RNA Tools for Optimization of Multi-Protein Genetic Systems

by

Amar Ghodasara

B.S., University of California, Davis (2008)

SUBMITTED TO THE DEPARTMENT OF BIOLOGICAL ENGINEERING
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY IN BIOLOGICAL ENGINEERING
AT THE
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

MAY 2017

[June 2017]

© 2017 Massachusetts Institute of Technology. All Rights Reserved.
This doctoral thesis has been examined by a committee of the Department of Biological Engineering as follows:

Signature redacted

Certified by . ..................................

Ron Weiss
Professor of Biological Engineering
Thesis Committee Chair

Signature redacted

Certified by . ..................................

Christopher A. Voigt
Professor of Biological Engineering
Thesis Advisor

Signature redacted

Certified by . ..................................

Kristala L. Jones Prather
Associate Professor of Chemical Engineering
Committee Member
RNA Tools for Optimization of Multi-Protein Genetic Systems

by

Amar Ghodasara

Submitted to the Department of Biological Engineering on May 2, 2017 in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in Biological Engineering

ABSTRACT

Balancing protein expression is critical when optimizing genetic systems. Typically, this requires construction of a library where variants of parts (e.g. promoters) are tried for each gene, which can be expensive and time-consuming. Here, we present an approach that leverages trans-acting RNA regulators to explore large gene expression spaces without de novo library construction.

First, we developed six sRNAs whose strengths have been optimized against a set of 15nt “target” sequences that can be inserted upstream of a ribosome-binding site to generate up to 175-fold repression when maximally expressed. By controlling sRNA expression, the targeted gene can be tunably repressed from 1.6- to 121-fold. We then built a pool where each of the six sRNAs was placed under the control of 16 promoters, yielding ~10^7 combinations. This pool can optimize up to six genes in any system. Only a single variant of the system is constructed, where a target sequence is placed upstream of each gene. This is then transformed with the pre-built sRNA pool and the resulting library is screened. The system is then rebuilt by rationally selecting parts that reproduce the optimal knockdown of each gene identified by the screen. We demonstrated the versatility of this tool by using the same pool to optimize a beta-carotene pathway and an XNOR circuit. In a second study, we developed tools to facilitate a similar approach in yeast using CRISPRi. We leveraged T7 RNA polymerase to produce guide RNAs (gRNA), and show that modulating gRNA levels with T7 promoters can regulate gene expression. As a proof of principle, we used this system to modulate flux in a carotenoid pathway.

Together, the tools presented in this thesis drastically reduce the time and cost to optimize multi-gene systems in a variety of organisms.

Thesis Supervisor: Christopher A. Voigt
Title: Professor of Biological Engineering
“Live as if you were to die tomorrow. Learn as if you were to live forever.”

Mahatma Gandhi
Acknowledgements

First and foremost, I would like to thank my advisor, Chris Voigt, for his mentorship, guidance, support, trust in my abilities, and for providing me the financial and equipment resources to fully focus on science. I am also very grateful to my committee members Ron Weiss and Kristala Prather for their advice and support.

As member of the Voigt Lab and the MIT BE community, I have had the great privilege of working with some of the most intelligent and helpful people these past few years. Many thanks to all of the members of the Voigt lab, past and present, who went out of their way to help me even when I had no idea how to clone. We have had great scientific discussions, shared late nights in the lab, and developed wonderful friendships.

I reserve special thanks for Barbara Karampalas, our outstanding lab manager, and Terry King, our excellent administrator. Their dedication and commitment made working in the lab enjoyable as they worked tirelessly so that I could concentrate on doing science.

To my parents, Navin and Jyotsna, and to my brother, Amit, I wish to express my deepest love and gratitude for all of their love and support all of these years. To my partner, Alicia DeFrancesco, thank you for riding the highs and lows of graduate school with me and always being there.

I am grateful to have had the opportunity to do an internship at DSM in the Netherlands. I would like to thank Hans Roubous and Rene Verwaal for their guidance during my few months at DSM.

I would also like to acknowledge the funding sources that made this work possible: the US Defense Advanced Research Projects Agency (DARPA) and the Office of Naval Research (ONR).
# Table of Contents

Abstract ........................................................................................................................................... 3  
Acknowledgements ............................................................................................................. 5  
Table of Contents .................................................................................................................. 6  
List of Figures ............................................................................................................................... 8  
List of Tables .................................................................................................................................. 9  
Preface ........................................................................................................................................... 10  
1. Introduction ............................................................................................................................. 11  
   1.1 Noncoding RNAs ........................................................................................................... 11  
   1.2 RNA-mediated gene regulation ..................................................................................... 12  
   1.3 Trans-acting RNAs and genetic system tuning ............................................................. 15  
2. Balancing gene expression without library construction via a reusable sRNA pool....... 17  
   2.1 Abstract ......................................................................................................................... 17  
   2.2 Introduction ................................................................................................................... 18  
   2.3 Materials and Methods .................................................................................................. 22  
      2.3.1 Strains and media ..................................................................................................... 22  
      2.3.2 sRNA repression assay and growth curves .............................................................. 23  
      2.3.3 Promoter strength measurements ........................................................................... 23  
      2.3.4 Flow cytometry ........................................................................................................ 24  
      2.3.5 Computational methods .......................................................................................... 24  
      2.3.6 sRNA library construction and sorting ................................................................. 25  
      2.3.7 sRNA pool construction .......................................................................................... 25  
      2.3.8 High-Throughput Sequencing .................................................................................. 26  
      2.3.9 β-carotene production quantification ...................................................................... 26  
      2.3.10 XNOR circuit assay and sorting ............................................................................ 27  
      2.3.11 Part replacement ..................................................................................................... 28  
      2.3.12 sRNA-mediated knockdown in other gram-negative bacteria .............................. 29  
   2.4 Results ............................................................................................................................... 30  
      2.4.1 Optimization of sRNA:target pairs ......................................................................... 30  
      2.4.2 sRNA function against chromosomal targets and diverse species ......................... 36  
      2.4.3 Combinatorial sRNA pool construction and validation .......................................... 40  
      2.4.4 Application to a metabolic pathway: optimization of β-carotene ........................... 43  
      2.4.5 Application to a genetic circuit ................................................................................. 46  
2.5 Discussion ........................................................................................................................... 49  
2.6 Supplementary Figures ......................................................................................................... 52
3. Tunable CRISPR-based regulation of multiple genes in eukaryotes using T7 RNA polymerase

3.1 Abstract

3.2 Introduction

3.3 Materials and Methods

3.3.1 Yeast strains and media

3.3.2 E. coli strains and media

3.3.3 Flow cytometry

3.3.4 gRNA repression assay

3.3.5 Promoter strength measurements

3.3.6 Termination Assay

3.3.7 Computational Methods

3.3.8 β-carotene production quantification

3.4 Results

3.4.1 Optimizing T7RNAp-mediated gRNA production

3.4.2 Modulating gene expression with T7 promoters

3.4.3 Applying dCas9-mediated repression to a metabolic pathway

3.5 Discussion

3.6 Supplementary Figures

4. Conclusions and Future Directions

5. Supplementary Tables for Chapter 2

6. Supplementary Tables for Chapter 3

7. Bibliography
List of Figures

Figure 1.1: RNA-mediated gene regulation mechanisms .......................................................... 14

Figure 2.1: Combinatorial repression of target genes with sRNAs ........................................... 21
Figure 2.2: Optimization of sRNA-mediated gene knockdown against target sequences ........ 38
Figure 2.3: Pool assembly from differentially expressed sRNAs ............................................. 42
Figure 2.4: Optimization of β-carotene production using the sRNA pool .............................. 45
Figure 2.5: Optimization of a XNOR circuit with the sRNA pool ........................................... 48

Figure 2.S1: sRNA knockdown mechanism and assay .......................................................... 52
Figure 2.S2: Plasmid Maps ........................................................................................................ 53
Figure 2.S3: β-carotene standard curve and LC analysis ......................................................... 55
Figure 2.S4: sRNA target position and length optimization ...................................................... 56
Figure 2.S5: sRNA-terminator chimera screening .................................................................. 57
Figure 2.S6: Genomic targeting and sRNA-mediated knockdown in other organisms ........... 58
Figure 2.S7: sRNA orthogonality plots ................................................................................. 59
Figure 2.S8: Δhfq strain and hfq titration characterization ....................................................... 60
Figure 2.S9: GFP / RFP knockdown correlation plots for sRNAs ............................................ 61
Figure 2.S10: sRNA pool context effects ................................................................................ 62
Figure 2.S11: Protein expression levels of sRNA-optimized β-carotene pathway .................... 63

Figure 3.1: dCas9-mediated repression with T7RNAp-produced gRNAs .............................. 74
Figure 3.2: Optimizing guide RNA targeting of yeast promoters ........................................... 76
Figure 3.3: T7 promoter library construction, characterization, and gRNA multiplexing .......... 79
Figure 3.4: Tuning β-carotene production by modulating guide RNA expression ................. 81

Figure 3.S1: Repression assay and growth curves ................................................................. 84
Figure 3.S2: T7 termination and its effect on gRNA repression efficiency .............................. 85
List of Tables

Table 2.1: sRNA sequences ........................................................................................................ 35
Table 2.S1: sRNA pool parts and promoter characterization .................................................. 89
Table 2.S2: T7 promoter characterization and selected promoters ........................................ 93
Table 2.S3: Genetic part sequences used in Chapter 2 .......................................................... 95
Table 2.S4: sRNA tag homologies with the *E. coli* genome .............................................. 103
Table 3.S1: T7 terminator strength and corresponding fold repression .............................. 106
Table 3.S2: Predicted nucleosome occupancy and fold repression for gRNA target sites .... 107
Table 3.S3: T7 promoter library sequences, strength, and corresponding fold repression .... 108
Table 3.S4: Genetic part sequences used in Chapter 3 .......................................................... 109
Preface

The scientific results described within this thesis have been recently submitted for peer-review. The bulk of the actual written content and figures presented in Chapter 2 is based on this manuscript, in accordance with the copyright policy of the publisher. Since writing of Chapter 2 was a joint effort with Chris Voigt, the words which communicate my thesis work are not solely my own.
1. Introduction

1.1 Noncoding RNAs

For several decades after Francis Crick proposed the central dogma of molecular biology in 1958, the primary role for RNA was thought be a mere messenger between DNA and protein (1). In the sixty years since Crick’s seminal work the perception of RNA has drastically changed as RNAs have been found and engineered to play a myriad of functions within cells in addition to those of mRNA, tRNA, and rRNA. These functions include transcription termination, catalyzing biochemical reactions (2), modifying DNA (3), gene regulation by switching translation and transcription on or off (4,5), sensing molecules (6), and as structural scaffolding for chemical reactions (7). The remarkable versatility and universality of this biological macromolecule has sparked great interest from synthetic biologists in recent years.

Noncoding RNAs (NC) have been found to play a major role in gene regulation in a wide range of organisms spanning multiple kingdoms of life (8). As highly potent regulators they present an attractive target for engineering and reprogramming for synthetic systems. One of the main reasons NCs have garnered so much interest is that they often impart specificity through their sequence as they find their target via Watson-Crick base pairing, though bulges, loops, and noncanonical base pairing are often found (9). Unknown binding interactions between nucleic acids are much easier to decipher than those involving proteins. Even compared to programmable proteins, such as zinc finger proteins, transcription activator-like effectors, and the Pumillo protein, whose amino acid to nucleic acid binding codes are known, NCs can be much more rapidly constructed and tested (10-12). Further, they are more amenable to folding
predictions and kinetic modeling as they are not translated and typically do not form complex tertiary structures (13).

1.2 RNA-mediated gene regulation

RNA elements and NCs have been found to be potent regulators that work at multiple stages of gene expression. At the transcriptional level, four mechanisms of RNA-mediated control are commonly used: (i) small RNA (sRNA); (ii) ligand-mediated (called a riboswitch) anti-terminator disruption (repression) (iii) or formation (activation); (iv) CRISPRi, which sterically hinders RNA polymerase and inhibits transcription (14,15) (Figure 1.1A and B). In the former mechanism, a highly structured RNA element is incorporated into the 5’UTR of the target gene. This element has two parts, the aptamer or binding region that will base pair with a specific sRNA or ligand, and a transcriptional terminator that occludes RNA polymerase and stops transcription. The two parts are adjacent to each other and the binding of the ligand or sRNA leads to a conformational change that disrupts (or forms) the anti-terminator (or terminator).

CRISPRi is an unnatural transcriptional regulation system engineered from the type II CRISPR system from Streptococcus pyogenes. In nature, CRISPR is a complex defense system in which Cas9, a nuclease, complexes with a small guide RNA (gRNA) that contains a region complementary to a specific target DNA sequence. In CRISPRi, the Cas9 is mutated to be catalytically null, but still able to bind DNA, and in many bacterial systems can impede RNA polymerase. CRISPRi has successfully been transferred to a variety of organisms, including a few eukaryotes (16,17), where it can also function as a transcriptional activator when fused to activation domains. Chapter 3 presents a novel method of producing gRNAs in eukaryotes to enable, simplify, and expand its use eukaryotes.
At the translational level, there are two widely used RNA-mediated gene regulation mechanisms: ribosome binding site (RBS) occlusion or release by an sRNA or riboswitch, and ligand-dependent RNA cleavage. Similar to transcriptional riboswitches, translational riboswitches change conformation in response to an sRNA or ligand, which then blocks or opens up access to the RBS (5). sRNAs typically block access to the RBS by directly binding to or near it with the aid of RNA chaperone protein, Hfq (18) (Figure 1.1C). Chapter 2 details the engineering of this type of sRNA and their applications in controlling multi-gene systems. The second mechanism uses an RNA element that has two components, an aptamer that responds to a specific ligand or sRNA, and a ribozyme, a catalytic RNA sequence that can cleave itself or other nucleic acids when in a specific conformation. Combined, these components yield an aptazyme, an RNA element that cleaves itself to prevent or enable translation in response to a ligand (19) (Figure 1.1D).
Figure 1.1: RNA-mediated gene regulation mechanisms.

(A) Schematic of riboswitch-mediated repression. An aptamer / anti-terminator pair is encoded in the 5'UTR of the target gene, and changes conformation to form a terminator in the presence of the corresponding ligand. (B) Schematic of CRISPRi repression. A gRNA specific to the target promoter is expressed along with dCas9. The gRNA allows dCas9 to bind the promoter and stop transcription. (C) Schematic of sRNA-mediated repression. An sRNA targeting the 5'UTR is expressed which then complexes with Hfq and binds the target mRNA. Once bound, ribosome docking on the mRNA is blocked. (D) Schematic of aptazyme-mediated repression. An aptamer / ribozyme is inserted in the 5'UTR of the target gene. In the presence of an appropriate ligand, a conformation change activates the ribozymes which then cleaves the mRNA.
1.3 Trans-acting RNAs and genetic system tuning

Of the RNA-mediated gene regulation mechanisms described above, those that involve trans-acting RNAs, NCs that are transcribed from another genetic locus and not encoded in the target mRNA, are particularly interesting as they can present a way of specifically and independently regulating any gene of interest. As I detail in Chapter 2, this feature can be extremely powerful when multiple genes need to be tuned as different expression levels of each gene can be explored without modify the original genetic system; the “tuning” NCs (sRNAs or gRNAs) can be expressed from a separate pre-constructed vector or pool of vectors, which enables the construction of a gene expression library without having to build one ad hoc for each system.

Tuning a genetic system typically involves trying regulatory elements, most commonly promoters and RBSs, which have been mined from nature, engineered, or computationally predicted to exhibit a certain activity (20). This involves physically assembling these parts combinatorially for each gene to create a library that explores a diverse gene expression space. Even with the recent advances in DNA assembly, this process can still be laborious and costly (21). Further, parts are extremely limited in many eukaryotic systems. A library of trans-acting RNAs presents a new way of exploring gene expression spaces. In Chapter 2, I present a method that enables sRNAs to be multiplexed and show that modulating their expression can finely tune the expression of target genes. In Chapter 3, I describe a method to produce gRNAs that could lead the way in developing a similar tuning pool-based tuning system in eukaryotes where tools to fine tune gene expression are currently severely lacking.

Overall, I aim to complement or offer an alternative to the current approaches to optimizing large genetic systems with the novel trans-acting regulator library approach presented
in this thesis. As new trans-acting regulators and organisms are discovered and characterized, I hope for this approach to prove even more value in precise gene expression control.
2. Balancing gene expression without library construction via a reusable sRNA pool

2.1 Abstract

Balancing protein expression is critical when optimizing genetic systems. Typically, this requires the construction of a library where variants of parts are tried for each gene, which can be expensive and time-consuming. Here, we develop six sRNAs whose strengths have been optimized against a set of 15nt “target” sequences that can be inserted upstream of a ribosome-binding site to generate up to 175-fold repression when maximally expressed. By controlling sRNA expression, the targeted gene can be tunably repressed from 1.6- to 121-fold. A pool is built where each sRNA is placed under the control of 16 promoters, yielding ~10⁷ combinations. This pool can optimize up to six genes in any system. Only a single variant of the system is constructed, where a target sequence is placed upstream of each gene. This is then transformed with the pre-built sRNA pool and the resulting library is screened. The system is then rebuilt by rationally selecting parts that reproduce the optimal knockdown of each gene identified by the screen. We demonstrate the versatility of this tool by using the same pool to optimize a pathway and circuit. By avoiding library construction, this approach drastically reduces the time and cost to optimize multi-gene systems.
2.2 Introduction

Expressing proteins at the wrong levels is a common reason why genetic designs fail. For a metabolic pathway, suboptimal enzyme concentrations could lead to many problems, including the accumulation of toxic intermediates, titration of co-factors, or overburdening the host (22-26). When building a genetic circuit, the expression of each constituent protein needs to be tuned so that signals can propagate through the layers to perform the desired computational function (27,28). The challenge is that the correct expression levels are often not known a priori and there is uncertainty in the function of genetic parts in new contexts (29,30). Further, the optimal expression of one gene often depends on the expression levels of other genes in the system, thus creating a rugged search space (31-34). For these reasons, optimization requires the creation of libraries, involving part substitution or random mutagenesis, and screening in order to identify the optimum.

There are many guided and unguided approaches in the literature to build libraries of multi-gene genetic systems. Modern approaches have been driven by improvements in methods to build large DNA fragments. Typically, many variants of the pathway are constructed where the genetic parts controlling each gene are substituted from a set, for example promoters, ribosome-binding sites (RBS), and RNA control elements that modulate mRNA stability (20,22,26,35-39). Beyond part substitution, libraries can be built that alter the gene order, orientation, and operon occupancy (21,40,41). These approaches are often blind searches that are hindered by assay throughput. To address this, mathematical modeling and combinatorial optimization algorithms have been applied to reduce the number of parameters that need to be tuned to reach design objectives, such as improved metabolic flux (42-47). When a biophysical model is not possible, multivariate statistical methods can be used, including multifactor design
of experiments, and Bayesian approaches (33,48-51). At the other extreme, completely random approaches (e.g., mutagenesis and recombination) can be applied across the entire construct or targeted to particular regions (20,52-56). These approaches are cheap because they do not require sequence verification of each construct but limited by low mutagenesis rates, which make co-optimization of genes difficult (57). Whether guided or unguided, the construction and validation of the library is a slow and expensive step that has to be repeated for each system to be optimized.

We sought to develop a method where the expression levels of multiple genes could be simultaneously tuned without having to rebuild a library for each system. Variation is achieved with a separate library, built once, that contains regulators expressed at different levels. Each regulator would control a different gene in the system such that changes in regulator expression would lead to changes in the target gene. This required identifying regulators that are:

- **Orthogonal** so that each regulator only controls its target
- **Wide in dynamic range** to be able to sweep across expression levels,
- **Programmable** so that the regulator can be targeted to different genes, and
- **Non-toxic** so that the regulators themselves do not influence the system.

We decided to focus on sRNAs (58-61) as they satisfy these criteria. sRNAs are potent negative regulators that do not require heterologous proteins to function (62). When expressed, sRNAs bind their cognate mRNA in the 5’ untranslated region (UTR) via Watson-Crick base pairing with assistance from Hfq, a ubiquitous RNA chaperone protein. Once bound, translation initiation is inhibited and RNA degradation by RNase E is accelerated (Figure 2.S1) (63,64). Most synthetic sRNAs studied to date have involved fusing new target sequences to natural sRNA scaffolds and minor scaffold modifications (60,65-67). The best engineered sRNAs have
been shown to achieve up to 85-fold repression without cross reactions and they are usually non-toxic to the host (60,65). However, since each RNA element targets a sequence specific to its cognate gene, the RNA must be designed ad hoc for each target and screening of multiple designs is often required (68,69).

To address programmability, we present a new approach where designed sRNA “target sequences” are placed in the 5’ untranslated region (5’-UTR) of the gene to be tuned (Figure 2.1A). Fixing the sRNA target site avoids the need to redesign the sRNA for each target gene and provides reliable and predictable knockdowns. The sRNA is rigorously optimized to increase the fold-repression against the target. This is repeated to create a set of 6 sRNA:target pairs that are orthogonal to each other and do not impact the host even when highly expressed. Further, it is demonstrated that the repression can be controlled in a graded manner by using different constitutive promoters to control the expression level of the sRNA (Figure 2.1B). All six sRNAs are used to build a combinatorial pool of ~10^7 differentially expressed sRNAs (Figure 2.1C). This pool is capable of tuning genetic systems with up to six different genes over a wide expression space (Figure 2.1D). As a proof of principle, we use the same sRNA pool to tune very different systems: a metabolic pathway (β-carotene) and a logic circuit (XNOR). This only requires building a single construct for each system, where the target sequence is placed upstream of each gene to be tuned. Once the optimal construct is identified, the reliance on sRNA expression can be removed by performing limited part substitution in order to recover the same expression levels discovered with the sRNA screen.
**Figure 2.1: Combinatorial repression of target genes with sRNAs.**

(A) Targeting of genes by engineered sRNAs. Raised arrows are promoters, circles on stems are ribozymes, hemispheres are RBSs, and “T”s are terminators. Colored rectangle with a white square represents the target sequence in the 5’ UTR of the target gene. The control promoter drives sRNA expression. (B) Effect of sRNA expression level on target knockdown. (C) Combinatorial pool of differentially expressed sRNAs. Up to six different sRNAs driven by 16 unique promoters are expressed from one plasmid. Sequences for pool parts are provided in Table 2.1. (D) The process of using the sRNA pool to optimize a genetic system. Optimization of a three gene metabolic pathway in which promoters (colored) are replaced to recapitulate optimal gene expression (identified from sRNAs expression levels) is shown.
2.3 Materials and Methods

2.3.1 Strains and media.

Escherichia coli DH1OB (F- mcrA Δ(mrr-hsdRMS-mcrBC) Φ80lacZΔM15 ΔlacX74 recA1
endA1 araD139 Δ(ara leu) 7697 galU galK rpsL nupG λ-) (70) was used in all cloning
procedures and experiments unless stated otherwise. Escherichia coli K-12 JW4130-1 (F-,
Δ(araD-araB)567, ΔlacZ4787(:::rrnB-3), λ-, rph-1, Δ(rhaD-rhaB)568, Δhfq-722::kan, hsdR514)
and BW25113 (F-, Δ(araD-araB)567, ΔlacZ4787(:::rrnB-3), λ-, rph-1, Δ(rhaD-rhaB)568,
hsdR514) from the KEIO collection (71) was used for hfq knockout experiments. MegaX
DH10B T1 Electrocomp Cells (ThermoFisher, C640003) were used for sRNA pool construction.
“Clonetegration” as described by St-Pierre and co-workers (72) was used to integrate sRNA-A-
tagged RFP at attP2. The following Pseudomonas and Salmonella strains were used:
Pseudomonas protegens Pf-5 (73), and Salmonella typhimurium LT2 (74). Cells were grown in
LB Miller broth (Difco, 90003-350) or SOC (SOB (Teknova, S0210) supplemented with 0.4%
glucose) for cloning. Supplemented M9 minimal media (1x M9 Salts (Sigma-Aldrich, M6030), 2
mM MgSO4, 100 μM CaCl2, 0.4% glucose, 0.2% casamino acids, 340 mg/L thiamine) was used
for knockdown assays and iron starvation assays. Chloramphenicol (34 μg/ml) (Alfa Aesar,
AAB20841-14), kanamycin (50 μg/ml) (GoldBio, K-120-10) or spectinomycin sulfate (50
μg/mL) (MP Biomedicals LLC, 158993) was supplemented where appropriate. IPTG (Sigma-
Aldrich, I6758) and aTC (37919) were used as inducers. Three fluorescent proteins, GFPmut3
(75), mRFP1 (76) and YFP (75), were used as reporters.
2.3.2 sRNA repression assay and growth curves.

Colonies (on three different days) were picked and inoculated into 300 μL LB with antibiotics and grown 16 h at 37°C and 900 RPM in a Multitron Pro incubator shaker (In Vitro Technologies). Three microliters of this overnight culture were then added to 147 μl of M9 media with antibiotics and grown for 3 h in an ELMi shaker (Elmi Ltd.) at 37°C and 1000 RPM, resulting in an OD_600 of 0.2-0.4. Twenty microliters of the cultures were diluted into 180 μl 1X phosphate buffered saline (PBS) with 2 mg/ml kanamycin. For sRNA optimization experiments, sRNAs were expressed using promoter BBa_J23119 and terminator BBa_B0010 from a pBR322 plasmid (pRNA) while the reporter (GFP or RFP) was expressed from a p15A plasmid using promoter BBa_J23101, RBS BBa_B0032, terminator BBa_B0012, and RiboJ, an insulating ribozyme (77) (pREP, Figure 2.S2). For growth curves, cells were grown for 18 h in LB, normalized to an OD_600 of 1, then diluted 1:200 in LB (1:50 into M9 for iron starvation assays) into a microtiter plate and growth was measured for 16 h in a plate reader.

2.3.3 Promoter strength measurements.

The strengths of all constitutive promoters and T7 promoters were characterized by using them to drive YFP expression from a reporter plasmid (pPROM, Figure 2.S2). Reported promoter strengths are the geometric mean of the measured fluorescence output. The sRNA repression assay was followed with the exception that cells were grown in LB instead of M9 media. T7 promoter characterizations were performed by co-transforming the T7RNAP expression plasmid N249 (78) with the reporter plasmid. For these assays colonies were grown overnight and diluted 1:200 into LB with antibiotics and 1 mM IPTG and grown for 6 h at 37°C, then analyzed on the flow cytometer.
2.3.4 Flow cytometry.

Fluorescence was measured using a MACSQuant VYB (Miltenyi Biotec) with a 488-nm laser for GFP and YFP excitation and 561-nm for RFP excitation. For each sample at least $5 \times 10^4$ events were recorded using a flow rate of 0.5 μL/s. FlowJo v10 (TreeStar Inc.) was used to analyze the data. All events were gated by forward scatter and side scatter. Fluorescence values are shown as the geometric mean. Fold-repression is calculated as the geometric mean of the fluorescence measured with an empty vector (pEMPTY, Figure 2.S2) divided by the geometric mean of the fluorescence measured with an sRNA expression vector.

2.3.5 Computational methods.

All Random DNA sequences were generated using a random DNA sequence generator (http://www.faculty.ucr.edu/~mmaduro/random.htm), with a GC content probability parameter of 0.5. All sRNA tags were evaluated against the *E. coli* genome using BLASTn (http://www.ecogene.org/ecoblast) (79); sRNA tags exhibiting more than 11bp chromosomal homology or homology within 200bp of annotated coding sequences on either strand were omitted. The search was performed on the *E. coli* K-12 MG1655 U00096.3 genome database with an expect threshold of 10 and the BLOSUM62 matrix. Nupack (www.nupack.org) (13) was used to predict RNA secondary structures. The following parameters were used for all RNA folding simulations: Nucleic acid type: RNA; Temperature: 37°C; Number of strand species: 1; Maximum complex size: 1; RNA energy parameters: “Mathews et al., 1999”; Dangle treatment: some. RBS sequences for the pathway and circuit were generated using the RBS Calculator (https://www.denovodna.com/software/) (80). For all calculations the following parameters were used unless otherwise noted: Pre-sequence: (insulating ribozyme sequence followed by an sRNA
2.3.6 sRNA library construction and sorting.

All degenerate nucleotide libraries were built using circular polymerase extension cloning (CPEC) (81) with oligos containing N’s (Integrated DNA Technologies) at appropriate positions. The CPEC product was column purified (DNA Clean & Concentrator-5, ZymoResearch) and transformed into cells harboring the appropriate reporter plasmid. After 1 h outgrowth in SOC, cells were diluted in to 5 ml LB with antibiotics and grown overnight at 37°C. The sRNA repression assay was performed with the following two modifications: 1) 100 µL of culture was used instead of 3 µL 2) after 3 h of growth in M9 media cells were pelleted and resuspended in 2 mL 1X PBS with 1 mg/ml chloramphenicol. Cells were sorted on a FACS Aria II (BD Biosciences). Sorting gates were set to sort cells with lower GFP expression than that achieved by the previous sRNA scaffold design. Sorting was stopped after 5 × 10⁴ cells were collected. Sorted cells were allowed to recover in SOC for 2 h then plated and allowed to form colonies. Colonies were then streaked out and individually assayed by flow cytometry.

2.3.7 sRNA pool construction.

The CombiGEM (82) protocol was used to assemble the sRNA pool (pPOOL, Figure 2.S2) with the following modifications. 1) The AvrII restriction site was replaced with XbaI. 2) Ligation products were transformed into MegaX DH10B T1 Electrocomp Cells and after 1 h outgrowth in
SOC, cells were diluted in to 5 ml LB with antibiotics and grown overnight at 37°C; the following morning, the culture was miniprepped and the process was iterated until all six sRNAs were cloned.

2.3.8 High-Throughput Sequencing.

sRNA pool sequencing barcodes were PCR amplified using 100 ng of the pool as template. Barcode PCR products were sent to the Massachusetts General Hospital DNA Core where the samples were processed and run on a Mi-Seq producing ~200,000 reads. Custom software was written to search the raw sequencing FASTQ file against the list of all 96 sequencing barcodes (Table 2.S1) to count the number of occurrences of each corresponding promoter in the pool.

2.3.9 β-carotene production quantification.

The β-carotene pathway construct (pCAROTENE, Figure 2.S2) was co-transformed with plasmid N249 (78). Cells were grown overnight, made electro-competent (83), and transformed with 100 ng of the sRNA pool and plated. Colonies harboring all three plasmids were grown in 1 ml LB with antibiotics at 30°C for 18 h. Cultures were then diluted to an OD_{600} of 0.2 into 1 ml Terrific Broth (TekNova, T7060) with appropriate antibiotics and 1 mM IPTG and grown for 48 h in a Multitron Pro at 30°C in 96-well deep plates (USA Scientific, 1896-2000). Cells were harvested for quantification by centrifugation at 4000g for 10 min. They were then washed with 0.6 ml water, and re-centrifuged. Cell pellets were then resuspended in 0.5 ml acetone and incubated at 55°C for 30 min with frequent vortexing. Pigment extraction mixtures were then centrifuged at 4000g for 15 min. 0.2 ml of supernatant was transferred to a polypropylene 96-well plate (Greiner, 655201) for quantification. Samples were measured using a H1 Synergy
plate reader (BioTek Instruments) at an absorbance wavelength of 470 nm. A standard curve (Figure 2.S3), using purified β-carotene (Sigma-Aldrich, 22040), was established to correlate absorbance to amount of β-carotene produced. The following equation was used: titer = 120.6 \frac{mg}{L} \times (\text{measured absorbance} - \text{background}). Background absorbance was measured from a strain harboring pEMPTY (Figure 2.S2) in place of the pathway construct. Samples from the top producing strains were further analyzed by liquid chromatography (1260 Infinity LC System, Agilent Technologies) using a reversed phase C18 column (Phenomenex, 00A-4462) to confirm absence of lycopene and other precursor metabolites and confirm sample purity (Supplementary Figure 2.S2).

2.3.10 XNOR circuit assay and sorting.

Cells harboring the XNOR circuit (pCIRCUIT) and output (pOUT) plasmids (Figure 2.S2) were grown overnight at 37°C in 1 mL LB then diluted 1:200 into 200 µL M9 with antibiotics and grown for 3 h at 37°C. The culture was then diluted 1:700 into 200 µL M9 with antibiotics and inducers and grown for 6 h. Twenty microliters of the cultures were diluted into 180 µl 1X phosphate buffered sulfate (PBS) with 2 mg/ml kanamycin and run on the cytometer. For the first input 1 mM IPTG was used for a ‘1’ input and no IPTG for a ‘0’ input. For the second input 2 ng/ml aTC was used for a ‘1’ input and no aTC for a ‘0’ input. The geometric mean of fluorescence in each state was used as the output. For XNOR optimization, cells harboring the XNOR circuit and output plasmid were grown overnight, made electro-competent (83), and transformed with 100 ng of the sRNA pool. After 1 h of outgrowth in SOC, cells were diluted in 5 ml LB with antibiotics and grown overnight at 37°C. The following day the culture was diluted 1:200 into 5 mL M9 with antibiotics and grown for 3 h at 37°C. Induction was commenced at
this point by diluting cultures 1:700 into 3.5 mL M9 with antibiotics and appropriate inducers (1 mM IPTG and/or 2 ng/ml aTC). After 6 h of growth, cells were pelleted and resuspended in 2 ml PBS supplemented with 1 mg/ml chloramphenicol and sorted. Sort gates were selected according to the desired output (YFP-high or YFP-low) for each input condition and were designed to include 5% of all cells. Sorting was stopped after $5 \times 10^4$ cells were collected. Sorted cells were allowed to recover in SOC for 2 hours then diluted into 5 mL LB with antibiotics and grown overnight. The induction and sorting process was repeated for all input conditions. After the final sort cells were allowed to recovery in SOC for 2 hours then plated and allowed to form colonies. Colonies were then streaked out and individually assayed by flow cytometry.

2.3.11 Part replacement.

For the pathway, the repression level of each gene in the top strain was identified by mapping sRNA promoter strengths to fold repression (Figure 2.3A). Fold repression was used to calculate optimal T7 promoter strengths for each gene using the following equation: $T7_{\text{optimal}} = \frac{T7_{\text{original}}}{\text{Fold Repression}}$. T7 promoters closest in strength to $T7_{\text{optimal}}$ were then selected from a list of previously characterized T7 promoters (Table 2.S2) and the pathway was reconstructed. For the circuit, the repression level of each repressor in the best circuit was identified by mapping sRNA promoter strengths to fold repression (Figure 2.3A). Fold repression was used to calculate optimal RBS strengths for each repressor using the following equation: $RBS_{\text{optimal}} = \frac{RBS_{\text{original}}}{\text{Fold Repression}}$. An RBS sequence corresponding to the optimal RBS strength was calculated using the RBS calculator with the Target Translation Initiation Rate set to $RBS_{\text{optimal}}$ and the circuit was reconstructed.
2.3.12 sRNA-mediated knockdown in other gram-negative bacteria.

The sRNA expression cassette from sRNA-A and from its corresponding reporter plasmid were cloned into a pBBR1 vector (84) (pBBR-RNA, Figure 2.S2). The plasmid was electroporated into Pseudomonas and Salmonella cells. Pseudomonas cells were subjected to the sRNA repression assay with the exception that cells were grown at 30°C in LB. Salmonella cells were subjected to the sRNA repression assay with the exception that cells were grown in LB.
2.4 Results

2.4.1 Optimization of sRNA:target pairs

Our approach is based on the ability to design a target sequence that can be inserted upstream of any gene to make it sensitive to a cognate sRNA. This required first identifying the optimal region within the 5'-UTR to place the target. Then, the sRNA was optimized to maximally reduce target gene expression. Multiple parameters, including sRNA tag length (60), Hfq scaffold sequence, terminator strength, and 5' standardization (85) and RNA stabilization (35) were optimized. Finally, a set of orthogonal sRNA:target pairs were created so that multiple genes can be simultaneously controlled. All sRNA sequences are listed in Supplementary Table 2.S3.

To evaluate sRNA designs, we developed an assay in which we constitutively express an sRNA and a target sequence upstream of a fluorescent protein. The sRNA and reporter are expressed from different plasmids, pRNA and pREP (Figure 2.S2), respectively, and target repression is quantified by flow cytometry.

We first identified the optimal region in the 5'-UTR to target with an sRNA. For our initial tests, ten 20nt sRNA tags with 50% GC content were generated using a random DNA sequence generator (Materials and Methods). Each tag was tested against the E. coli genome using EcoBlast (79); a tag with no homology to the genome was selected. This tag was fused to the wild-type MicC sRNA scaffold as described by Na and co-workers (60) (sRNA-AG1) while its cognate targeting sequence (reverse complement of tag) was inserted at multiple positions in the 5'UTR of GFP and RFP (Figure 2.S4). We found that placing the targeting sequence closer to the RBS on the 5' side improved gene repression while inserting it on the 3' of the RBS resulted in reduced repression. These results corroborate similar studies and align with the
mechanism of sRNA-mediated gene repression (63,86). We avoided directly targeting the RBS or coding sequences to ensure that the target can be reused to control different genes and prevent crosstalk by targeting similar sequences.

Once the optimal target position was identified, we tested different tags and target lengths to further improve repression. In E. coli, which has a genome size of ~4.6Mbp, 12nt is the minimum number of nucleotides necessary to ensure perfect orthogonality (without RNA bulges or loops) to the genome. Therefore, we designed tags and corresponding targeting sequences ranging from 12nt to 20nt in length. Repression was maximized with a 15nt tag (Figure 2.5). As tag length decreases, binding energy increases and reduces RNA hybridization. However, as tag length increases the probability of alternate secondary conformations for the sRNA and target mRNA increases, which may reduce repression (87). We standardized sRNA tag/target length at 15nt as a compromise between these different constraints.

Next, we evaluated different Hfq scaffolds and sought to engineer one for strong repression. Six well-studied Hfq scaffolds from natural sRNAs in E. coli were fused to the 15nt tag described above and assayed: DsrA (88), GcvB (89), MicC (90), MicF (91), SgrS (92), and Spot42 (93). All tested sRNAs exhibited activity, of which the Spot42-derived sRNA was the strongest achieving 68-fold repression (Figure 2.2A).

Most natural sRNAs include a rho-independent terminator in their Hfq scaffolds as it is essential for Hfq recruitment (18,94). Prior studies have shown that this terminator can be exchanged for other rho-independent terminators without abrogating sRNA-mediated repression (66,95). Since the secondary structure of sRNAs is essential for Hfq-binding (96), and consequently gene repression, we hypothesized that replacing the natural terminator found in the Spot42 Hfq scaffold with a stronger terminator would improve repression efficiency as this
would result in fewer sRNAs with extended 3' ends that could interfere with sRNA folding. The best, T(BF1176), was selected as the core scaffold for further optimization (Figure 2.2A and 2.85).

A prior structural study showed that all but six nucleotides (not including the terminator stem-loop) of the Spot42 scaffold are involved in Hfq-binding, and consequentially essential for sRNA-mediated repression (97). Therefore, we created a targeted mutagenesis library in which these six nucleotides were randomized resulting in a 4096-variant library (Figure 2.2C, arrows). Potent scaffold variants were identified by fluorescence-activated cell sorting (FACS) (Materials and Methods). The best one (sRNA-AG4) yields 115-fold repression with all positions mutated (Figure 2.2A). Folding predictions of these high-performing scaffold variants revealed that the predicted secondary structure of the Spot42 Hfq-binding region was preserved (Materials and Methods).

A challenge with using promoters to control sRNA expression is that they can generate a distribution of transcription start sites (20,98). The initial attempt to control gene expression using different strength promoters produced a poor correlation between promoter strength and repression (Figure 2.2B, top panel). Using the published transcription start sites of the promoters (20), we performed folding predictions after adding the 5' overhang to the sRNA and found that 10/16 disrupted folding. Ribozymes were chosen to standardize the 5' end of the sRNA. Using a hammerhead ribozyme (HHR) scaffold (99), we designed a random library in which 18 non-essential nucleotides of the ribozyme (Figure 2.2D, purple nucleotides with arrows) were replaced with degenerate bases resulting in a theoretical library of $10^{10}$ variants fused to sRNA-AG4. This library was then screened using FACS. Approximately 3% of the screened variants exhibited improved repression from which the top-performing HHR/sRNA design, sRNA-A,
capable of 175-fold repression was discovered. The repression achieved with this variant was found to strongly correlate with promoter strength (Figure 2.2B, lower panel).

Once the sRNA was optimized for repression, we then created a set of sRNAs that target different tags to enable the control of multiple genes within a cell. Re-designing the sRNA scaffold for a new target simply requires changing the 15nt tag on the sRNA and inserting the reverse complement of the cognate target sequence preceding the SD sequence on the gene to be controlled. A bioinformatic approach was pursued to generate a list of tags that avoid off-target effects in E. coli (Materials and Methods). In addition, secondary structure predictions were performed to ensure that the addition of the target sequence to the sRNA would not disrupt folding of the Hfq-binding region. From this list, 25 tags were selected, fused to sRNA-A, HHR appropriately modified for each, and tested for repression against corresponding reporter constructs. Those were eliminated that exhibited toxicity (9), crosstalk (3), or produced less than 100-fold repression (7). This led to a set of 6 sRNAs that efficiently target different sequences without impacting the host. The top ten homologies against the genome for each are listed in Table 2.S4.

The optimized sRNA scaffold (without the tag) is 55nt and the HHR is 42nt. Composing all six sRNAs presents a problem as sequence repeats of >25bp, especially in RecA+ strains, can lead to genetic instability due to recombination (100). To overcome this, we diversified the scaffold of each sRNA by mutagenizing 11 non-essential positions (6 in the Hfq-binding stem and 5 in the terminator loop) along with the 18 positions in the ribozyme and screening by FACS, as before (Figure 2.2D, arrows). The best performing sRNA for each tag was selected, and all sRNAs have a maximum of 19nt of continuous shared sequence identity with one another (Table 2.1). The six final sRNAs exhibited similar repression against their cognate target
sequence ranging from 120-fold to 175-fold and showed near perfect orthogonality to each other (Figures 2.2E and Figure 2.S7).
### Table 2.1: sRNA sequences

| sRNA     | DNA sequence
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>sRNA-A</td>
<td>TTTGACCTGATGAGTCCGTAGACGAGCTAGCTCGTCGTCAA&lt;br&gt;AGTTTGTATTATTTGTAGAATAATTTATTTCGCCCCCGGAAGATCATATTCCCGG&lt;br&gt;GGGCTTTTTTATT</td>
</tr>
<tr>
<td>sRNA-B</td>
<td>GATGAGCTGATGAGCCCGACACGGCAGAACTCAGCAGTCTCCCTCAT&lt;br&gt;CATTAAATTATTTGTATATTTATTTACGCGCCCCCGGAAGGCCTTCCG&lt;br&gt;GGGCTTTTTTATT</td>
</tr>
<tr>
<td>sRNA-C</td>
<td>CTACACTGATGAGGTAGACAACCAGAATCTGCCTCTCTCTGCTCTGTA&lt;br&gt;TAGTGTGAAATTGTTATAAATATTAAAACGCCCCCGGAATTACGTTCCGG&lt;br&gt;GGGCTTTTTTATT</td>
</tr>
<tr>
<td>sRNA-D</td>
<td>CGGTTACTGATGAGTCGCTCTTTACCGAACCCTGAGCCCTCGTCTAAACGC&lt;br&gt;CTTAAATAATTGCTATTATATATTAGCGGCCCCCGGAATTACGTTCCGG&lt;br&gt;GGGCTTTTTTATT</td>
</tr>
<tr>
<td>sRNA-E</td>
<td>AGTTTACTGATGAGCCCGACACGAAACGCAACGGCTCCGTCGCTCAACGC&lt;br&gt;TACCAGGAATATTGCTAATAATTATTAGCGGCCCCCGGAATTACGTTCCGG&lt;br&gt;GGGCTTTTTTATT</td>
</tr>
<tr>
<td>sRNA-F</td>
<td>GCCAAACCTGATGAGCGACCGGTCTCGAAACGGTAGTCTCTCGTCTCTTT&lt;br&gt;GGCTGTGAAATTGACTAATAATTATTAGCGGCCCCCGGAACAACTTTCC&lt;br&gt;GGGCGCTTTTTTATT</td>
</tr>
</tbody>
</table>

*a. Underline indicates ribozyme, bold indicates sRNA tag, and italics indicates sRNA scaffold.*
In order to evaluate any toxicity resulting from the sRNAs, all six were expressed at high levels simultaneously and the fitness measured (Figure 2.2F). A significant growth defect was not observed. Additionally, synthetic sRNA dependence on Hfq, which is essential for the activity of natural sRNAs, was evaluated by testing sRNA-A in an hfq knockout strain (Figure 2.S8). No repression was observed suggesting that Hfq is necessary for repression from our engineered sRNAs. Since Hfq plays multiple important roles in the cell, including stress response, we were concerned that constitutive expression of sRNAs may titrate Hfq and cause growth defects in conditions where Hfq is crucial. Since Hfq plays a major role in gene regulation during iron starvation, we tested whether highly expressing all six of our sRNAs simultaneously would decrease fitness in cells grown in minimal media without iron (Figure 2.S8). A significant growth defect was not observed corroborating models presented in the literature that propose that Hfq-sRNA interactions are transient and allow rapid Hfq recycling as its dose does not affect growth or repression efficiency in these conditions (101-104).

2.4.2 sRNA function against chromosomal targets and diverse species

We evaluated the potential of our engineered sRNAs to knockdown chromosomal genes, in which the tag of sRNA-A was replaced with a 24nt sequence that exhibits perfect reverse complementarity to the coding sequence of the target (60) and the HHR appropriately modified. We knocked down expression of LacZ (β-galactosidase) and MalT, a regulator that indirectly decreases production of LamB, which decreases the susceptibility of E. coli to lambda phage infection (Figure 2.S6) (105). Both sRNAs targeting chromosomal genes exhibited activity, however, the magnitude of repression was not as great as of the sRNA targeting a characterized
target sequence. Despite reduced repression efficiency, our engineered sRNAs provide a quick way of targeting chromosomal genes without genome modification.

To better understand the effect of target copy number on sRNA-mediated repression, we tested our engineered sRNAs against a chromosomally integrated target. A reporter strain in which an sRNA-A-tagged RFP is constitutively expressed from the genome was constructed. Repression levels for the chromosomal target were consistent with those of a high copy target at low sRNA expression levels. After a critical point, repression saturates and increasing sRNA dose does not yield greater efficiency. This deviation from predicted repression levels occurs as RFP expression approaches background level, though RFP expression is not completely abrogated (Figure 2.2H).

As Hfq is conserved in most gram-negative bacteria (106), we evaluated the activity of sRNA-A in other gram-negative organisms (Figure 2.2G). The sRNA-A expression cassette and its corresponding GFP reporter gene were cloned into a broad host range plasmid and assayed for activity (pBBR-RNA, Figure 2.S2). sRNA activity was observed in both Pseudomonas and Salmonella, producing knockdowns of 16-fold and 28-fold notwithstanding that the sRNA was optimized for E. coli and expressed from a non-native promoter and terminator. These data suggest that the approach presented here for engineering sRNAs and building a tuning sRNA pool holds potential in other bacteria. Such a potent, reprogrammable regulator could be useful in a variety of organisms, particularly those in which protein-based regulators are limited.
Figure 2.2: Optimization of sRNA-mediated gene knockdown against target sequences. (A) Results of sRNA scaffold engineering. The reporter and sRNA are expressed from different plasmids as shown in Figure 2.S2. Sequences for all sRNA variants are provided in Table 2.1 and Table 2.S3. (B) Repression response of sRNAs expressed without (top panel) and with (lower panel) a ribozyme from 16 constitutive promoters of increasing strength. Sequences of all promoters are provided in Table 2.S1. (C) Sequence and structure of sRNA T(BF1176). The sRNA tag is shown in blue and underlined. Bold nucleotides are essential for Hfq-binding (97). Red nucleotides indicate the terminator stem loop of terminator BF1176. Arrows indicate nucleotides that were randomized to screen for improved scaffolds. (D) Design of library used to identify and diversify ribozymes and scaffolds for each sRNA tag. Purple nucleotides indicate the ribozyme; dashed line represents ribozyme cleavage. ‘X’ denotes sRNA tag sequence while ‘X’ denotes ribozyme nucleotides complementary to the tag. Arrows indicate nucleotides that were randomized to screen for improved, diverse ribozymes and sRNA scaffolds. (E) Orthogonality matrix for all six sRNAs and their target sequences. (F) Representative growth curves for DH10B E. coli in LB media
harboring an empty vector (circle) or a plasmid highly expressing all six sRNAs. (G) sRNA-A activity in other gram-negative bacteria. For this assay, the reporter and sRNA are expressed from the same plasmid as shown in Figure 2.S2. (H) Repression response of sRNAs targeting a chromosomally integrated reporter. The dashed line denotes complete gene repression (background level). For all data in this figure error bars correspond to the standard deviation of three experiments performed on different days.
2.4.3 Combinatorial sRNA pool construction and validation

An essential feature of the sRNA pool is its ability to explore a wide gene expression range, which is achieved by modulating sRNA expression. Ninety-six promoters (16 different promoters for each of 6 sRNAs) from a published set (29) were selected. The promoters were re-characterized in our plasmid and found to span strengths of 18 to 40,000 au (Table 2.S1). Each of the 96 promoter-sRNA combinations were evaluated for repression of both GFP and RFP to capture any variability due to the target gene sequence (Figure 2.3A and 2.S9, the Riboj insulator was included upstream of both targets (77)). For each sRNA, there is a strong correlation between promoter strength and repression and a graded response where the repression can be fine-tuned at intermediate promoter strengths (Figure 2.3A).

The next step was to build the sRNA pool by randomly combining the promoter-sRNA pairs so that each member of the pool expresses the six sRNAs to a different level. The sRNA pool was constructed using CombiGEM assembly, which is an iteration-based barcoded high-order combinatorial library assembly method that is amenable to high-throughput sequencing (82). Each of the 96 sRNA assembly plasmids (pPOOL, Figure 2.S2) contain strategically placed restriction sites that encompass a 10nt sequencing barcode associated with a specific promoter/sRNA pair, unique constitutive promoter, sRNA, and a strong terminator (107) (Figure 2.3B, Table 2.S1). The assembly reactions are then performed, the result of which is a barcode string that can be used to determine the promoter identity at each sRNA position with a 90bp sequencing read (Figure 2.3C, Materials and Methods). Deep sequencing of the final pool resulted in the recovery of 80/96 promoters from ~200,000 reads and an even distribution of promoter strengths (Figure 2.3D).
Recently Yeung and co-workers showed that the expression orientation of genes on a plasmid can have a large effect on gene expression levels (108). We evaluated any positional effects on the sRNA pool to see if the characterized repression levels were valid in the pool plasmid. A dual reporter and six sRNA pool plasmids, which express two corresponding sRNAs at three different levels from two extreme positions on the plasmid, were constructed (Figure 2S10). Repression did not significantly deviate from expected levels in either configuration. This may be attributed to the strong terminators, the short length of the full construct compared to one with multiple coding sequences, or the lack of translation/transcription coupling for sRNAs.
Figure 2.3: Pool assembly from differentially expressed sRNAs.

(A) Repression response for all six sRNAs against GFP (circles) and RFP (triangles) genes containing targets in their 5'-UTR. Each sRNA is expressed from a set of 16 promoters unique to each sRNA. Sequences of all promoters are provided in Table 2.S1. Error bars correspond to the standard deviation of three experiments performed on different days. (B) Example of pool assembly plasmids with restriction sites for CombiGEM assembly. The barcode symbol depicts a unique 10nt sequencing barcode corresponding to each promoter in the pool. Sequences for promoters and barcodes are provided in Table 2.S1. (C) Iterative assembly example of a six sRNA pool variant. (D) Deep sequencing results for promoter coverage in the assembled pool containing 6 RNAs (1% pool coverage with ~200,000 reads); '*' signifies that the promoter was not detected.
2.4.4 Application to a metabolic pathway: optimization of β-carotene

The β-carotene pathway was selected as a model system to test our sRNA pool because it is known that balanced enzyme levels increases yield (26,56,109-111). A single construct was built that includes the four genes to β-carotene (*crtE, crtB, crtI, crtY*) and the *E. coli* genes *dxs* and *idi*, which have been shown to improve yields when overexpressed (Figure 2.4A and B) (56,112). Each gene was placed under the control of a strong wild-type T7 promoter. Each enzyme gene is preceded by a unique insulating ribozyme, strong designed RBS and sRNA target sequence (A-F) (Figure 2.4B). Six strong terminators were placed after each gene. The plasmid map (pCAROTENE) and part names/sequences are provided in Figure 2.5S2 and Table 2.5S3. T7 RNA polymerase (RNAP) is expressed from a separate plasmid under IPTG-inducible control (78).

The sRNA pool was transformed into a strain harboring the pathway, generating a library of variants. More than 1200 colonies were randomly picked, induced with 1 mM IPTG, and assayed for β-carotene production (Materials and Methods). Approximately half of the variants improved yield from a starting titer of 43±7 mg/L, with the best one producing a titer of 204±25 mg/L (Figure 2.4C).

sRNA tuning plasmids from the top ten strains (161-204 mg/L) were isolated and barcodes sequenced. From this, the repression of each gene can be elucidated for these strains (Figure 2.4D). When a system is highly affected by the expression level of a gene, the repression levels across these genes are more similar. This suggests that the pathway is most sensitive to expression levels of Dxs and CrtY. In all ten high titer strains, Dxs, an enzyme known to cause growth defects when overexpressed (113), is lowly expressed (>10-fold repression), while CrtY, the final enzyme in the pathway, is highly expressed (<2.5-fold repression). Protein expression
levels of all genes in the pathway, measured by fusing YFP to the first 30 codons of each gene in the sRNA-optimized system, corroborate predicted expression levels (Figure 2.S11).

The analysis of the top strain reveals the optimal repression of each gene. Literally using this strain for production would require carrying the sRNA plasmid and wasting cellular resources on mRNA production of the genes. Instead, we sought to recapitulate this balance without sRNAs by constructing a pathway with T7 promoters that correspond to the optimal expression level for each enzyme. A T7 RNAP promoter library had been previously constructed and we selected the promoter for the optimal expression level for each gene (Figure 2.4E and Table 2.S2). After these substitutions and with no additional tuning, the pathway produces 161±19 mg/L β-carotene (Figure 4.4F).
Figure 2.4: Optimization of β-carotene production using the sRNA pool.

(A) The metabolic pathway for β-carotene production is shown. Enzymes that are essential or known to improve β-carotene production are circled and colored. (B) The genetic design for the β-carotene production pathway including the insertion of sRNA targets. The plasmid map and part sequences are provided in Figure 2.S2 and Table 2.S3. (C) β-carotene production for 1200 screened variants. The orange line represents the titer from the un-tuned β-carotene pathway where an empty vector (pEMPTY) is used in place of a sRNA pool plasmid. Quantification of titers is described in Materials and Methods. (D) The repression of each gene for the ten highest β-carotene producing strains (in order from left to right). (E) Calculated optimal T7 promoter strength (Materials and Methods) for each gene (left axis). The box to the right shows T7 promoters available in the library. Solid lines extended across the graph represent T7 promoters from the library that were selected to hard-code the optimal expression levels. T7 promoter sequences are provided in Table 2.S2. (F) β-carotene production for the top sRNA-tuned pathway and its corresponding reconstructed pathway with substituted T7 promoters. Error bars correspond to the standard deviation of three experiments performed on different days.
2.4.5 Application to a genetic circuit

To evaluate the potential of the sRNA pool in tuning circuits, a two-input XNOR circuit was designed and constructed by layering transcriptional NOT and NOR gates (Figure 2.5A, B, and C). Our paradigm for circuit construction requires fixed promoters, so we tune the expression levels of repressors using RBSs (28,55). The circuit was constructed with repressors harboring RBSs generated with the RBS Calculator (Materials and Methods) each preceded by an sRNA target sequence (targets A-D) and an insulating ribozyme. The plasmid maps (pCIRCUIT and pOUT) and part names/sequences are provided in Figure 2.S2 and Table 2.S1. This “blindly” constructed XNOR circuit does not produce the correct logic operation (Figure 2.5D).

Because the circuit output is a fluorescent protein, FACS can be used to rapidly screen for improvements. The sRNA pool was co-transformed with the circuit and then the library was serially sorted for correct output under the different combinations of inputs (1 mM IPTG and 2 ng/ml aTc) (Figure 2.5E, Materials and Methods). Ninety-six colonies were individually assayed for the XNOR circuit, and 21 exhibited XNOR functionality. The high false-positive rate could be attributed to low sort purity. The best performing circuit exhibited the correct XNOR logic and 6-fold activation between the highest OFF state and lowest ON state (Figure 2.5F). The optimal repression of each gene was determined by sequencing the barcode on the sRNA tuning plasmid (Figure 2.5G).

While the circuit now works, it is not ideal to have to continue to carry the sRNA construct. Therefore, RBS substitution was carried out to recapitulate the optimal levels. The RBS calculator was used to predict the strength of each starting RBS and then this is divided by the desired fold-repression in order to get the desired strength of the new RBS (Figure 2.5H,
Materials and Methods). The circuit was reconstructed with the new RBSs and was found to be functional with 4-fold activation (Figure 2.51).
Figure 2.5: Optimization of a XNOR circuit with the sRNA pool.

(A) The circuit diagram for the XNOR function is shown. Inputs are sensor promoters (P_lac and P_ara) that are activated by IPTG and aTC, respectively. (B) Truth table for a two-input XNOR gate. (C) The genetic design for the XNOR circuit is shown including the insertion of sRNA targets. Blunt ended lines represent the repression of promoters. The plasmid map and part sequences are provided in Figure 2.S2 and Table 2.S3. (D) Response of an untuned XNOR circuit where empty vector, pEMPTY, is used in place of a sRNA pool plasmid. Inputs correspond to the absence or presence of 1 mM IPTG (left ‘0’ / ‘1’) and aTC (2 ng/ml; right ‘0’ / ‘1’). (E) Sort gates (shown in gray) used for each input. The sorting procedure is described in Materials and Methods. (F) Response of the best sRNA-tuned XNOR circuit. (G) The sRNA-mediated knockdown of each repressor gene in the tuned circuit. (H) Calculated RBS strengths (Materials and Methods) for the original untuned circuit (black bars) and for the reconstructed (new RBSs) tuned circuit (white bars). (I) Response of the reconstructed circuit. For all data, error bars correspond to the standard deviation of three experiments performed on different days.
2.5 Discussion

We developed a method to tune genetic systems by exploring large gene expression spaces without having to build ad hoc libraries. This method can be used to tune a variety of multi-gene systems, and presents a novel approach to genetic system optimization by using trans-acting RNA regulators that target standardized elements. Further, once a functional variant is discovered from the sRNA pool, the sRNA tuning plasmid, which may be undesirable to maintain in a production context, can be used to extract information on relative gene expression levels that can be used to hard-code these levels using characterized genetic parts.

By simply co-transforming our sRNA pool, a library is generated that can be immediately used for screening. Compared to other tuning methods, such as RBS or promoter screening, a library of constructs in which different parts are tried for each gene does not need to be assembled—only a single construct harboring sRNA targeting sequences is built. Even with modern DNA assembly techniques, such as Golden Gate assembly, library assembly requires multiples cloning steps for high assembly efficiency. For example, a combinatorial library for six-gene system with one variable part per gene (i.e. RBS or promoter) requires joining at least 13 pieces (6 variable parts, 6 transcription units, and a backbone), and would need to be performed in multiple stages to ensure sufficient efficiency and coverage (21,114). Additional steps equates to more reagents, labor, and time before the library can be screened. With our method, the arduous assembly process only needs to be performed once as the sRNA pool is not specific to a particular genetic system. Moreover, we show that the pool can explore a tuning range of about 50-fold for each gene. Though promoter and RBS tuning can achieve expression levels spanning more than two orders of magnitude, such vast ranges are not necessary for first-pass tuning, especially when starting levels are already high (20,115).
Trans-acting tuning approaches that are currently used, including zinc-finger proteins (ZFP) and CRISPRi, require expressing additional proteins, and can vary widely in efficiency \((116,117)\). Using ZFPs to regulate six genes would require expressing six ZFPs, which can burden cells, particularly when expressed at levels necessary for strong repression. Additionally, the tuning range with ZFPs is limited as Hsia and co-workers recently reported only 12-fold induction in a simple ZFP invertor circuit \((118)\). Our sRNA approach leverages standardized targeting to achieve predictable repression. We systematically validate this by targeting two different genes with all six sRNAs over a wide range of expression levels as shown in Figure 2.3A and Figure 2.310. Standardized targeting may be applied to other methods that use trans-acting regulators, such as CRISPRi, however, targeting must provide a sufficient tuning range that can be explored via differential expression of the regulator (i.e. guide RNAs) or other means. In the case of CRISPRi, dCas9, the protein necessary for repressor, may be limiting in large systems as dCas9 is not recycled and high expression leads to toxicity \((119,120)\). Our system averts a regulatory protein bottleneck as Hfq-sRNA interactions are transient and allow the reuse of Hfq.

One limitation of our approach is that genes can only be repressed, as opposed to being up-regulated. This constrains the expression space to the maximum level of the initial promoters. This can be problematic if any of the genes are toxic. This was averted in the β-carotene pathway by using T7 promoters and the expression of T7 RNAP on a separate plasmid. This avoids the problem of toxicity because the genes will not be expressed until T7 RNAP is induced. For the circuit, this was a more challenging problem as over expression of some of the repressors was toxic, and since the promoters are specific to the system, we could not replace them with weaker ones. Therefore, we had to select weaker RBSs so that the starting construct could be made
before co-transforming with the pool. Up-regulating sRNAs have been discovered (121,122) and may be amenable to a tuning pool; however, systematic characterization and engineering of these sRNAs is still required. Another limitation of our approach is dynamic range. With approximately a 50-fold window, low to moderate expression levels are not explored when starting with high expression. This may be remedied by weakening RBS strength of the target, but this may not be apparent until initial screening is performed. Furthermore, low copy targets, such as those on the genome, have a reduced tuning range as repression was observed to saturate after a certain sRNA expression level. This is not an issue for systems expressed from multi-copy plasmids, but it may reduce the utility of the pool for chromosomally integrated systems.

The reusable trans-acting RNA pool approach presented in this work establishes a new paradigm for genetic system optimization. Combinatorial optimization in genetic engineering has been very successful, but the libraries are expensive and technically challenging to build. It is particularly challenging if the constructs have to be inserted chromosomally as transferring a construct from a plasmid to the chromosome often requires re-tuning. Our approach is to build a library of regulators once and then use this library to rapidly optimize systems. It is sufficiently simple to be routinely or systematically applied when building constructs for new pathways. For example, if pathways are being mined en masse from microbiome but the required expression levels of the targets are unknown. The size of the system to be optimized is more limited by screening capacity than the sRNAs. Here, 6 sRNAs leads to a library size of $\sim 10^7$, which is appropriate for FACS. Optimizing a larger genetic system just requires breaking the genes into smaller sets for co-optimization. This method can be applied simply, without the need for fancy robotics and automation platforms.
2.6 Supplementary Figures

Figure 2S1: sRNA knockdown mechanism and assay.

(a) A schematic of sRNA-mediated knockdown via translation inhibition and accelerated mRNA degradation is shown. (b) The two plasmid knockdown assay used. Promoter J23119, a strong promoter (25,388 au, arbitrary units, Table 2.S1), and terminator B0010 were used to drive sRNA expression from a pBR322 plasmid while GFP or RFP was expressed from a p15A plasmid with promoter J23101 (17,681 au, Table 2.S1), RBS B0032, terminator B0012, and RiboJ, an insulating ribozyme (77).
Figure 2.S2: Plasmid Maps.
(a) The general sRNA expression plasmid that was used for all sRNA optimization experiments. Only the sRNA sequence was modified to test different designs. (b) The general reporter plasmid that was used for all sRNA optimization experiments. (c) The empty vector that was used for fold repression normalization for all sRNA knockdown assays, background subtraction for the pathway, and as an untuned control for the circuit. (d) The general promoter characterization plasmid. Only the promoter was replaced with constitutive or T7 promoters to calculate its strength (Materials and Methods). Promoter sequences and strengths are provided in Tables 2.S1 and 2.S2. (e) The general sRNA pool assembly plasmid. Part sequences for all 96 sRNA assembly plasmids are provided in Table 2.S1. (f) The β-carotene pathway plasmid. For the reconstructed pathway T7 promoters were substituted for those listed in Table 2.S2. The genetic design and part sequences are provided in Figure 2.4A and Table 2.S3. (g) The XNOR circuit and output plasmids. The genetic design and part sequences are provided in Figure 2.5A and Table 2.S3. (h) The broad host plasmid used for knockdown assays in Salmonella and Pseudomonas. Only this plasmid was transformed for the assays. Pseudomonas cells were grown at 30°C. (i) The dual reporter plasmid that was used for pool context experiments.
Figure 2S3: β-carotene standard curve and LC analysis.

(a) Standard curve used for β-carotene quantification using spectrophotometry. Background absorbance has been subtracted. Pigment extraction is detailed in Materials and Methods. Error bars correspond to the standard deviation of three experiments performed on different days. (b) Liquid chromatography analysis of the β-carotene standard. The trace shown is for the standard at a concentration of 10 mg/ml in acetone measured at 470 nm on the diode array detector. (c) Liquid chromatography analysis of the extract from the top β-carotene strain. *E. coli* cells harboring the pathway, T7RNAp, and the sRNA plasmids were harvested after 48 h of growth at 30°C in terrific broth.
Figure 2.S4: sRNA target position and length optimization.

(a) A schematic depicting target sequence placement (numbered with respect to the 5' end of the RBS) on the reporter (GFP or RFP) mRNA is shown. The RBS is shown as a filled semi-circle and is defined as the 35nt preceding the start codon. ‘*’ denotes the position immediately before the RBS. (b) Fold repression of GFP (squares) and RFP (circles) harboring a 20nt target sequence placed at varying positions with respect to the RBS. (c) A schematic depicting the placement of target sequences of different lengths in the 5' UTR of the gene. (d) Fold repression of GFP (squares) and RFP (circles) harboring a target sequence with different lengths positioned immediately before the RBS. For all data, error bars correspond to the standard deviation of three experiments performed on different days.
Figure 2.S5: sRNA-terminator chimera screening.

(a) Fold repression (black bars) and termination strength (gray bars) of sRNA-terminator chimeras. The termination assay described in Chen et al. was used to calculate termination strength (107). For all data, error bars correspond to the standard deviation of three experiments performed on different days.
Figure 2.S6: Genomic targeting and sRNA-mediated knockdown in other organisms. (a) The lambda phage infectivity response is shown for an sRNA targeting MalT, where infectivity is measured by the number of lambda phage plaques formed on a bacterial lawn on an agar plate. (b) LacZ activity for an sRNA targeting genomic lacZ. Error bars represent the standard deviation of three independent experiments on different days.
Figure 2.S7: sRNA orthogonality plots.

Cross reactivity for all sRNA tag / targeting sequence combinations is shown. Fold repression less than 1 corresponds to up-regulation. For all data, error bars correspond to the standard deviation of three experiments performed on different days.
Figure 2.S8: Δhfq strain and hfq titration characterization.

(a) Knockdown efficiency for sRNA-A in a Δhfq strain (*E. coli* JW4130-1) and its wild-type parent strain (*E. coli* BW25113) is shown. Error bars correspond to the standard deviation of three experiments performed on different days. (b) Representative growth curves for MG1655 *E. coli* in M9 media (no iron) harboring an empty vector (circles) or a plasmid highly expressing all six sRNAs (triangles).
Figure 2.S9: GFP / RFP knockdown correlation plots for sRNAs

A correlation plot for GFP and RFP knockdown levels for each sRNA / promoter combination in the tuning pool is shown. The line depicts perfect correlation (slope of 1). The following slopes and $R^2$ values were calculated for each correlation: sRNA-A: slope = $1.183 \pm 0.02070$, $R^2 = 0.9960$; sRNA-B: slope = $1.022 \pm 0.05668$, $R^2 = 0.9587$; sRNA-C: slope = $0.9815 \pm 0.09532$, $R^2 = 0.8934$; sRNA-D: slope = $0.8695 \pm 0.04369$, $R^2 = 0.9659$; sRNA-E: slope = $1.069 \pm 0.04634$, $R^2 = 0.9743$; sRNA-F: slope = $1.016 \pm 0.03569$, $R^2 = 0.9830$. 
Figure 2.S10: sRNA pool context effects

(a) A schematic depicting the dual reporter (pDUAL) and the two sRNA plasmid configurations.

(b) GFP and RFP repression levels resulting from sRNAs expressed from weak ($P_{A5}$ and $P_{B3}$), medium ($P_{A12}$ and $P_{B10}$), and strong ($P_{A16}$ and $P_{B16}$) promoters in configuration 1 (black bars) and configuration 2 (white bars) are shown. Error bars correspond to the standard deviation of three experiments performed on different days.
Figure 2.S11: Protein expression levels of sRNA-optimized β-carotene pathway

Measured repression of all genes in the sRNA-optimized β-carotene pathway (black bars) and their predicted repression (white bars) are shown. Predicted repression levels are reported as the average knockdown level of GFP and RFP (Figure 2.3A) for each promoter and sRNA tag combination. For all data, error bars correspond to the standard deviation of three experiments performed on different days.
3. Tunable CRISPR-based regulation of multiple genes in eukaryotes using T7 RNA polymerase

3.1 Abstract

CRISPR-based gene regulation is a powerful strategy for perturbing and engineering virtually any genetic system. Guide RNAs (gRNA) that specifically and efficiently target and up- or down-regulate genes of interest can be useful tools for tuning genetic circuits and metabolic pathways. Here, we develop a method to produce gRNAs, using T7 RNA polymerase (T7RNAp), that obviates the need for characterized eukaryotic RNA Pol III promoters (R3p), allows gene expression tuning, and enables easy multiplexing. We show that T7RNAp can produce highly levels of functional gRNAs in vivo in Saccharomyces cerevisiae. We optimize gRNA targeting and T7RNAp-produced gRNAs are compared to those produced using native R3p promoters. We then show that modulating the production of gRNAs using T7 promoters of different strengths can be used to tune gene expression, and that multiple genes can be targeted simultaneously with gRNAs produced from different T7 promoters. Finally, we apply this tool to a carotenoid pathway and show that we can control metabolic flux and titer using these gRNAs.
3.2 Introduction

A plethora of genetic tools to control gene expression in prokaryotes have been developed, including engineered promoters and ribosome-binding sites \((20,80)\), transcriptional activators \((123)\) and repressors \((55)\), small RNAs \((124)\), RNA stability elements \((22)\), ribo-regulators \((125)\), and protein degradation tags \((126)\). However, few characterized tools for precise gene regulation exist in eukaryotes, particularly mammalian, plant, and fungal cells due to their complex transcription and translation mechanisms. Though native gene regulation methods, including microRNAs, have been adapted for synthetic applications, their efficiency greatly varies and not all organisms have such machinery \((127,128)\). Efforts to import gene regulation tools, such as repressor / operator pairs, developed in prokaryotes have also been pursued; however, low efficiency and a limited number of parts that are functional in eukaryotes have hindered their use \((129,130)\).

CRISPR-based gene editing is revolutionizing genome engineering as it enables rapidly programmable precise DNA cleavage. In particular, the Cas9 CRISPR system (from the bacteria *Streptococcus pyogenes*) has been shown to be almost universal as it is functional in a variety of organisms spanning all kingdoms of life \((131,132)\). In this system, Cas9, an endonuclease, is expressed along with a short synthetic RNA called a guide RNA (gRNA) that encodes a hairpin that binds Cas9 and a targeting region (>13nt) that determines what DNA sequence Cas9 will cut; the DNA target must be adjacent to a short sequence called a protospacer adjacent motif (PAM) \((133,134)\). The ability for such machinery to function in so many different organisms is quite remarkable as intracellular conditions can vary vastly.

In 2012 Qi and co-workers developed a catalytically dead Cas9 (dCas9), in which two point mutations to the protein were made. dCas9 is unable to cleave DNA, but it is still able to
bind DNA (15). Since the binding affinity for Cas9-gRNA to DNA is very strong (~1 nM) and the protein is relatively large (160 kDa), it can function as a potent transcriptional repressor by physically blocking RNA polymerase and inhibiting transcription (135). Further, protein fusions of dCas9 to eukaryotic repressor domains, such as Mxi1, and transcriptional activation domains, such as VP64, have resulted in even more effective gene regulation (16,17).

Early studies have proven the robustness and efficacy of dCas9 for gene regulation in eukaryotic cells; however, it lacks many features that could immensely expand its utility in eukaryotes, such as the following:

1) Simple guide RNA production without having to find and characterize RNA Pol III promoters (R3p),

2) Ease of targeting multiple genes, and

3) Tunability of gene repression or activation.

We sought to address these issues by focusing on gRNAs and their production instead of the dCas9 protein. Currently, gRNAs must be transcribed with Rp3 as this polymerase does not attach 5’ caps, which are used by the ribosome for docking, to transcripts (132,133). However, in most organisms few if any Rp3 promoters have been characterized (knowing the precise transcript start site is particularly important for gRNA production) (136). Further, once an appropriate Rp3 promoter is found to produce functional gRNAs, it must be re-used to produce multiple gRNAs that target different genes. This is not ideal as Rp3 promoters are typically >200 bp and can lead to difficulties in assembly and plasmid instability when many repeats of it are used in one construct (137,138). Finally, gRNA expression levels are fixed when using Rp3 promoters. Similar to transcriptional repressors and RNA regulators, such as small RNAs, modulating gRNA levels may be an effective method for tuning gene expression.
We hypothesized that we could address all of these issues by using an orthogonal transcription system to produce gRNAs. We selected T7 RNA polymerase (T7RNAp) for this application as it is well characterized and has been shown to be functional in many different organisms including eukaryotes (78,139). T7RNAp functions without any additional cofactors and exhibits very high processivity (140). Additionally, T7RNAp promoters are quite short in length (~23 bp) and are tunable (78,141). To explore the potential of T7RNAp-transcribed gRNAs in eukaryotes, we developed a knockdown assay in *Saccharomyces cerevisiae* in which we express dCas9, T7RNAp, and gRNAs (produced from a T7 promoter) targeting a strong constitutive promoter driving a fluorescent reporter (Figure 3.1).
3.3 Materials and Methods

3.3.1 Yeast strains and media.

Saccharomyces cerevisiae CEN.PK 13-7D (MATa ura3-52 trp1-289 leu2-3_112 his3 Δ1/Δ1 MAL2-8C SUC2) was modified and used for all experiments. YPD broth (Sigma-Aldrich, Y1375) was used for general growth and cloning. YEP-GAL broth (YEP: MPbio, 114004032; 2% D-Galactose: Sigma-Aldrich, G0750) was used for all assays. Yeast growth curves were measured on a BioLector I (m2p-labs). Biomass was measured at 620 nm in a 48-well FlowerPlate (m2p-labs) at 30°C and 750 rpm.

3.3.2 E. coli strains and media.

Escherichia coli DH10b (F- mcrA Δ(mrr-hsdRMS-mcrBC) Φ80lacZΔM15 ΔlacX74 recA1 endA1 araD139 Δ(ara leu) 7697 galU galK rpsLupG λ−) was used for all plasmid construction, and T7 promoter and terminator characterization. E. coli cells were grown in LB Miller broth (Difco, 90003-350) or SOC (Teknova, S0210) supplemented with 0.4% glucose) for cloning. Supplemented M9 minimal media (1x M9 Salts (Sigma-Aldrich, M6030), 2 mM MgSO4, 100 μM CaCl2, 0.4% glucose, 0.2% casamino acids, 340 mg/L thiamine) was used for T7 promoter and termination characterization. Chloramphenicol (34 μg/ml) (Alfa Aesar, AAB20841-14) was supplemented where appropriate. IPTG (Sigma-Aldrich, I6758) was used to induce T7RNAp expression.

Yeast genome modifications, plasmids, and transformations. Genome modification and plasmid transformations were performed using the Zymo Frozen EZ Yeast Transformation II Kit (ZymoResearch, T2001). For genome integrations all pieces to be inserted (all harboring >50bp
homology to adjacent pieces), including the 5’ and 3’ flanks with homology to chromosome XV, and a hygromycin B expression cassette, were transformed and plated on YPD plates with Hygromycin B (200 mg/mL) (ThermoFisher, 10687010). Resultant colonies were screened and confirmed by PCR. All yeast plasmids were constructed in E. coli and then transformed into yeast and plated on YPD plates with G418 sulfate (200 mg/mL) (ThermoFisher, 10131035) and/or Nourseothricin sulfate (50 mg/mL) (GoldBio, N-500-100).

3.3.3 Flow cytometry.

Fluorescence was measured using a MACSQuant VYB (Miltenyi Biotec) with a 488-nm laser for GFP and YFP excitation and 561-nm for RFP excitation. For each sample at least 5 × 10⁴ events were recorded using a flow rate of 0.5 μL/s. FlowJo v10 (TreeStar Inc.) was used to analyze the data. All events were gated by forward scatter and side scatter. Fluorescence values are shown as the geometric mean. Fold-repression is calculated as the geometric mean of the fluorescence measured with an empty vector divided by the geometric mean of the fluorescence measured with a gRNA expression vector.

3.3.4 gRNA repression assay.

Yeast colonies (on three different days) were picked and inoculated into 1mL YPD with selection in 14mL round bottom culture tubes (FisherSci, 14-959) in a Multitron Pro incubator shaker (In Vitro Technologies) and grown for 24 h at 30°C. These cells were then diluted 1:200 into YEP-GAL with selection and grown for 24 h at 30°C. Twenty microliters of the culture was diluted into 180 μl 1X phosphate buffered saline (PBS) then analyzed on the flow cytometer.
3.3.5 Promoter strength measurements.

T7 promoter characterizations were performed by co-transforming the T7RNAp expression plasmid N249 (78) with a GFP reporter plasmid (Figure 3.82). Reported promoter strengths are the geometric mean of the measured fluorescence output. For these assays colonies were grown overnight and diluted 1:200 into LB with antibiotics and 1 mM IPTG and grown for 6 h at 37°C, then analyzed on the flow cytometer.

3.3.6 Termination Assay

The termination strength assay described in Chen et al. (107) was used to measure terminator strength with the following modifications: 1) The pGR test plasmid was modified such that the pBAD E. coli promoter, which drives the expression of the operon, was replaced by a wild-type T7 promoter. 2) T7RNAp expression plasmid N249 was co-transformed with pGR. 3) Cells were grown overnight, diluted 1:200 in LB with antibiotics and 1 mM IPTG and grown for 6 h at 37°C, then analyzed on the flow cytometer.

3.3.7 Computational Methods.

All nucleosome occupancy predictions were performed using a web-based tool (https://genie.weizmann.ac.il/software/nucleo_prediction.html). The following parameters were used for all predictions: Model Version: Version 3.0 (Kaplan et al. Nature 2008); Nucleosome concentration: 0.1; (Inverse) Temperature: 1.
3.3.8 β-carotene production quantification.

The gRNA repression assay was used for β-carotene quantification with the exception that the final growth in YEP-GAL was scaled up to 5mL. After 24 h induction in YEP-GAL cells were pelleted and washed with water. Cells were then resuspended in 1.5 mL acetone (in 2 mL polypropylene microfuge tubes (FisherSci, 21-403-192)) and incubated at 55°C for 30 min with frequent vortexing. Pigment extraction mixtures were then centrifuged at 14000g for 10 min. 0.2 ml of supernatant was transferred to a polypropylene 96-well plate (Greiner, 655201) for quantification. Samples were measured using a H1 Synergy plate reader (BioTek Instruments) at an absorbance wavelength of 470 nm. A standard curve, using purified β-carotene (Sigma-Aldrich, 22040), was established to correlate absorbance to amount of β-carotene produced. Background absorbance was measured from a strain harboring no carotenoid genes.
3.4 Results

3.4.1 Optimizing T7RNAp-mediated gRNA production

One of our goals was to obviate the need to use native Rp3 promoters for dCas9-mediated gene regulation. Therefore, we first wanted to recapitulate the repression achieved by Rp3-transcribed gRNAs. For our experiments, we integrated T7RNAp and GFP into the genome (Figure 3.S1). T7RNAp is expressed from GAL1, a galactose inducible promoter, and GFP is expressed from TEF1, a strong constitutive promoter. dCas9 is constitutively expressed (TDH3 promoter) from a low-copy plasmid while gRNAs targeting the GFP promoter were expressed from a high-copy plasmid. In an effort to maximize the magnitude of repression, the repressor domain from the mammalian gene MxiI was fused to the C-terminus of dCas9 as it has been shown to increase repression efficiency by 3-fold (17).

We reasoned that high expression levels of gRNAs would result in strong repression of the target as is observed with other trans-acting regulators, such as the sRNAs described in Chapter 2. Therefore, for our initial experiments we used the strongest T7 promoter (wild-type) to produce gRNAs. A known problem with the T7RNAp expression system is inefficient termination resulting in undesired long 3’ ends on transcripts (142). Run-on transcription may hinder the formation of crucial gRNA secondary structures and proper binding to dCas9 and target DNA. To address this issue, we assayed multiple terminators and found that strong E. coli terminators (used to stop T7 transcription) produced the most potent gRNAs (Figure 3.S2).

After finding a suitable terminator to produce gRNAs, we compared the repression efficiency of T7RNAp-produced gRNA to that of the same gRNA produced from the native yeast R3p, SNR52 (Figure 3.1) (143). The T7 promoter achieved repression comparable to that achieved by SNR52. Though encouraging, we expected greater repression efficiency from the
T7RNAp system, as it should produce many more transcripts than the SNR52 promoter. This discrepancy can be explained by variety of reasons including, inefficient T7RNAp expression, improper gRNA folding and targeting due addition of a ‘G’ to the 5’ end by T7RNAp (140), or run-on transcription despite terminator optimization.
Figure 3.1: dCas9-mediated repression with T7RNAp-produced gRNAs.

(A) Schematic of guide RNA production from yeast RNA polymerase III promoter, SNR52, and tRNA terminator, SUP4. Raised arrows denote promoters and “T”s denote terminators. Positive regulation is indicated with arrows, and negative regulation is indicated with flat-headed arrows.

(B) Schematic of guide RNA production from a T7 promoter and T7 terminator. (C) Knockdown efficiency of different promoters is shown. Error bars correspond to the standard deviation of three experiments performed on different days.
Once we were able to match the repression efficiency of gRNAs produced from SNR52, we explored ways to modulate repression. Since DNA in eukaryotes is organized in nucleosomes, access to DNA is often occluded by histones (144). Horbleck and colleagues have recently shown that nucleosome occupancy, the probability that a specific region of DNA will be wrapped around a histone, affects CRISPR targeting efficiency (145). Many computational models have been developed to predict occupancy associated transcriptional efficiency (146,147). We used the yeast nucleosome occupancy model developed by Kaplan and coworkers (148) to predict dCas9 targeting efficiency. The model predicted one major occupancy peak which we validated experimentally by targeting every possible site (adjacent to an ‘NGG’ PAM) on the promoter (Figure 3.2). A strong correlation between low nucleosome occupancy and dCas9-mediated repression efficiency was observed.
Figure 3.2: Optimizing guide RNA targeting of yeast promoters.

(A) Schematic showing guide RNA target sites on the output promoter. Stars depict targets sites (containing the required ‘NGG’ PAM sequence for dCas9). Stars on top of the line represent target sites on the + strand while stars below it on the – strand. (B) Repression efficiency (circles, left axis) of all target sites on the promoter is shown along with the predicted nucleosome occupancy (line, right axis) at each site. Black circles represent target sites on the + strand while white circles represent sites on the – strand. Error bars correspond to the standard deviation of three experiments performed on different days.
Though the correlation is not perfect, the model provides a way to tune gene expression by targeting different sites to achieve varying levels of repression. T7RNAp enables another tuning knob for repression, promoter strength. Unlike most eukaryotic promoters, T7 promoters are short and the precise nucleotides responsible for T7RNAp specificity and strength are known (141). Using the wild-type T7 promoter as a starting point, degenerate promoter libraries in which the promoter strength determining nucleotides are randomized are commonly used to find optimal gene expression levels in prokaryotic systems (78,149) (Figure 3.3A). In order to test if modulating gRNA expression levels would affect target gene expression, a 4^5 variant library of T7 promoters to express gRNAs was constructed. More than 100 variants, producing ~2- to 40-fold repression, were screened from which 16 were selected and characterization.

3.4.2 Modulating gene expression with T7 promoters

The selected promoters were sequenced and cloned into an E. coli vector in front of GFP and used to quantify promoter strength (Figure 3.3B). Since T7RNAp does not cap transcripts, which is essential for efficient translation in eukaryotes, the strengths of the T7 promoters were measured in E. coli. T7RNAp is a completely orthogonal system; therefore, we reasoned that relative T7 promoter strength measured in E. coli would be preserved regardless of the host. A strong correlation was observed between T7 promoter strength (measured in E. coli) and dCas9-mediated gene repression (Figure 3.3C). This confirmed our hypothesis of controlling gene expression by modulating gRNA levels.

T7RNAp is widely used in E. coli not only due to the tunability of T7 promoters, but the ability to simultaneously produce multiple transcripts from different promoters. In eukaryotes, R3ps are limited so they are often reused which can lead to genetic instability as large sequence
homology can lead to recombination. To alleviate this, we sought to multiplex gRNA production using multiple T7 promoters. An expression vector with two T7 promoters driving gRNAs targeting promoters driving GFP and RFP was constructed along with a construct in which the same gRNAs were expressed from two copies of SNR52. Interestingly, dual gRNA expression from T7 promoters performed on par with the dual R3p construct (Figure 3.3D).
Figure 3.3: T7 promoter library construction, characterization, and gRNA multiplexing.

(A) Wild-type T7 promoter sequence and library design is shown. Nucleotides that determine polymerase specificity are underlined while those that determine promoter strength are shown in bold. ‘N’ denotes a degenerate nucleotide (i.e. ‘A’, ‘T’, ‘G’, or ‘C’).

(B) Plasmids used for T7 promoter assay in E. coli. (C) Effect of T7 promoter strength (characterized in E. coli) on repression is shown. T7 promoters were used to express guide RNAs targeting a promoter driving the expression of GFP. Line shows linear regression of the relationship between promoter strength and repression. (D) Multiplex repression of GFP (black bars) and RFP (gray bars) from different promoters. Error bars correspond to the standard deviation of three experiments performed on different days.
3.4.3 Applying dCas9-mediated repression to a metabolic pathway

Once the efficacy of these new tools for dCas9-mediated repression was established, we wanted to explore their potential in engineering metabolic pathways. The well-studied carotenoid production was selected as our test bed. We constructed the carotenoid pathway from *Xanthophyllomyces dendrorhous* as described by Verwaal and coworkers (150) (Figure 3.4A). CrtYB is the final enzyme in the pathway to β-carotene and must be highly expressed for a high titer, therefore, pigment production is very sensitive to its expression levels. In our pathway construct, we express crtYB from TEF1, a promoter for which we already have an efficient gRNA with characterized knockdown levels that we can explore from different T7 promoters (Figure 3.4B). Using our set of characterized T7 promoters, we were able to precisely control β-carotene production (Figure 3.4C).
Figure 3.4: Tuning β-carotene production by modulating guide RNA expression.

(A) Schematic of β-carotene production in yeast using the carotenogenic pathway from *Xanthophyllomyces dendrorhous.* (B) Strategy to tune β-carotene production using dCas9 is depicted. (C) Effect of CrtYB repression on β-carotene production is shown. CrtYB repression is achieved by using T7 promoters of varying strengths to target the promoter for CrtYB. Line shows linear regression of the relationship between CrtYB repression and β-carotene production. Error bars correspond to the standard deviation of three experiments performed on different days.
3.5 Discussion

We developed a method to expand the utility of CRISPRi in eukaryotes by producing gRNAs from an orthogonal RNAp. By using T7RNAp instead of an endogenous R3p, the number of organisms in which CRISPRi can be used is drastically increased, dCas9-mediated regulation of multiple genes is simplified, and tuning using different strength T7 promoters is enabled. Once dCas9 and T7RNAp expression constructs are built and verified, a vector expressing gRNAs from a WT T7 promoter is used to find efficient target sites on the gene of interest. After the optimal site is discovered, a library of T7 promoters can be built to explore different expression levels. We show that a wide expression range can be achieved, large multi-gene systems can be regulated, and that multiple genes can be targeted concurrently.

The ability to precisely control gRNA expression and achieve graded target expression levels holds great potential. Large genetic systems, including genetic circuits and metabolic pathways, often require each constituent gene to be precisely tuned to function properly or maximize yield. The ability to explore a large gene expression space is challenging even in well-studied organisms, such as *S. cerevisiae*. Currently, promoter libraries are constructed to vary the expression level of each gene. With each additional gene, library construction becomes increasingly difficult and inefficient. Since gRNAs are trans-acting regulators that yield a graded response when differentially expressed, a pool-based tuning approach as described in Chapter 2 can be pursued. Additionally, as gene up-regulation can be achieved with dCas9 fusions (16,17), the explorable gene expression space is potentially even larger.

A limitation of our method is the need to express two heterologous proteins, dCas9 and T7RNAp, which requires parts native to the host for expression. Further, since the promoter sequence for T7RNAp is only 18 bp, rogue transcripts are possible in larger genomes, which can
potentially lead to toxicity. Additionally, this work only evaluates the efficiency of two gRNAs concurrently produced by T7RNAp. As the number of gRNAs increases, T7RNAp or dCas9 may be titrated resulting in reduced efficiency. One solution could be to use another or multiple orthogonal transcription systems simultaneously.

The novel gRNA expression method presented in this work could establish a new paradigm for engineering eukaryotic systems. The optimization of small genetic systems in well-characterized eukaryotes has been successful, but the lack of precise gene regulation tools and limited choice of organisms are major impediments. Our method addresses these problems by importing an established prokaryotic transcription system to empower CRISPRi.
Figure 3.S1: Repression assay and growth curves.

(A) Construct integrated in reporter strain. Parent strain is CEN.PK113-7D. Selection was not used after strain construction. (B) Repression assay plasmids. dCas9 is constitutively expressed from a CEN (low-copy) plasmid (top). gRNAs are expressed from a separate 2-micron (high-copy) plasmid. (C) Growth curves for the parent strain (CEN.PK113-7D, black circles) and the reporter (white circles) grown in YEP-GAL (inducing media for \( P_{GAL1} \)).
Figure 3.S2: T7 termination and its effect on gRNA repression efficiency.

(A) Plasmids for *E. coli* termination assay. T7RNAP is induced with 1mM IPTG. The red ‘T’ denotes the terminator being tested. Terminator strength is calculated as described in Chen et al. (107). (B) Effect of T7 terminator strength (measured in *E. coli*) on gRNA production / dCas9-mediated repression (measured in yeast). Error bars correspond to the standard deviation of three experiments performed on different days.
4. Conclusions and Future Directions

Precise control over gene expression is the holy grail of synthetic biology as it is one of the most important variables in a genetic system. A plethora of tools have been developed to regulate gene expression; however, as the number of genes in a system increases, the utility of most of these tools diminishes as large, costly combinatorial libraries need to be built de novo for each system. The reason libraries have to be built for each system is that most gene regulation tools are cis-acting (i.e. encoded on the same construct as they actuate). The sRNA pool described in this work is the first demonstration of a trans-acting tuning system that can be reused. With this, we have established a new paradigm for genetic system tuning and precise gene regulation.

Due to the lack of highly efficient trans-acting regulators and methods to use them for tuning, much of this work revolved around the engineering, troubleshooting, and characterization of these regulators. The systematic engineering approach used to develop a potent sRNA trans-acting regulator for *E. coli* provided insights on how to tackle similar efforts in other organisms and for other RNA regulators. Establishing an assay that was amenable to high throughput screening was essential for sRNA engineering. Fluorescence activated cell sorting allowed us to screen large degenerate libraries; since small RNAs are relatively short, typically consisting of a few core regions, random screens can be tractable and effective. With all of the recent efforts to mine genomes for novel CRISPR systems, such an approach potentially holds value in synthetic guide RNA design and optimization.

Another tool we used to engineer highly efficient sRNAs is an insulating ribozyme. Due to their short sequences, extraneous nucleotides (from imprecise or run-on transcription) at the 5’ and 3’ ends of sRNAs can disrupt proper folding, and reduce targeting specificity and sRNA
activity. To precisely define transcripts, the transcription start and termination sites must be known for the promoter / terminator combination used. Since these sites are typically not known and not always precise, two different promoters of the same strength can produce sRNAs that widely differ in activity. By incorporating ribozymes, which have a precise cut site and very high cleavage efficiency, sRNA activity can be decoupled from the promoter and terminator. For our *E. coli* sRNAs we had a range of strong terminators to choose from to ensure precise termination at the 3’ end. However, for other organisms and systems where characterized terminators are not available, a 3’ cutting ribozyme, such as the Hepatitis Delta Virus Ribozyme (151), may be valuable. Since ribozymes have been shown to work with high efficiency in a wide range of intracellular environments (ribozymes only need Mg$^{2+}$ to function) and organisms this is potentially a universal tool (152). Further, our approach to engineer short RNA elements with cell sorting provides a way to diversify and optimize ribozymes for different contexts.

In addition to sRNA design and insulation tools, we present a new approach to overexpress these regulators. Orthogonal transcription systems enable the decoupling of RNA regulator production from host machinery. For regulators that are not translated and are immediately active after transcription, such as gRNAs, this decoupling presents a reliable method for gene expression tuning through differential expression of the regulator. Since transcription from T7RNAP and many other bacteriophage RNA polymerases, including SP6 (153), are single protein systems that can be controlled by promoter engineering, a range of regulators levels can be achieved. Further, as these transcription systems and CRISPRi are functional in a wide variety of organisms, this approach may be expandable to even relatively uncharacterized and less genetically tractable organisms, such as those in the plant kingdom. Moreover, since CRISPRi in eukaryotes can be used for gene up-regulation in addition to
repression, the potential gene expression range is even greater. Other gene regulation approaches that rely on RNA regulators, such as microRNAs, may also benefit from these features of orthogonal transcription.

Beyond tools to facilitate gene regulation with RNA regulators, we present a new framework for optimizing multi-gene systems that is potentially applicable to all trans-acting regulators. We show that targeting of these regulators to a standardized, characterized site on a target gene allows the decoupling of the regulator and target system while preserving the expected response (regulator efficiency). This enables the construction of a reusable tuning pool that explores a wide gene expression space, and is not hardcoded on the target system. Though we only show this for RBSs with sRNAs in *E. coli*, and promoters with CRISPRi in yeast, this approach may be expandable to other trans regulation approaches, such as ZFPs, TALEs, and microRNAs. Future work will benefit from expanding the dynamic range of the pool, and utilizing it with new trans regulation mechanisms that are discovered.

In all, we have described a new paradigm for gene regulation and multi-gene system optimization. The ability to rapidly optimize large systems with a reusable pool of regulators will empower researchers to explore previously intractable systems.
5. Supplementary Tables for Chapter 2

Table 2.S1: sRNA pool parts and promoter characterization

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Promoter Sequence</th>
<th>Strength (AU)</th>
<th>Sequencing Barcode</th>
<th>Terminator</th>
</tr>
</thead>
<tbody>
<tr>
<td>pAG-A1</td>
<td>TCGACATTTATCCCTTGCGGCGAATACT</td>
<td>18</td>
<td>AAAGAGTTGT</td>
<td>ECK120033737</td>
</tr>
<tr>
<td>pAG-A2</td>
<td>ATGATAGCTGCTAGTCAGTCTAGGAG</td>
<td>27</td>
<td>CTCTATACAG</td>
<td>ECK120033737</td>
</tr>
<tr>
<td>pAG-A3</td>
<td>TCGACATTAATCATCAGGCTGCCGTA</td>
<td>540</td>
<td>CAGGAAAATT</td>
<td>ECK120033737</td>
</tr>
<tr>
<td>pAG-A4</td>
<td>CAGGAGCTGCTGCTAGTCAGTCTAGGAG</td>
<td>1,100</td>
<td>GGTCTGTCGC</td>
<td>ECK120033737</td>
</tr>
<tr>
<td>pAG-A5</td>
<td>CAGGAGCTGCTGCTAGTCAGTCTAGGAG</td>
<td>3,100</td>
<td>GATAGGCGAG</td>
<td>ECK120033737</td>
</tr>
<tr>
<td>pAG-A6</td>
<td>TCGACACATTAATCATCAGGCTGCCGTA</td>
<td>4,400</td>
<td>CACGAAATAC</td>
<td>ECK120033737</td>
</tr>
<tr>
<td>pAG-A7</td>
<td>CAGGAGCTGCTGCTAGTCAGTCTAGGAG</td>
<td>5,400</td>
<td>GTATGATTAC</td>
<td>ECK120033737</td>
</tr>
<tr>
<td>pAG-A8</td>
<td>TCGACATTTATCCCTTGCGGCGAATACT</td>
<td>6,600</td>
<td>GAGGATTTAG</td>
<td>ECK120033737</td>
</tr>
<tr>
<td>pAG-A9</td>
<td>CAGGAGCTGCTGCTAGTCAGTCTAGGAG</td>
<td>8,200</td>
<td>ACCGGAATTC</td>
<td>ECK120033737</td>
</tr>
<tr>
<td>pAG-A10</td>
<td>TCGACACATTAATCATCAGGCTGCCGTA</td>
<td>8,900</td>
<td>ATCTTCTGCT</td>
<td>ECK120033737</td>
</tr>
<tr>
<td>pAG-A11</td>
<td>TCGACACATTAATCATCAGGCTGCCGTA</td>
<td>11,000</td>
<td>CATAGGATTAC</td>
<td>ECK120033737</td>
</tr>
<tr>
<td>pAG-A12</td>
<td>TCGACACATTAATCATCAGGCTGCCGTA</td>
<td>12,000</td>
<td>ATCAGCTCAG</td>
<td>ECK120033737</td>
</tr>
<tr>
<td>pAG-A13</td>
<td>TCGACACATTAATCATCAGGCTGCCGTA</td>
<td>13,000</td>
<td>AGCAGCTCAG</td>
<td>ECK120033737</td>
</tr>
<tr>
<td>pAG-A14</td>
<td>TCGACACATTAATCATCAGGCTGCCGTA</td>
<td>14,000</td>
<td>CATAGGATTAC</td>
<td>ECK120033737</td>
</tr>
<tr>
<td>pAG-A15</td>
<td>TCGACACATTAATCATCAGGCTGCCGTA</td>
<td>18,000</td>
<td>GTATGATTAC</td>
<td>ECK120033737</td>
</tr>
<tr>
<td>pAG-A16</td>
<td>TCGACACATTAATCATCAGGCTGCCGTA</td>
<td>21,000</td>
<td>GTATGATTAC</td>
<td>ECK120033737</td>
</tr>
<tr>
<td>pAG-B1</td>
<td>TCGACATTTATCCCTTGCGGCGAATACT</td>
<td>18</td>
<td>TGAGCAGCTG</td>
<td>L3S2P11</td>
</tr>
<tr>
<td>pAG-B2</td>
<td>TCGACATTTATCCCTTGCGGCGAATACT</td>
<td>46</td>
<td>CATAGCTGCT</td>
<td>L3S2P11</td>
</tr>
<tr>
<td>pAG-B3</td>
<td>TCGACATTTATCCCTTGCGGCGAATACT</td>
<td>750</td>
<td>TACTGAGCTG</td>
<td>L3S2P11</td>
</tr>
<tr>
<td>pAG-B4</td>
<td>TCGACATTTATCCCTTGCGGCGAATACT</td>
<td>1,500</td>
<td>CGCATAGCTG</td>
<td>L3S2P11</td>
</tr>
<tr>
<td>pAG-B5</td>
<td>TCGACATTTATCCCTTGCGGCGAATACT</td>
<td>3,400</td>
<td>CCAGTACGCTG</td>
<td>L3S2P11</td>
</tr>
<tr>
<td>pAG-B6</td>
<td>TCGACATTTATCCCTTGCGGCGAATACT</td>
<td>4,700</td>
<td>TATAATGCTG</td>
<td>L3S2P11</td>
</tr>
<tr>
<td>pAG-B7</td>
<td>TCGACATTTATCCCTTGCGGCGAATACT</td>
<td>5,500</td>
<td>ATCAGCTGCT</td>
<td>L3S2P11</td>
</tr>
</tbody>
</table>
pAG-B8  CACGTTGTTAGCATAAAGTCTAAACCTA  6,700  CAACATTACG  L3S2P11
TAGGATACCTTAGCCTATCCCGT
pAG-B9  TTGACGGCTAGCTCAGCTTGTTATG  8,500  CCGGCTATCC  L3S2P11
TGCTAGCGCTCG
pAG-B10  TTGACAATATATCCGCTCGTATG  9,000  ACCATGCTTG  L3S2P11
TTTGTTGATGTTG
pAG-B11  AAAAGATATATGCTTTATCCTCCTGAG  11,000  TACGTAAC  L3S2P11
GGGACATATTACCCGCTGAGA
pAG-B12  TTGACATAAGTCTAAACCTATGATGAG  12,000  CCGGAAGAGA  L3S2P11
TTTACGAGTATTTCTGATACAT
pAG-B13  CACGGTTAGCATAAAGTCTAAACCTA  13,000  GAGCGCCTTG  L3S2P11
TAGGATAATGTCGCTCCTCG
pAG-B14  TTGACAGCTAGCTCAGCTCTAGGTACTG  14,000  TTATCAACGG  L3S2P11
TGCTAGCACCT
pAG-B15  TTGACAGCTAGCTCAGCTCTAGGTATTG  18,000  ACCACAATTC  L3S2P11
TGCTAGCTATGCA
pAG-B16  TTGACATCGAGAAGAAATTTTTTCTGATAC  21,000  TCAGCTGGTT  L3S2P11
TACACCCATACCTA
pAG-C1  CACGGTTAGCATAAAGTCTAAACCTA  19  GGGCAGTATT  L3S3P21
CATAGTAATGTCGCTCCTCG
pAG-C2  AAAAGATATATGCTTTATCCTCCTGAG  56  TGTTCCCTCA  L3S3P21
GGGACATATTACCCGCTGAGA
pAG-C3  AAAAGATATATGCTTTATCCTCCTGAG  830  TCGACTCTGC  L3S3P21
TTTACGAGTATTTCTGATACAT
pAG-C4  TTGACAGCTAGCTCAGCTCTAGGTACTG  1,600  GAGTTACATG  L3S3P21
TTTACGAGCTAGCTCAGCTCTAGGTACTG
pAG-C5  TCGACATCGAGAAGAAATTTTTTCTGATAC  3,500  GAGCGCCTTG  L3S3P21
TTTACGAGCTAGCTCAGCTCTAGGTACTG
pAG-C6  TTGACATAAATCATCCTGCTCGTATG  4,800  TTGACATGG  L3S3P21
TTTACGAGCTATG
pAG-C7  TTTACGAGCTAGCTCAGCCCTAGATTATA  5,600  GCATTCTATA  L3S3P21
TTTACGAGCTATG
pAG-C8  TTGACATCGAGAAGAAATTTTTTCTGATAC  7,200  CTGGGCTTTC  L3S3P21
TTGACATATATGCTTTCTGATACAT
pAG-C9  TCGACATTTATCCTCCTGCGCAGTATAAA  8,500  CGGCAATACCG  L3S3P21
TTGCTAGTATACCC
pAG-C10  AAAAGATATATGCTTTATCCTCCTGAG  9,500  CGGCAATACCG  L3S3P21
GGGACATATTACCCGCTGAGA
pAG-C11  CCCCAGGCTTTACACTTTATGCTTCCCGG  12,000  GACTAGCTATG  L3S3P21
CTCGATATAAAGTGCTGAGAATTGGAGTG
pAG-C12  AAAAGATATATGCTTTATCCTCCTGAG  12,000  GCGTGTTATT  L3S3P21
TTGACATATATGCTTTCTCCTGATAC
pAG-C13  TTGACATATATGCTTTATCCTCCTGATAC  14,000  AAAACAGTAC  L3S3P21
TTGCTAGTATACCC
pAG-C14  TTGACATATATGCTTTATCCTCCTGATAC  14,000  GGTGTTCGCA  L3S3P21
TTGCTAGTATACCC
pAG-C15  TTGACATATATGCTTTATCCTCCTGATAC  18,000  CCCAGGAGAC  L3S3P21
TTGCTAGTATACCC
pAG-C16  TTGACATATATGCTTTATCCTCCTGATAC  21,000  CCTGGGTGAG  L3S3P21
TTGACATATATGCTTTCTCCTGATAC
pAG-D1  TTGACATATATGCTTTATCCTCCTGATAC  19  GTCTGGTGCT  L3S3P22
TTGACATATATGCTTTCTCCTGATAC
pAG-D2  TTGACATATATGCTTTATCCTCCTGATAC  130  CACTACGCTA  L3S3P22
TTGACATATATGCTTTCTCCTGATAC
pAG-D3  TTGACATAAAGTCTAAACCTATGATGAG  1,000  TATGGCAGA  L3S3P22
TTGACATAAAGTCTAAACCTATGATGAG
<table>
<thead>
<tr>
<th></th>
<th>Sequence</th>
<th>Score</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>pAG-D4</td>
<td>ATGTGTGGATGCTGCG</td>
<td>1,800</td>
<td>L33SP22</td>
</tr>
<tr>
<td>pAG-D5</td>
<td>TGGCAATTTAATCATCGGCTCCTAGGA</td>
<td>3,800</td>
<td>L33SP22</td>
</tr>
<tr>
<td>pAG-D6</td>
<td>TGGTGGAGGGAC</td>
<td>5,000</td>
<td>L33SP22</td>
</tr>
<tr>
<td>pAG-D7</td>
<td>TGGCAATTTAATCATCGGCTCCTAGGA</td>
<td>5,700</td>
<td>L33SP22</td>
</tr>
<tr>
<td>pAG-D8</td>
<td>TGGCAATTTAATCATCGGCTCCTAGGA</td>
<td>7,600</td>
<td>L33SP22</td>
</tr>
<tr>
<td>pAG-D9</td>
<td>TGGCAATTTAATCATCGGCTCCTAGGA</td>
<td>8,500</td>
<td>L33SP22</td>
</tr>
<tr>
<td>pAG-D10</td>
<td>TGGCAATTTAATCATCGGCTCCTAGGA</td>
<td>9,500</td>
<td>L33SP22</td>
</tr>
<tr>
<td>pAG-D11</td>
<td>TGGCAATTTAATCATCGGCTCCTAGGA</td>
<td>12,000</td>
<td>L33SP22</td>
</tr>
<tr>
<td>pAG-D12</td>
<td>TGGCAATTTAATCATCGGCTCCTAGGA</td>
<td>13,000</td>
<td>L33SP22</td>
</tr>
<tr>
<td>pAG-D13</td>
<td>TGGCAATTTAATCATCGGCTCCTAGGA</td>
<td>14,000</td>
<td>L33SP22</td>
</tr>
<tr>
<td>pAG-D14</td>
<td>TGGCAATTTAATCATCGGCTCCTAGGA</td>
<td>14,000</td>
<td>L33SP22</td>
</tr>
<tr>
<td>pAG-D15</td>
<td>TGGCAATTTAATCATCGGCTCCTAGGA</td>
<td>18,000</td>
<td>L33SP22</td>
</tr>
<tr>
<td>pAG-D16</td>
<td>TGGCAATTTAATCATCGGCTCCTAGGA</td>
<td>22,000</td>
<td>L33SP22</td>
</tr>
<tr>
<td>pAG-E1</td>
<td>TGGCAATTTAATCATCGGCTCCTAGGA</td>
<td>21</td>
<td>L33SP00</td>
</tr>
<tr>
<td>pAG-E2</td>
<td>TGGCAATTTAATCATCGGCTCCTAGGA</td>
<td>290</td>
<td>L33SP00</td>
</tr>
<tr>
<td>pAG-E3</td>
<td>TGGCAATTTAATCATCGGCTCCTAGGA</td>
<td>1,100</td>
<td>L33SP00</td>
</tr>
<tr>
<td>pAG-E4</td>
<td>TGGCAATTTAATCATCGGCTCCTAGGA</td>
<td>2,500</td>
<td>L33SP00</td>
</tr>
<tr>
<td>pAG-E5</td>
<td>TGGCAATTTAATCATCGGCTCCTAGGA</td>
<td>3,800</td>
<td>L33SP00</td>
</tr>
<tr>
<td>pAG-E6</td>
<td>TGGCAATTTAATCATCGGCTCCTAGGA</td>
<td>5,000</td>
<td>L33SP00</td>
</tr>
<tr>
<td>pAG-E7</td>
<td>TGGCAATTTAATCATCGGCTCCTAGGA</td>
<td>6,500</td>
<td>L33SP00</td>
</tr>
<tr>
<td>pAG-E8</td>
<td>TGGCAATTTAATCATCGGCTCCTAGGA</td>
<td>7,900</td>
<td>L33SP00</td>
</tr>
<tr>
<td>pAG-E9</td>
<td>TGGCAATTTAATCATCGGCTCCTAGGA</td>
<td>8,600</td>
<td>L33SP00</td>
</tr>
<tr>
<td>pAG-E10</td>
<td>TGGCAATTTAATCATCGGCTCCTAGGA</td>
<td>10,000</td>
<td>L33SP00</td>
</tr>
<tr>
<td>pAG-E11</td>
<td>TGGCAATTTAATCATCGGCTCCTAGGA</td>
<td>12,000</td>
<td>L33SP00</td>
</tr>
<tr>
<td>pAG-E12</td>
<td>TGGCAATTTAATCATCGGCTCCTAGGA</td>
<td>13,000</td>
<td>L33SP00</td>
</tr>
<tr>
<td>pAG-E13</td>
<td>TGGCAATTTAATCATCGGCTCCTAGGA</td>
<td>14,000</td>
<td>L33SP00</td>
</tr>
<tr>
<td>pAG-E14</td>
<td>TGGCAATTTAATCATCGGCTCCTAGGA</td>
<td>15,000</td>
<td>L33SP00</td>
</tr>
<tr>
<td>pAG-E15</td>
<td>TTGACATTTTATCCCTGCGCGACATAA</td>
<td>18,000</td>
<td>GTGGCGCATCA</td>
</tr>
<tr>
<td>pAG-E16</td>
<td>TTGACATTTATCCCTGCGCGATATAA</td>
<td>24,000</td>
<td>CTCTTCTACATAC</td>
</tr>
<tr>
<td>pAG-F1</td>
<td>AAAAAGATTTGACTTTCCGATCTTTTT</td>
<td>25</td>
<td>ATGTGATGAG</td>
</tr>
<tr>
<td>pAG-F2</td>
<td>AAAAAGATTTGACTTTTTATCCGATGCT</td>
<td>340</td>
<td>CTTACTCCGTACG</td>
</tr>
<tr>
<td>pAG-F3</td>
<td>TTGACATCGCATCTTTTTGCTACCTAAAA</td>
<td>1,100</td>
<td>GCATTAAGGACG</td>
</tr>
<tr>
<td>pAG-F4</td>
<td>AAAAATTTTTATTTTATTACCTTTTGG</td>
<td>2,600</td>
<td>TTGGCCTACA</td>
</tr>
<tr>
<td>pAG-F5</td>
<td>TTGACATTTAATCATCCGCTGCTTATAA</td>
<td>4,300</td>
<td>CTGCAGGTA</td>
</tr>
<tr>
<td>pAG-F6</td>
<td>TTGACATCGCATCTTTTTTGTACCCATAA</td>
<td>5,100</td>
<td>GAACTCTCCA</td>
</tr>
<tr>
<td>pAG-F7</td>
<td>TTGACATTTAATCATCCGCTGCTTATAA</td>
<td>6,500</td>
<td>GGAATTACCTCA</td>
</tr>
<tr>
<td>pAG-F8</td>
<td>AAAAAGATTTGACTTTCCGATCTTTTT</td>
<td>8,000</td>
<td>TGATTGAGGACG</td>
</tr>
<tr>
<td>pAG-F9</td>
<td>TTGACATTTAATCATCCGCTGCTGAGATG</td>
<td>8,700</td>
<td>AGAACATCTT</td>
</tr>
<tr>
<td>pAG-F10</td>
<td>AAAAAGATTTGACTTTAATCTAACC</td>
<td>11,000</td>
<td>GCCTAAGGGCAACG</td>
</tr>
<tr>
<td>pAG-F11</td>
<td>AAAAAGATTTGACTTTAATCTAACC</td>
<td>12,000</td>
<td>CTACCAGAT</td>
</tr>
<tr>
<td>pAG-F12</td>
<td>AAAAATTTTTATTTTATTACCTTTTGG</td>
<td>13,000</td>
<td>GGAATTACCTCA</td>
</tr>
<tr>
<td>pAG-F13</td>
<td>AAAAATTTTTATTTTATTACCTTTTGG</td>
<td>14,000</td>
<td>GGAATTACCTCA</td>
</tr>
<tr>
<td>pAG-F14</td>
<td>TTGACATTTAATCATCCGCTGCTTATAA</td>
<td>16,000</td>
<td>CGGCGCAGGCAACG</td>
</tr>
<tr>
<td>pAG-F15</td>
<td>TTGACATTTAATCATCCGCTGCTTATAA</td>
<td>20,000</td>
<td>ACAAGCAACC</td>
</tr>
<tr>
<td>pAG-F16</td>
<td>AAAAAGATTTGACTTTTATTACCTTTTGG</td>
<td>24,000</td>
<td>CGGCGCAGGCAACG</td>
</tr>
<tr>
<td>Promoter</td>
<td>Sequence</td>
<td>Strength (AU)</td>
<td>β-carotene Gene</td>
</tr>
<tr>
<td>----------</td>
<td>-------------------------</td>
<td>---------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>T7-1</td>
<td>TAATAGACTCACTATAGGGGAC</td>
<td>40,000</td>
<td>crtY</td>
</tr>
<tr>
<td>T7-2</td>
<td>TAATAGACTCACTATAGGAAAG</td>
<td>30,000</td>
<td></td>
</tr>
<tr>
<td>T7-3</td>
<td>TAATAGACTCACTCTGGGAGA</td>
<td>26,000</td>
<td></td>
</tr>
<tr>
<td>T7-4</td>
<td>TAATAGACTCACTATGGGAGA</td>
<td>22,000</td>
<td></td>
</tr>
<tr>
<td>T7-5</td>
<td>TAATAGACTCACTCTGGGAGA</td>
<td>21,000</td>
<td></td>
</tr>
<tr>
<td>T7-6</td>
<td>TAATAGACTCACTCGTGAAGA</td>
<td>20,000</td>
<td></td>
</tr>
<tr>
<td>T7-7</td>
<td>TAATAGACTCACTCAAGGGGA</td>
<td>19,000</td>
<td>crtE</td>
</tr>
<tr>
<td>T7-8</td>
<td>TAATAGACTCACTCTGGGAGA</td>
<td>16,000</td>
<td></td>
</tr>
<tr>
<td>T7-9</td>
<td>TAATAGACTCACTCTGGGAGA</td>
<td>14,000</td>
<td></td>
</tr>
<tr>
<td>T7-10</td>
<td>TAATAGACTCACTCTAGATAAG</td>
<td>12,000</td>
<td></td>
</tr>
<tr>
<td>T7-11</td>
<td>TAATAGACTCACTCTATAGGA</td>
<td>11,000</td>
<td>idi</td>
</tr>
<tr>
<td>T7-12</td>
<td>TAATAGACTCACTCTAAGGCTA</td>
<td>9,700</td>
<td>crtI</td>
</tr>
<tr>
<td>T7-13</td>
<td>TAATAGACTCACTCTTAAGGCA</td>
<td>9,100</td>
<td></td>
</tr>
<tr>
<td>T7-14</td>
<td>TAATAGACTCACTCTAGGCTA</td>
<td>9,000</td>
<td></td>
</tr>
<tr>
<td>T7-15</td>
<td>TAATAGACTCACTCTAGGCTA</td>
<td>8,100</td>
<td></td>
</tr>
<tr>
<td>T7-16</td>
<td>TAATAGACTCACTCTAGGCTA</td>
<td>7,900</td>
<td></td>
</tr>
<tr>
<td>T7-17</td>
<td>TAATAGACTCACTCTAGGCTA</td>
<td>7,300</td>
<td>crtB</td>
</tr>
<tr>
<td>T7-18</td>
<td>TAATAGACTCACTCTAGGCTA</td>
<td>7,100</td>
<td></td>
</tr>
<tr>
<td>T7-19</td>
<td>TAATAGACTCACTCTAGGCTA</td>
<td>7,100</td>
<td></td>
</tr>
<tr>
<td>T7-20</td>
<td>TAATAGACTCACTCTAGGCTA</td>
<td>7,100</td>
<td></td>
</tr>
<tr>
<td>T7-21</td>
<td>TAATAGACTCACTCTAGGCTA</td>
<td>7,100</td>
<td></td>
</tr>
<tr>
<td>T7-22</td>
<td>TAATAGACTCACTCTAGGCTA</td>
<td>7,100</td>
<td></td>
</tr>
<tr>
<td>T7-23</td>
<td>TAATAGACTCACTCTAGGCTA</td>
<td>7,100</td>
<td></td>
</tr>
<tr>
<td>T7-24</td>
<td>TAATAGACTCACTCTAGGCTA</td>
<td>7,100</td>
<td></td>
</tr>
<tr>
<td>T7-25</td>
<td>TAATAGACTCACTCTAGGCTA</td>
<td>7,100</td>
<td></td>
</tr>
<tr>
<td>T7-26</td>
<td>TAATAGACTCACTCTAGGCTA</td>
<td>7,100</td>
<td></td>
</tr>
<tr>
<td>T7-27</td>
<td>TAATAGACTCACTCTAGGCTA</td>
<td>7,100</td>
<td></td>
</tr>
<tr>
<td>T7-28</td>
<td>TAATAGACTCACTCTAGGCTA</td>
<td>7,100</td>
<td></td>
</tr>
<tr>
<td>T7-29</td>
<td>TAATAGACTCACTCTAGGCTA</td>
<td>7,100</td>
<td></td>
</tr>
<tr>
<td>T7-30</td>
<td>TAATAGACTCACTCTAGGCTA</td>
<td>7,100</td>
<td></td>
</tr>
<tr>
<td>T7-31</td>
<td>TAATAGACTCACTCTAGGCTA</td>
<td>7,100</td>
<td></td>
</tr>
<tr>
<td>T7-32</td>
<td>TAATAGACTCACTCTAGGCTA</td>
<td>7,100</td>
<td></td>
</tr>
<tr>
<td>T7-33</td>
<td>TAATAGACTCACTCTAGGCTA</td>
<td>7,100</td>
<td></td>
</tr>
<tr>
<td>T7-34</td>
<td>TAATAGACTCACTCTAGGCTA</td>
<td>7,100</td>
<td></td>
</tr>
<tr>
<td>T7-35</td>
<td>TAATAGACTCACTCTAGGCTA</td>
<td>7,100</td>
<td></td>
</tr>
<tr>
<td>T7-36</td>
<td>TAATAGACTCACTCTAGGCTA</td>
<td>7,100</td>
<td></td>
</tr>
<tr>
<td>T7-37</td>
<td>TAATAGACTCACTCTAGGCTA</td>
<td>7,100</td>
<td></td>
</tr>
<tr>
<td>T7-38</td>
<td>TAATACGACTCACTGGGCGATGA</td>
<td>1,000</td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>-------------------------</td>
<td>-------</td>
<td></td>
</tr>
<tr>
<td>T7-39</td>
<td>TAATACGACTCACTAGCCCTTG</td>
<td>930</td>
<td></td>
</tr>
<tr>
<td>T7-40</td>
<td>TAATACGACTCCTGAACAGAGA</td>
<td>810</td>
<td></td>
</tr>
<tr>
<td>T7-41</td>
<td>TAATACGACTCCTGCGCGCCG</td>
<td>800</td>
<td></td>
</tr>
<tr>
<td>T7-42</td>
<td>TAATACGACTCCTCTAGGCAGGC</td>
<td>620</td>
<td></td>
</tr>
<tr>
<td>T7-43</td>
<td>TAATACGACTCCTATGGCAAGG</td>
<td>600</td>
<td></td>
</tr>
<tr>
<td>T7-44</td>
<td>TAATACGACTCCTAACCAGCG</td>
<td>590</td>
<td></td>
</tr>
<tr>
<td>T7-45</td>
<td>TAATACGACTCCTCCGAAAGGA</td>
<td>490</td>
<td></td>
</tr>
<tr>
<td>T7-46</td>
<td>TAATACGACTCCTTGGGGGCA</td>
<td>480</td>
<td></td>
</tr>
<tr>
<td>T7-47</td>
<td>TAATACGACTCCTTACGAGTA</td>
<td>380</td>
<td></td>
</tr>
<tr>
<td>T7-48</td>
<td>TAATACGACTCCTTTGGGGTCCG</td>
<td>270</td>
<td></td>
</tr>
<tr>
<td>Part name</td>
<td>Type</td>
<td>DNA sequence</td>
<td>Note</td>
</tr>
<tr>
<td>------------</td>
<td>--------</td>
<td>------------------------------------------------------------------------------</td>
<td>---------------</td>
</tr>
<tr>
<td>sRNA-AG1</td>
<td>sRNA</td>
<td>GTCAAAGTTTTGTTTATGATCCTTTCTGTGGGCCATTGCATTTGCTGCCTTTTTT</td>
<td>Figure 2A</td>
</tr>
<tr>
<td>MicC</td>
<td>sRNA</td>
<td>GTCAAAGTTTTGTTTAATTTTTTTAAAGCTTTCTGCTCTCAAGAAGCACAGCCGACAGCTGTCGCCTTTTTTTT</td>
<td>Figure 2A</td>
</tr>
<tr>
<td>DsrA</td>
<td>sRNA</td>
<td>GTCAAAGTTTTGTTTATACCATGGTAGCAAAGCTACCTTTTTTTC</td>
<td>Figure 2A</td>
</tr>
<tr>
<td>GcvB</td>
<td>sRNA</td>
<td>GTCAAAGTTTTGTTTATTTTTTTAGTGCTTCTTGCTTCAATGTTAATTTTGTGGTAAATACCCCGAACAGGAT</td>
<td>Figure 2A</td>
</tr>
<tr>
<td>MicC</td>
<td>sRNA</td>
<td>GTCAAAGTTTTGTTTATTTTTTTAGTGCTTCTTGCTTCAATGTTAATTTTGTGGTAAATACCCCGAACAGGAT</td>
<td>Figure 2A</td>
</tr>
<tr>
<td>MicF</td>
<td>sRNA</td>
<td>GTCAAAGTTTTGTTTATTTTTTTAGTGCTTCTTGCTTCAATGTTAATTTTGTGGTAAATACCCCGAACAGGAT</td>
<td>Figure 2A</td>
</tr>
<tr>
<td>SgrS</td>
<td>sRNA</td>
<td>GTCAAAGTTTTGTTTATTTTTTTAGTGCTTCTTGCTTCAATGTTAATTTTGTGGTAAATACCCCGAACAGGAT</td>
<td>Figure 2A</td>
</tr>
<tr>
<td>Spot42</td>
<td>sRNA</td>
<td>GTCAAAGTTTTGTTTATTTTTTTAGTGCTTCTTGCTTCAATGTTAATTTTGTGGTAAATACCCCGAACAGGAT</td>
<td>Figure 2A</td>
</tr>
<tr>
<td>T(B1006)</td>
<td>sRNA</td>
<td>GTCAAAGTTTTGTTTATTTTTTTAGTGCTTCTTGCTTCAATGTTAATTTTGTGGTAAATACCCCGAACAGGAT</td>
<td>Figure 2A</td>
</tr>
<tr>
<td>T(MicF)</td>
<td>sRNA</td>
<td>GTCAAAGTTTTGTTTATTTTTTTAGTGCTTCTTGCTTCAATGTTAATTTTGTGGTAAATACCCCGAACAGGAT</td>
<td>Figure 2A</td>
</tr>
<tr>
<td>T(SgrS)</td>
<td>sRNA</td>
<td>GTCAAAGTTTTGTTTATTTTTTTAGTGCTTCTTGCTTCAATGTTAATTTTGTGGTAAATACCCCGAACAGGAT</td>
<td>Figure 2A</td>
</tr>
<tr>
<td>T(BF1176)</td>
<td>sRNA</td>
<td>GTCAAAGTTTTGTTTATTTTTTTAGTGCTTCTTGCTTCAATGTTAATTTTGTGGTAAATACCCCGAACAGGAT</td>
<td>Figure 2A</td>
</tr>
<tr>
<td>sRNA-AG4</td>
<td>sRNA</td>
<td>GTCAAAGTTTTGTTTATTTTTTTAGTGCTTCTTGCTTCAATGTTAATTTTGTGGTAAATACCCCGAACAGGAT</td>
<td>Figure 2A</td>
</tr>
<tr>
<td>gfp</td>
<td>gene</td>
<td>ATGGCGGAAGAGAGAAGAACCTTTTCACTGAGGTTGTGCCCAATTTCTGTTGTAATAGGGCAATATATTCT</td>
<td>reporter</td>
</tr>
<tr>
<td>rfp</td>
<td>gene</td>
<td>ATGGCTTCCTCCGAGAAGACGTTTATCAAAAGAGCTCATCCTGCTGTTTCT</td>
<td>reporter</td>
</tr>
</tbody>
</table>
GCTTGGGAGATCCCTGCTCCCCGCAGTTCCAGTACGGTTCCAAA
GCTTACGTTAAACACCCGGCTGACATCCCGGACTACCTGAAA
CTGTCCTTCCCGGAAGGTTTCAAATGGGAACGTGTTATGAAC
TTCGAAGACGGTGGTGTTGTTACCGTTACCCAGGACTCCTCC
CTGCAAGACGGTGAGTTCATCTACAAAGTTAAACTGCGTGGT
ACCAACTTCCCGTCGAGCCTGGTTATGCAGAAAAAAACC
ATGGGTTGGGAAGCTTCCACCGAACGTATGTACCCGGAAGAC
GGTGCTCTGAAAGGTGAAATCAAAATGCGTCTGAAACTGAAA
GACGGTGGTCACTACGACGCTGAAGTTAAAACCACCTACATG
GCTAAAAACCGTTCAGCTGCCGGGTGCTTACAAAACCGAC
ATCAAACTGGACATCACCTCCCACAACGAAGACTACACCATC
GTTGAACAGTACGAACGTGCTGAAGGTCGTCACTCCACCGGT
GCTTAATAA

terminator

CCAATTATTGAAGGCCGCTAACGCGGCCTTTTTTTGTTTCTG
GTCTCCC

terminator

CCAATTATTGAAGGGGAGCGGGAAACCGCTCCCCTTTTTTTG
TTTCTGGTCTCCC

terminator

GACGAACAATAAGGCCTCCCTAACGGGGGGCCTTTTTTATTG
ATAACAAAA

terminator

CCAATTATTGAAGGCCTCCCTAACGGGGGGCCTTTTTTTGTT
TCTGGTCTCCC

gene

ATGCCGAAACTCTGGGTATGCAAGACATTTCTGCTGCTGCACTG
ATGATGCAACCCTGGAAGCAATTAATGAAGTTGGTATGCAT
GATGCAACCCTGGAAGCAATTAATGAAGTTGGTATGCAT

gene

ATGCCGAAACTCTGGGTATGCAAGACATTTCTGCTGCTGCACTG
ATGATGCAACCCTGGAAGCAATTAATGAAGTTGGTATGCAT

gene

ATGCCGAAACTCTGGGTATGCAAGACATTTCTGCTGCTGCACTG
ATGATGCAACCCTGGAAGCAATTAATGAAGTTGGTATGCAT

gene

ATGCCGAAACTCTGGGTATGCAAGACATTTCTGCTGCTGCACTG
ATGATGCAACCCTGGAAGCAATTAATGAAGTTGGTATGCAT
tetR

tetR gene

AAAGATCAGTTTATGGAACGTCGTCGTGAGATGCCGAAAAAATGCTCAGAATGATATACACACGCGATTA
GAGCTGCTTTATAGGCTCGGATACGAGGTTAAACACACTGAA
AAACTCGCCCAAGAGCTAGGTTGGAGCATGAGGCTAACATTG
TGGCATGTAACAAATAGCAGGAAAGGGTCTTGGCTCGACGCCCTAGG
ATTGAGATGTTAGATAGGCACCATACTCTTCTTTTTTCTGGTATCGCCTGCTGACCGAACAGCTGACCGTTGAACAGGAT
ATTGAAGAATTTACCTTCCTGCTGATTAATGGTGTTTGTCCG
GGTACACAGCGTTAA
gene ATGTCCAGATTAGATAAAAGTAAAGTGATTAACAGCGCATTA
GAGCTGCTTTATAGGCTCGGATACGAGGTTAAACACACTGAA
AAACTCGCCCAAGAGCTAGGTTGGAGCATGAGGCTAACATTG
TGGCATGTAACAAATAGCAGGAAAGGGTCTTGGCTCGACGCCCTAGG
ATTGAGATGTTAGATAGGCACCATACTCTTCTTTTTTCTGGTATCGCCTGCTGACCGAACAGCTGACCGTTGAACAGGAT
ATTGAAGAATTTACCTTCCTGCTGATTAATGGTGTTTGTCCG
GGTACACAGCGTTAA

RiboJ57
ingulator

XNOR circuit

ECK120033736
terminator

XNOR circuit

L3S3P11
RBS

XNOR circuit

phIF-50K
RBS

XNOR circuit

bm3rl-50K
RBS

XNOR circuit

betl-25K
RBS

XNOR circuit

hylrII-25K
RBS

XNOR circuit

B0064
RBS

XNOR circuit / reporter

yfp
gene

XNOR circuit

XNOR circuit / reporter

XNOR circuit / sRNA pool

XNOR circuit / β-carotene pathway

98
SarJ  insulator  AGACTGTGGCAGATGTGATCCGACCTGAGATGGCCCAAA AGGGCCGAACACCTCTACAAATAATTGTGGTTAA
RiboJ51  insulator  AGTAGTCACCGGCTGTGCTTGCGCTGATGAGCCTGTGAGGGCGGAACAGCCTCTACAAATAATTTGTTAAG
L352P21  terminator  CTTTTTCTGTTTTGGTCC AGCTGCACCGGCTGTGCTTGCCGGTCTGATGAGCCTGTGAGGGCGGAACAGCCTCTACAAATAATTTGTTAAG
RiboJ  insulator  AGCTGTCACCGGCTGTGCTTGCCGGTCTGATGAGCCTGTGAGGGCGGAACAGCCTCTACAAATAATTTGTTAAG
ECK120033737  terminator  GGAACACAGAAGAAAAAGGCAGCAACCTGAGATCGGCTCGTTTTTTGGACACCGAAAG

crtB  gene  ATGAGCAGCGAGGCAGGCTGCTTTGAACACCGCAACGCAAGCAGACCATG GCTGAACCCCGGCTGCTGCTTGACCACGCCACGCAGACCATGGCCAAACAGCACGCTCTACAAATAATTTGTTAAG

crtE  gene  ATGGTGAGCAGATGGTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAAAGCGGGCGTTTCGCCTCATCGCAAAAGATCGAGTTAA9
CTGGCCCTGACGCGCCGTCGCAAATTGTCGCGGCGATTAAA
TCTCCGCGCCGATCAACTGGGTGCCAGGCTGGTGTTGTCGATG
GTAGAAGCAGCGGCGTCAAGCTCTAAAGCAGCGGTGCAC
AACTCCTCTCAGGACGCTACTGCTGTGCTGACGCTACTGACT
CGCTGCTGCTGCTGATCTATAATTCTCATCGCAATCAAAATT
CACGGTACGAGGGAAGCGGACTGAGTCTGCCACACTCTCT
GCTGTGGAGCATCTGGTCGCATTGGGTCACCAGCAAATCTG
CGCTGTTAGCGGGCCCATTAAGTTCTGTCTCGGCGCGTCTG
CGTCTGGCTGGCTGGCATAAATATCTCACGCAATCAAAATT
CAGACCGATCGCAGTCTGTTGCCAAACGACGAGATGCGCTGGGC
GCAATGGCGGGCCATTACCGAGTCCGGGCTGCGGTGTTGTCG
GATAATCCTGGTATGGGATACGAGATACCGAAGACAGCTCA
TGTTATATCCCGCGCTCAACCCACCATCAACAGGATTTTCCG
CTGGCTGGCTGGCTGGCTGGCATAAATATCTCATCGCAATCA
CTGTCCTGCCCGGGGCCGATCCGACGATGGCTGGGGCAATG
CTGTAATGCAAAGAAACACCACCCTCGGGCCCAATACGACAC
GCTCTCCCAGGCGGGTTGCGCGATTCATTAATGCAGCTGGCA
CGACAGTTTCCAGACTGGAAGCCGCGTCTGCTCA

CTGGCCCTGACGCGCCGTCGCAAATTGTCGCGGCGATTAAA
RiboJ10
insulator
AGGCTCTCAAGGCTTGCTTCCGGTCTGATGAGTCCGCTGAG
GACGNAAGCCTCTCAAAATATTTTGTTTAA
BydvJ
insulator
AGGCTCTCAAGGCTTGCTTCCGGTCTGATGAGTCCGCTGAG
GACGNAAGCCTCTCAAAATATTTTGTTTAA
ECK120010818
terminator
GTCGGATTCAAGTTTCTGGACGTGCTGATGCGTCCGAAGAG
TTTTACTTTTTG
ECK120015440
terminator
TCCGGCAATTAAAAAGCGGGCTAACCACGCCGCTTTTTTAC
GTCGCA
L3S2P55
terminator
CTCGGTACCAAAAGACGACGATAGACGCTGAAAGCGGTCTT
TTTTCTGTTTTGCTC
L3S2P24
terminator
CTCGGTACCAAAAGACGACGATAGACGCTGAAAGCGGTCTT
TTTTCTGTTTTGCTC
dxs-50K
RBS
CCACCAAGCAGAAGAGAGAGAAG
idi-50K
RBS
TCTCCAGGATTAGGGGAGATAGATAGGGACGAGAT
crtE-50K
RBS
GGATCCTCAGAAGCTTAATATATAGAGGAAGTA
crtB-50K
RBS
GCCCAGACAATCCCGCATTAGAGAAATATAT
crtI-50K
RBS
AATAAAGAACAATATATAGAGGAAGA
crtY-50K
RBS
GCTTAAGATATATAGGGACGAGGAGAGAGAAG

a. Underline indicates sRNA tag.
<table>
<thead>
<tr>
<th>sRNA tag</th>
<th>Rank (E-value)</th>
<th>Homology Overlap (bp)</th>
<th>Gene</th>
<th>Region</th>
<th>Correct sense for sRNA-mediated repression?</th>
<th>Gene function</th>
<th>Essential?</th>
</tr>
</thead>
<tbody>
<tr>
<td>A 1</td>
<td>TCAAGT</td>
<td>12</td>
<td>ChaA</td>
<td>CDS</td>
<td>no</td>
<td>Na+/K+:H+ antiporter ChaA</td>
<td>no</td>
</tr>
<tr>
<td>A 2</td>
<td>CAAAGTT</td>
<td>11</td>
<td>ybiO</td>
<td>CDS</td>
<td>no</td>
<td>mechanosensitive channel</td>
<td>no</td>
</tr>
<tr>
<td>A 3</td>
<td>AAAGTTT</td>
<td>11</td>
<td>selD</td>
<td>CDS</td>
<td>yes</td>
<td>selenide, water dikinase</td>
<td>no</td>
</tr>
<tr>
<td>A 4</td>
<td>AAGTTT</td>
<td>11</td>
<td>yejM</td>
<td>CDS</td>
<td>no</td>
<td>cardiolipin transport protein</td>
<td>no</td>
</tr>
<tr>
<td>A 5</td>
<td>AAGTTTG</td>
<td>11</td>
<td>yqiG</td>
<td>CDS</td>
<td>yes</td>
<td>predicted outer membrane usher protein</td>
<td>no</td>
</tr>
<tr>
<td>A 6</td>
<td>GTCAAAG</td>
<td>11</td>
<td>yqiK</td>
<td>intergenic</td>
<td>yes</td>
<td>inner membrane protein; heptose 7-phosphate kinase/heptose 1-phosphate adenyltransferase</td>
<td>no / no</td>
</tr>
<tr>
<td>A 7</td>
<td>TCAAGT</td>
<td>11</td>
<td>sokA</td>
<td>CDS</td>
<td>no</td>
<td>psuedogene</td>
<td>unknown</td>
</tr>
<tr>
<td>A 8</td>
<td>CAAAGTT</td>
<td>11</td>
<td>yicL</td>
<td>CDS</td>
<td>yes</td>
<td>predicted inner membrane protein</td>
<td>no</td>
</tr>
<tr>
<td>A 9</td>
<td>GTCAAAG</td>
<td>10</td>
<td>yaaU</td>
<td>CDS</td>
<td>yes</td>
<td>putative transport protein</td>
<td>no</td>
</tr>
<tr>
<td>A 10</td>
<td>AGTTTGT</td>
<td>10</td>
<td>intF</td>
<td>CDS</td>
<td>yes</td>
<td>predicted phage integrase</td>
<td>no</td>
</tr>
<tr>
<td>B 1</td>
<td>CTCACTCA</td>
<td>14</td>
<td>acrS</td>
<td>CDS</td>
<td>yes</td>
<td>DNA-binding transcriptional repressor</td>
<td>no</td>
</tr>
<tr>
<td>B 2</td>
<td>TCACTAT</td>
<td>13</td>
<td>queA</td>
<td>CDS</td>
<td>no</td>
<td>S-adenosylmethionine:tRNA ribosyltransferase-isomerase</td>
<td>no</td>
</tr>
<tr>
<td>B 3</td>
<td>CATCATT</td>
<td>13</td>
<td>ygbT</td>
<td>CDS</td>
<td>no</td>
<td>multifunctional nuclease</td>
<td>no</td>
</tr>
<tr>
<td>B 4</td>
<td>TCATCAT</td>
<td>13</td>
<td>ysaB</td>
<td>CDS</td>
<td>yes</td>
<td>putative protein</td>
<td>no</td>
</tr>
<tr>
<td>B 5</td>
<td>TCACTCA</td>
<td>12</td>
<td>paaG</td>
<td>CDS</td>
<td>no</td>
<td>predicted ring 1,2-epoxyphenylacetyl-CoA isomerase (oxepin-CoA forming) serine/threonine kinase HipA</td>
<td>no</td>
</tr>
<tr>
<td>B 6</td>
<td>CATCATT</td>
<td>12</td>
<td>hipA</td>
<td>CDS</td>
<td>no</td>
<td>SerA DNA-binding transcriptional dual regulator primosome factor N'</td>
<td>no</td>
</tr>
<tr>
<td>B 7</td>
<td>CATCATT</td>
<td>12</td>
<td>lrhA</td>
<td>CDS</td>
<td>no</td>
<td>LrHA DNA-binding transcriptional dual regulator primosome factor N'</td>
<td>no</td>
</tr>
<tr>
<td>B 8</td>
<td>CATCATT</td>
<td>12</td>
<td>priA</td>
<td>CDS</td>
<td>no</td>
<td>SecD/SecF/YaJ/C/YidC complex</td>
<td>no</td>
</tr>
<tr>
<td>B 9</td>
<td>TCACTAT</td>
<td>12</td>
<td>lacZ</td>
<td>CDS</td>
<td>yes</td>
<td>member of the SecD/SecF/YaJ/C/YidC complex</td>
<td>no</td>
</tr>
<tr>
<td>B 10</td>
<td>TCACTAA</td>
<td>11</td>
<td>yajC</td>
<td>3' UTR</td>
<td>N/A</td>
<td>putative transport protein, ABC superfamily</td>
<td>no</td>
</tr>
<tr>
<td>C 1</td>
<td>TATAGTG</td>
<td>11</td>
<td>mleB</td>
<td>CDS</td>
<td>no</td>
<td>nucleoprotein/polynucleotide-associated enzyme</td>
<td>no</td>
</tr>
<tr>
<td>C 2</td>
<td>AGTAGTG</td>
<td>10</td>
<td>yaiL</td>
<td>CDS</td>
<td>yes</td>
<td>CP4-6 prophage; partial regulator</td>
<td>no</td>
</tr>
<tr>
<td>C 3</td>
<td>TATAGTG</td>
<td>9</td>
<td>insN</td>
<td>5' UTR</td>
<td>no</td>
<td>CP4-6 prophage; partial regulator</td>
<td>no</td>
</tr>
<tr>
<td>C</td>
<td>4</td>
<td>TAGTGTG</td>
<td>ecpR</td>
<td>CDS</td>
<td>no</td>
<td></td>
<td></td>
</tr>
<tr>
<td>----</td>
<td>-----</td>
<td>--------</td>
<td>------</td>
<td>------</td>
<td>----</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>5</td>
<td>TAGTGTG</td>
<td>artQ</td>
<td>CDS</td>
<td>no</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>6</td>
<td>TAGTGTG</td>
<td>focA</td>
<td>CDS</td>
<td>yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>7</td>
<td>TATAGTG</td>
<td>gadC</td>
<td>CDS</td>
<td>no</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>8</td>
<td>ATAGTGT</td>
<td>fliH</td>
<td>CDS</td>
<td>no</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>9</td>
<td>ATAGTGT</td>
<td>yedW</td>
<td>CDS</td>
<td>no</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>10</td>
<td>ATAGTGT</td>
<td>yehM</td>
<td>CDS</td>
<td>no</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>1</td>
<td>CGCTTA</td>
<td>copA</td>
<td>CDS</td>
<td>no</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>2</td>
<td>TAACGC</td>
<td>putP</td>
<td>CDS</td>
<td>yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>3</td>
<td>ACGCCT</td>
<td>rmb</td>
<td>CDS</td>
<td>no</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>4</td>
<td>ACGCCT</td>
<td>yfjW</td>
<td>CDS</td>
<td>no</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>5</td>
<td>TAACGC</td>
<td>yjbG</td>
<td>CDS</td>
<td>yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>6</td>
<td>TAACGC</td>
<td>insM</td>
<td>5' UTR</td>
<td>no</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>7</td>
<td>ACGCCT</td>
<td>rna</td>
<td>CDS</td>
<td>yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>8</td>
<td>TAACGC</td>
<td>mukB</td>
<td>CDS</td>
<td>no</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>9</td>
<td>TAACGC</td>
<td>REP180a extragenic</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>10</td>
<td>TAACGC</td>
<td>ISSR</td>
<td>extragenic no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>1</td>
<td>AACTAC</td>
<td>puuP</td>
<td>CDS</td>
<td>no</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>2</td>
<td>AACTACA</td>
<td>yehI</td>
<td>3' UTR</td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>3</td>
<td>AACTAC</td>
<td>yraJ</td>
<td>CDS</td>
<td>yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>4</td>
<td>AACTAC</td>
<td>yjiC</td>
<td>CDS</td>
<td>no</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>5</td>
<td>AACTACA</td>
<td>puuD</td>
<td>CDS</td>
<td>no</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>6</td>
<td>ACTACCA</td>
<td>lpxT</td>
<td>CDS</td>
<td>yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>7</td>
<td>CTACCAG</td>
<td>mazG</td>
<td>CDS</td>
<td>no</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>8</td>
<td>AACTACC</td>
<td>aas</td>
<td>CDS</td>
<td>yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>9</td>
<td>AACTACC</td>
<td>tufA / B</td>
<td>CDS</td>
<td>no</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

of insertion element IS911A
MatA DNA-binding transcriptional dual regulator
L-arginine ABC transporter - membrane subunit
formate channel FocA

glutamic acid:4-aminobutyrate antiporter
flagellar biosynthesis protein

DNA-binding response regulator in two-component system with YedV
putative protein

Cu⁺ export ATPase

proline:Na⁺ symporter

ribonuclease II

CP4-57 prophage; predicted inner membrane protein
part of yhbEFGH complex

psuedogene

RNase I

cell division protein involved in chromosome partitioning
repetitive extragenic palindromic element

insertion sequence

putrescine:H⁺ symporter PuuP

conserved protein

predicted outer membrane export usher protein
putative protein

γ-glutamyl-γ-aminobutyrate hydrolase
Kdo₂-lipid A phosphotransferase

nucleoside triphosphate pyrophosphohydrolas 2-
acylglycerophosphoethanolamine acyltransferase / acyl-ACP synthetase
elongation factor Tu

104
| E  | 10 | A A A C T A C  | 11 | methH | CDS | no | cobalamin-dependent methionine synthase | no |
| F  |  1 | G T T T G G C  | 12 | evgS | CDS | no | EvgS sensory histidine kinase | no |
| F  |  2 | T G T A G T | 11 | nei | CDS | yes | endonuclease VIII | no |
| F  |  3 | T T T G G C T G T A G | 11 | nusA | CDS | yes | transcription termination/antitermination L factor | no (in enriched LB) |
| F  |  4 | G T T T G G C  | 11 | gspE | CDS | no | putative protein secretion protein for export | no |
| F  |  5 | G T T T G G C  | 11 | xdhA | CDS | no | xanthine dehydrogenase subunit | no |
| F  |  6 | G C T G T A G | 10 | mtn | CDS | yes | 5'-methylthioadenosine/S-adenosylhomocysteine nucleosidase | no |
| F  |  7 | G T T T G G C  | 10 | lacY | CDS | no | lactose/melibiose:H\(^+\) symporter | no |
| F  |  8 | T T T G G C T G A | 10 | dtpD | CDS | yes | dipeptide:H\(^+\) symporter DtpD | no |
| F  |  9 | T G G C T G T | 10 | hcp | CDS | no | nitric oxide reductase | no |
| F  | 10 | G G C T G T A G | 10 | paaZ | CDS | no | oxepin-CoA hydrolase/3-oxo-5,6-dehydrosuberyl-CoA semialdehyde dehydrogenase | no |
6. Supplementary Tables for Chapter 3

Table 3.S1: T7 terminator strength and corresponding fold repression

<table>
<thead>
<tr>
<th>Part name</th>
<th>Terminator strength (E. coli)</th>
<th>Fold Repression</th>
</tr>
</thead>
<tbody>
<tr>
<td>BBa_B0061</td>
<td>10.2</td>
<td>17.6</td>
</tr>
<tr>
<td>BBa_B0057</td>
<td>15.2</td>
<td>24.7</td>
</tr>
<tr>
<td>BBa_K088008</td>
<td>20.5</td>
<td>27.8</td>
</tr>
<tr>
<td>ECK120026481</td>
<td>35.2</td>
<td>29.1</td>
</tr>
<tr>
<td>ECK120023928</td>
<td>40.1</td>
<td>31.5</td>
</tr>
<tr>
<td>BBa_B0010</td>
<td>67.2</td>
<td>32.6</td>
</tr>
<tr>
<td>pheA-1</td>
<td>180.4</td>
<td>33.9</td>
</tr>
<tr>
<td>ECK120029600</td>
<td>220.8</td>
<td>35.7</td>
</tr>
</tbody>
</table>
Table 3.S2: Predicted nucleosome occupancy and fold repression for gRNA target sites

<table>
<thead>
<tr>
<th>Target site</th>
<th>Strand</th>
<th>PAM</th>
<th>Predicated nucleosome occupancy</th>
<th>Fold repression</th>
</tr>
</thead>
<tbody>
<tr>
<td>-369</td>
<td>-</td>
<td>AGG</td>
<td>0.35</td>
<td>7.4</td>
</tr>
<tr>
<td>-355</td>
<td>-</td>
<td>TGG</td>
<td>0.43</td>
<td>4.9</td>
</tr>
<tr>
<td>-353</td>
<td>+</td>
<td>CGG</td>
<td>0.44</td>
<td>17.0</td>
</tr>
<tr>
<td>-338</td>
<td>-</td>
<td>CGG</td>
<td>0.57</td>
<td>20.0</td>
</tr>
<tr>
<td>-328</td>
<td>-</td>
<td>CGG</td>
<td>0.62</td>
<td>11.1</td>
</tr>
<tr>
<td>-323</td>
<td>-</td>
<td>TGG</td>
<td>0.63</td>
<td>14.3</td>
</tr>
<tr>
<td>-310</td>
<td>-</td>
<td>GGG</td>
<td>0.64</td>
<td>9.6</td>
</tr>
<tr>
<td>-309</td>
<td>-</td>
<td>TGG</td>
<td>0.64</td>
<td>13.7</td>
</tr>
<tr>
<td>-287</td>
<td>-</td>
<td>GGG</td>
<td>0.65</td>
<td>16.4</td>
</tr>
<tr>
<td>-286</td>
<td>-</td>
<td>GGG</td>
<td>0.65</td>
<td>16.0</td>
</tr>
<tr>
<td>-285</td>
<td>-</td>
<td>AGG</td>
<td>0.65</td>
<td>17.0</td>
</tr>
<tr>
<td>-279</td>
<td>+</td>
<td>AGG</td>
<td>0.65</td>
<td>4.9</td>
</tr>
<tr>
<td>-278</td>
<td>+</td>
<td>GGG</td>
<td>0.65</td>
<td>3.8</td>
</tr>
<tr>
<td>-275</td>
<td>-</td>
<td>AGG</td>
<td>0.65</td>
<td>8.2</td>
</tr>
<tr>
<td>-254</td>
<td>-</td>
<td>GGG</td>
<td>0.65</td>
<td>7.3</td>
</tr>
<tr>
<td>-253</td>
<td>+</td>
<td>AGG</td>
<td>0.65</td>
<td>9.1</td>
</tr>
<tr>
<td>-253</td>
<td>-</td>
<td>CGG</td>
<td>0.65</td>
<td>10.5</td>
</tr>
<tr>
<td>-248</td>
<td>+</td>
<td>TGG</td>
<td>0.5</td>
<td>2.4</td>
</tr>
<tr>
<td>-221</td>
<td>-</td>
<td>CGG</td>
<td>0.3</td>
<td>24.3</td>
</tr>
<tr>
<td>-218</td>
<td>-</td>
<td>AGG</td>
<td>0.28</td>
<td>20.5</td>
</tr>
<tr>
<td>-201</td>
<td>+</td>
<td>AGG</td>
<td>0.13</td>
<td>25.6</td>
</tr>
<tr>
<td>-121</td>
<td>-</td>
<td>AGG</td>
<td>0.04</td>
<td>27.7</td>
</tr>
<tr>
<td>-118</td>
<td>-</td>
<td>GGG</td>
<td>0.04</td>
<td>27.4</td>
</tr>
<tr>
<td>-117</td>
<td>-</td>
<td>TGG</td>
<td>0.04</td>
<td>30.0</td>
</tr>
<tr>
<td>-103</td>
<td>+</td>
<td>CGG</td>
<td>0.039</td>
<td>32.0</td>
</tr>
<tr>
<td>Promoter sequence</td>
<td>Measured strength</td>
<td>Fold repression</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------------------------------</td>
<td>-------------------</td>
<td>----------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TAATACGACTCACTATAGGGGAGC</td>
<td>40000</td>
<td>35.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TAATACGACTCACTATAGGAAG</td>
<td>30000</td>
<td>28.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TAATACGACTCACTGCGGGGAGA</td>
<td>26000</td>
<td>20.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TAATACGACTCACTATGGGGAGA</td>
<td>22000</td>
<td>7.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TAATACGACTCACTGCGGGGAGA</td>
<td>21000</td>
<td>26.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TAATACGACTCACTCGTAGAGA</td>
<td>20000</td>
<td>21.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TAATACGACTCACTACAGGGGAGA</td>
<td>19000</td>
<td>22.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TAATACGACTCACTATGGAGAGA</td>
<td>16000</td>
<td>26.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TAATACGACTCACTGTCGGCAAGA</td>
<td>14000</td>
<td>10.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TAATACGACTCACTGAGATAAGA</td>
<td>12000</td>
<td>6.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TAATACGACTCACTAAGATAAGA</td>
<td>11000</td>
<td>17.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TAATACGACTCACTGTCGGGGGA</td>
<td>9700</td>
<td>1.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TAATACGACTCACTAGAGGAGAG</td>
<td>9100</td>
<td>7.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TAATACGACTCACTATGGGCTCA</td>
<td>9000</td>
<td>8.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TAATACGACTCACTTGGGTTATA</td>
<td>8100</td>
<td>1.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TAATACGACTCACTTAAGAGAGA</td>
<td>7100</td>
<td>3.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Part name</td>
<td>Type</td>
<td>DNA sequence</td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------</td>
<td>-----------</td>
<td>-------------------------------------------------------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BBa_B0061</td>
<td>E. coli</td>
<td>AAGTCAAAAAGCCTCCGGTGGAGGCTTTTGACTTTT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BBa_B0057</td>
<td>E. coli</td>
<td>CAGAAATCATCTTTAGCAGAAAGCTAAGGATTTTTTTATCTG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BBa_K088008</td>
<td>E. coli</td>
<td>AGAAACAGCAGAAACTCCAACAGCAGGGCTGTTTTCTGTCTGTTTTCTCTGCTT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ECK120026481</td>
<td>E. coli</td>
<td>TACCACGTCAAAAAACACGCGGCTTTTTAGCGCGGTCTTTTTATTTTTCAAA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ECK120023928</td>
<td>E. coli</td>
<td>GCTGATGCGAAGGCTCTGAATTCAGGCGCCCTTTTTATCTGAGAT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BBa_B0010</td>
<td>E. coli</td>
<td>CCAGGCAATCAATAAAACGAAAGGCTCAGCGAAGACGGGCTTTTTCTGTCTG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pheA-1</td>
<td>E. coli</td>
<td>GACGACAAATAGCGCCCTCCAAAATCGGGGGCCCTTTTTATATGATAACAAAA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ECK120029600</td>
<td>E. coli</td>
<td>TTTCCAGCCAAAAACACTACTAGCTTCTACTGAGAGCTTGGGTCTTTCTCTTCTCA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T7RNAp</td>
<td>CDS</td>
<td>ATGAAACACAAATCTATTTGCAAAAAATGCGCAGAAAGGCTAAGGCTATTTGAGCGTTTTTTTTTTTTTGGCTG</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(codon-optimized)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

109
<table>
<thead>
<tr>
<th>gRNA structural component</th>
<th>R3p terminator</th>
</tr>
</thead>
<tbody>
<tr>
<td>SUP4</td>
<td>Yeast promoter</td>
</tr>
<tr>
<td>TEF1</td>
<td></td>
</tr>
<tr>
<td>GAL1</td>
<td></td>
</tr>
<tr>
<td>STE</td>
<td></td>
</tr>
<tr>
<td>TDH3</td>
<td></td>
</tr>
</tbody>
</table>

Yeast promoter

integration flank

Chr XV 5'
Chrv XV 3' integration flank

CrtE CDS

GTTCCTTTTGGAAAGATCTAGCTCAGTCATTATTAGAACCAGGGAGGTCCA
GGCTTTGCTGGTGGGAGAGAAAGCTTATGAAGCTGGGGTTGCAGATTTGT
CGATTGGTCGCCAGTACACAGTTTTAAAAAGTCAGAGAATGTAGAGAAGT
ATGGATCTTTGAAACCCT
ATTTTACTTTTTTAGAATGACCTGTTCCCGACACTATGTAAGATCTAGC
TTTTAACATATTATGGAAACCTGAAATGTAAAATCTGAATTTTTGTATAT
GTGTATTATATTTGGGTAGTTCTTTTGAGGAAAGCATGCATAGACTTGCTG
TACGAACTTTATGTGACTTGTAGTGACGCTGTTTCATGAGACTTTAGCCC
TTTGAACATATTATCATATCTCAGCTTGAAATACTATAGATTTACTTTTG
CAGCCATTTCTGTGCTCCAAAGTTGGTGCGTATCTATTACTTTACTTCT
GTCCTTTGCACAGTTTCTGGCAAGGTTGTGCGTATCTATTACTTTCTTTCT
GTCCTTGCCAAGTTTTGCAGCAGGGCGGTCACAA
ATGGACTACGCTAACATCTTGACTGCCATTCCTTTGGAATTCACCCCACA
AGATGACATTGTCTTGTTGGGAACCATACCACTACTTATGGAAACACGAG
GTTACCTTTTCTGGTATGTTGAGTCTTCTGTCTTGCTCTGCTCTGGAA
TACGCTCAAGATGTCTTGGTAAGACCAGTGTTGTTTGCGAGAATTTGCTG
CAGATGAGATGAGTCAGATCTAGCTAATGGAAATTGCTCTGCTCTGGAA
GAGCTCTCTGTCTTCTCCCAACTCCACACTTACCACGGTACTACAGTAA
ACCTTTCTCTTTTCTCTTGGCTTACCAAGAAATCTTCAAATTGCG
TCCAACTCCAATTCTCAAACTTACCGTCATATCCTGCTCACTTCTCCT
ATGGTGGTTACTTCTGCTGCAATCCTGCAAGAGCTTTGTCTCTCGCTG
TCTTCTTATTGACCAACTTGATGATCGTTTTGGGTTTGTCTGCCTGTGAC
CACACTCAAGCCTTGTACTTGTTGCACGGTAGAACTATCTACGGTAACAA
GAAGATGCCATCTTCTTTCCCATTAATCACTCCACCAGTTTTGTCCTTGT
TCTTCTCCTCCAGACCATACTCCTCCCAACCAAAGAGAGATTTGGAATTG
GCTGTCAAGTTGTTGGAAAAGAAGTCCAGATCTTTCTTCGTTGCTTCTGC
CGGTTTCCCATCTGAAGTCAGAGAAAGATTGGTTGGTTTGTACGCTTTCT
GTCGTGTCACCGATGACTTGATTGACTCTCCAGAAGTTTCCTCCAACCCA
CACGCTACCATTGGAAGATGTCTTGGCAAGTCTTCTGACTTTATATTTCGGTCC
TCCAATTGCACCATCTCTCAACAGACAAGATTTTGTCTTCTCCATTATTAC
7. Bibliography


Xanthophyllomyces dendrorhous. *Applied and environmental microbiology*, 73, 4342-4350.

