCP-$SoxS_{R96A}$ and P_{tet}-sgRNA9A-rbc373 integrated at NS1D.	This study
7002-A01-r-338	7002-P_{rbc}-*ypet*-P*tet*-dCas9 with P_{tet}-MCP-$SoxS_{R96A}$ and P_{tet}-sgRNA9A-rbc338 integrated at NS1D.	This study
7002-A01-r-297	7002-P_{rbc}-*ypet*-P*tet*-dCas9 with P_{tet}-MCP-$SoxS_{R96A}$ and P_{tet}-sgRNA9A-rbc297 integrated at NS1D.	This study

Strain Name	Description	Source
7002-A01-r-254	7002-P_{rbc}-*ypet*-P*tet*-dCas9 with P_{tet}-MCP-$SoxS_{R96A}$ and P_{tet}-sgRNA9A-rbc254 integrated at NS1D.	This study
7002-A01-r-241	7002-P_{rbc}-*ypet*-P*tet*-dCas9 with P_{tet}-MCP-$SoxS_{R96A}$ and P_{tet}-sgRNA9A-rbc241 integrated at NS1D.	This study
7002-A01-r-168	7002-P_{rbc}-*ypet*-P*tet*-dCas9 with P_{tet}-MCP-$SoxS_{R96A}$ and P_{tet}-sgRNA9A-rbc168 integrated at NS1D.	This study
7002-A01-r-162	7002-P_{rbc}-*ypet*-P*tet*-dCas9 with P_{tet}-MCP-$SoxS_{R96A}$ and P_{tet}-sgRNA9A-rbc162 integrated at NS1D.	This study
7002-A01-r-142	7002-P_{rbc}-*ypet*-P*tet*-dCas9 with P_{tet}-MCP-$SoxS_{R96A}$ and P_{tet}-sgRNA9A-rbc142 integrated at NS1D.	This study
7002-A01-r-47	7002-P_{rbc}-*ypet*-P*tet*-dCas9 with P_{tet}-MCP-$SoxS_{R96A}$ and P_{tet}-sgRNA9A-rbc47 integrated at NS1D.	This study
7002-A02-r-cont	7002-P_{rbc}-*ypet*-P*tet*-dCas9 with P_{tet}-MCP-SigC and P_{tet}-sgRNA9A-cont integrated at NS1D.	This study
7002-A02-r-373	7002-P_{rbc}-*ypet*-P*tet*-dCas9 with P_{tet}-MCP-SigC and P_{tet}-sgRNA9A-rbc373 integrated at NS1D.	This study
7002-A02-r-338	7002-P_{rbc}-*ypet*-P*tet*-dCas9 with P_{tet}-MCP-SigC and P_{tet}-sgRNA9A-rbc338 integrated at NS1D.	This study
7002-A02-r-297	7002-P_{rbc}-*ypet*-P*tet*-dCas9 with P_{tet}-MCP-SigC and P_{tet}-sgRNA9A-rbc297 integrated at NS1D.	This study
7002-A02-r-241	7002-P_{rbc}-*ypet*-P*tet*-dCas9 with P_{tet}-MCP-SigC and P_{tet}-sgRNA9A-rbc241 integrated at NS1D.	This study

5.4.3. Cultivation and characterization of CRISPRa strains

Standard cultivation conditions for the 7002 CRISPRa strains include growth on A+ medium[102] in an Innova 42R shaking incubator at 30°C and 150 rpm (for a 1 inch orbital) with a photosynthetic light bank containing alternating cool white and plant fluorescent lights at an intensity of approximately 60 μmol photons m^{-2} s^{-1}. For CRISPRa characterization experiments, swabs from three PCR-positive transformant patches on agar plates were inoculated into 16mm test tubes with vented caps containing 4 mL of A+ medium with appropriate antibiotics (40 μg/mL spectinomycin hydrochloride and 50 μg/mL kanamycin monosulfate). The test tubes were cultivated under standard conditions for 4 days. Cell density of the cultures were measured by taking OD_{730} measurements on a LambdaBio spectrophotometer, and the cultures were transferred to two test tubes containing A+ medium with antibiotics so that OD730 = 0.3 and the total culture volume was 4 mL. Anhydrotetracycline (aTc) was added to one test tube at a final concentration of 1 μg/mL.

Test tubes were randomized in the test tube rack within the incubator to address any potential variability due to light gradients within the incubator. After 48 hours of growth under standard conditions, cell density and Ypet fluorescence measurements were acquired. Cell density was estimated based on absorbance at 730 nm of a 200 μL sample in a black 96-well plate with clear bottom (VWR catalog no. 89091-010) using a BioTek Synergy H4 microplate reader. Ypet fluorescence was measured by placing a 200 μL sample in a black 96-well plate (Corning 3631) and measuring fluorescence on a BioTek Synergy H4 microplate reader using an excitation of 500 nm, emission of 530 nm, a gain of 100, and a read height of 8 mm.

This page left blank.

6. TARGETED CRISPRI, PROTEOMICS, AND 13C-MFA IN *SYNECHOCOCCUS* SP. PCC 7002

Contributing Authors: John A. Morgan, Nathaphon Yu Hing King, Melissa Marsing, Tessa Dallo, and Anne M. Ruffing

6.1. Introduction

While the ultimate goal of the CERES project is to conduct high-throughput gRNA library screening of 7002 for enhanced growth phenotypes, we first attempted to generate 7002 strains with enhanced growth phenotypes using targeted CRISPRi. This part of the project was conducted as part of a two-year Academic Alliance (AA) LDRD project with Dr. John Morgan's laboratory at Purdue University.

During the first year of the AA LDRD project, the Morgan laboratory investigated carbon flux changes associated with circadian rhythms in the parental strain, 7002-dCas9. Circadian rhythms are known to influence the gene expression and hence metabolic processes in cyanobacteria.[103-105] Cyanobacterial metabolism is highly variable depending on both the amount of incident light and the time of day, as cyanobacteria must convert the solar energy into chemical energy or dissipate it as photochemical energy, fluorescence, or heat. To investigate the impact of a day-night rhythm, the parental strain 7002-dCas9 was characterized using proteomics and 13C-MFA during two different timepoints in the diurnal light cycle.

The SNL team constructed and characterized CRISPRi strains of 7002-dCas9 with altered growth phenotypes in the first year of this collaboration. Ten 7002 CRISPRi strains were constructing with gRNAs targeting genes predicted to lead to improved growth phenotypes in another model cyanobacterium, *Synechocystis* sp. PCC 6803. A CRISPRi strain with altered growth phenotype was then sent to the Morgan laboratory for 13C-MFA during the second year of the AA LDRD project. This report includes preliminary results from this analysis.

6.2. Results and Discussion

6.2.1. Proteomics and 13C-MFA of 7002 and 7002-dCas9

To probe the impact of the change in incident light during a day-night rhythm, a spectinomycin resistant 7002-dCas9 control strain was grown in an airlift photobioreactor in A+ media with 40 μg/mL spectinomycin and exposed to a 12-hour sinusoidal period of alternating light and dark with a peak light intensity of 500 μmol m^{-2} s^{-1}. Proteomic analysis was conducted on triplicate samples of 7002-dCas9, at 3 hours and 6 hours after the beginning of the light cycle. At 3h, the light intensity was ~350 μmol m^{-2} s^{-1}. The metabolic flux analysis experiments were performed on the same bioreactor cultures immediately after sampling for proteomics.

From the proteomic results, we identified proteins that were significantly different between the 3 h and 6 h time periods (i.e. likely to be related to the circadian rhythm). To determine the proteins that had the most change, we performed ANOVA and filtered for q values < 0.05. This resulted in 101 proteins with significant changes between 3h and 6h. However, this list of proteins did not include any of the initial CRISPRi target genes.

Table 6-1 shows the top 8 proteins in terms of fold change between the 3h and 6 h time points. Several of the highest decreases in protein levels were proteases and a heat shock protein. The protein that went up the most was a putative CO_2 hydration protein.

Table 6-1. List of highest changes in ratio of absolute protein concentrations between 3 hours and 6 hours after initiation of light. The 6 hour represents the highest light intensity of the cycle.

Protein name	Gene name	Fold Change (6h/3h)
Metalloendoprotease, M23/M37 family	SYNPCC7002_A1644	0.016
Uncharacterized protein	SYNPCC7002_A2678	0.03
Uncharacterized protein	SYNPCC7002_A2008	0.036
Small heat shock protein	hspA SYNPCC7002_A0654	0.04
Putative carboxyl-terminal protease	SYNPCC7002_A1844	0.06
Uncharacterized protein	SYNPCC7002_A0092	0.06
CO_2 hydration protein (Putative CO2-hydration protein)	chpY cupA SYNPCC7002_A0174	6.4
Uncharacterized protein	SYNPCC7002_A0175	4.4

Additionally, we conducted isotopically non-stationary flux analysis (INST-MFA) at the 3h and 6 h time points. The flux maps strongly resemble previous flux maps of *Synechocystis* sp. PCC 6803 grown under similar environmental conditions. There was no statistically significant difference in the metabolic fluxes between the 3h and 6h time points. The flux map is shown in Figure 6-1.

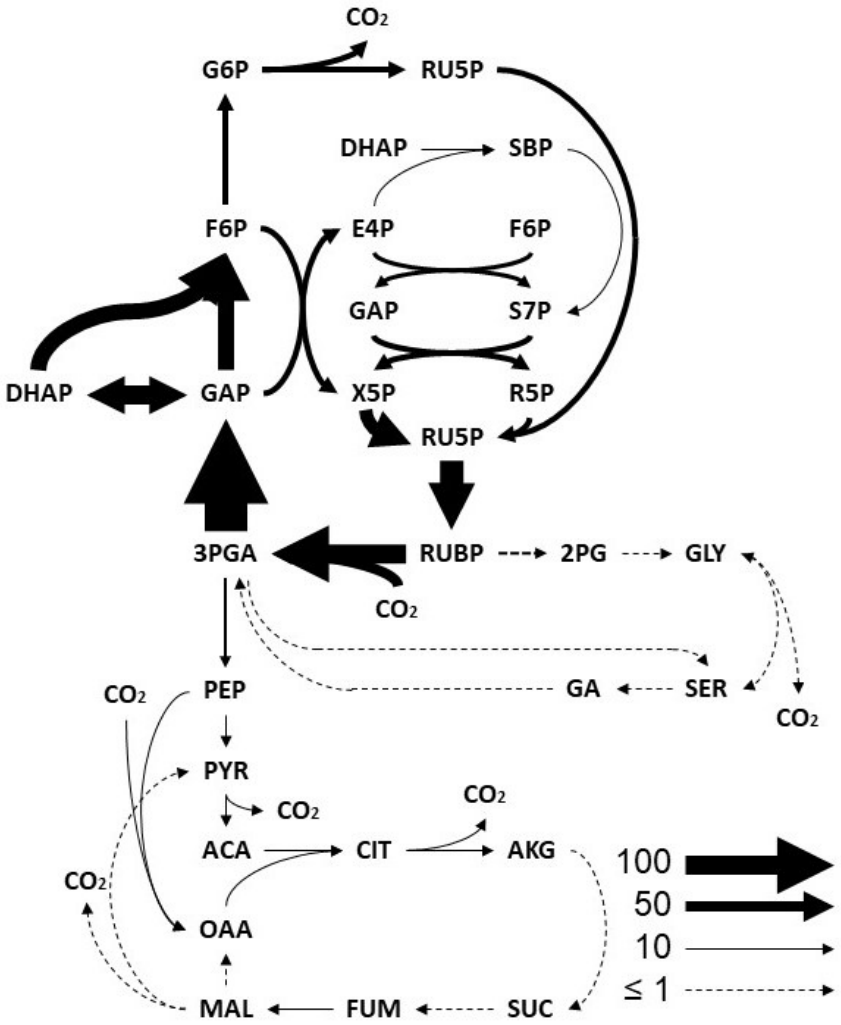


Figure 6-1. Metabolic flux map of 7002-dCas9 strain grown under sinusoidal light conditions. Map corresponds to peak light intensity 500 µmol m^{-2} s^{-1}.

6.2.2. Targeted CRISPRi strain construction and characterization

Previous CRISPR gRNA library screening identified five gRNA targets leading to enhanced growth phenotypes in the freshwater cyanobacterium *Synechocystis* sp. PCC 6803.[52] Additionally, the Morgan laboratory had previously conducted a proteomics study with *Synechocystis* sp. PCC 6803 that identified six genes that were negatively correlated with growth (unpublished data). Using BlastP, we identified homologs in 7002 for ten of these eleven genes (Table 6-2). CRISPRi strains for each gene target in 7002 were constructed by designing gRNAs targeting within the first 60 coding nucleotides from the start codon (Table A-3) and integrating the gRNA construct into the genome of 7002-dCas9 at NSD1. The resulting CRISPRi strains were tested under both continuous and diurnal light conditions to identify changes in growth phenotype.

Table 6-2. *Synechocystis* sp. PCC 6803 genes shown or predicted to improve growth and the corresponding homologs in 7002.

6803 gene locus	7002 homolog	E-value	Function
sll1968	Synpcc7002_A0056	2e-35	Regulates photosystem stoichiometry (photomixotrophic growth related protein)
slr1916	Synpcc7002_A1594	7e-104	Esterase
sll1969	Synpcc7002_A1441	8e-66	Triacylglycerol lipase
slr1340	No homolog	N/A	Predicted acetyl-transferase
ssl2982	Synpcc7002_A0614	6e-29	ω subunit of RNA polymerase
slr0943	Synpcc7002_A0010	N/A	Fructose bisphosphate aldolase class I
slr1106	Synpcc7002_A2606	7e-123	Prohibitin
slr0455	Synpcc7002_A1240	7e-29	Conserved hypothetical protein
slr1641	Synpcc7002_A2353	0.0	Endopeptidase Clp ATP-binding chain B1
slr0435	Synpcc7002_A0052	1e-45	Acetyl-CoA carbosylase, biotin carboxyl carrier protein
sll1021	Synpcc7002_A2510	4e-152	Hypothetical membrane protein

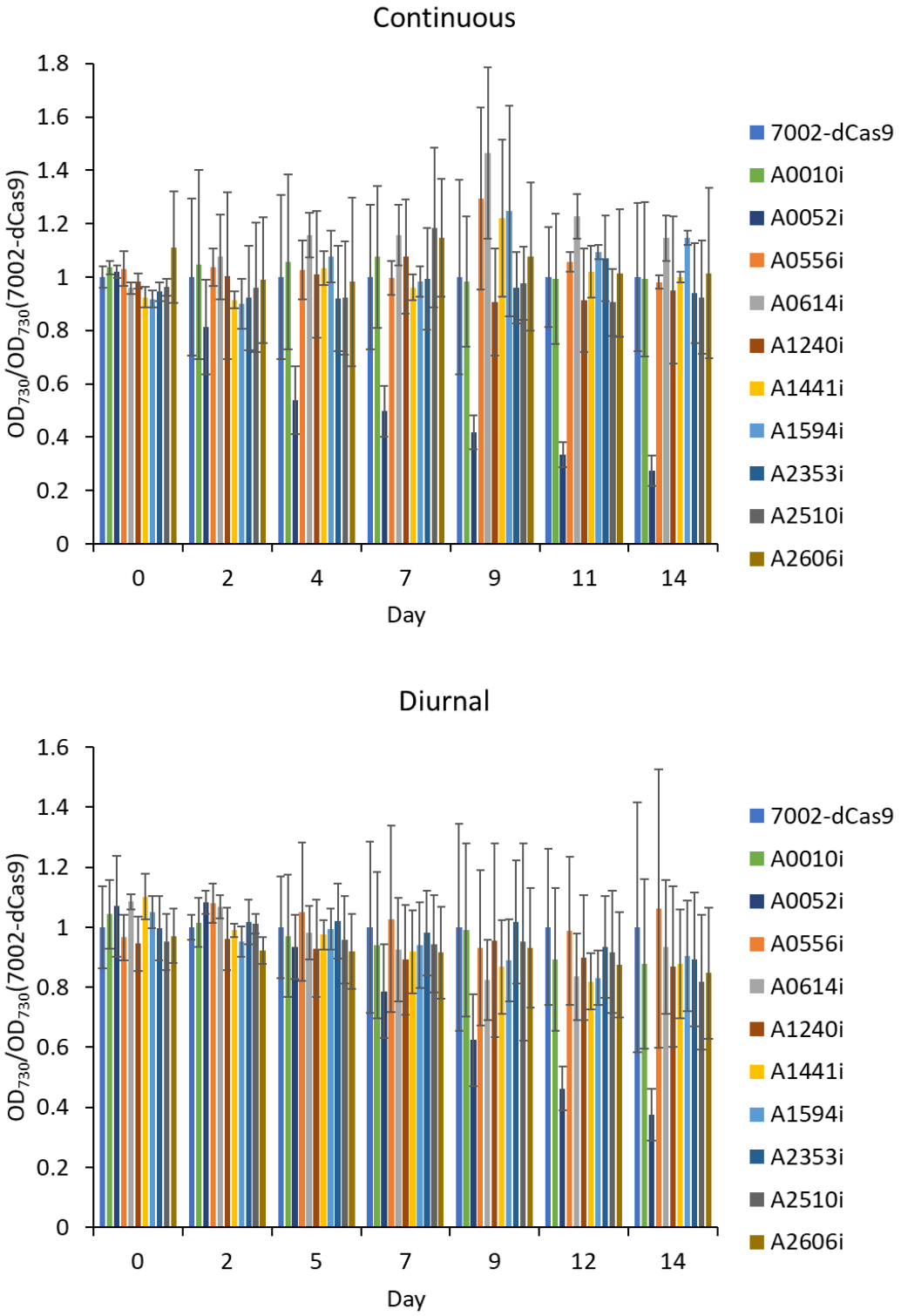


Figure 6-2. Relative growth of targeted CRISPRi strains and parental strain (7002-dCas9) under continuous (top) and 12:12 diurnal (bottom) light conditions. The OD_{730} measurement for each strain was normalized to the OD_{730} of the parental strain. Data are averages from two biological transformants with error bars indicating the standard deviation.

No CRISPRi strain showed significant or repeatable enhanced growth relative to the parental strain under either continuous or diurnal light conditions (Figure 6-2). In fact, one strain, A0052i, showed significantly reduced growth under both conditions. The gRNA in A0052i targets acetyl-CoA carboxylase biotin carboxyl carrier protein, which is involved in fatty acid biosynthesis. These CRISPRi strains were sent to the Morgan laboratory at Purdue for characterization and 13C-MFA.

6.2.3. Characterization of CRISPRi strains at Purdue and 13C-MFA

We studied the growth of several CRISPRi engineered strains provided by Dr. Ruffing at SNL. At Purdue, we grew the strain in shake flasks at 30°C and 200 rpm under 200 μmol m^{-2} sec^{-1} of cool white light. The growth was measured by absorbance under both continuous and diurnal light conditions. The results in Figure 6-3 show that under continuous light, the inhibition of A0556i, has reduced growth. The A0556 target gene encodes for *pmgA*, a protein implicated in regulating photosystem stoichiometry under photomixotrophic growth. Under the same environmental conditions, we grew the cells under 12-hour diurnal light/dark cycles. However, as compared to the continuous light, all Syn 7002 CRISPRi strains grew at a slightly lower rate than the wild-type, but with no statistical differences between them (Figure 6-4). This conflicts with growth studies conducted at SNL (Figure 6-2) and could possibly be due to instability of the dCas9 or gRNA integration (preliminary results of instability obtained by SNL).

At the time of writing this report, the ^{13}C labeling experiment was completed for the strain 7002-A0556i, but the results are still being analyzed. If these results provide insight into the reduced growth of strain 7002-A0556i, they will be included in a future publication. Figure 6-5 shows the quantification of the M0 peak for Fructose 6-P, a central metabolite in the Calvin cycle. As expected, the fraction of the unlabeled isotopomer decreases rapidly after 13C bicarbonate was added.

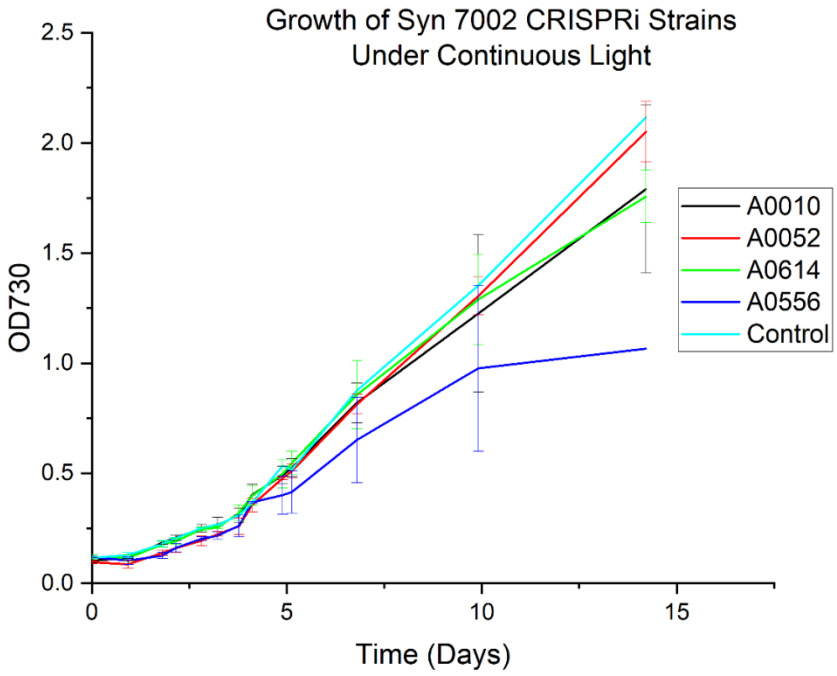


Figure 6-3 Shake flask growth of *Syncechococcus* sp. PCC 7002 CRISPRi lines. The cells were grown at 30°C and 200 RPM under 200 µmol m^{-2} sec^{-1} of cool white light. The results are from duplicate flasks.

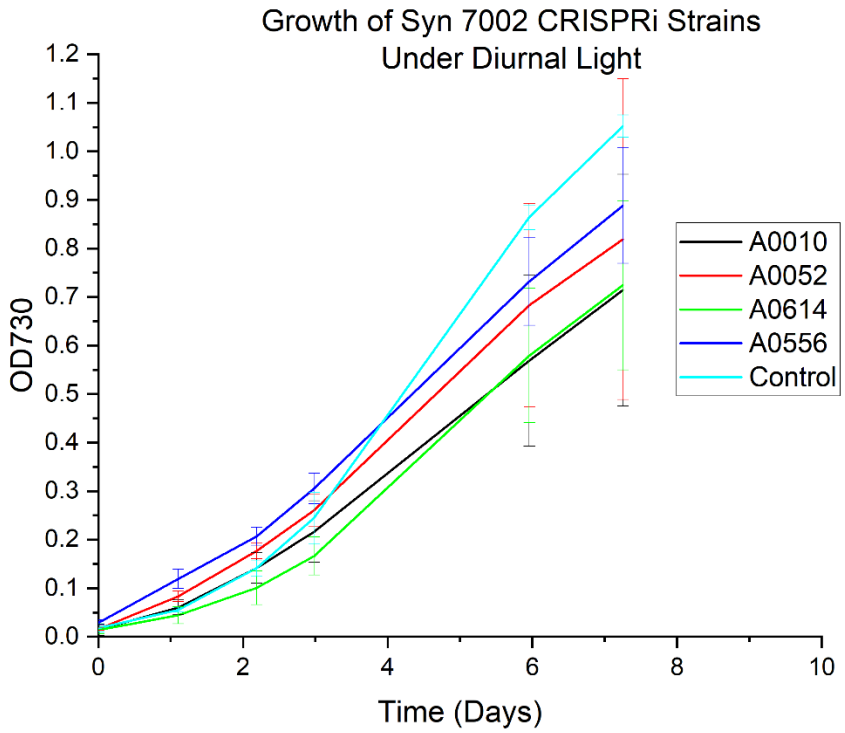


Figure 6-4. Growth of Syn 7002 CRISPRi strains under diurnal light conditions. The cells were grown at 30°C and 200 RPM under 200 µmol m^{-2} sec^{-1} of cool white light. The results are from duplicate flasks.

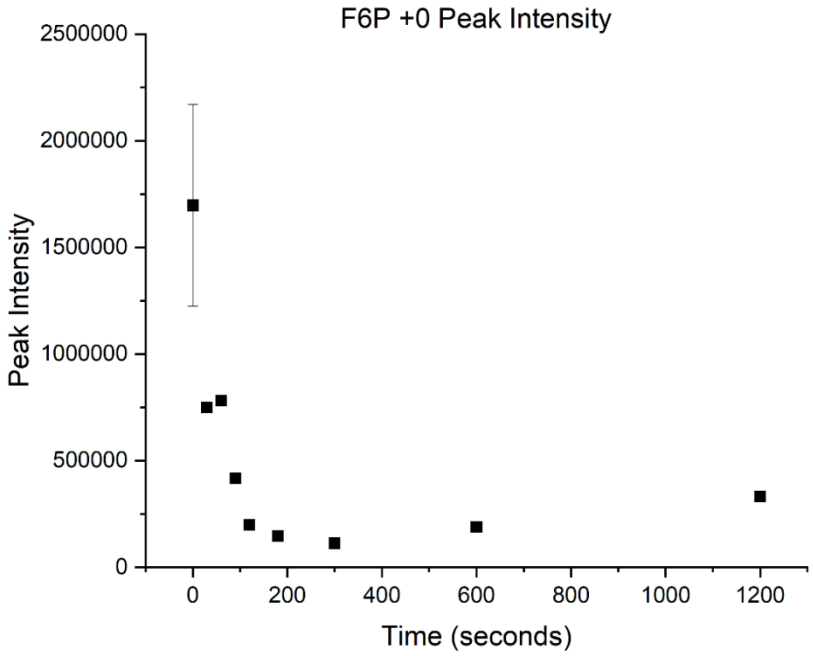


Figure 6-5. The time course for the mass isotopomer of unlabeled F6P after addition of ^{13}C bicarbonate. The y-axis is in arbitrary units representing the area under the curve of F6P.

6.3. Conclusions

In this AA LDRD collaboration, we identified several proteins in 7002 that change significantly between the 3h and 6h timepoint after lights are initiated in a 12:12 light:dark cycle. Significantly downregulated proteins include proteases and a heat shock protein, which may be needed at the 3h timepoint as the cells adapt to light conditions after the dark period. Under peak light conditions (6h), a putative CO_2 hydration protein was significantly up-regulated; this protein may contribute to high photosynthetic CO_2 fixation under these high light conditions relative to the 3h timepoint. Despite these changes in protein levels, no significant change in carbon flux was detected using INST-MFA.

We also investigated whether gene knockdowns shown or predicted to improve growth in *Synechocystis* sp. PCC 6803 would lead to enhanced growth in 7002. None of the ten CRISPRi strains showed improved growth relative to the 7002-dCas9 parental strain. In fact, knockdown of Synpcc7002_A0052 and Synpcc7002_A0556 resulted in reduced growth. As Synpcc7002-A0052 encodes the acetyl-CoA carboxylase biotin carboxyl carrier protein involved in fatty acid biosynthesis, knockdown of this gene would therefore be expected to reduce growth, for fatty acids are essential for cell and thylakoid membrane biosynthesis. Synpcc7002_A0556 encodes a gene involved in regulating photosystem stoichiometry; therefore, it is surprising that growth inhibition was detected under continuous light conditions and not diurnal light conditions. INST-MFA results for A0556i are currently being analyzed to provide insight into the mechanism for reduced growth in this strain. Unexpectedly, growth phenotypes of the CRISPRi strains varied between the two different sites (SNL and Purdue). Repeated growth experiments at SNL resulted in loss of the dCas9

construct for some of the CRISPRi strains, as verified through PCR. Therefore, strain instability may explain the variability between growth at the different sites.

6.4. Materials and Methods

6.4.1. Proteomics sample preparation

Immediately prior to 13C labelling experiments, a 20 mL aliquot of culture was withdrawn and pelleted for use in non-targeted proteomics. Cell pellets were washed twice in 20 mM PBS and then stored at -80 °C until sample preparation for proteomic analysis. Cell pellets were re-suspended in 400 μL of 100 mM ammonium bicarbonate containing 1 mM phenylmethylsulfonyl fluoride and homogenized at high pressure (20,000 psi) at 6500 rpm for 90 seconds (Bertin Technologies SAS). The homogenate was transferred to new tubes and centrifuged at 13,500 rpm for 15 min at 4 °C. The supernatant was transferred to a new tube and treated as a soluble fraction, and the pellets were treated as insoluble fraction. Proteins in the soluble fractions were precipitated overnight at -20°C using 4 volume of cold (-20°C) acetone and then were pelleted by centrifugation. The soluble and insoluble fractions were dissolved in 8 M urea at room temperature for 1 h with continuous vortexing. The protein concentration was determined by Bicinchoninic Acid (BCA assay and 50 μg of the protein (equivalent volume) was used for proteomics sample preparation. Samples were incubated in 10 mM dithiothreitol at 37 °C for an hour for reduction followed by incubation in 20 mM iodoacetamide for an hour at room temperature in the dark for alkylation. Samples were digested using Trypsin/LysC protease mix at 1:25 (enzyme to substrate) ratio at 37°C. After 3 h of digestion, additional Trypsin/LysC protease mix was added at 1:50 (enzyme to substrate) ratio and digestion was allowed to proceed overnight at 37°C. Peptides were desalted using C18 micro spin desalting columns (The Nest Group, Inc.). Eluted peptides were dried and re-suspended in 3% acetonitrile/0.1% formic acid to a final concentration of 1 μg/μL and 1 μL (1 μg) was loaded used for LC-MS/MS analysis.

6.4.2. LC-MS/MS analysis for peptide sequencing

Peptides were analyzed in an Dionex UltiMate 3000 RSLC nano System coupled on-line to Q Exactive Orbitrap High Field Hybrid Quadrupole Mass Spectrometer (Thermo Fisher Scientific, Waltham, MA, USA) as described previously.[106] Briefly, reverse phase peptide separation was accomplished using a trap column (300 μm ID × 5 mm) packed with 5 μm 100 Å PepMap C18 medium coupled to a 50-cm long × 75 μm inner diameter analytical column packed with 2 μm 100 Å PepMap C18 silica (Thermo Fisher Scientific). The column temperature was maintained at 50°C. Sample was loaded to the trap column at a flow rate of 5 μL/min and eluted from the analytical column at a flow rate of 300 nL/min using a 120-min LC gradient. The column was washed and equilibrated by using three 30 min LC gradient before injecting next sample. Precursor ion (MS1) scans were collected at a resolution of 120,000 and MS/MS scans at a resolution of 15,000 at 200 m/z in data dependent acquisition mode. LC-MS/MS data were analyzed using MaxQuant (version 1.6.3.4) against the *Synechococcus* sp. PCC 7002 database downloaded from the NCBI (www.ncbi.nlm.nih.gov). We edited the following parameters for our search: precursor mass tolerance of 10 ppm; enzyme specificity of trypsin/Lys-C enzyme allowing up to 2 missed cleavages; oxidation of methionine (M) as a variable modification and carbamidomethylation (C) as a fixed modification. False discovery rate (FDR) of peptide spectral match (PSM) and protein identification was set to 0.01. Proteins with LFQ > 0 and MS/MS (spectral counts) ≥ 3 were considered as true identification and used for further downstream analysis.

To compare data between the 3h and 6h time point, the label free quantification (LFQ) intensities were used. The intensities of the soluble and insoluble fractions were combined for each replicate. InfernoRDN (omics.pnl.gov/software/infernordn) was used to perform a Kruskal-Walis test to determine p-values and associated q-values to correct for background false discovery rate.

6.4.3. Metabolic Flux Analysis

Isotopically nonstationary metabolic flux analysis (INST-MFA) is a method for estimation of flux distribution throughout photoautotrophic central carbon metabolism by tracking the transient incorporation of ^{13}C into cellular metabolism. $NaH^{13}CO_3$ in solution was injected into the bioreactor to a total concentration of 1 g $NaH^{13}CO_3$/L to quickly introduce ^{13}C into cellular metabolism during exponential phase. Culture samples (10 mL) were removed and quenched at separate time intervals (0, 0.5, 1, 1.5, 2, 3, 5, 10 and 20 minutes) into 40 mL pre-chilled 60% methanol in a dry ice-ethanol bath and rapidly pelleted at 8000 g, -20 °C and then extracted. Pellets were extracted using 500 µL chloroform/methanol (3:7 v/v) under -20 °C for 2 hours followed by two rounds of 500 µL 0.2 M NH_4OH. The upper methanol-water layer was collected. Samples were dried under N_2 and then reconstituted in 60 µL ultrapure water. An QTRAP 5500 (AB Sciex) linked to HPLC (Shimadzu) was used to quantify the mass isotopomers of central carbon metabolites using an established ion-pairing method[107] using a Polaris C18 column (150 x 2 mm) (Agilent, CA). The MATLAB-based software INCA v1.8 was used to estimate pathway fluxes at each light condition.[108] Flux estimates were only accepted if the calculated χ^2 goodness of fit is within the expected 95% confidence interval. Upper and lower bounds on fluxes were estimated using the parameter continuation method.

6.4.4. CRISPRi strain construction

The targeted CRISPRi strains listed in **Error! Reference source not found.** were constructed using the 7002-dCas9 parental strain and genome integration vector for gRNA expression (pNS1D-sgRNA9-SpR) developed in Section 0. The oligos used to construct each gRNA plasmid are listed in Table A-3, and the golden gate assembly method described in Ran *et. al.* was used to anneal the oligos and integrate them into the gRNA expression plasmid.[101] Successful integration of the gRNA sequence was confirmed using Sanger sequencing from Eurofins Genomics. The gRNA plasmids were linearized by digestion with AatII, and transformed into 7002-dCas9 using the method described previously in Ruffing *et. al.* with slight modification.[86] 7002-dCas9 was grown to an OD_{730} of 0.5, measured using a PerkinElmer LambdaBio spectrophotometer. This culture was concentrated to an OD_{730} of 2.0 in 1 mL volume, and 1 µg of digested plasmid was added to the concentrated culture. The transformation culture was placed in a 16 mm test tube with vented cap and incubated at 30°C and 173 rpm with 60 µmol photons m^{-2} s^{-1} of light in an Innova 42R shaking incubator with photosynthetic light bank and ¾ inch orbital. After 24 hours of incubation, the transformation culture was concentrated and spread on medium A+ agar plates with 40 µg/mL of spectinomycin. Transformation plates were incubated under the same light and temperature conditions until colonies appeared, typically within 5 to 7 days. Each colony was restreaked on solid selection media until homozygous transformants were obtained, as verified by PCR screening of the NS1D site. Two PCR-positive transformants for each strain were used in the subsequent growth characterization experiments.

Table 6-3. Targeted CRISPRi strains constructed in this study.

Strain	Description	Reference
7002-dCas9	*Synechococcus* sp. PCC 7002 with dCas9 integrated at *ascA* under control of the aTc inducible *tet* promoter	Section 0 (this report)
7002-dCas9-cont	7002-dCas9 with a random control gRNA integrated at NS1D under control of the aTc inducible *tet* promoter	Section 0 (this report)
A0010i	7002-dCas9 with a gRNA targeting Synpcc7002_A0010 integrated at NS1D under control of the aTc inducible *tet* promoter	This study
A0052i	7002-dCas9 with a gRNA targeting Synpcc7002_A0052 integrated at NS1D under control of the aTc inducible *tet* promoter	This study
A0556i	7002-dCas9 with a gRNA targeting Synpcc7002_A0556 integrated at NS1D under control of the aTc inducible *tet* promoter	This study
A0614i	7002-dCas9 with a gRNA targeting Synpcc7002_A0614 integrated at NS1D under control of the aTc inducible *tet* promoter	This study
A1240i	7002-dCas9 with a gRNA targeting Synpcc7002_A1240 integrated at NS1D under control of the aTc inducible *tet* promoter	This study
A1441i	7002-dCas9 with a gRNA targeting Synpcc7002_A1441 integrated at NS1D under control of the aTc inducible *tet* promoter	This study
A1594i	7002-dCas9 with a gRNA targeting Synpcc7002_A1594 integrated at NS1D under control of the aTc inducible *tet* promoter	This study
A2353i	7002-dCas9 with a gRNA targeting Synpcc7002_A2353 integrated at NS1D under control of the aTc inducible *tet* promoter	This study
A2510i	7002-dCas9 with a gRNA targeting Synpcc7002_A2510 integrated at NS1D under control of the aTc inducible *tet* promoter	This study

Strain	Description	Reference
A2606i	7002-dCas9 with a gRNA targeting Synpcc7002_A2606 integrated at NS1D under control of the aTc inducible *tet* promoter	This study

6.4.5. CRISPRi strain characterization

The CRISPRi strains were characterized for changes in growth phenotype compared to the control strain containing a randomized gRNA sequence (7002-dCas9-cont). For each strain, two PCR positive transformants were inoculated in a test tube containing 4 mL of medium A+ with 40 μg/mL of spectinomycin. The seed cultures were incubated at 30°C and 173 rpm with 60 μmol photons m^{-2} s^{-1} in an Innova 42R incubator with ¾ inch orbital and photosynthetic light bank. After 4 to 7 days of cultivation, the seed cultures were diluted to an OD_{730} of 0.3 in fresh medium A+ with antibiotics and 1 μg/mL of the aTc inducer. The cultures were incubated at the aforementioned conditions, and cell density was estimated from OD_{730} measurements of 200 μL of culture (diluted to maintain within the linear range) in a black clear-bottom 96-well microplate analyzed on a BioTek Cytation5 microplate reader.

7. CONCLUSIONS AND FUTURE WORK

Contributing Author: Anne M. Ruffing

7.1. CERES project conclusions

The CERES project developed three tools to enable future efforts for CRISPR gRNA library screening in 7002, and these tools are likely to have impact beyond just strain development in 7002.

First, Operon-SEQer was developed to improve operon prediction in bacteria using short-read RNA-seq data. Operon-SEQer may be used to improve operon prediction in 7002 for designing genome-wide gRNA libraries, and it is broadly applicable to improve operon prediction in any bacterium. By using long-read RNA-seq data available for *E. coli*, we demonstrated that Operon-SEQer predicts operon gene pairs that are missed by the standard methods for operon prediction. The code for Operon-SEQer was deposited in GitHub so that it is available for community use, and additional details on Operon-SEQer may be found in our manuscript on bioRxiv, which is currently under review at PLOS Computational Biology.[109]

Second, we generated gRNA design rules for CRISPRi in 7002 using high-density gRNA tiling to evaluate three reporter operons in 7002 and correlation analysis of the resulting datasets to identify important features. The datasets were also used to predict gRNA effectiveness using machine learning. These machine learning algorithms may be applied to improve the design of gRNA libraries for CRISPRi in 7002. Additionally, the gRNA design rules are likely to be applicable to other cyanobacteria and prokaryotes in general. This study explored gRNAs targeting regions that have largely been ignored in previous studies, including the region upstream of the promoter, the promoter, and the terminator. Interestingly, we found that CRISPRi can activate gene expression if the gRNA is targeted upstream of the promoter region, although the level of activation is low. This study advances our understanding of gRNA effectiveness for CRISPRi in prokaryotes.

Third, we developed a shuttle vector for recombinant gene expression in 7002. This *E. coli*-7002 shuttle vector may be used to express gRNA libraries. Moreover, we have shown that gene expression from the shuttle vector is higher than genome integration of the expression construct. Therefore, the shuttle vector may be used in future 7002 strain development efforts in which high recombinant gene expression is desirable. Furthermore, this plasmid-based expression system enables more efficient strain construction in 7002 compared to the traditional approach of genome integration, as it does not require additional growth to isolate homozygous transformants.

7.2. Future Work

The tools established in CERES will enable future studies of CRISPRi gRNA library screening in 7002. Operon-SEQer and machine learning algorithms using the high-density gRNA tiling data will be used to design effective gRNA libraries. The crRNA oligonucleotides will be inserted into the *E. coli*-7002 shuttle vector, and after transformation of the gRNA library into 7002-dCas9, the cultures will be screened under different environmental conditions to identify gene knockdown targets that improve growth under these conditions.

In addition to high-throughput screening for gene knockdown targets, gRNA libraries may also be used to screen for gene activation targets if CRISPRa can be established in 7002. Based on preliminary work in the CERES project on attempting to establish CRISPRa, we propose a high-throughput, combinatorial screening of both transcriptional activators and gRNA locations to identify candidates for CRISPRa in 7002. When CRISPRa is established, it may be used

independently or simultaneously with the CRISPRi system to perform high-throughput screening in 7002. By developing a high-throughput CRISPRi/a screening platform in 7002, this marine cyanobacterium may be optimized to enable large-scale photosynthetic conversion of CO_2 into useful biofuels and biochemicals. This cyanobacterial production system will contribute to the U.S. bioeconomy, advance energy independence, and reduce greenhouse gas emissions to mitigate climate change.

8. REFERENCES

1. Deng, M.-D. and J.R. Coleman, *Ethanol synthesis by genetic engineering in cyanobacteria.* Applied and environmental microbiology, 1999. 65(2): p. 523-528.
2. Dexter, J. and P. Fu, *Metabolic engineering of cyanobacteria for ethanol production.* Energy & Environmental Science, 2009. 2(8): p. 857-864.
3. Gao, Z., H. Zhao, Z. Li, X. Tan, and X. Lu, *Photosynthetic production of ethanol from carbon dioxide in genetically engineered cyanobacteria.* Energy & Environmental Science, 2012. 5(12): p. 9857-9865.
4. McEwen, J.T., M. Kanno, and S. Atsumi, *2, 3 Butanediol production in an obligate photoautotrophic cyanobacterium in dark conditions via diverse sugar consumption.* Metabolic engineering, 2016. 36: p. 28-36.
5. Nozzi, N.E. and S. Atsumi, *Genome engineering of the 2, 3-butanediol biosynthetic pathway for tight regulation in cyanobacteria.* ACS Synthetic Biology, 2015. 4(11): p. 1197-1204.
6. Nozzi, N.E., A.E. Case, A.L. Carroll, and S. Atsumi, *Systematic approaches to efficiently produce 2, 3-butanediol in a marine cyanobacterium.* ACS Synthetic Biology, 2017. 6(11): p. 2136-2144.
7. Oliver, J.W., I.M. Machado, H. Yoneda, and S. Atsumi, *Cyanobacterial conversion of carbon dioxide to 2, 3-butanediol.* Proceedings of the National Academy of Sciences, 2013. 110(4): p. 1249-1254.
8. Oliver, J.W., I.M. Machado, H. Yoneda, and S. Atsumi, *Combinatorial optimization of cyanobacterial 2, 3-butanediol production.* Metabolic engineering, 2014. 22: p. 76-82.
9. Savakis, P.E., S.A. Angermayr, and K.J. Hellingwerf, *Synthesis of 2, 3-butanediol by Synechocystis sp. PCC6803 via heterologous expression of a catabolic pathway from lactic acid-and enterobacteria.* Metabolic engineering, 2013. 20: p. 121-130.
10. Kageyama, H., R. Waditee-Sirisattha, S. Sirisattha, Y. Tanaka, A. Mahakhant, and T. Takabe, *Improved alkane production in nitrogen-fixing and halotolerant cyanobacteria via abiotic stresses and genetic manipulation of alkane synthetic genes.* Current microbiology, 2015. 71(1): p. 115-120.
11. Wang, W., X. Liu, and X. Lu, *Engineering cyanobacteria to improve photosynthetic production of alka (e) nes.* Biotechnology for biofuels, 2013. 6(1): p. 1-9.
12. Liu, X., J. Sheng, and R. Curtiss III, *Fatty acid production in genetically modified cyanobacteria.* Proceedings of the National Academy of Sciences, 2011. 108(17): p. 6899-6904.
13. Ruffing, A.M., *RNA-Seq analysis and targeted mutagenesis for improved free fatty acid production in an engineered cyanobacterium.* Biotechnology for biofuels, 2013. 6(1): p. 113.
14. Ruffing, A.M., *Borrowing genes from Chlamydomonas reinhardtii for free fatty acid production in engineered cyanobacteria.* Journal of applied phycology, 2013. 25(5): p. 1495-1507.
15. Ruffing, A.M., *Improved free fatty acid production in cyanobacteria with Synechococcus sp. PCC 7002 as host.* Frontiers in bioengineering and biotechnology, 2014. 2: p. 17.
16. Ruffing, A.M. and H.D. Jones, *Physiological effects of free fatty acid production in genetically engineered Synechococcus elongatus PCC 7942.* Biotechnology and bioengineering, 2012. 109(9): p. 2190-2199.
17. Bentley, F.K., A. Zurbriggen, and A. Melis, *Heterologous expression of the mevalonic acid pathway in cyanobacteria enhances endogenous carbon partitioning to isoprene.* Molecular plant, 2014. 7(1): p. 71-86.
18. Chaves, J.E. and A. Melis, *Engineering isoprene synthesis in cyanobacteria.* FEBS letters, 2018. 592(12): p. 2059-2069.
19. Chaves, J.E., P. Rueda-Romero, H. Kirst, and A. Melis, *Engineering isoprene synthase expression and activity in cyanobacteria.* ACS Synthetic Biology, 2017. 6(12): p. 2281-2292.
20. Davies, F.K., V.H. Work, A.S. Beliaev, and M.C. Posewitz, *Engineering limonene and bisabolene production in wild type and a glycogen-deficient mutant of Synechococcus sp. PCC 7002.* Frontiers in bioengineering and biotechnology, 2014. 2: p. 21.

21. Gao, X., F. Gao, D. Liu, H. Zhang, X. Nie, and C. Yang, *Engineering the methylerythritol phosphate pathway in cyanobacteria for photosynthetic isoprene production from CO 2.* Energy & Environmental Science, 2016. 9(4): p. 1400-1411.
22. Kiyota, H., Y. Okuda, M. Ito, M.Y. Hirai, and M. Ikeuchi, *Engineering of cyanobacteria for the photosynthetic production of limonene from CO2.* Journal of biotechnology, 2014. 185: p. 1-7.
23. Lin, P.-C., R. Saha, F. Zhang, and H.B. Pakrasi, *Metabolic engineering of the pentose phosphate pathway for enhanced limonene production in the cyanobacterium Synechocysti s sp. PCC 6803.* Scientific reports, 2017. 7(1): p. 1-10.
24. Lin, P.-C., F. Zhang, and H.B. Pakrasi, *Enhanced limonene production in a fast-growing cyanobacterium through combinatorial metabolic engineering.* Metabolic Engineering Communications, 2021. 12: p. e00164.
25. Lindberg, P., S. Park, and A. Melis, *Engineering a platform for photosynthetic isoprene production in cyanobacteria, using Synechocystis as the model organism.* Metabolic engineering, 2010. 12(1): p. 70-79.
26. Abramson, B.W., J. Lensmire, Y.-T. Lin, E. Jennings, and D.C. Ducat, *Redirecting carbon to bioproduction via a growth arrest switch in a sucrose-secreting cyanobacterium.* Algal Research, 2018. 33: p. 248-255.
27. Du, W., F. Liang, Y. Duan, X. Tan, and X. Lu, *Exploring the photosynthetic production capacity of sucrose by cyanobacteria.* Metabolic engineering, 2013. 19: p. 17-25.
28. Ducat, D.C., J.A. Avelar-Rivas, J.C. Way, and P.A. Silver, *Rerouting carbon flux to enhance photosynthetic productivity.* Applied and environmental microbiology, 2012. 78(8): p. 2660-2668.
29. Kirsch, F., Q. Luo, X. Lu, and M. Hagemann, *Inactivation of invertase enhances sucrose production in the cyanobacterium Synechocystis sp. PCC 6803.* Microbiology, 2018. 164(10): p. 1220-1228.
30. Lin, P.-C., F. Zhang, and H.B. Pakrasi, *Enhanced production of sucrose in the fast-growing cyanobacterium Synechococcus elongatus UTEX 2973.* Scientific reports, 2020. 10(1): p. 1-8.
31. Ansari, S. and T. Fatma, *Cyanobacterial polyhydroxybutyrate (PHB): screening, optimization and characterization.* PLoS One, 2016. 11(6): p. e0158168.
32. Arisaka, S., N. Terahara, A. Oikawa, and T. Osanai, *Increased polyhydroxybutyrate levels by ntcA overexpression in Synechocystis sp. PCC 6803.* Algal Research, 2019. 41: p. 101565.
33. Costa, J.A.V., J.B. Moreira, B.F. Lucas, V.d.S. Braga, A.P.A. Cassuriaga, and M.G.d. Morais, *Recent advances and future perspectives of PHB production by cyanobacteria.* Industrial Biotechnology, 2018. 14(5): p. 249-256.
34. Hondo, S., M. Takahashi, T. Osanai, et al., *Genetic engineering and metabolite profiling for overproduction of polyhydroxybutyrate in cyanobacteria.* Journal of bioscience and bioengineering, 2015. 120(5): p. 510-517.
35. Roh, H., J.S. Lee, H.I. Choi, Y.J. Sung, S.Y. Choi, H.M. Woo, and S.J. Sim, *Improved CO2-derived polyhydroxybutyrate (PHB) production by engineering fast-growing cyanobacterium Synechococcus elongatus UTEX 2973 for potential utilization of flue gas.* Bioresource technology, 2021. 327: p. 124789.
36. Yashavanth, P., M. Das, and S.K. Maiti, *Recent progress and challenges in cyanobacterial autotrophic production of polyhydroxybutyrate (PHB), a bioplastic.* Journal of Environmental Chemical Engineering, 2021: p. 105379.
37. Fontana, J., C. Dong, C. Kiattisewee, V.P. Chavali, B.I. Tickman, J.M. Carothers, and J.G. Zalatan, *Effective CRISPRa-mediated control of gene expression in bacteria must overcome strict target site requirements.* Nature Communications, 2020. 11(1): p. 1-11.
38. Ho, H.I., J.R. Fang, J. Cheung, and H.H. Wang, *Programmable CRISPR-Cas transcriptional activation in bacteria.* Molecular systems biology, 2020. 16(7): p. e9427.

39. Peters, J.M., B.-M. Koo, R. Patino, et al., *Enabling genetic analysis of diverse bacteria with Mobile-CRISPRi.* Nature Microbiology, 2019. 4(2): p. 244-250.
40. Todor, H., M.R. Silvis, H. Osadnik, and C.A. Gross, *Bacterial CRISPR screens for gene function.* Current opinion in microbiology, 2021. 59: p. 102-109.
41. Wang, T., C. Guan, J. Guo, et al., *Pooled CRISPR interference screening enables genome-scale functional genomics study in bacteria with superior performance.* Nature Communications, 2018. 9(1): p. 1-15.
42. Wang, T., J. Guo, C. Guan, et al., *Pooled CRISPR interference screens enable high-throughput functional genomics study and elucidate new rules for guide RNA library design in Escherichia coli.* bioRxiv, 2017.
43. Zhang, R., W. Xu, S. Shao, and Q. Wang, *Gene Silencing Through CRISPR Interference in Bacteria: Current Advances and Future Prospects.* Frontiers in Microbiology, 2021. 12: p. 567.
44. Behler, J., D. Vijay, W.R. Hess, and M.K. Akhtar, *CRISPR-based technologies for metabolic engineering in cyanobacteria.* Trends in Biotechnology, 2018. 36(10): p. 996-1010.
45. Gordon, G.C., T.C. Korosh, J.C. Cameron, A.L. Markley, M.B. Begemann, and B.F. Pfleger, *CRISPR interference as a titratable, trans-acting regulatory tool for metabolic engineering in the cyanobacterium Synechococcus sp. strain PCC 7002.* Metabolic engineering, 2016. 38: p. 170-179.
46. Higo, A., A. Isu, Y. Fukaya, S. Ehira, and T. Hisabori, *Application of CRISPR interference for metabolic engineering of the heterocyst-forming multicellular cyanobacterium Anabaena sp. PCC 7120.* Plant and Cell Physiology, 2018. 59(1): p. 119-127.
47. Huang, C.-H., C.R. Shen, H. Li, L.-Y. Sung, M.-Y. Wu, and Y.-C. Hu, *CRISPR interference (CRISPRi) for gene regulation and succinate production in cyanobacterium S. elongatus PCC 7942.* Microbial cell factories, 2016. 15(1): p. 1-11.
48. Li, H., C.R. Shen, C.-H. Huang, L.-Y. Sung, M.-Y. Wu, and Y.-C. Hu, *CRISPR-Cas9 for the genome engineering of cyanobacteria and succinate production.* Metabolic engineering, 2016. 38: p. 293-302.
49. Ungerer, J. and H.B. Pakrasi, *Cpf1 Is a versatile tool for CRISPR genome editing across diverse species of cyanobacteria.* Scientific reports, 2016. 6: p. 39681.
50. Wendt, K.E., J. Ungerer, R.E. Cobb, H. Zhao, and H.B. Pakrasi, *CRISPR/Cas9 mediated targeted mutagenesis of the fast growing cyanobacterium Synechococcus elongatus UTEX 2973.* Microbial cell factories, 2016. 15(1): p. 1-8.
51. Yao, L., I. Cengic, J. Anfelt, and E.P. Hudson, *Multiple gene repression in cyanobacteria using CRISPRi.* ACS Synthetic Biology, 2016. 5(3): p. 207-212.
52. Yao, L., K. Shabestary, S.M. Björk, J. Asplund-Samuelsson, H.N. Joensson, M. Jahn, and E.P. Hudson, *Pooled CRISPRi screening of the cyanobacterium Synechocystis sp PCC 6803 for enhanced industrial phenotypes.* Nature Communications, 2020. 11(1): p. 1-13.
53. Griese, M., C. Lange, and J. Soppa, *Ploidy in cyanobacteria.* FEMS microbiology letters, 2011. 323(2): p. 124-131.
54. Knoot, C.J., S. Biswas, and H.B. Pakrasi, *Tunable repression of key photosynthetic processes using Cas12a CRISPR interference in the fast-growing Cyanobacterium Synechococcus sp. UTEX 2973.* ACS Synthetic Biology, 2019. 9(1): p. 132-143.
55. Dong, C., J. Fontana, A. Patel, J.M. Carothers, and J.G. Zalatan, *Synthetic CRISPR-Cas gene activators for transcriptional reprogramming in bacteria.* Nature Communications, 2018. 9(1): p. 1-11.
56. Kiattisewee, C., C. Dong, J. Fontana, W. Sugianto, P. Peralta-Yahya, J.M. Carothers, and J.G. Zalatan, *Portable bacterial CRISPR transcriptional activation enables metabolic engineering in Pseudomonas putida.* Metabolic engineering, 2021. 66: p. 283-295.
57. Liu, Y., X. Wan, and B. Wang, *Engineered CRISPRa enables programmable eukaryote-like gene activation in bacteria.* Nature Communications, 2019. 10(1): p. 1-16.
58. Bervoets, I. and D. Charlier, *Diversity, versatility and complexity of bacterial gene regulation mechanisms: opportunities and drawbacks for applications in synthetic biology.* FEMS Microbiol Rev, 2019. 43(3): p. 304-339.

59. Bundalovic-Torma, C., G.B. Whitfield, L.S. Marmont, P.L. Howell, and J. Parkinson, *A systematic pipeline for classifying bacterial operons reveals the evolutionary landscape of biofilm machineries.* PLoS Comput Biol, 2020. 16(4): p. e1007721.
60. Dar, D. and R. Sorek, *Extensive reshaping of bacterial operons by programmed mRNA decay.* PLoS Genet, 2018. 14(4): p. e1007354.
61. Osbourn, A.E. and B. Field, *Operons.* Cell Mol Life Sci, 2009. 66(23): p. 3755-75.
62. Saenz-Lahoya, S., N. Bitarte, B. Garcia, et al., *Noncontiguous operon is a genetic organization for coordinating bacterial gene expression.* Proc Natl Acad Sci U S A, 2019. 116(5): p. 1733-1738.
63. Jacob, F., D. Perrin, C. Sanchez, and J. Monod, *[Operon: a group of genes with the expression coordinated by an operator].* C R Hebd Seances Acad Sci, 1960. 250: p. 1727-9.
64. Mao, X., Q. Ma, C. Zhou, et al., *DOOR 2.0: presenting operons and their functions through dynamic and integrated views.* Nucleic Acids Res, 2014. 42(Database issue): p. D654-9.
65. Taboada, B., R. Ciria, C.E. Martinez-Guerrero, and E. Merino, *ProOpDB: Prokaryotic Operon DataBase.* Nucleic Acids Res, 2012. 40(Database issue): p. D627-31.
66. Dehal, P.S., M.P. Joachimiak, M.N. Price, et al., *MicrobesOnline: an integrated portal for comparative and functional genomics.* Nucleic Acids Res, 2010. 38(Database issue): p. D396-400.
67. Yan, B., M. Boitano, T.A. Clark, and L. Ettwiller, *SMRT-Cappable-seq reveals complex operon variants in bacteria.* Nat Commun, 2018. 9(1): p. 3676.
68. Conway, T., J.P. Creecy, S.M. Maddox, et al., *Unprecedented high-resolution view of bacterial operon architecture revealed by RNA sequencing.* mBio, 2014. 5(4): p. e01442-14.
69. Tjaden, B., *A computational system for identifying operons based on RNA-seq data.* Methods, 2020. 176: p. 62-70.
70. Niu, S.Y., B. Liu, Q. Ma, and W.C. Chou, *rSeqTU-A Machine-Learning Based R Package for Prediction of Bacterial Transcription Units.* Front Genet, 2019. 10: p. 374.
71. Fortino, V., O.P. Smolander, P. Auvinen, R. Tagliaferri, and D. Greco, *Transcriptome dynamics-based operon prediction in prokaryotes.* BMC Bioinformatics, 2014. 15: p. 145.
72. Price, M.N., K.H. Huang, E.J. Alm, and A.P. Arkin, *A novel method for accurate operon predictions in all sequenced prokaryotes.* Nucleic Acids Res, 2005. 33(3): p. 880-92.
73. Fortino, V., R. Tagliaferri, and D. Greco, *CONDOP: an R package for CONdition-Dependent Operon Predictions.* Bioinformatics, 2016. 32(20): p. 3199-3200.
74. Zaidi, S.S.A. and X. Zhang, *Computational operon prediction in whole-genomes and metagenomes.* Brief Funct Genomics, 2017. 16(4): p. 181-193.
75. Peters, J.M., M.R. Silvis, D. Zhao, J.S. Hawkins, C.A. Gross, and L.S. Qi, *Bacterial CRISPR: accomplishments and prospects.* Curr Opin Microbiol, 2015. 27: p. 121-6.
76. Villegas Kcam, M.C., A.J. Tsong, and J. Chappell, *Rational engineering of a modular bacterial CRISPR-Cas activation platform with expanded target range.* Nucleic Acids Res, 2021. 49(8): p. 4793-4802.
77. Fontana, J., C. Dong, C. Kiattisewee, V.P. Chavali, B.I. Tickman, J.M. Carothers, and J.G. Zalatan, *Effective CRISPRa-mediated control of gene expression in bacteria must overcome strict target site requirements.* Nat Commun, 2020. 11(1): p. 1618.
78. Liu, T., H. Luo, and F. Gao, *Position preference of essential genes in prokaryotic operons.* PLoS One, 2021. 16(4): p. e0250380.
79. Lim, H.N., Y. Lee, and R. Hussein, *Fundamental relationship between operon organization and gene expression.* Proc Natl Acad Sci U S A, 2011. 108(26): p. 10626-31.
80. Hu, X.F., I, *Identifying Core Operons in Metagenomic Data.* biorxiv, 2019.
81. Liu, G., Y. Zhang, and T. Zhang, *Computational approaches for effective CRISPR guide RNA design and evaluation.* Computational and structural biotechnology journal, 2020. 18: p. 35-44.

82. Wang, D., C. Zhang, B. Wang, et al., *Optimized CRISPR guide RNA design for two high-fidelity Cas9 variants by deep learning.* Nature Communications, 2019. 10(1): p. 1-14.
83. Wang, J., X. Xiang, L. Cheng, X. Zhang, and Y. Luo, *CRISPR-GNL: an improved model for predicting CRISPR activity by machine learning and featurization.* bioRxiv, 2019: p. 605790.
84. Calvo-Villamañán, A., J.W. Ng, R. Planel, H. Ménager, A. Chen, L. Cui, and D. Bikard, *On-target activity predictions enable improved CRISPR–dCas9 screens in bacteria.* Nucleic Acids Research, 2020. 48(11): p. e64-e64.
85. Wang, L. and J. Zhang, *Prediction of sgRNA on-target activity in bacteria by deep learning.* BMC bioinformatics, 2019. 20(1): p. 1-14.
86. Ruffing, A.M., T.J. Jensen, and L.M. Strickland, *Genetic tools for advancement of Synechococcus sp. PCC 7002 as a cyanobacterial chassis.* Microbial cell factories, 2016. 15(1): p. 190.
87. Moore, K.A., J.W. Tay, and J.C. Cameron, *Multi-generational analysis and manipulation of chromosomes in a polyploid cyanobacterium.* bioRxiv, 2019: p. 661256.
88. Taton, A., F. Unglaub, N.E. Wright, et al., *Broad-host-range vector system for synthetic biology and biotechnology in cyanobacteria.* Nucleic acids research, 2014. 42: p. e136.
89. Ikeda, K., M. Ono, H. Akiyama, T. Onizuka, S. Tanaka, and H. Miyasaka, *Transformation of the fresh water cyanobacterium Synechococcus PCC7942 with the shuttle-vector pAQ-EX1 developed for the marine cyanobacterium Synechococcus PCC7002.* World Journal of Microbiology and Biotechnology, 2002. 18: p. 55-56.
90. Akiyama, H., S. Kanai, M. Hirano, and H. Miyasaka, *Nucleotide sequence of plasmid pAQ1 of marine cyanobacterium Synechococcus sp. PCC7002.* DNA research : an international journal for rapid publication of reports on genes and genomes, 1998. 5: p. 127-9.
91. Roberts, T.M. and K.E. Koths. *The Blue-Green Alga Agmenellum quadruplicatum Contains Covalently Closed DNA Circles*. in *Cell*. 1976.
92. Kimura, A., T. Hamada, E.H. Morita, and H. Hayashi, *A high temperature-sensitive mutant of Synechococcus sp. PCC 7002 with modifications in the endogenous plasmid, pAQ1.* Plant & cell physiology, 2002. 43: p. 217-23.
93. Chen, Y., A. Taton, M. Go, R.E. London, L.M. Pieper, S.S. Golden, and J.W. Golden, *Self-replicating shuttle vectors based on pANS, a small endogenous plasmid of the unicellular cyanobacterium Synechococcus elongatus PCC 7942.* Microbiology (Reading, England), 2016. 162: p. 2029-2041.
94. Ruffing, A.M., T.J. Jensen, and L.M. Strickland, *Genetic tools for advancement of Synechococcus sp. PCC 7002 as a cyanobacterial chassis.* Microbial Cell Factories, 2016. 15.
95. Akiyama, H., S. Kanai, M. Hirano, and H. Miyasaka, *A novel plasmid recombination mechanism of the marine cyanobacterium Synechococcus sp. PCC7002.* DNA research : an international journal for rapid publication of reports on genes and genomes, 1998. 5: p. 327-34.
96. Bikard, D., W. Jiang, P. Samai, A. Hochschild, F. Zhang, and L.A. Marraffini, *Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system.* Nucleic Acids Research, 2013. 41(15): p. 7429-7437.
97. Imamura, S. and M. Asayama, *Sigma Factors for Cyanobacterial Transcription.* Gene Regulation and Systems Biology, 2009. 3: p. GRSB.S2090.
98. Imashimizu, M., K. Tanaka, and N. Shimamoto, *Comparative study of cyanobacterial and E. coli RNA polymerases: misincorporation, abortive transcription, and dependence on divalent cations.* Genetics research international, 2011. 2011.
99. Xie, W.-Q., K. Jäger, and M. Potts, *Cyanobacterial RNA polymerase genes rpoC1 and rpoC2 correspond to rpoC of Escherichia coli.* Journal of bacteriology, 1989. 171(4): p. 1967-1973.
100. Lee, D.-k., Y.H. Kim, J.-S. Kim, and W. Seol, *Induction and characterization of taxol-resistance phenotypes with a transiently expressed artificial transcriptional activator library.* Nucleic Acids Research, 2004. 32(14): p. e116-e116.

101. Ran, F.A., P.D. Hsu, J. Wright, V. Agarwala, D.A. Scott, and F. Zhang, *Genome engineering using the CRISPR-Cas9 system.* Nature protocols, 2013. 8(11): p. 2281-2308.
102. Stevens Jr, S., C.P. Patterson, and J. Myers, *The Production of Hydrogen Peroxide by Blue-Green Algae: A Survey 1.* Journal of Phycology, 1973. 9(4): p. 427-430.
103. Cohen, S.E. and S.S. Golden, *Circadian rhythms in cyanobacteria.* Microbiology and Molecular Biology Reviews, 2015. 79(4): p. 373-385.
104. Johnson, C.H., T. Mori, and Y. Xu, *A cyanobacterial circadian clockwork.* Current Biology, 2008. 18(17): p. R816-R825.
105. Pattanayak, G. and M.J. Rust, *The cyanobacterial clock and metabolism.* Current opinion in microbiology, 2014. 18: p. 90-95.
106. Mohallem, R. and U.K. Aryal, *Regulators of TNFa mediated insulin resistance elucidated by quantitative proteomics.* Scientific reports, 2020. 10(1): p. 1-15.
107. Shastri, A.A. and J.A. Morgan, *A transient isotopic labeling methodology for 13C metabolic flux analysis of photoautotrophic microorganisms.* Phytochemistry, 2007. 68(16-18): p. 2302-2312.
108. Young, J.D., *INCA: a computational platform for isotopically non-stationary metabolic flux analysis.* Bioinformatics, 2014. 30(9): p. 1333-1335.
109. Krishnakumar, R. and A.M. Ruffing, *OperonSEQer: A set of machine-learning algorithms with threshold voting for detection of operon pairs using short-read RNA-sequencing data.* bioRxiv, 2021.

APPENDIX A. SUPPLEMENTAL INFORMATION

A.1. Section 0 Supplemental Information

Table A-1. Primers and DNA oligonucleotides used for plasmid construction

Name	Sequence	Description
MS2F	TCAACTTGAAAAAGTGGCACatgaggatcacccatgTGCTTTTTTTGAAGCTTGGG	Oligos for MS2 RNA stem loop
MS2R	CCCAAGCTTCAAAAAAAGCAcatgggtgatcctcatGTGCCACTTTTTCAAGTTGA	
pgRNA9-MS2_F	TGCTTTTTTTGAAGCTTGGG	For amplification of pNS1D-sgRNA9-SpR to insert MS2 stem loop via Gibson assembly
pgRNA9-MS2_R	GTGCCACTTTTTCAAGTTGATAAC	
MCP_gRNA_F	TCTTCAGCATGCAGTCACACTGGCTCACCTTCG	For amplification of MCP-SoxS$_{R96A}$ and TetR operons from pJF104B
MCP_gRNA_R	ATCTGAGCATGCGTAGCGAGTCAGTGAGCGAGG	
rbc_162a_F	CACCataaaaataagtcgaggctg	For inserting crRNA targeting P$_{rbc}$-*ypet* 162 bp upstream of start codon
rbc_162a_R	AAACcagcctcgacttatttttat	
rbc_168a_F	CACCtaaaagataaaaataagtcg	For inserting crRNA targeting P$_{rbc}$-*ypet* 168 bp upstream of start codon
rbc_168a_R	AAACcgacttatttttatctttta	
rbc_241a_F	CACCgcatttctaatgctgtttga	For inserting crRNA targeting P$_{rbc}$-*ypet* 241 bp upstream of start codon
rbc_241a_R	AAACtcaaacagcattagaaatgc	
rbc_297a_F	CACCcccataaattaagactcgag	For inserting crRNA targeting P$_{rbc}$-*ypet* 297 bp upstream of start codon
rbc_297a_R	AAACctcgagtcttaatttatggg	
rbc_338a_F	CACCctgtacctctactgcaatag	For inserting crRNA targeting P$_{rbc}$-*ypet* 338 bp upstream of start codon
rbc_338a_R	AAACctattgcagtagaggtacag	
rbc_373a_F	CACCgcgataagggtggcccttgc	For inserting crRNA targeting P$_{rbc}$-*ypet* 373 bp upstream of start codon
rbc_373a_R	AAACgcaagggccacccttatcgc	
rbc_447a_F	CACCgtctgtcgatattccaagtt	For inserting crRNA targeting P$_{rbc}$-*ypet* 447 bp upstream of start codon
rbc_447a_R	AAACaacttggaatatcgacagac	
rbc_482a_F	CACCttttccccttttcaagctct	For inserting crRNA

Name	Sequence	Description
rbc_482a_R	AAACagagcttgaaaaggggaaaa	targeting P_{rbc}-*ypet* 482 bp upstream of start codon
rbc_564a_F	CACCagacgggcaacaatcgaaaa	For inserting crRNA targeting P_{rbc}-*ypet* 564 bp upstream of start codon
rbc_564a_R	AAACttttcgattgttgcccgtct	
rbc_625a_F	CACCataaagcgtaagaaagttaa	For inserting crRNA targeting P_{rbc}-*ypet* 625 bp upstream of start codon
rbc_625a_R	AAACttaactttcttacgctttat	
rbc_749a_F	CACCttcttcgttctttagattta	For inserting crRNA targeting P_{rbc}-*ypet* 749 bp upstream of start codon
rbc_749a_R	AAACtaaatctaaagaacgaagaa	
A0614_F1	attcCCTCAGCaatcgcagcaaactccggcatctacggtggcggaggtagcATGCAAAAACGTAGCTCCTTC	For amplification of A0614 and fusion to MCP in pNS1D-sgRNA9A01
A0614_R1	cgtggcggccgcCTATTCGCCAATAATTTCAGGC	
A2014_F1	attcCCTCAGCaatcgcagcaaactccggcatctacggtggcggaggtagcATGACCCAAGCGACGAACCCC	For amplification of A2014 and fusion to MCP in pNS1D-sgRNA9A01
A2014_R1	cgtggcggccgcCTAGCGGATATATTCTTTCAGGATG	
A0364_F1	attcCCTCAGCaatcgcagcaaactccggcatctacggtggcggaggtagcATGTTCGACACAACCACAACC	For amplification of A0364 and fusion to MCP in pNS1D-sgRNA9A01
A0364_R1	cgtggcggccgcTCAATACAGCCAACTTTTGAGG	
A1202_F1	attcCCTCAGCaatcgcagcaaactccggcatctacggtggcggaggtagcATGCCTAAAGTAAATCTTCAACAAG	For amplification of A1202 and fusion to MCP in pNS1D-sgRNA9A01
A1202_R1	cgtggcggccgcTTAGCTGGCAACCAGATATTCAC	
A0270_F1	attcCCTCAGCaatcgcagcaaactccggcatctacggtggcggaggtagcATGACCACCGCCAAGCAGTCC	For amplification of A0270 and fusion to MCP in pNS1D-sgRNA9A01
A0270_R1	cgtggcggccgcTTAGGACAATGTGTCCAAGTAATCCCG	
A1924_F1	attcCCTCAGCaatcgcagcaaactccggcatctacggtggcggaggtagcATGGAACCTACGACGACAAAAG	For amplification of A1924 and fusion to MCP in pNS1D-sgRNA9A01
A1924_R1	cgtggcggccgcCTAATCACCACCGAGGATTTTC	
A1970_F1	attcCCTCAGCaatcgcagcaaactccggcatctacggtggcggaggtagcATGAACCGATCCCTTTCTATCCCC	For amplification of A1970 and fusion to MCP in pNS1D-sgRNA9A01
A1970_R1	cgtggcggccgcCTAGTCCGCCAAATAGGGCTGG	

Name	Sequence	Description
A0641_F1	attcCCTCAGCaatcgcagcaaactccggcatctacggtggcggaggtagc ATGGCTACGGAAAAGACGGC	For amplification of A0641 and fusion to MCP in pNS1D-sgRNA9A01
A0641_R1	cgtggcggccgcTCATGAATCTGAGGTAGCGGAC	
A1229_F1	attcCCTCAGCaatcgcagcaaactccggcatctacggtggcggaggtagc ATGTTTCAGCAGCGATCGCC	For amplification of A1229 and fusion to MCP in pNS1D-sgRNA9A01
A1229_R1	cgtggcggccgcCTAAGCCAAGACCCCATGGAG	
A1232_F1	attcCCTCAGCaatcgcagcaaactccggcatctacggtggcggaggtagc ATGGACGCTGCGGAAAAAATTTG	For amplification of A1232 and fusion to MCP in pNS1D-sgRNA9A01
A1232_R1	cgtggcggccgcCTAAAGTCGAGGCAGGGGGAGG	

Table A-2. Primers used for DNA sequencing.

Name	Sequence	Description
AscA5F	GATTACGAAGCCGCCAATG	For sequencing integration of pCas2C at AscA
AscA3R	GGACTTGCGACGACAACTGC	
NS1D_5F	GTCTCCACAGTGTTCCTGACTCC	For sequencing integration of sgRNA CRISPRi and CRISPRa constructs at NS1D
NS1D_3R	CGACTATGATGACTATACCTATTACTACCAC	
TetR_F	GGCGAGTTTACGGGTTGTTA	For sequencing insertion of MS2 RNA stem loop into pNS1D-sgRNA9-SpR
SoxS_F	CAACGAATGTTCCGCACG	For sequencing insertion of MCP-$SoxS_{R96A}$ operon into pNS1D-sgRNA9A
SoxS_Frc	CGTGCGGAACATTCGTTG	
pCas34CmR	GCTACGGCGTTTCACTTC	For sequencing crRNA insertion in pNS1D-sgRNA9A01 and pNS1D-sgRNA9A02

A.2. Section 6 Supplemental Information

Table A-3. Oligos used to construct gRNA plasmids for targeted CRISPRi strains

Name	Oligo Sequence	Description
A0010_F	CACCAATTGATTTTTCGGCAGAAA	Construction of pNS1D-sgRNA-A0010
A0010_R	AAACTTTCTGCCGAAAAATCAATT	
A0052_F	CACCCGATGGTGGATAGAAGTTCT	Construction of pNS1D-sgRNA-A0052
A0052_R	AAACAGAACTTCTATCCACCATCG	
A0556_F	CACCGAGGGTAGAGGCAAAACTGA	Construction of pNS1D-sgRNA-A0556
A0556_R	AAACTCAGTTTTGCCTCTACCCTC	
A0614_F	CACCCCTGCATTAGATCTTCAGCC	Construction of pNS1D-sgRNA-A0614
A0614_R	AAACGGCTGAAGATCTAATGCAGG	
A1240_F	CACCGGGAAGCGGCCCCGACCCCA	Construction of pNS1D-sgRNA-A1240
A1240_R	AAACTGGGGTCGGGGCCGCTTCCC	
A1441_F	CACCGGGGTTTAAACACAGTGTGG	Construction of pNS1D-sgRNA-A1441
A1441_R	AAACCCACACTGTGTTTAAACCCC	
A1594_F	CACCCTTCTCCGGCAGGATCTGCG	Construction of pNS1D-sgRNA-A1594
A1594_R	AAACCGCAGATCCTGCCGGAGAAG	
A2353_F	CACCGTAAATTGATTCGGGTTAGT	Construction of pNS1D-sgRNA-A2352
A2353_R	AAACACTAACCCGAATCAATTTAC	
A2510_F	CACCTTCAGGCAGAGATTGGAGCA	Construction of pNS1D-sgRNA-A2510
A2510_R	AAACTGCTCCAATCTCTGCCTGAA	
A2606_F	CACCGAACGCTCAAGACACCGGCT	Construction of pNS1D-sgRNA-A2606
A2606_R	AAACAGCCGGTGTCTTGAGCGTTC	

DISTRIBUTION

Email—Internal

Name	Org.	Sandia Email Address
Kamlesh Patel	01923	kdpatel@sandia.gov
Anthe George	08622	angeorg@sandia.gov
Stephen Casalnuovo	08630	sacasal@sandia.gov
Amanda Barry	08631	anbarry@sandia.gov
Anne Ruffing	08631	aruffin@sandia.gov
Technical Library	01977	sanddocs@sandia.gov

Email—External (encrypt for OUO)

Name	Company Email Address	Company Name
John Morgan	jamorgan@purdue.edu	Purdue University

This page left blank

Sandia National Laboratories

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.