Development of Environmentally Responsive Synthetic Promoters for Application in Soil

by

Matthew A. Bruno

A Thesis

Submitted to the Faculty

of the

WORCESTER POLYTECHNIC INSTITUTE

In partial fulfillment of the requirements for the

Degree of Master of Science

in

Biology and Biotechnology

April 2022

APPROVED:

_____

Dr. Natalie Farny, Advisor

_____

Dr. Luis Vidali, Committee Chair

_____

Dr. Scarlet Shell, Committee Member

_____

Dr. Eric Young, Committee Member

Abstract

Butanol biosensors have been developed in previous studies (Dietrich et al., 2013; Yu et al., 2019) but have fallen short in terms of their reproducibility and inducibility. This work aims to characterize existing sensors with a series of butanol inductions and to improve the existing butanol biosensors with a wide variety of synthetic biology tools and techniques. More specifically, this study aims to design and implement synthetic promoters for use in both the biosensors and for general genetic circuit design. Possible promoters were discovered by validating a transcriptomic data meta-analysis of genetic expression levels of *Pseudomonas putida* under a wide range of stressful conditions with qPCR. One possible promoter was tested in a reporter assay, and another was identified through comparing genetic loci in *P. putida* to analogous loci in *E. coli*. In conclusion, degradation of a vital activator protein was observed in the butanol biosensor. The promoter tested with a reporter assay did not perform as expected. This work confirms the need to validate all aspects of a microbial biosensor circuit to ensure its proper function, and to validate meta-analysis results with benchtop experiments to ensure transcriptomic information is transferrable to genetic circuit design.

Acknowledgements

Completing this thesis meant the input and help from many people along the way. First and foremost, I would like to thank the Farny Lab and everyone in it for being supportive over last two years. Specifically, I would like to thank Natalie for being everything we could ask for in a PI, and Felipe for answering all my questions in the learning process. I would also like to thank the Young Lab, who part of this project was in collaboration with, and Nilesh for teaching me how to use new instruments. A special thanks to Ally, Alli, and Julia not just for being in lab, but for being friends outside of it as well. Lastly, I would like to thank my partner Luisa for helping me understand statistics and for the years of continued moral support along with my friends and family who have always been there.

Contents

Table of Figures

List of Tables

List of Equations

Chapter 1, Introduction

1.1 Overview of Synthetic Biology

Broadly, synthetic biology can be described as the search for interchangeable biological parts, either natural or manufactured, that can be assembled to perform specific functions (Benner & Sismour, 2005). Synthetic biology got its start in the 1960's when the presence of regulatory systems that control cellular activity were proposed in response to studying the *lac* operon in *E. coli* (Monod & Jacob, 1961). Over time, the development of synthetic biology tools and techniques has allowed for the characterization, design, and engineering of microbial models, the most common of which is *Escherichia coli* (Cameron et al., 2014). The field grew rapidly with the advent of polymerase chain reaction (PCR), molecular cloning, sequencing, and the development of more higher throughput assays (Cameron et al., 2014). These new techniques allowed researchers to design and implement a wide array of novel genetic circuits based around promoter function to act as repressors, switches, protein producers, and more (Benner & Sismour, 2005).

The tools, techniques, and products from the field are used for in a wide range of applications from medical diagnostics to industrial production. However, manipulation of biological systems does not come without its challenges. While individual proteins or nucleic acids may be very well understood, their introduction into and interactions within an organism may not be more complex and often require validation (M. C. Y. Chang & Zhao, 2015). Experimental conditions, such as temperature or carbon sources, play a role in the ability to engineer and predict an organism's performance by altering the efficiency of engineered and natural constructs. This work aims to use synthetic biology approaches to design and implement bacterial promoter constructs for use under environmental stresses, specifically aimed toward soil stresses.

1.2 Microbial Biosensors

E. coli is a common Gram-negative bacterium that was first discovered by Theodor Escherich in 1884 while studying gut microbiomes (Escherich, n.d.). Over the following decades *E. coli* would become one of the most widely studied organisms as it was used as the primary model organism for anything and everything in microbiology (Blount, 2015). Because of its widespread use, many discoveries about basic biological mechanisms and processes were made using *E. coli* as the model. The use of *E. coli* was so common because of how easy it is to work with. It is fast growing, noninfectious, easy to isolate, and robust, making it advantageous for research purposes (Blount, 2015).

In addition to the wild type parent, modified strains of *E. coli* were developed for more specialized use. Two strains relevant to this work, whose function and purpose will become more apparent when discussing butanol and butanol biosensors, are the DH1 Δ*adhE* strain and the BL21 DE3 strain. The DH1 Δ*adhE* strain is a knockout strain which lacks the gene *adhE*, a gene encoding for an alcohol dehydrogenase. Alcohol dehydrogenases are a group of enzymes, often found in the liver and stomach, that play a main role in the degradation and clearance of alcohols, such as ethanol or butanol (Jörnvall, 1994). The BL21 DE3 strain is derived from an *E. coli* strain simply named B, which was originally used for phage research. BL21 DE3 expresses a T7 RNA polymerase and is protease deficient, making it less likely to degrade proteins and a good candidate for protein production and purification processes (Daegelen et al., 2009). These strains, along with a multitude of others, have made *E. coli* able to perform a wide range of specialized functions.

One of these more specialized functions enabled by synthetic biology tools is the creation of biosensors. Analytical systems using biological parts had been created in the 1960's and 1970's for detecting oxygenation levels in blood with an enzyme electrode (Renneberg et al., 2007). However, this system required extra analysis and reagents for use, and it was not until 1997 that biosensors became more well defined, shaping how current research looks at biosensors (Renneberg et al., 2007). The new description of biosensors defined them as being integrated into a self-contained system and able to report quantitative information using biological elements (Thévenot et al., 2001). Microbial biosensors are an even more specific group within the realm of biosensors. These sensors, which are more relevant to this work, focus more on genetic modification of the sensing microbe to allow it to monitor conditions in its environment

(Nakamura et al., n.d.). Microbial sensors have the advantages of longevity, cost effectiveness, and tolerance to a wide range of environmental measuring conditions (Nakamura et al., n.d.). The wide range of measuring conditions becomes apparent when looking at the number of uses for microbial biosensors. These sensors can be used in food applications (fermentation, glucose measurement, vitamin sensors, etc.), clinical applications (mutagen sensing, diagnostics, antibiotic measuring systems, etc.), and environmental analysis applications (oxygen level detection, toxicity measurement, organic molecule sensing, etc.) (Nakamura et al., n.d.). Microbial biosensors, due to their functionality under a wide range of conditions, allow for the detection of an array of molecules.

Transcription-based microbial biosensors fall under two main categories, repressors and activators. In repressors, a transcription factor is bound to its operator site preventing transcription unless the analyte, the molecule being sensed, is present and induces a change in the transcriptional repressor to release it from the DNA operator and allow transcription (Figure 1-1, A). Repressors can also function in combination with the analyte to repress transcription, only allowing transcription when there is no analyte present (Figure 1-1, B) (Fernandez-López et al., 2015). One example of a common repressor system in bacteria is the tetracycline (Tet) operon. In this system, the repressor protein TetR occupies an operator within the promoter region which prevents transcription, but when the effector molecule tetracycline is added it releases TetR and allows transcription (T. Das et al., 2016). Another highly studied repressor is the lactose (lac) operon in which the LacI repressor prevents transcription under certain concentrations of glucose and lactose (Marbach & Bettenbrock, 2012). The other main type of microbial biosensor is the activator. Activators, when bound to their effector molecule, recruit RNA polymerase and allow for transcription (Figure 1-1, C) (Fernandez-López et al., 2015). A butanol biosensor is an example of an activator that is relevant to this work. In this sensor, a protein called BmoR, when in the presence of butanol, is reported to bind butanol and activate a promoter by recruiting sigma factors and RNA polymerase.

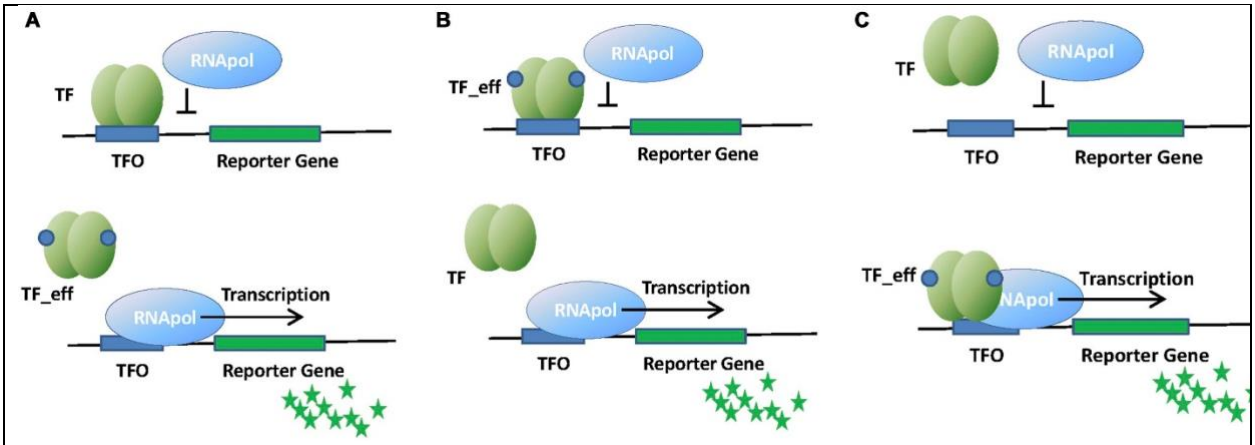


Figure 1-1, types of biosensors: Schematic representation of whole cell biosensors (WCBs) based on repressor (A), aporepressor (B), or activator (C) transcription factors (TFs) (Fernandez-López et al., 2015).

1.3 Limitations of Microbial Biosensors

Despite the advances made in recent years regarding the creation of microbial biosensors there remain limitations. One major limitation of microbial biosensors is low sensitivity to the effector molecule (Lim et al., 2015). Low sensitivity can be related to both bacterial population size and optical signal produced. The sensitivity and dynamic range of microbial biosensors are limited in culture growth because of limited resources, meaning that the full potential fluorescence will not be observed. Also, the depletion of analytes for induction over time may additionally limit output (Kim et al., 2015). Poor selectivity of the effector molecule is another issue biosensors face (Lim et al., 2015). The ability of a biosensor to detect the presence of the effector molecule is imperative to its function and if it cannot efficiently and correctly select the appropriate molecule, its ability to produce a signal will be hampered. Cellular heterogeneity, originating from slight cellular differences even when selecting from clonal populations, can also limit the functionality of some biosensors. Slight variations in nutrient availability or cellular components can lead to variation in levels of gene expression and therefore the amount of fluorescence produced from cell to cell can vary (Swain et al., 2002). The last issue with microbial biosensor is the leakiness of some genetic circuits. Some basic promoters, such as the T7 promoter, can have inherent leakiness and express genes cloned into the genetic circuit without activation from an effector molecule

(Namdev et al., 2019). Leakiness can be especially troublesome because a false positive signal will be observed, and it is difficult to turn off.

1.4 Transcriptomics Mining for New Biosensor Parts

Analysis of large amounts publicly available transcriptomics data can be used to overcome the limitations of microbial biosensors by identifying genetic parts, such as promoters, which behave in a desired way for integration into synthetic constructs. Promoters are important in bacterial transcription because gene expression is dependent on promoter recognition by RNA polymerases (RNAP) in order for them to bind to the DNA and for transcription to start (Chevez-Guardado & Peña-Castillo, 2021). Typically, bacterial promoters will have two short, conserved, and identifiable consensus boxes which lay 10 and 35 base pairs upstream from the transcriptional start site. RNAP, after binding to sigma factors (σ), form a holoenzyme which can then bind to the consensus sites in promoters to mediate transcription (Davis et al., 2017). Often multiple genes are regulated together with a single promoter in a multi gene operon. The power of a promoter to initiate expression of its gene is what makes them so important. Selecting the right promoter can allow the expression of a target protein even under stressed conditions (Browning & Busby, 2016). There are a multitude of promoters with varying strengths and functions, many of which have not been characterized, spread across bacterial genomes which makes promoters good candidates for identification through deep transcriptomic analysis.

The transcriptome is a good measure of promoter function because genetic transcription can be directly related to promoter strength. Comparing the transcriptome under multiple conditions gives an even better indication of how the promoter for a specific gene functions, which is where large scale transcriptomic analysis comes in. A common way of quantifying the transcriptome is with RNA-seq which amplifies total RNA extracted from a sample and then performs high throughput next generation sequencing to determine quantities of individual transcript present (Kukurba & Montgomery, 2015). The real power from transcriptomic comes not from one individual RNA-seq study, but from multiple studies analyzed under one pipeline (Caldas & Vinga, 2014). Meta-analysis takes advantage of vast amounts of publicly available transcriptomics studies and breaks them down into a set of specific phenotypes and their expression values, which can then be aggregated into one large study. Once together, comparisons

can be made between the phenotypes and expression levels from all the individual studies (Caldas & Vinga, 2014). Results derived from the assembly of multiple studies into one gives more phenotypic comparisons to define behavior of the transcriptome, providing more information than would be possible from one study. This type of deep transcriptome analysis can determine genes that have high levels of expression and small amounts of change across a wide range of conditions, and therefore have promoters that would be interest for synthetic biology applications (Kukurba & Montgomery, 2015).

There are some significant drawbacks when looking at RNA-seq data from both individual studies as well as data combined into a meta-analysis. Variability can be introduced into RNA-seq data with factors such as study size, sample pooling, statistical pipeline, and general experimental conditions (Bruning et al., 2015). It was determined that expression levels of up to 20% of genes identified through RNA-seq may be non-concordant with qPCR validation, meaning the values obtained from each assay do not mirror each other as they should (Coenye, 2021). Of these genes, 93% of them have fold change values less than two (Coenye, 2021). Results can be confounded further when comparing RNA-seq data across multiple studies as in bioinformatic meta-analysis. Challenges include accounting for complexities of biological phenotypes, small variation in experimental conditions of independent studies, and differences in microarray platforms used to gather the RNA-seq data (Lim et al., 2015). Lastly, meta-analyses of transcriptomic data have trouble inferring global transcriptome trends across multiple studies and can only compare data from studies where similar phenotypes were tested (Lim et al., 2015).

1.5 This Work

The overall subject of this study is to characterize and improve existing biosensors for detection of butanol, with the help of a new promoter found through validation of comparative transcriptomic analysis. Butanol biosensors have been developed to detect the presence of butanol in the environment using constructs based off the BmoR protein its corresponding inducible promoter P_{bmo} (Dietrich et al., 2013; Yu et al., 2019). However, these sensors do not have high reproducibility or high levels of inducibility. First, this work aims to further characterize the function of the existing butanol biosensors from the literature as well as apply a broad range of synthetic biology tools and techniques to improve the capabilities of the sensors. Second, this study

aims to use qPCR to validate a previous bioinformatic meta-analysis which determined genes with strong expression and low fold change under a wide range of stresses in *P. putida* (Harding et al., 2021). Finally, this work aims to identify the promoter region of one of the validated genes and use it for genetic engineering applications, such as integration with the butanol biosensor.

Chapter 2, Butanol Biosensors

2.1 Background

Butanol is a basic four carbon alcohol used mainly as a solvent or as an intermediate step in chemical synthesis. Mainly, butanol is produced through chemical synthesis which requires the uses of a petroleum-based feedstock called propylene (Nielsen et al., 2009). Recently, butanol has been identified as a possible next generation biofuel and possible gasoline replacement because of its compatibility in non-modified internal combustion engines along with a higher energy content compared to ethanol, which is already added to gasoline (Bokinsky et al., 2011). The attractiveness of butanol as a biofuel has led to an increased interest to produce butanol through alcohol-butanol-ethanol (ABE) fermentation, a process which can be facilitated by engineered microbes fermenting renewable food sources such as glucose (Dietrich et al., 2013). Butanol production through ABE fermentation, which can be carried out in organisms such as *E. coli* or *S. cerevisiae*, highlights the need to create a biosensor that can monitor the production of butanol within a sample (Shi et al., 2017). There have been some biosensors already developed which aim at screening for butanol producers and at monitoring butanol production, such as the work carried out by Yu et al., 2019 Dietrich et al., 2013. These biosensors are similar in their architecture, which use a system of promoters, regulatory proteins, and a reporter gene to signify the presence of butanol in E. coli (Figure 2-1).

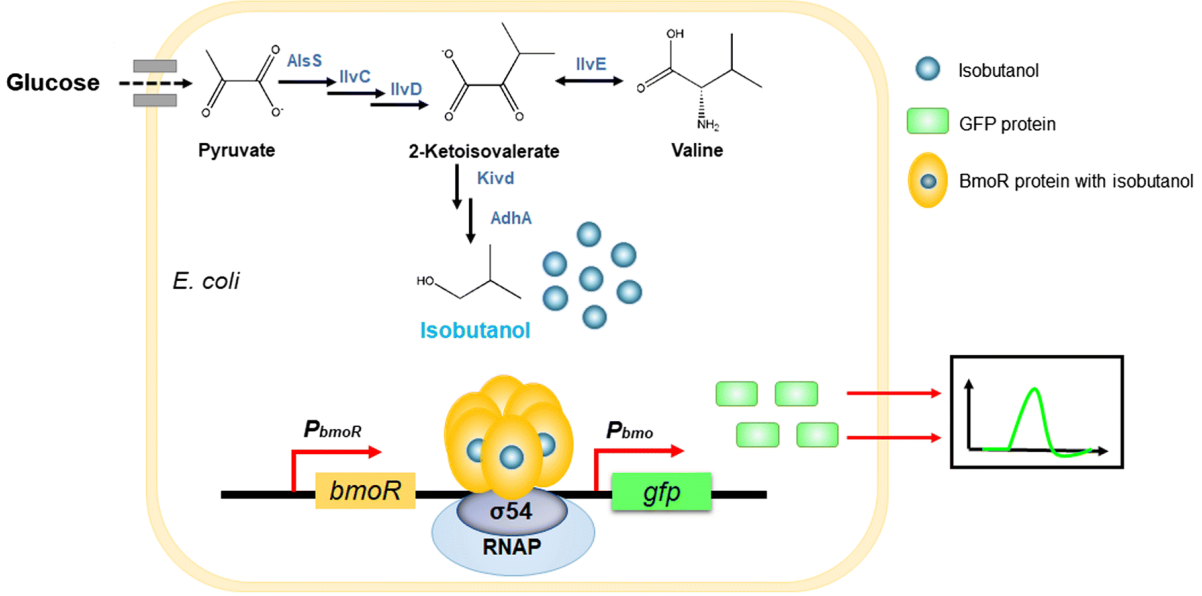


Figure 2-1, genetic circuit of a butanol biosensor from Yu et al., 2019.

The constructs rely on a constitutive promoter, called P_{bmoR}, to drive the expression of a regulatory protein called BmoR (Figure 2-1). BmoR, which should always be expressed due to its constitutive promoter, binds to butanol from the environment and should recruit the σ^{54} and RNAP holoenzyme to the inducible promoter P_{bmo} (Yu et al., 2019). The complex thereby induces the transcription of a reporter, in this case GFP, whose fluorescence can be quantified (Dietrich et al., 2013; Yu et al., 2019). Ideally, only in an environment containing butanol can the system be induced, and the reporter expressed. However, the butanol biosensors in the literature were found to have low levels of induction along with leakiness upon attempted reproduction of results (Dietrich et al., 2013). It is also difficult to quantify the fluorescence observed in previous studies as controls were not included (Dietrich et al., 2013; Yu et al., 2019). The overall lack of sensor functionality may be due to the assumption that the genetic circuit was performing as expected without verifying the functionality of the BmoR protein.

Herein are described a series of induction experiments using a published butanol biosensor in comparison to a butanol sensor designed within our laboratory. Three different techniques were used in parallel to measure biosensor function: plate reader, flow cytometry, and western blotting. Included were control constructs lacking the BmoR regulator and lacking a GFP gene. The results reveal no evidence of inducibility by any of the genetic circuits tested. The circuits were redesigned and rebuilt to monitor BmoR expression, and it was found that BmoR is likely unstable in *E. coli*.

The work indicates that BmoR-based butanol biosensors are likely not functional, and that prior published studies using the system may be artifacts produced by assay conditions and lack of appropriate controls.

2.2 Methods

Plasmid and Strain Construction

Escherichia coli strains DH1 Δ*adhE* and BL21 DE3 were used as hosts for experimental trials. To create competent cells from the host strains, starter cultures of each host strain were grown in Difco LB Miller Broth containing tryptone at 10g/L, yeast extract at 5g/L, and sodium chloride at 10g/L and were incubated overnight at 37℃ and at 220 RPM on a rotary shaker. The next day 125 mL of fresh media was inoculated with 1 mL of the overnight cultures and grown at 37℃ and 220 RPM for about 2 hours, until OD_{600} of 0.3 was achieved. The cultures were transferred to chilled centrifuge tubes and spun at 3000g and 4℃ for 10 minutes, then resuspended in a total of 40 mL cold CCMB80 buffer containing 10 mM KOAc, 80 mM $CaCl_2 \cdot 2H_2O$, 20 mM $MnCl_2 \cdot 4H_2O$, 10 mM $MgCl_2 \cdot 6H_2O$, and 10% glycerol. The resuspensions were spun again at 3000g and 4℃ for 10 minutes and resuspended again in a total of 5 mL each of cold CCMB80 buffer. 250 µL aliquots of the competent cells were created and stored at -80℃. Five plasmids were transformed into each of the two competent host strains for a total of ten plasmid/host combinations. The plasmids were obtained by growing a culture of *E. coli* containing the plasmid overnight and using the GeneJet Plasmid MiniPrep Kit (ThermoFisher) per the manufacturer's protocol. The plasmids used (listed in Table 2-1) are a negative control pUC19 with Amp^R, positive control *taclac*-GFP with Kan^R, Keasling PBMO#1 with Amp^R, SHEOL PJ-0204 Butanol Sensor with Kan^R, and PBMO-GFP with Kan^R.

Table 2-1, butanol biosensor construct summary

Butanol Sensor Strains and Plasmids		
E. coli Strain Name	Description	Source
BL21 DE3	Protease deficient, less likely to degrade BmoR	New England Biolabs
DH1 Δ*adhE*	Alcohol dehydrogenase deficient, less likely to degrade butanol	A gift from Dr. Jay Keasling, University of California Berkeley

Plasmid Name	Description	Source
pUC-19	Negative control, no GFP	New England Biolabs
Taclac-GFP	Positive control, constitutive GFP	Constructed by Dr. Andres Felipe Carrillo, WPI
Keasling PBMO#1	P_{BmoR}-*bmoR*, P_{BMO}-*gfp* (native promoter)	(Dietrich et al., 2013)
SHEOL Sensor	PJ_{23119}-*bmoR*, P_{BMO}-*gfp* (non-native promoter)	Constructed by Dr. Andres Felipe Carrillo, WPI
PBMO-GFP	P_{BMO}-*gfp* (does not contain BmoR)	(Dietrich et al., 2013)
Modified SHEOL Sensor	$P_{Library}$-*bmoR*, P_{BMO}-mCardinal	This Work

A summary and description of all the strains and plasmids used for the butanol sensor constructs.

An aliquot of competent DH1 Δ*adhE* and BL21 DE3 host strains were thawed on ice. 1µL of each plasmid was added to separate microcentrifuge tubes containing 50 µL of host cells. The host and plasmid were allowed to incubate on ice for 30 minutes then warmed at 42℃ for 45 seconds. After 5 more minutes on ice 250 µL of LB was added and the cultures were incubated at 37℃ on a rotator for 2 hours. Cultures were streaked onto sterile Lennox L Agar (Invitrogen) plates containing the appropriate antibiotic and allowed to incubate overnight at 37℃. The antibiotics used for this particular experiment were kanamycin sulfate at a stock concentration of 50 mg/mL and ampicillin sodium salt at a stock concentration of 100 mg/mL. The next day colonies from each of the plasmid/host combinations were picked and grown in liquid culture comprised of LB and the appropriate amount of corresponding antibiotic. 500 µL of each overnight culture was added to 500 µL of a sterile solution containing 50% glycerol in 50% dH_20. Glycerol stocks were stored at -80℃ for future use in inductions.

Induction Assays

A series of assays was used to determine the inducibility of the butanol sensor and other plasmids in the BL21 DE3 and DH1 Δ*adhE* strains (Figure 2-2). Cultures containing 5 mL LB and 5 µL of the appropriate antibiotic were inoculated with each plasmid/parent combination and allowed to grow overnight at 37℃ and 220 RPM. These cultures were then back diluted into four new tubes containing 10 mL fresh LB and 10 µL appropriate antibiotic by adding 500 µL of the

dense overnight culture. These cultures, 40 total (20 each host type and 4 each plasmid type), were allowed to incubate for 2 hours at 37℃ and 220 RPM. After 2 hours the cultures were induced with butanol. 1-Butanol (Sigma Life Science) was added to three out of four of the cultures per plasmid to a final concentration of 10 mM, 20 mM, and 40 mM. The cultures were again incubated at 37℃ and 220 RPM but for 24 hours before readings were taken. Each culture was analyzed with three different assays and each induction was run three times, resulting in three replicates per assay. For both plate reader and flow cytometry assays a two-way ANOVA test was performed to determine any significance between each plasmid construct and butanol concentration.

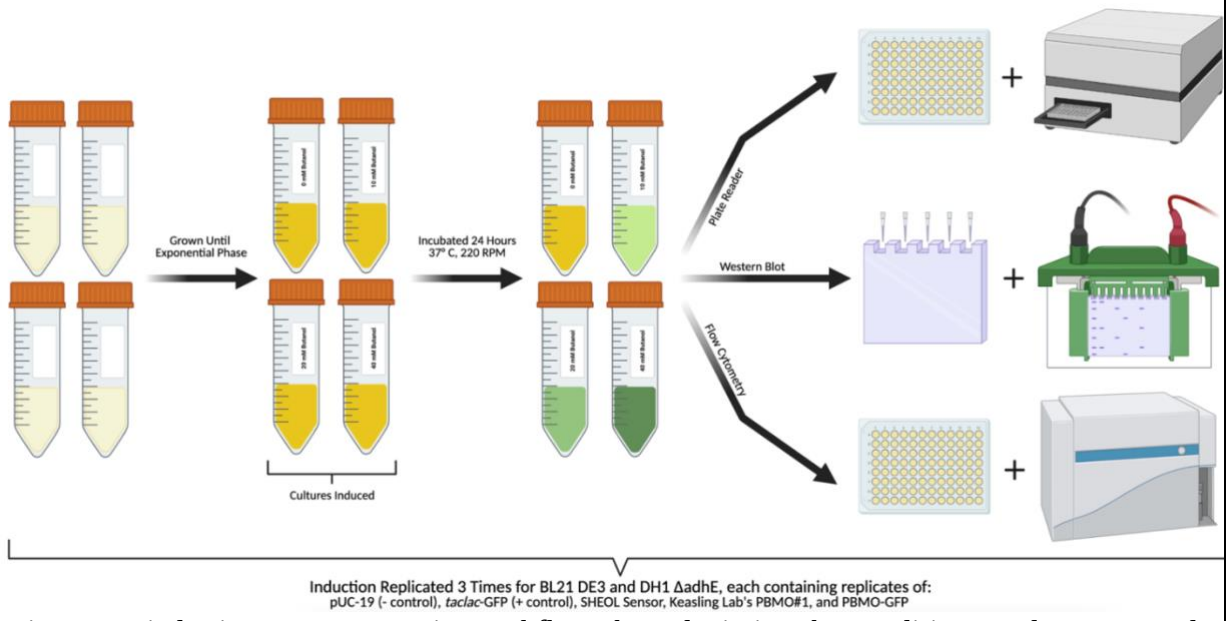


Figure 2-2, induction assays: Experimental flow chart depicting the conditions and assays used for butanol inductions.

Plate Reader:

A 100 µL sample was taken from each culture and loaded onto a Costar 96 well flat bottom plate (Corning) and an optical density reading at 600 nm (OD_{600}) was taken for each sample on a Perkin Elmer VICTOR3 plate reader. A ratio of ODs was then calculated, and samples were diluted with fresh LB according to that ratio so that a 1 mL sample of culture was created. These 1 mL cultures each had the same or very similar ODs. Three replicates of 100 µL from the normalized OD cultures, along with LB blanks, were then loaded onto a new Costar 96 well flat bottom plate

and analyzed on the Perkin Elmer VICTOR3 plate reader for OD_{600} and for GFP fluorescent intensity with an excitation wavelength of 535 nm and an emission wavelength of 485 nm. By averaging the values for fluorescence and OD the relative fluorescence compared to the LB blank was then calculated. The equation used to calculate relative fluorescence is shown in Equation 2-1 below (*Part:BBa R0051 - Parts.Igem.Org*, n.d.).

$$Relative\ Fluorescence = \frac{Flu_{Bacteria} - Flu_{Blank}}{Flu_{Blank}} \div (OD600_{Bacteria} - OD600_{Blank}$$

Equation 2-1, relative fluorescence: Formula used for calculation relative fluorescence relative to an LB blank.

Flow Cytometry:

A clear, flat bottom 96 well plate was used for analysis with flow cytometry. A master mix comprised of 5 mL H_20 and 20 µL of kanamycin sulfate stock (50 mg/mL) was created. 195 µL of master mix was distributed into all the wells needed for analysis. 5 µL of culture, each with three replicates, were then added to the master mix in the plate. The plate was then read on a Beckman Coulter CytoFLEX S flow cytometer recording a maximum of 50,000 events. The data from the flow cytometry was analyzed on FlowJo analysis software version 10.7.1

Western Blot:

The remaining cultures with normalized OD were spun down at 14,000 RCF for 10 minutes, the supernatant was removed, and the pellets were resuspended in 100 µL 1 X SDS + 1 mg/mL dithiothreitol (DTT). Samples were then heated at 80°C for 10 minutes and stored at -20°C for later analysis. When ready for analysis, samples were removed from the freezer and thawed at 37°C for 10 minutes, vortexed for 5 seconds, and centrifuged at 14,000 RCF for 1 minute. The positive control taclac-GFP samples were diluted at a ratio of 1:500 to prevent overexposure during imaging. 10 µL of each sample was then loaded into Invitrogen Novex 4-20% Tris-Glycine gels. Gels were run in the Invitrogen Mini Gel Tank with 1 X Tris-Glycine SDS Page run buffer. The gel was then transferred to a nitrocellulose membrane in cold 0.5 X Tris-Glycine in 20% methanol buffer. After a series of washes with PBS-Tween-20 blot wash, the membrane was probed for GFP

with GFP tagged mouse monoclonal antibody (Invitrogen catalog number MA5-15256) resuspended at a ratio of 1:1000 in 5% BSA. Anti-mouse IgG HRP linked secondary antibody (Cell Signaling Technology catalog number 7074) diluted at a ratio of 1:20,000 was added after the primary antibody. After another series of blot washes, 1mL of SuperSignal West Pro Plus Luminol and Enhancer (Thermo Scientific) were added to the membrane. Membranes were imaged on an Azure Biosystems c600, adjusting for white balance and contrast.

Building an Improved Butanol Biosensor

gBlock Gene Fragment Preparation:

In order to improve the butanol biosensor, it was determined that swapping out GFP for mCardinal would reduce background in fluorescence and creating a promoter library for the promoter driving bmoR could help with inducibility. The plasmid that would be modified was the original SHEOL sensor (Figure 2-3). A gBlock Gene Fragment from IDT containing an HA tagged mCardinal sequence, restriction sites for promoter library insertion (Table 2-2), and a FLAG tag for BmoR was ordered. Synthetic DNA gBlocks and oligos were purchased from Integrated DNA Technologies (IDT, Coralville, Iowa).

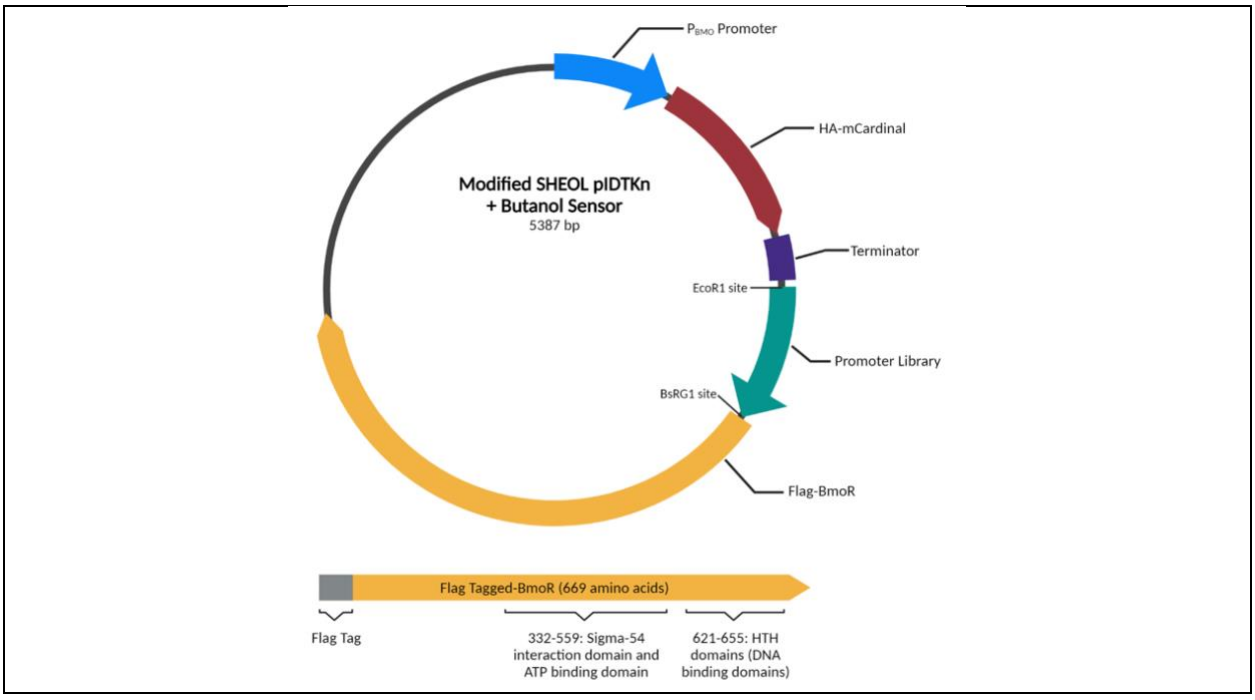


Figure 2-3, modified butanol biosensor construct: Construction of the modifications made to the SHEOL biosensor to improve its inducibility.

Table 2-2, promoter library sequences

Butanol Sensor Proposed Promoter Library	
Promoter Name	Sequence (5'-3')
J109	TTTACAGCTAGCTCAGTCCTAGGGACTGTGCTAGC
J117	TTGACAGCTAGCTCAGTCCTAGGGATTGTGCTAGC
J114	TTTATGGCTAGCTCAGTCCTAGGTACAATGCTAGC
J115	TTTATAGCTAGCTCAGCCCTTGGTACAATGCTAGC
J106	TTTACGGCTAGCTCAGTCCTAGGTATAGTGCTAGC
J101	TTTACAGCTAGCTCAGTCCTTGGTATTATGCTAGC

Sequence of the promoters to be cloned into the promoter library position in the new biosensor construct. Nucleotides highlighted red indicate where the weaker promoters differ from J101, the strongest constitutive promoter (Wan et al., 2019).

The gBlock was centrifuged at >3000g for 30 seconds to ensure the powered DNA was at the bottom of the tube. 50 µL of molecular biology grade H_20 was added to create a solution with a concentration of 20 ng/µL. The solution was vortexed and then incubated at 50°C for 20 minutes. The solution was then vortexed and centrifuged once again and then stored at -20°C for later use. In order to ensure large amounts of high-quality product, the gBlock was PCR amplified with Q5 High-Fidelity 2X Master Mix DNA Polymerase (New England Biolabs). A 50 µL reaction was assembled on ice according to the manufacturer's specifications using 1 µL of the gBlock as the template DNA and an annealing temperature of 50°C to allow the polymerase to attach to the DNA more efficiently. To determine if the PCR reaction amplified the gBlock correctly, 5 µL of reaction product was mixed with 4 µL of H_20 and 1 µL of 10X FastDigest Green Buffer (ThermoFisher). Gel electrophoresis was run with the reaction product mixture in a 50 mL gel containing 1% UltraPure Agarose (Invitrogen) in 1X Tris acetate EDTA (TAE) buffer and SYBR Safe DNA Gel Stain (Invitrogen). The gel was then imaged on a Bio-Rad Gel Doc XR. After it was determined that the reaction was successful, the remaining reaction volume was purified using the GeneJet Gel Extraction and DNA Cleanup Micro Kit (Thermo Scientific) and the manufacturer provided procedure B: PCR cleanup, dimers removal protocol. The purified PCR product was then digested for at least 30 minutes at 37°C with FastDigest NdeI (ThermoFisher) enzyme and diluted 10X FastDigest Green Buffer, then run on a gel. The digested insert was then excised from the gel and

purified using the same Thermo Scientific kit as before but with procedure C: DNA extraction from gel protocol and stored at -20℃ for later use.

Cloning Procedure:

Miniprepped SHEOL plasmid was being used as the vector and the digested PCR amplified gBlock serves as the insert. First, the vector was digested with NdeI for 30 minutes at 37℃, spun down, and then incubated at 80℃ for 20 minutes to inactivate the restriction enzymes. The ligation kit used was the Rapid DNA Dephosphorylation and Ligation Kit (Roche) per the manufacturer's instructions. The ligation product was transformed into SIG10 Chemically Competent Cells (Sigma-Aldrich) and plated to LB agar plates containing kanamycin and left to incubate at 37℃ overnight.

Screening Clones:

Colonies that grew from the transformation needed to be screened both for insert and insert direction. 10 colonies were picked from the transformation plate and grown in liquid LB culture with kanamycin overnight. The plasmids were extracted from the cultures with GeneJET Plasmid Miniprep Kit (ThermoFisher). NdeI was used to screen for vector insert because it is the restriction site flanking the gBlock insert in the plasmid (Figure 2-13). When digesting with NdeI and running the product on a gel, a band of DNA at ~4500 base pairs and a band at ~950 base pairs should be seen. Out of the ten colonies screened only six had the correct restriction pattern. The six colonies that had the correct restriction pattern must be screened again but this time for insert direction to ensure the gBlock had not been cloned in backwards. The restriction enzymes used to do this were BamHI and EcoRI. Clones with insert and correct insert orientation would yeild a restriction patternt with two bands, one at ~4000 and one at ~1300. These clones were also screened with BsrGI because there is a unique restriction site for this enzyme found only in the gBlock. Restrictions patters imaged on an agarose gel are shown in Figure 2-13. Five out of the original ten clones screened had the correct restricition pattern and these clones were sent for sequencing to undoubtedly confirm that the clones containted the gBlock and contained it in the correct oritentation. Primer sets that would be used to sequence the insert were ordered as oligos from

IDT. The primers were resuspended to a stock solution of 100 µM; the working stock solution for primers was 5 µM. 5 µL of the primer working solution was combined with 10 µL of each of the plasmids. Plasmid concentration was also quantified using a NanoDrop Spectrophotometer (ThermoFisher). The primer and plasmid mixtures were sent for Sanger sequencing at QuintaraBio (Cambridge, MA, USA). After sequencing it was determined that 3 of the original 10 colonies screened contained the correct insert.

Biosensor Verification Via Butanol Induction:

The plasmid from each of the clones that were determined to have the correct insert were transformed once again into DH1 Δ*adhE* and BL21 DE3, however this time new stocks of competent cells were used. These stocks were created using the *Mix & Go! E. coli* Transformation Buffer Set (Zymo Research) and the manufacturer provided protocol. Inductions were performed the same as the previous inductions except only 0 mM and 40 mM butanol conditions were tested. After the induction, samples were run on a BioTek Synergy H1 plate reader, taking a reading for OD_{600} and a reading with the fluorescence spectrum set at an excitation wavelength of 604 nm and an emission wavelength of 659 nm. Different antibodies were also used for the Western blots to probe for the new FLAG tag and HA tag. Blots were first probed for the HA tag using HA Tag Rabbit PolyAb (Proteintech catalog number 51064-2-AP) diluted at a ratio of 1:1000 in 5% BSA. After imaging and washing with PBS-Tween-20 blot wash, the membranes were probed again, this time for the FLAG tag with a DYKDDDDK Tag Mouse McAb (Proteintech catalog number 66008-3-Ig) diluted at the same ratio. DH1 Δ*adhE* induction samples were also run on another SDS gel, then stained to visualize total protein. The staining solution contains 0.1% Coomassie Blue in 30% methanol and 5% acetic acid. The gel was placed in a small microwave safe box and enough staining solution was added to cover the gel. The gel was then microwaved four times in 20 second intervals. The gel was destained by boiling in dH_20 for 6 minutes or until the excess blue stain has been removed. The gel could then be sealed in plastic and imaged.

2.3 Results

Induction and Attempted Validation of Published Butanol Biosensors

Quantitative assessment of butanol circuit induction by plate reader assay.

In order to validate previous observations of butanol sensor function, we obtained the PBMO#1 plasmid (Table 2-1 and Figure 2-4) from Dr. Jay Keasling's lab (Dietrich et al., 2013) and performed an induction trial as per Yu et al., 2019. *E. coli* DH1 Δ*adhE*, a host strain lacking an alcohol dehydrogenase gene that is meant to inhibit the metabolism of butanol (Jörnvall, 1994), was used for the induction trials. DH1 Δ*adhE* cultures expressing PBMO#1 were exposed to 0, 10, 20, and 40 mM butanol for 24 hours, then samples were measured for GFP fluorescence and OD_{600}. Results were reported by calculating the ratio of GFP fluorescence over the OD_{600} measurement. Using this method, we observed a similar apparent induction of GFP upon addition of butanol up to 20 mM, with no additional fluorescence obtained at the 40 mM concentration (Figure 2-5, A). The results are similar to the observations published by Yu et al., 2019 (Figure 2-5, B).

The fluorescence per OD metric can be misleading for a few reasons. The GFP/OD values are not being compared to a positive or negative control meaning that their value is somewhat arbitrary. Also, since the fluorescence values are large and the OD values are relatively small, slight changes in the OD can result in large changes in the total fluorescence per OD values along with high error. Further, the results published by Yu et al., 2019 do not contain controls that describe the constitutive expression of the genetic circuits, nor do they measure negative controls that do not express GFP.

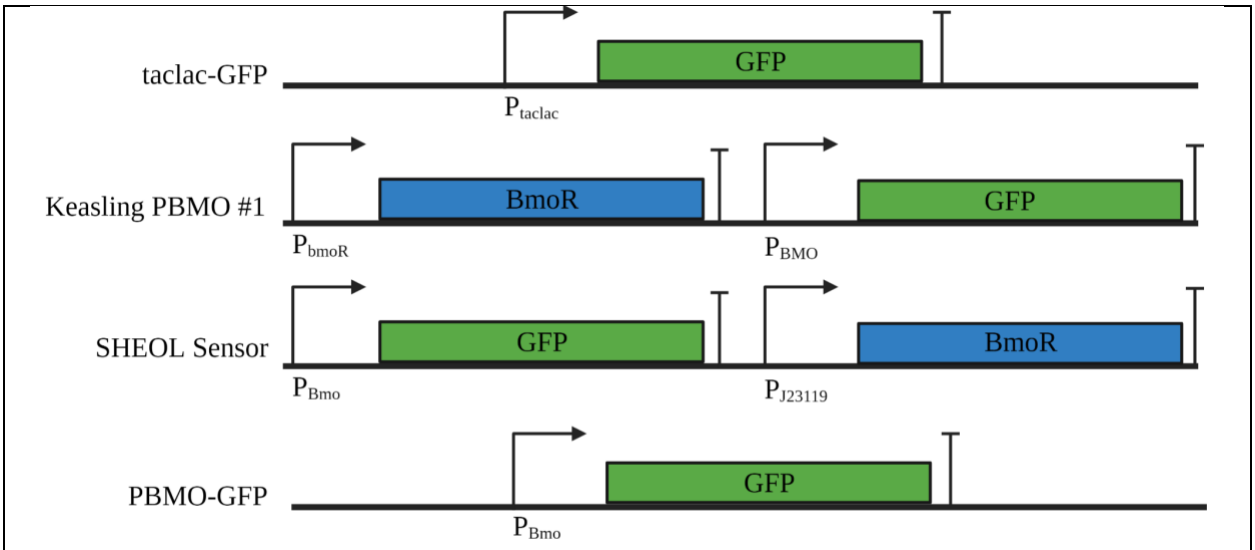


Figure 2-4, butanol biosensor schematics: Schematic representation of biosensors and constructs used. Information about these constructs appears in Table 2-1.

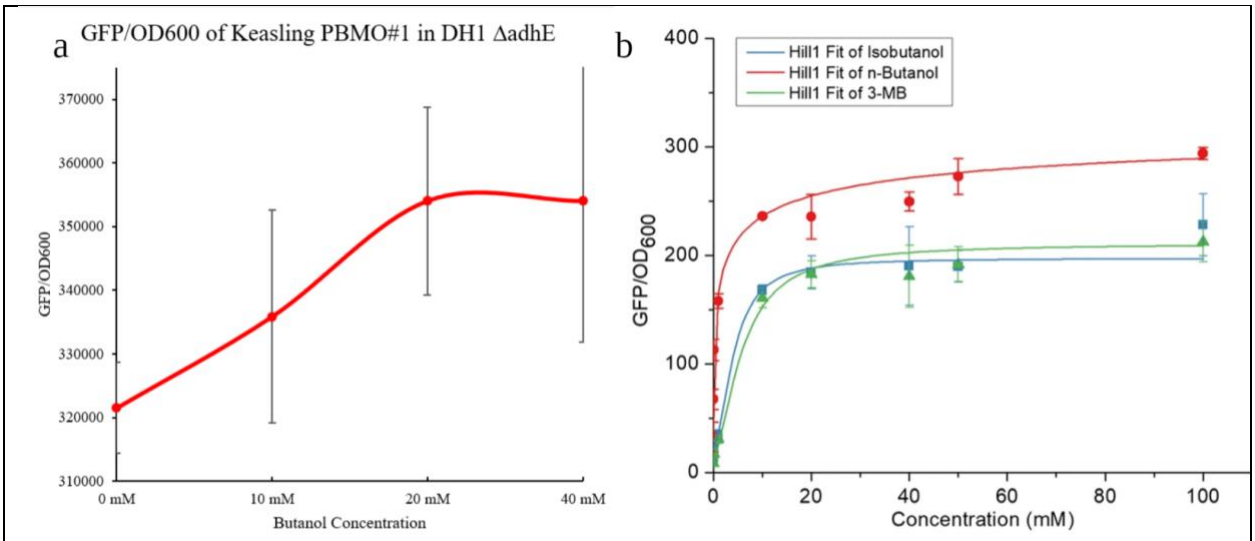


Figure 2-5, GFP/OD of existing butanol biosensor: Comparison of GFP fluorescence per OD of samples obtained in this study (a) and those from the literature (b) (Yu et al., 2019). *The GFP/OD values seen in the data from this figure are from inductions with DH1 ΔadhE but are representative of BL21 DE3 inductions as well. No notable increase was seen in GFP levels from cells induced with more than 40 mM of butanol, therefore for this work only 0 mM, 10 mM, 20 mM, and 40 mM butanol concentrations were tested. Fluorescence values of the LB blank are subtracted out. Error bars are +/- S.E.M. with n=3.*

To remedy the issues associated with using a direct fluorescence per OD ratio, and to measure additional controls, the experiment in Figure 2-5 A was repeated with a broader range of samples. A version of the BmoR biosensor was created by Dr. Andres Felipe Carrillo that contains the BmoR regulator under the control of a strong constitutive promoter (Figure 2-4, SHEOL sensor), as opposed to the Keasling PBMO#1, which expresses BmoR under the control of its endogenous promoter (Figure 2-4, PBMO#1). To measure the constitutive strength of the P_{BMO} promoter, Dr. Carrillo created a plasmid lacking the BmoR regulator (Figure 2-4, PBMO-GFP). pUC19 was used as a negative control lacking GFP. These plasmids were transformed into both the *E. coli* DH1 Δ*adhE* strain, and into BL21 DE3, then induced with butanol. The data were analyzed using a relative fluorescence calculation (Equation 2-1). This technique calculates the fluorescence relative to a blank, in this case fresh LB media was used. Relative fluorescence has the advantage of comparing all the samples and control to a blank therefore allowing a direct comparison of the samples relative to each other. Another advantage of relative fluorescence is a reduction of background fluorescent noise because the fluorescence and OD of the LB blank are subtracted out from the samples. When re-analyzing the fluorescence and OD values using relative fluorescence and appropriate controls, there is no significant induction of either the SHEOL or the PBMO#1 sensors by butanol in either strain background (Figures 2-6 and 2-7). Further, we observe no difference between the SHEOL sensor with or without the BmoR regulator. A two-way ANOVA test was performed that revealed there is no significance ($p \leq 0.05$ indicating significance) difference in fluorescence caused by induction between the three butanol sensors and the pUC-19 control in either DH1 Δ*adhE* (p=.984) or BL21 DE3 (p=1.00). Overall, the promoter P_{BMO} is a weak driver of GFP expression, as compared to the taclac promoter.

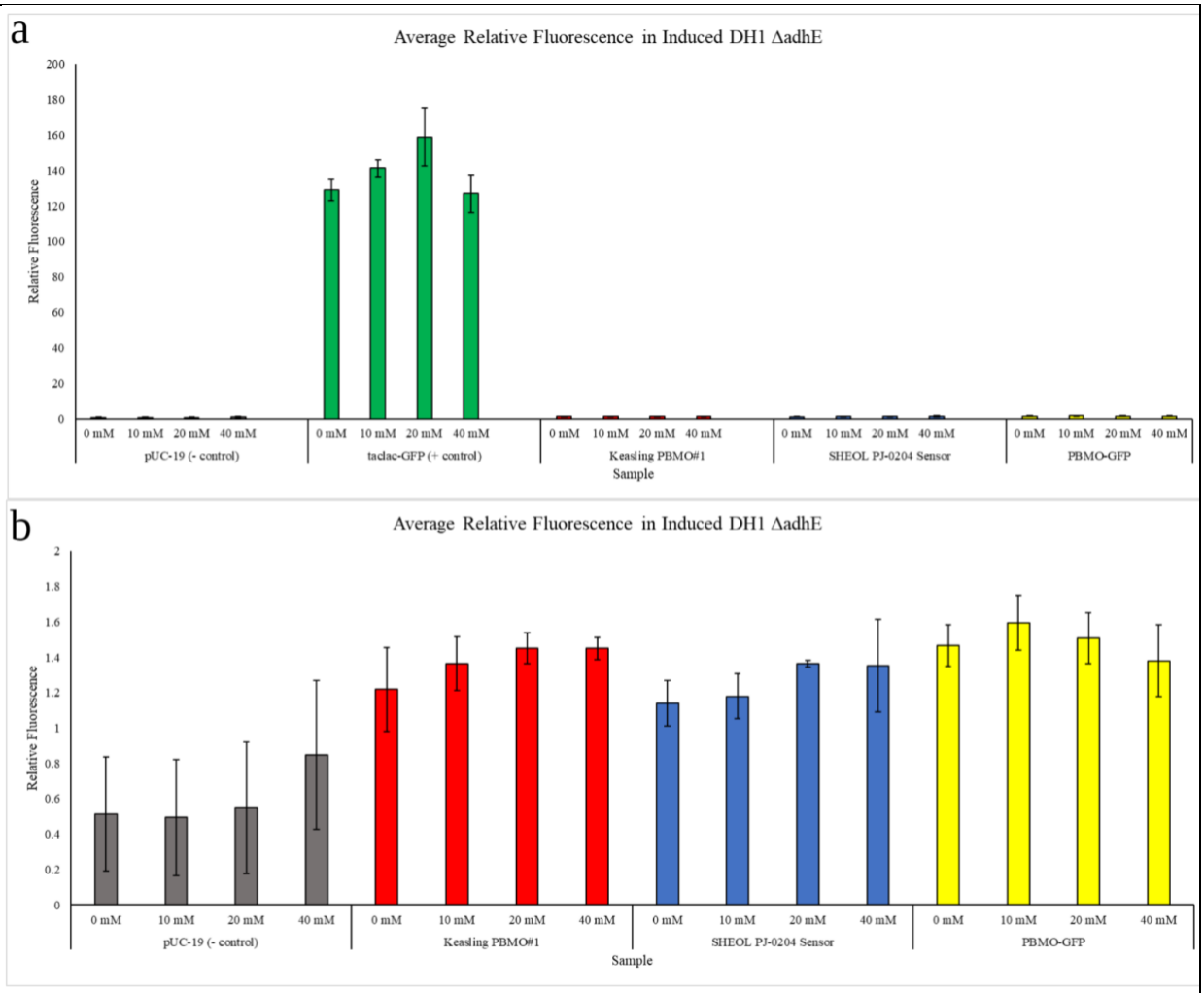


Figure 2-6, PR DH1 ΔadhE inductions: Plate reader data from DH1 ΔadhE inductions analyzed with relative fluorescence. Fluorescence values are shown both in comparison to the positive control (a) and without the positive control (b) to illustrate how any induction observed did not produce nearly as much fluorescence as the positive control. Fluorescence is relative to fresh LB media, which has a relative fluorescence of 0. Error bars are +/- S.E.M. with n=3

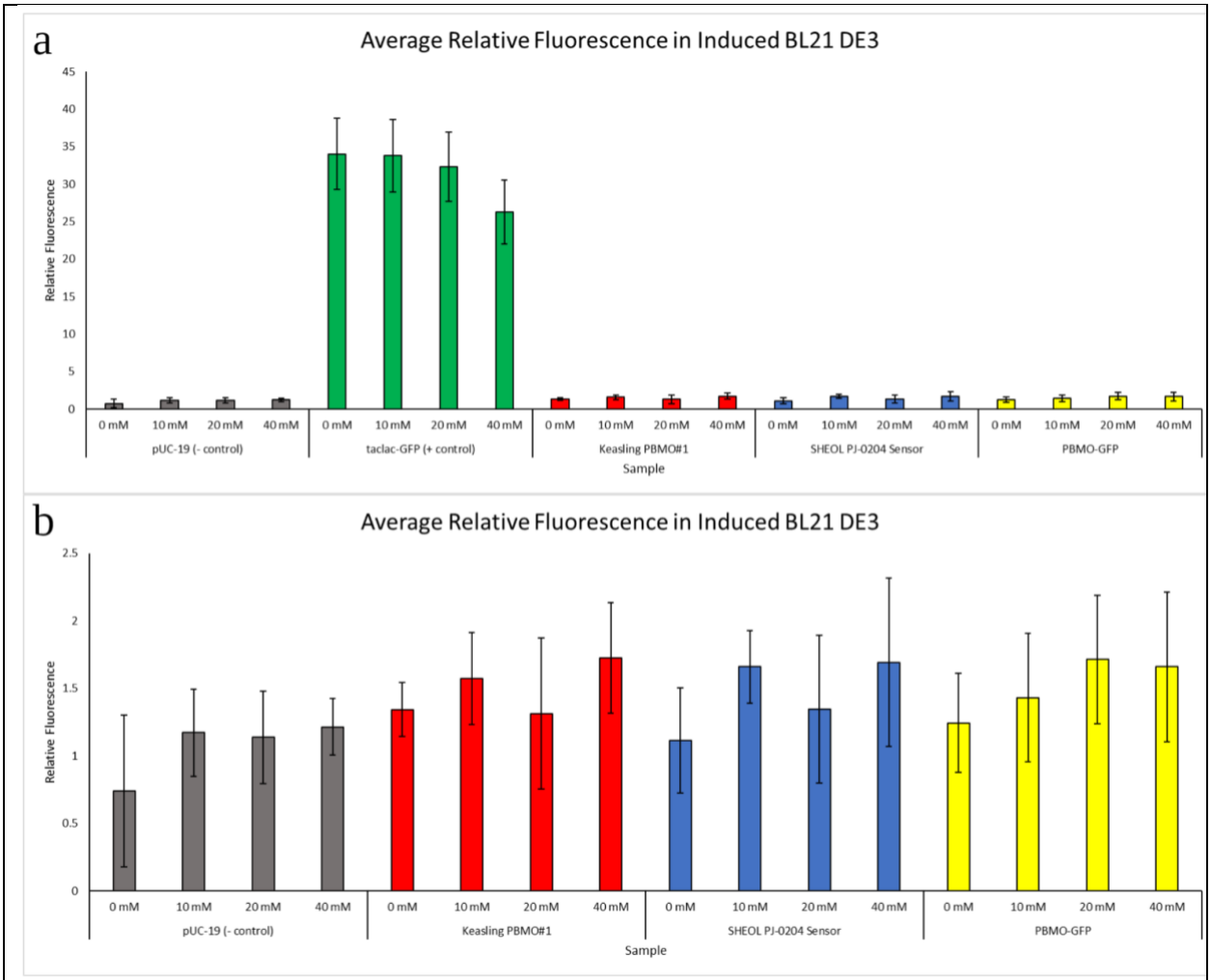


Figure 2-7, PR BL21 DE3 inductions: Plate reader data from BL21 DE3 inductions analyzed with relative fluorescence. Fluorescence values are shown both in comparison to the positive control (a) and without the positive control (b) to illustrate how any induction observed did not produce nearly as much fluorescence as the positive control. Fluorescence is relative to fresh LB media, which has a relative fluorescence of 0. Error bars are +/- S.E.M. with n=3

Quantitative assessment of butanol circuit induction by flow cytometry.

There is significant variability in the data generated from the plate reader, as evidenced by the large error bars associated with these samples. Therefore, to validate our plate reader observations, a more robust assay was needed to quantify small amounts of fluorescence produced to better characterize induction patterns. Flow cytometry was chosen because of its ability to determine fluorescent signal on the scale of individual cells, making it ideal to detect induction,

particularly in sub-populations of cells. PBMO#1, SHEOL sensor, and PBMO-GFP constructs were transformed into *E. coli* DH1 Δ*adhE* and BL21 DE3 for induction with 0 mM, 10 mM, 20 mM, and 40 mM butanol. Taclac-GFP was used as the positive control and pUC-19 was used as the negative control and both were induced alongside the biosensor constructs. After 24 hours in butanol, samples were adjusted to a common OD_{600} before diluting and plating in a solution containing high levels of kanamycin to fix the cells. Plates were then run on the flow cytometer to detect fluorescence emitted by GFP. The lack of induction observed with the plate reader assay is especially apparent when looking at the flow cytometry data. No pattern of induction was observed in relation to increasing butanol concentration in either DH1 Δ*adhE* (Figure 2-8, A and Figure 2-9) or BL21 DE3 (Figure 2-8, B and Figure 2-10) across all the plasmids. A two-way ANOVA test was performed and revealed that there is no significant ($p \leq 0.05$ indicating significance) difference in fluorescence caused by induction between the three butanol sensors and the pUC-19 control in either DH1 Δ*adhE* (p=1.00) or BL21 DE3 (p=.270). It is also important to note that there was still fluorescence observed in the pUC-19 negative control plasmid, likely due to low levels of background autofluorescence emitted in the green spectrum by the LB media in which all the cultures were grown and but non fluorescent bacteria. Figure 8, A and B highlight how the levels of fluorescence observed in the biosensor strains are comparable to the background fluorescence levels while the mean fluorescence of the positive control is three orders of magnitude greater than the other constructs.

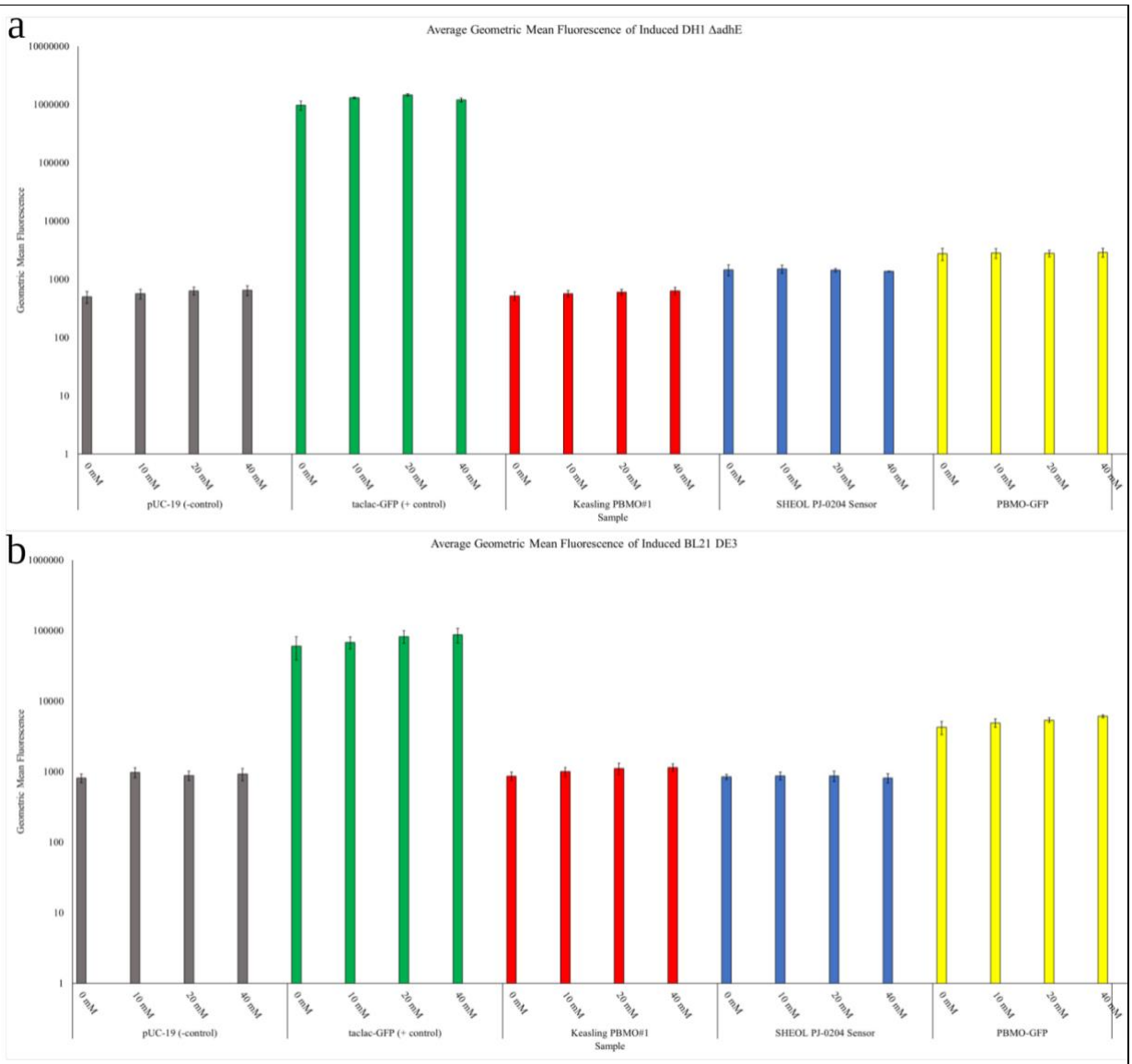


Figure 2-8, flow Bl21 and DH1 ΔadhE: Flow cytometry data from induced DH1 ΔadhE (a) and induced Bl21 DE3 (b). Mean fluorescence is compared against butanol concentrations for all the plasmid/host strain combinations. Fluorescent values are shown in a logarithmic scale and error bars are +/- S.E.M. with n=3.

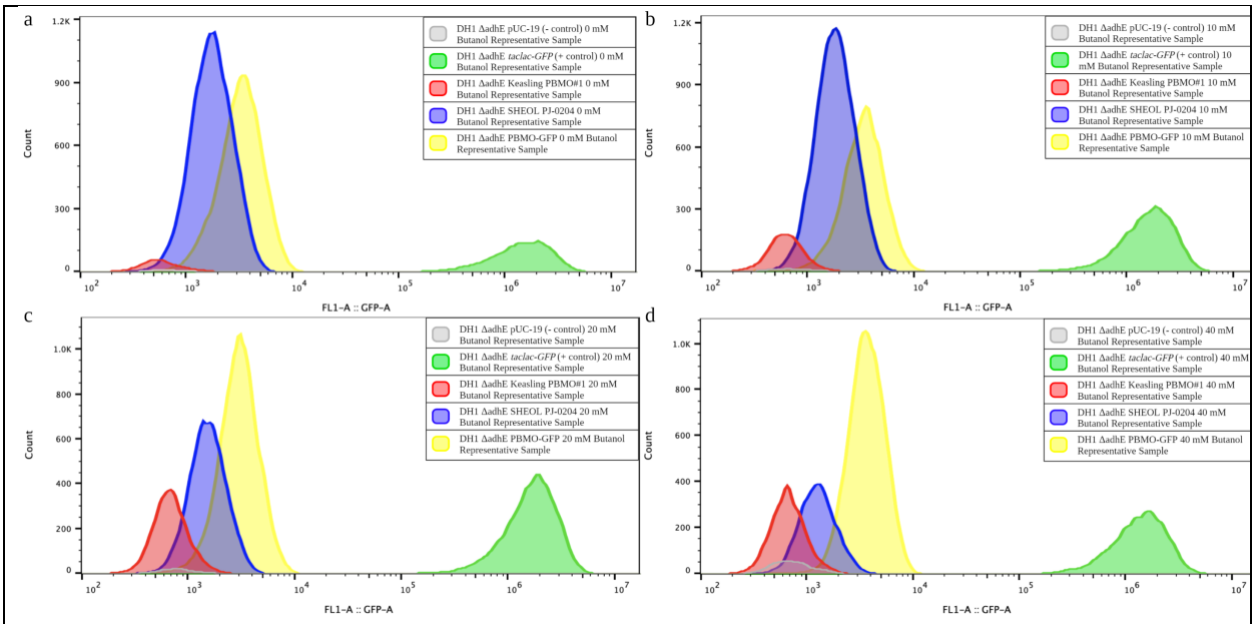


Figure 2-9, DH1 ΔadhE flow: Alternative depiction of flow cytometry data using a representative sample from the DH1 ΔadhE inductions. Number of cells and their detected fluorescence from induction with 0 mM butanol (a), 10 mM butanol (b), 20 mM butanol (c) and 40 mM butanol (d) are shown. The more GFP fluorescence detected, the further right the peak will move, something that would be expected with increasing butanol concentration but is not observed in the biosensors tested.

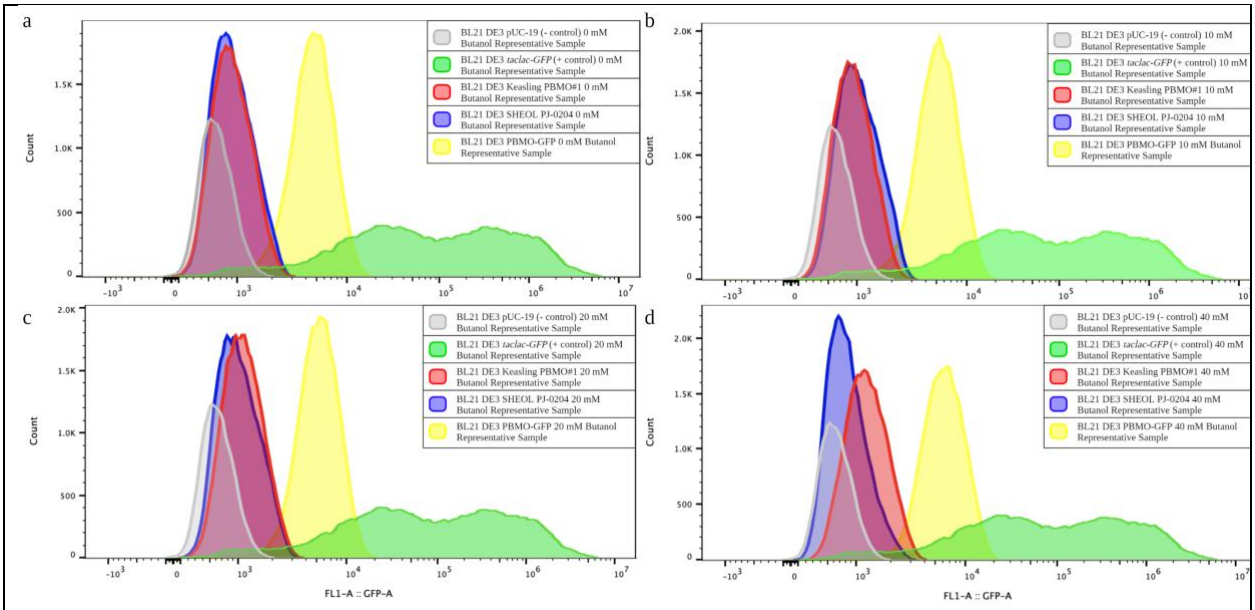


Figure 2-10, Bl21 DE3 flow: Alternative depiction of flow cytometry data using a representative sample from the BL21 DE3 inductions. Number of cells and their detected fluorescence from

induction with 0 mM butanol (a), 10 mM butanol (b), 20 mM butanol (c), and 40 mM butanol (d) are shown. The more GFP fluorescence detected, the further right the peak will move, something that would be expected with increasing butanol concentration but is not observed in the biosensors tested.

Qualitative assessment of butanol circuit induction with western blot.

Western blot was used to detect any production of fluorescent protein caused by butanol induction in the absence of background fluorescence generated by growth media. A western blot was chosen because it detects the presence of the GFP protein itself, not the fluorescence given off by a culture as in previous assays. Cultures used for flow cytometry also provided the samples for western blots. After OD_{600} values were normalized, samples were taken for western blot analysis, centrifuged, and resuspended in 1XSDS containing DTT. Samples were then heated and ran on an SDS gel to separate out proteins in the sample. The proteins in the gel were transferred to a nitrocellulose membrane that was probed using a GFP tagged antibody. Results from the plate reader assay and from the flow cytometry were mirrored in the western blots. No strong induction pattern was observed in any of the sensor constructs (Figure 2-11). Slight GFP protein expression was observed in the PBMO#1 and PBMO-GFP constructs, but these did not increase significantly with increasing butanol.

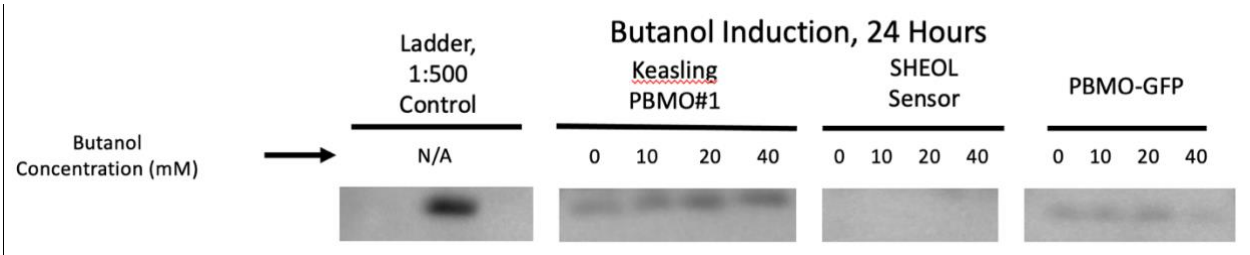


Figure 2-11, butanol biosensor representative western blot: A representative Western Blot showing the amount of GFP produced in each plasmid construction across the four butanol concentrations. The taclac-GFP positive control is diluted by a factor of 500 to prevent overexposure on the rest of the blot when imaging. The blot reinforces conclusions made plate reader and flow cytometry data, that no strong induction pattern is seen in any of the plasmid

constructions and none of the biosensors produce as much GFP as the positive control. Full western blot shown in Appendix B.

The results from the blots strongly suggest that the BmoR protein is not regulating the P_{BMO} promoter. However, the presence of the BmoR regulator itself, and its expression, have not been directly confirmed. Only being able to detect fluorescent protein makes it difficult to troubleshoot the biosensors when no induction is observed. Determining expression levels of BmoR would supply more information about what is happening in the cells, and allow for better hypotheses to be made about how to improve the biosensors.

Building and Testing an Improved Butanol Biosensor

Designing and building a new biosensor construct.

To improve the inducibility and detection capabilities of butanol biosensors, a new construct was designed (Figure 2-3). This biosensor's design was aimed at improving the issues seen in previous induction assays in a few major ways. GFP would be replaced with mCardinal tagged with an HA tag, serving to reduce the background fluorescence observed in the green spectrum. A FLAG tag was added to BmoR so the presence of the protein can be directly measured with western blot. Lastly, the construct included restriction sites for cloning in various promoters driving the production of BmoR (Table 2-2), allowing for the fine tuning of BmoR production. A gBlock was designed to include these changes, however it did not pass the manufacturer's quality and concentration standards so it needed to be PCR amplified (Figure 2-12) before cloning could begin (see Methods section for cloning details).

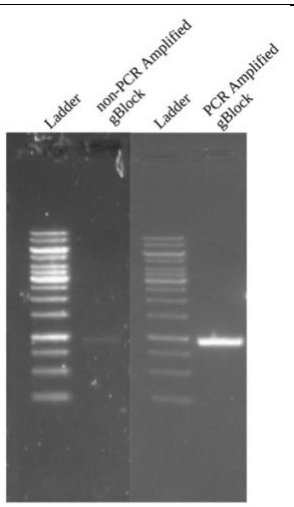


Figure 2-12, PCR amplification of gBlock: The gBlock was determined by IDT to not be up to all their quality control standards. In order to make more of the gBlock and to amplify the correct sequence a PCR was performed. This gel shows the before and after, highlighting the effect that the PCR had on the gBlock. The higher concentration of the gBlock, which will be the cloning insert, the higher chance of success for the cloning ligations.

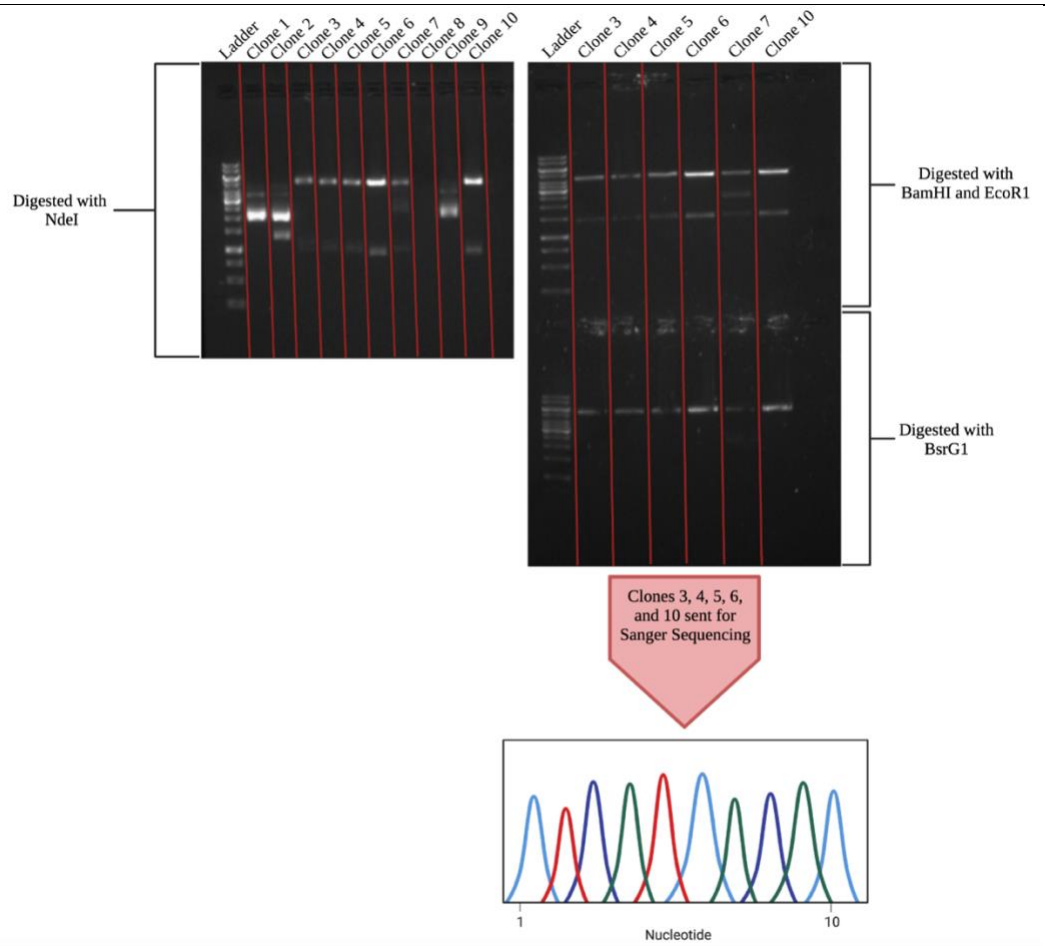


Figure 2-13, restriction pattern of biosensor clones: Restriction patterns from multiple enzymatic digestions of ten selected clone's plasmids. All ten were digested with NdeI, clones 3 through 7 and 10 had the correct restriction pattern and were digested again but with BamHI and EcoRI to test for insert direction. Clones 3 through 6 and 10 were sent for sequencing where it was finally determined that clones 3, 4, and 5 had the correct insert in the correct direction.

Assessment of functionality of new biosensor construct.

Before testing the promoter library, the function of the new biosensor was first examined. It was expected that there would be little fluorescence, as observed with previous inductions, and that any fluorescent protein expression could be detected with the HA tag attached to mCardinal. It was also expected that constitutive expression of BmoR would be observed and, for the first time, be confirmed via western blot probing for the flag tag attached to BmoR. Three of the sequence verified clones were transformed into DH1 Δ*adhE* and BL21 DE3 so a direct comparison

of previous inductions could be made. The clones were induced with either 0 mM or 40 mM butanol for 24 hours and samples were prepared for western blot as before. Western blots, probing for BmoR and mCardinal, as well as a Coomassie stain were performed on these samples (Figure 2-14 and Figure 2-15).

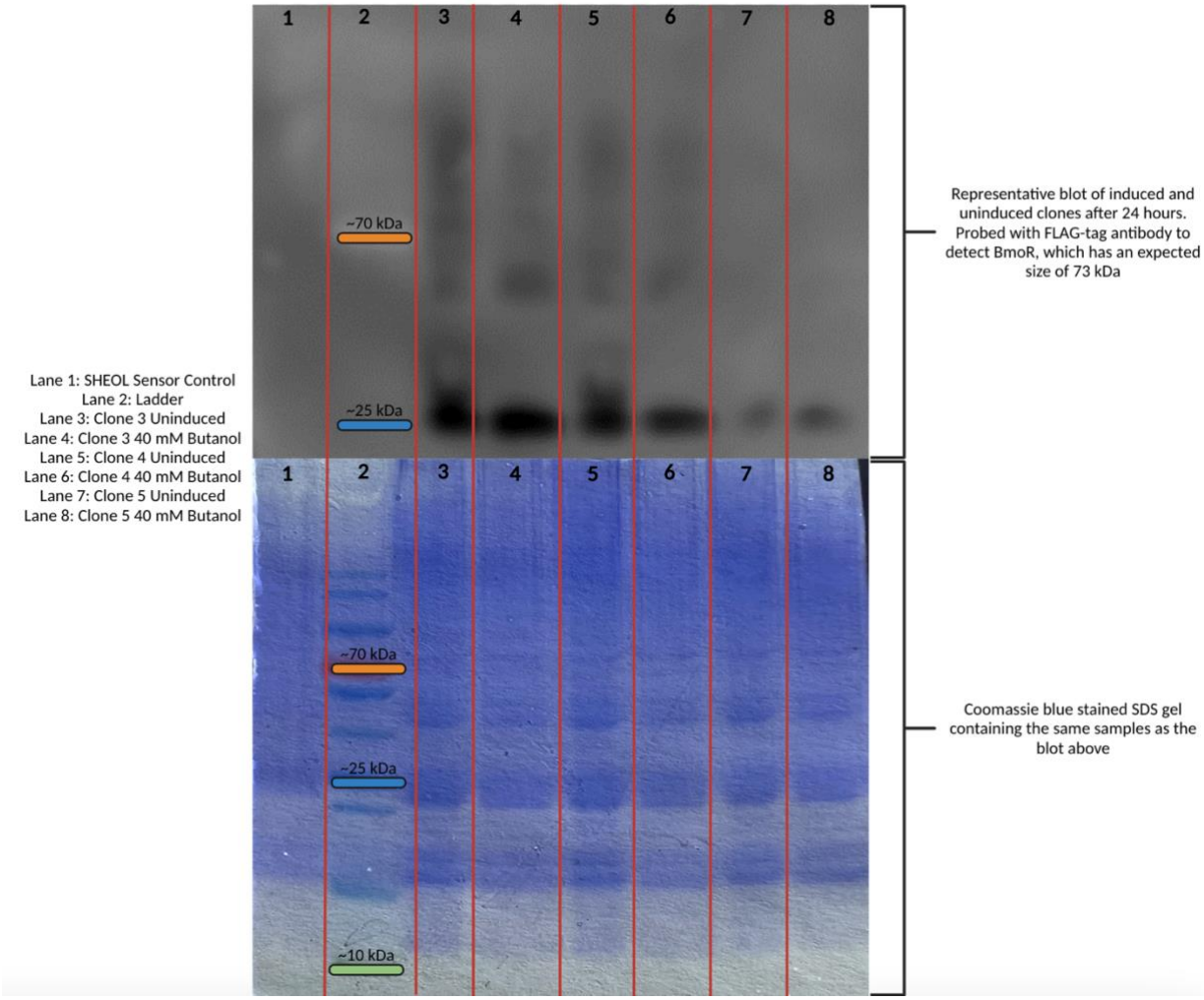


Figure 2-14, BmoR probed western blot and Coomassie stain: Western blot and Coomassie blue stained SDS-Page gels with induced and uninduced clones containing the new biosensor construct after 24 hours. These samples are from an induction of DH1 ΔadhE but are representative of the inductions in the BL21 DE3 strain. The samples run on each of the gels are from the same stock sample. Two complete biological replicates of this experiment were performed.

This gel was expected to contain a clear band at the size of 73kDa, the expected size of BmoR. Instead, the gels shows most antibody binding at ~25kDa, with some less prominent bands at larger sizes. There was no unspecific binding of the FLAG tag antibody in the control lane (Figure 2-14, lane 1), indicating that the binding seen in the western blot represents an accumulation of a degradation product of BmoR, and is not a non-specific signal. There was little to no fluorescence observed on the blot when probing for the HA tag on mCardinal, as seen in Figure 2-15. The fact that any observed fluorescence does not change with butanol concentration indicates that the promoter driving the expression of mCardinal, the P_{BMO} promoter in this case, had some leakiness in its function (Figure 2-15). This observation is consistent with the levels of fluorescence observed in previous inductions using the same promoter. The sum of the experiments performed here suggest that prior published observations of butanol biosensor function using the BmoR system may have been artifacts of plate reader data, and that BmoR is unstable in *E. coli* and does not regulate the promoter P_{BMO}. Given these observations, no further steps were taken to build the promoter library to regulate BmoR.

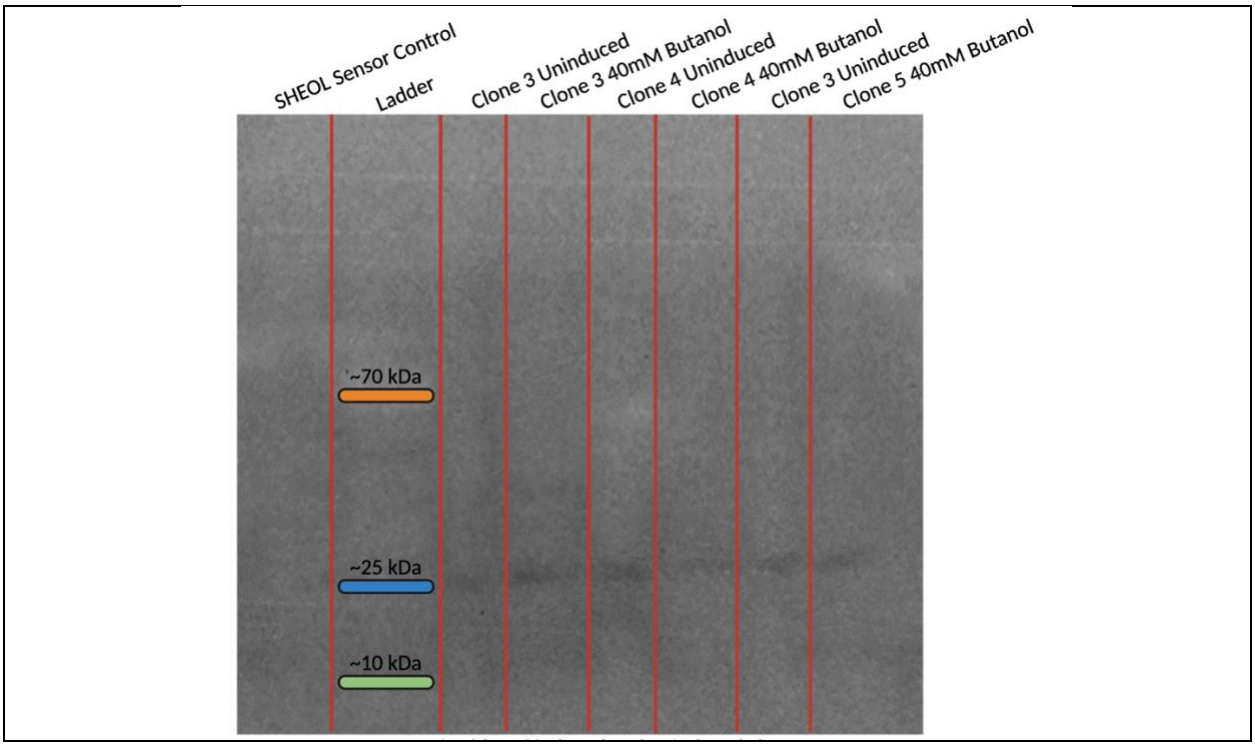


Figure 2-15, mCardinal probed western blot: Representative DH1 ΔadhE blot of induced and uninduced clones after 24 hours. Probed with HA-tag antibody to detect mCardinal, which has a size of 28 kDa.

2.4 Discussion

The aims of this project were to characterize existing butanol biosensors and to make improvements to make the sensors more sensitive and robust. Towards that goal, this work represents the first time that the expression of BmoR was directly measured using western blot. Previous works (Dietrich et al., 2013; Yu et al., 2019) assumed that the BmoR regulator was stably expressed without direct verification. Quantifying BmoR expression has allowed for new insights into how the biosensors function, and additionally, may explain the fluorescent patterns observed. All the versions of the butanol biosensors tested had small amounts of fluorescence regardless of butanol concentration. The fluorescence observed is likely due to the leakiness of the P_{bmo} promoter attached to the reporter instead of the P_{BMO} promoter being induced by butanol bound BmoR. The leakiness is most apparent in the P_{BMO}-*gfp* construct which does not contain a sequence for the BmoR protein. Without BmoR present, there should be no activation of the P_{BMO} promoter, yet the P_{BMO}-*gfp* construct still displays fluorescence observable in all three assays. Leakiness, combined with degradation seen in the western blot probing for tagged BmoR, suggest that BmoR is not functioning as a transcriptional activator in response to butanol. Since BmoR is not stable, it is not able to activate the P_{BMO} promoter to induce the biosensor construct. It is unlikely that the fluorescence pattern observed in the western blot is a result of partial transcription opposed to degradation because BmoR was sequence verified to not contain a premature stop codon. It is possible that the codon optimization of BmoR introduced a cleavage site, as observed when using ExPasy's PeptideCutter tool. If a cleavage site was introduced, one band should be observed on the blot in the size of cut protein attached to the flag tag. However, we observed multiple bands on the blot detecting the flag tag, indicating that the protein is likely being produced and then broken down.

BmoR's degradation has a wider impact on published butanol biosensors as a whole. Failure to validate BmoR expression in previous works means that induction observed in biosensors employing BmoR as a regulator may be a result of background artifacts of plate reader assays. When comparing biosensor constructs against the proper controls and across more robust assays it is apparent that there are no fully validated functional biosensors for butanol in *E. coli*. If reliable inducibility and high fluorescence were observed in these biosensors it is likely that they would be used in industrial applications, such as butanol production via fermentation. Since only low, if any, levels of fluorescence and induction have been observed in research settings, and no further reports of butanol biosensors have recently been published, it is not likely that the optimal butanol biosensor construct has been discovered.

In future studies, to assess the function of BmoR and to characterize its interaction with the P_{BMO} promoter and with butanol, it will first be necessary to find a method to stabilize BmoR. An easy first attempt to generate more protein may be transferring the butanol biosensor constructs to a different strain of *E. coli*. A strain originating from the K-12 line called HMS174 DE3 has been shown to greatly increase product amount over BL21 DE3 when grown in certain medias (Hausjell et al., 2018). Along with attempting expression in a different strain there are a few other techniques that may allow for more stable BmoR production. Reducing the temperature of the growing culture may result in improved protein folding, reduced protease activity, and aversion of occlusion bodies for the target protein (Sørensen & Mortensen, 2005). Another possible avenue is to use a fusion protein which will be expressed alongside the protein of interest. Fusing the target protein to a chaperone protein can help protect the target from proteases while also increasing its solubility (Ryan & Henehan, 2001). If it is not possible to produce stable BmoR protein in bacterial models, it also may instead be possible to produce with a mammalian cell culture model. The advantage of using mammalian cells to produce a target protein has to do with the additional complexity of mammalian cells over bacteria. Mammalian cells are able to perform post translational modifications that allow for increased protein stabilization (Gray, 1997). Mammalian cells also have quality control measures to ensure that incomplete or misfolded proteins are broken down and not secreted (Gray, 1997). A common cell type that has been used as a host for expressing stable proteins, and a good candidate for attempting to express BmoR, is Chinese hamster ovary cells (CHO).

After establishing a reliable way to produce BmoR, the next steps would be to purify it and then to characterize its interactions with the putative promoter region, and with butanol. Purification conditions such as temperature, pH, protease inhibitors, and preservatives can be tuned to yield the greatest amount of stable protein during the purification process (Ragoonanan & Aksan, 2007). To characterize the binding of BmoR to a target DNA region and better define its operator and essential regulatory sequences, a gel shift assay could be performed. Initial experiments were attempted to purify BmoR from *E. coli* using his-tagged proteins, but these attempts were not successful. Knowing now that BmoR is not stable in BL21, purification of BmoR from another strain or species would be necessary to attempt these experiments.

A thermal shift assay is a good assay to test protein stability and butanol binding in a high throughput way (Niesen et al., 2007). A thermal shift assay can be performed in existing qPCR instrumentation with a fluorescent dye which performs similar to that of dyes used in qPCR. Instead of the dye binding to generated nucleic acids as in qPCR, the dye in a thermal shift assay binds to hydrophobic areas of the purified proteins. The samples are then heated, and a fluorescent curve is generated that can be used to calculate melting temperature and protein stability levels similar to how cycle thresholds and expression levels are calculated in qPCR. This can help develop a more robust purification process because many different buffer conditions can be tested at once to determine how to obtain the highest quality protein. A thermal shift assay can also be used to quantify ligand binding interactions with a target protein (Huynh & Partch, 2015). Butanol can be added to BmoR and changes in protein stability or structure that result from binding with BmoR can be quantified in fluorescence output. Being able to determine the conditions BmoR is most stable under and if butanol is binding to BmoR will greatly aid the ability to troubleshoot and improve the biosensor.

Chapter 3, Validating Promoter Function Under Environmental Stresses

3.1 Background

Pseudomonas putida is a common bacteria found in soil and water, and its adaptability along with its vitality under stressful environmental conditions has made it a good chassis organism for expressing a microbial biosensor (Martin-Pascual et al., 2021). The most well studied strain, and the one used in this work, of *P. putida* is KT2440 which has been used in a wide range of research applications. *P. putida* has been used to degrade environmental pollutants because of its ability to catabolize natural or manufactured toxic compounds as a carbon source (Kivisaar, 2020). Its versatile metabolism, along with stress tolerance and ability to sustain redox reactions, also make *P. putida* a viable candidate for industrial biocatalysts (Kivisaar, 2020). The stability of *P. putida* is precisely why it is a good candidate to host a biosensor for stressful conditions such as growth in soil. However, preliminary data from our laboratory has shown that classic well defined inducible biosensor systems, such as the tet and lac operons, do not function when grown in soil (A. Carrillo, personal communication, March 2022). Therefore, understanding what changes occur in the transcriptome of *P. putida* under stressful environments such as soil enable the design of better promoters that function as desired even in soil.

To quantify the changes of *P. putida*'s transcriptome under a range of conditions, a meta-analysis of RNA-seq data was performed (Harding et al., 2021). In this study, transcriptomic data from five separate studies were combined into the same analysis pipeline, shown in Figure 3-1, using a bioinformatics tool called KBase. Briefly, RNA-seq data from studies in which *P. putida* was grown under stressful conditions were combined, sorted for quality, aligned to the reference genome, and an expression matrix containing fragments per kilobase of transcript per million mapped reads (FPKM) expression values was created. FPKM values were used because it is the best calculated value to use when comparing the expression levels between genes in transcriptional analysis (Abbas-Aghababazadeh et al., 2018).

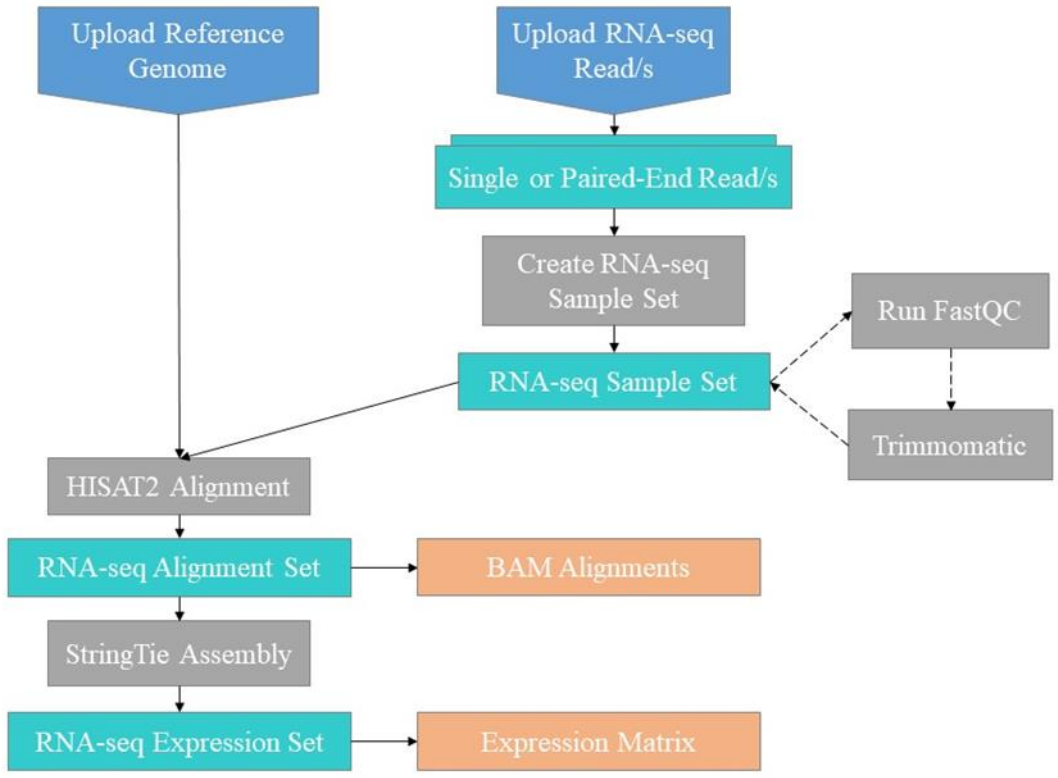


Figure 3-1, transcriptome analysis workflow: RNA sequence data and reference genome are uploaded first, followed by the creation of an RNA-seq Sample Set containing the imported single/paired end reads. The sample set is then checked for read quality using FastQC, where controversial runs may be trimmed by Trimmomatic tool. Combining the reference genome and RNA-seq Sample Set with the HISAT2 Alignment tool producing BAM alignments and an RNA-seq Alignment Set. This alignment set is run using the StringTie Assembly tool producing FPKM and TPM expression matrices. From Harding, 2021.

After completing the transcriptome analysis, five genes were identified to have high expression and low fold change across multiple environmental conditions (Table 3-1). This analysis predicts that the expression of these genes, and the activity of their promoters, is likely to be stable and strong across a broad range of conditions. However, validation of expression levels with qPCR is needed to ensure that the bioinformatic predictions are robust and reproducible. Here we show an attempt at qPCR validation for the bioinformatics observations described in Table 3-1. The issue of internal normalization to housekeeping genes, and its appropriate application when non-laboratory conditions are used, was also explored. After validation, a putative promoter region

of the *infC* gene was identified and cloned into a reporter circuit to see if its function mirrored the qPCR observations. The results and additional investigations of the homologous region in *E. coli* revealed complex transcriptional regulation of *infC* from different promoters. The work presented here provides insights into the complex considerations and analyses required in order to apply transcriptomic observations to the construction of genetic circuits.

Gene names/ locus	Protein names	Glucose	Citrate	Ferulic Acid	Serine	Oleic Acid	Sodium Gluconate	10°C	30°C	Average log2 fold change
infC PP_2466	Translation initiation factor IF-3	12.3218	12.0788	12.2729	12.3812	12.1528	11.8296	13.4897	13.1376	0.04180
atpI uncI PP_5420	ATP synthase protein I	10.2041	10.2204	10.43	10.2846	9.3949	9.8346	10.5233	10.4122	0.04368
galU PP_3821	UTP--glucose-1-phosphate uridylyltransferase (UDP-glucose pyrophosphorylase)	9.4304	9.9838	9.5635	9.4026	8.9248	9.2904	9.1788	8.5536	0.04114
PP_2088	RNA polymerase sigma factor SigX	9.2245	9.6387	9.2692	9.2425	9.1228	9.5478	9.2591	8.6865	0.02404
fabG PP_1914	3-oxoacyl-[acyl-carrier-protein] reductase	9.2175	9.219	8.7856	8.8501	9.0206	9.6118	9.1476	9.0028	0.03845

Table 3-1, meta-data analysis results: Table identifying genes with low fold change and high expression from Harding et al., 2021. *P. putida KT2440 gene and protein nomenclature, FPKM values for potential promoters with high FPKM values and low average log2 fold change. Analysis across growth conditions included different carbon sources (glucose, citrate, ferulic acid, serine, oleic acid, sodium gluconate) and temperatures (10°C, 30°C). FPKM values are color scaled from white (minimum) to green (maximum). Average log2 fold change is calculated with glucose as control, log2 fold change aggregated across all growth conditions. From Harding, 2021.*

3.2 Methods

RNA Extraction and cDNA Synthesis

A glycerol stock of *P. putida* KT2440 was inoculated to liquid LB and grown overnight at 30°C and 220 RPM. The next day the culture was back diluted into fresh LB and allowed to grow to an OD_{600} of 0.5, as measured on an Eppendorf BioPhotometer. 1 mL of culture was taken off

and spun down at least 5000g for 10 minutes. The supernatant was decanted, and cell pellets were frozen at -80°C until ready for RNA extraction using the RNeasy Plus Mini Kit (50) (Qiagen) but with a modified procedure (Hay et al., 2017). An overview of the protocol is shown in Figure 3-2.

RNA Extraction Protocol:

The reagents supplied with the RNeasy kit were used per the modified RNA extraction protocol. An additional required lysozyme reagent that was not included in the kit had to be made. Proteinase K and β-mercaptoethanol were also not included in the kit and needed to be purchased. First, all surfaces and instruments were cleaned with RNA Zap (Invitrogen) prior to RNA extraction. The lysozyme reagent used for lysing the cells was created by dissolving powered lysozyme (ThermoFisher) into TE buffer pH 8.5 (30 mM Tris-HCl and 1 mM EDTA) at a concentration of 15 mg/mL. A cell pellet was removed from the freezer and 200 µL of lysozyme in TE buffer along with 10 µL of Proteinase K (20 mg/mL) (Ambion) were added. The sample was vortexed for 10 seconds and allowed to incubate at room temperature for 10 minutes with frequent vortexing. 700 µL of Qiagen supplied buffer RLT was added along with 7 µL of β-mercaptoethanol (Gibco) and vortexed. The sample was then transferred to a supplied gDNA Elimination Column and centrifuged at >8000g for 30 seconds. 500 µL of 100% ethanol was added to the flow through and the cell lysate was transferred to a supplied RNeasy Mini Spin Column and spun at least 8000g for 30 seconds, discarding the flow through. 700 µL of supplied Buffer RW1 was added to the column and centrifuged again, discarding the flow through. Next, the column membrane was washed twice with supplied Buffer RPE, supplemented with ethanol, by centrifugation. The column was moved to an RNase free collection tube and the RNA was eluted into 50 µL of RNase free H_2O. 1 µL of eluted RNA was taken and run on a 1% agarose gel, the remainder of the RNA sample were be stored at -80°C. To ensure samples do not contain any gDNA a DNase digestion was carried out with TURBO DNase (ThermoFisher) and a slightly modified procedure. A 50 µL final reaction volume was set up using 1 µL of TURBO DNase enzyme, 5 µL of the manufacturer provided 10X TURBO DNase Buffer, and RNA at a final concentration of 100 ng/µL diluted with RNase free H_2O. The reaction was incubated at 37°C for 20 minutes and then the enzyme was inactivated by adding EDTA to a final concentration of 15 mM and heating at 75°C for 10 minutes. Samples were again stored at -80°C for future use.

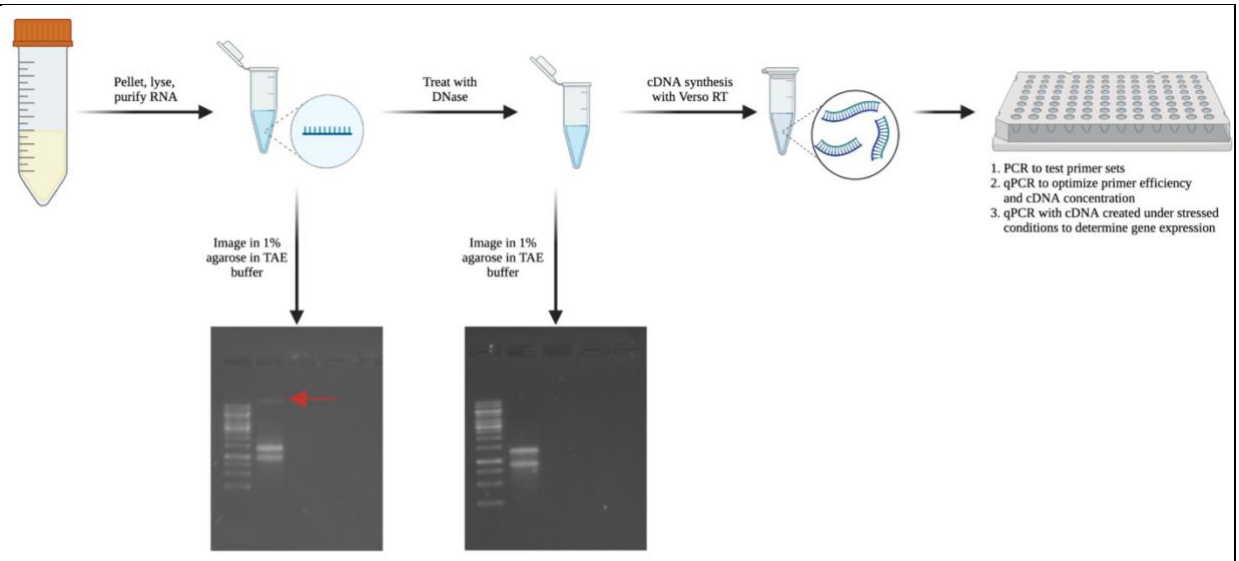


Figure 3-2, RNA extraction overview: Flow chart depicting the process of RNA extraction and cDNA synthesis.

cDNA Synthesis:

RNA samples were thawed on ice and their concentrations were measured on the NanoDrop's RNA setting. The Verso cDNA Synthesis Kit (ThermoFisher) was used, per the manufacturer's protocol, using 7.5 µg of input RNA, and random hexamers for cDNA priming. For each sample of RNA, two cDNA synthesis reactions were prepared, one with the reverse transcriptase (RT) enzyme mixture and one without. After the synthesis reaction, samples that included RT enzymes were stored at -20℃ and samples that did not include RT enzymes were stored at -80℃ to prevent their degradation.

PCR Test of Primer Sets

Primer sets were ordered for the genes of interest, genes identified to have up and down regulation in cold stress (D'Arrigo et al., 2019), and a control gene (Q. Chang et al., 2009). Primers were ordered as oligos from IDT and resuspended in molecular biology water as a stock solution with a concentration of 100 µM. Primers used in this study are listed in Table 3-2.

Table 3-2, qPCR primers

Promoter Validation Gene Primers				
Gene	Locus	Function	Sequence (5'-3')	Source
Genes with Expected Consistent Expression Under Stress				
sigX	PP_2088	RNA polymerase sigma factor sigX	F: AGAGGCGTCCGAAGACAAGGCT R: ACGCAGCACCAGAATTTCCCGG	This Work
aptI	PP_5420	ATP synthase protein 1	F: CCTTTCCATCGCTGGGCGGTTT R: GGGCAGCCAGGCAATCAAACCT	This Work
fabG	PP_1914	3-oxoacyl-[acyl-carrier-protein] reductase	F: TCGTTGGTGCCATGGGTAACGC R: CGGTGATACCACGCGAACCCAC	This Work
galU	PP_3821	UTP-glucose-1 phosphate uridylyltransferase	F: ACGCGCCAGACCGAAATGAAGG R: CTTCCGGGTTGACGCACAGGTC	This Work
infC	PP_2466	Translation initiation factor IF3	F: ACCGATGGGGACAAGGCCAAGA R: GCTTCGACCCGCTTCAACAGCT	This Work
Genes with Known Cold Stress Up Regulation				
pprA	PP_0185	LytTR family two component trx regulator	F: CTTGCGCAAGGCCGAAAAACCC R: TCAGTTCGATGCCTTTGCGCGT	This Work
N/A	PP_1619	Hypothetical protein	F: ACGCAAGGGCAGCTTCAGCTTT R: CTGCTCGTGGTACAGGCGTTCG	This Work
Genes with Known Cold Stress Down Regulation				
hptG	PP_4179	Heat Shock Protein 90	F: ACAACTCGCTGCTGTACGTGCC R: AACGACTCGGCCTGGTCCATGA	This Work
N/A	PP_5232	Hypothetical protein	F: AATTCCTTTCTCGCGGCGGCAA R: GCGGCTGCCGTACTTGTTGTCT	This Work
Internal Control Genes				
RpoD	PP_0191	RNA polymerase sigma factor *RpoD*	F: GAGATCAACCCACGGATCAACGACA R: TCATGCAACAACCCGCCCAAT	(Q. Chang et al., 2009)

A description of all the genes and their primers used in qPCR experiments.

A working solution was made by diluting the stock by a factor of 10 to a concentration of 10 µM. Two PCR reactions were prepped on ice for each of the primer sets, one reaction using the cDNA and the other using the no RT control. The polymerase used was OneTaq Quick-Load 2X Master Mix with Standard Buffer (New England Biolabs (NEB), Ipswich, MA) and reactions were prepared according to NEB's specifications. Reactions were placed in an Eppendorf Mastercycler X50s and were cycled according to the polymerase specifications. After the PCR reaction was completed, each sample was run on a 1% agarose gel.

qPCR Primer Set and cDNA Concentration Test

Having confirmed that the primer sets amplified their target genes, and that the cDNA was not contaminated with genomic DNA, primer efficiency was then assessed. A series of cDNA dilutions was created by diluting the cDNA stock to 100 pg/µL, 50 pg/µL, 20 pg/µL, and 5 pg/µL. A qPCR experiment was carried out that combines each dilution concentration with each primer set (three replicates for each primer set/cDNA concentration pair), being sure to also include no template controls (NT) and no reverse transcriptase controls (NRT). A new working solution for the primers was created that contains each primer set at a concentration of 5 µM. The polymerase used was PowerTrack SYBR Green Master Mix (Applied Biosystems). Each reaction consists of 1 µL of primer mix, 5 µL polymerase, 2 µL H_2O, and 2 µL template. The reactions were prepared on ice in a MicroAmp Optical 96-Well 0.2 mL Reaction Plate (Applied Biosystems) and sealed with MicroAmp Optical Adhesive Film (Applied Biosystems). The qPCR plate was run on ThermoFisher's QuantStudio 6 qPCR System using settings for SYBR reagents and a standard curve assay. The data was adjusted for C_T threshold and baseline cycle parameters then exported to Excel for further analysis. Primer efficiency was calculated by plotting the C_T vs. $\log_{10}$(cDNA concentration) and reporting the slope of the line of best fit. The efficiency of each primer set was equal to $10^{(-1/\text{slope})-1}$.

KT2440 WT Stressed Condition Validation

Temperature Stress:

Cold stress was applied to *P. putida* KT2440 using a method adapted from Frank et al., 2011. A glycerol stock of wild type *P. putida* was sampled and incubated in liquid LB overnight at 30°C and 220 RPM. 500 µL of culture was back diluted into 20 mL of fresh media. This media was minimal M9 broth (Teknova) supplemented with 15 mM sodium succinate dibasic hexahydrate (Sigma Aldrich). The culture was grown at 30°C and 220 RPM until mid-exponential phase, an OD_{600} of 0.8, was reached. Once achieved, three 1 mL samples with an OD_{600} reading of 0.5 worth of culture were pulled off, centrifuged, and had their supernatant removed. Three separate pellets were saved at each time point as backups in case the RNA from one pellet was

degraded. The pellets were stored at -80℃ until RNA extraction could be completed. The remaining culture was moved to an incubator set at 10℃ and allowed to incubate for 2 hours. Three more samples were pulled off, pelleted, and frozen for later RNA extraction. The frozen pellets had their RNA extracted and cDNA synthesized using the same procedure as before. A qPCR experiment was run using the synthesized cDNA at a concentration of 80 pg/µL and the no RT control, this time ensuring 4 replicates of each primer set per temperature condition. The qPCR experiments were performed using 0.2 mL MicroAmp Optical 96-Well Reaction Plates (Applied Biosystems) sealed with MicroAmp Optical Adhesive Film (Applied Biosystems) and run on QuantStudio 6 (ThermoFisher). This procedure was repeated three times and after each trial the raw C_T data was exported to excel and analyzed using the $2^{-\Delta\Delta C_T}$ method (Livak & Schmittgen, 2001).

Soil Extract Stress:

A glycerol stock of wild type *P. putida* was sampled and incubated in liquid LB overnight at 30℃ and 220 RPM. 500 µL of culture was back diluted into 20 mL of fresh LB. The culture was allowed to grow until an OD_{600} of 0.8 was reached. Once achieved, three 1 mL samples with an OD_{600} reading of 0.5 worth of culture were collected, centrifuged, and the supernatant removed. The pellets were stored at -80℃ until RNA extraction could be completed. The rest of the culture was centrifuged at 3000g for 5 minutes and resuspended in sterile soil-extracted solubilized organic and inorganic matter (SESOM). The culture was then allowed to incubate for 2 hours at 30℃ and 220 RPM, and additional samples were collected. The remaining culture was then allowed to incubate at 30℃ and 220 RPM for 24 hours. After 24 hours three final samples were taken from the culture and prepped for RNA extraction. RNA extraction, cDNA synthesis, and qPCR on the QuantStudio 6 were performed as described above. Upon completion of qPCR experiments a one-way ANOVA test was performed to determine the significance of the data across genes and conditions.

Testing *thrS* and Cloning *infC*

thrS Characterization and qPCR:

The gene coding for *infC* was determined to be the candidate most likely for success for conducting a reporter assay (Harding et al., 2021). The locus around the *infC* gene is shown in Figure 3-3 below.

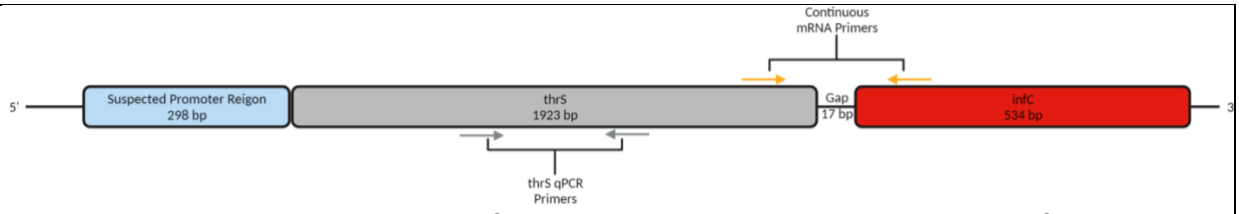


Figure 3-3, genetic locus surrounding infC: Relevant genetic region surrounding infC, including general positions of primer sets ordered.

Separating *infC* from the region suspected of being its promoter is a gene called *thrS* and a small gap of 17 base pairs. A primer set was ordered in which the forward primer was inside of *thrS*, and the reverse primer was inside *infC*, meaning any amplification would have to span the gap between the two genes. A reverse transcriptase PCR (RT-PCR) using DreamTaq Hot Start Green PCR Master Mix (Thermo Scientific) with this primer set and leftover LB control cDNA as the template was carried out, using the manufacturer's specified time and temperature conditions. The PCR product was then run on an agarose gel to check that *thrS* and *infC* are in fact on the same mRNA strand. To ensure that their expression levels were the same more qPCR experiments would need to be carried out. qPCR primer sets were ordered from IDT for *thrS*, and their efficiency was tested same as previously. Using cDNA created from the soil extract experiments, data was generated for *thrS* expression levels across the three trials.

infC Promoter Cloning:

A gBlock was ordered which contained the suspected promoter region followed by Flag tagged mCardinal codon optimized for *P. putida*. The gBlock would be cloned into the pJH0204 plasmid (Figure 3-4) constructed by Dr. Andres Felipe Carrillo, using FastDigest BamHI and XhoI.

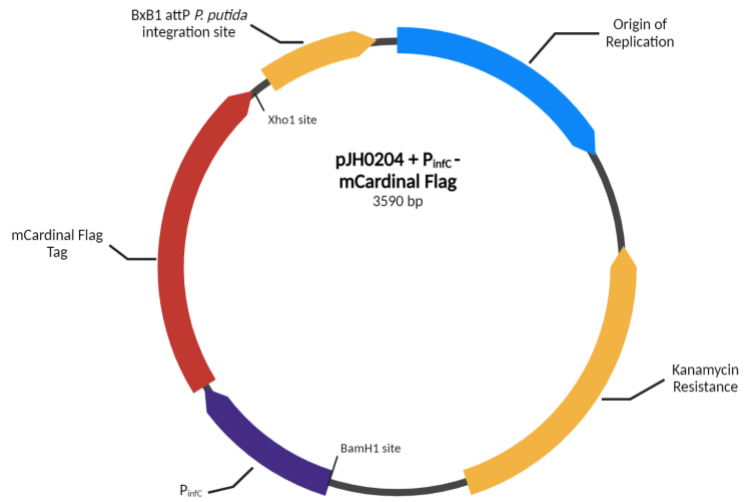


Figure 3-4, PinfC reporter construct: Constructed plasmid containing the promoter for infC attached to mCardinal as a reporter.

The Rapid DNA Dephosphorylation and Ligation Kit (Roche) was used for cloning, following the manufacturer's instructions. The ligation product was transformed into SIG10 Chemically Competent Cells and plated to LB agar plates containing kanamycin and left to incubate at 37°C overnight. Five colonies were picked the following day, miniprepped and then screened with BamHI and XhoI to test for correct insert size. All five clones were determined to have the correct restriction digest pattern and therefore were sent for sequencing to confirm the sequence was correct. The plasmid then needs to be integrated into the genome of *P. putida* so reporter expression could be compared to that of the temperature and soil extract stress experiments. When transforming into *P. putida* two plasmids need to be used, one is the target plasmid which needs to be integrated and the other is a plasmid containing the integrase needed to place the target plasmid into *P. putida's* genome (Elmore et al., 2017). A 50 µL stock of electrocompetent *P. putida* cells were thawed on ice and 1 µL of the pJH0204 + P_{infC}-mCardinal Flag target plasmid with a concentration of 72 ng/µL was added along with 1 µL of pGW31 integrase plasmid with a concentration of 66 ng/µL. The solution was transferred to a 0.1 cm electroporation cuvette and tamped down against the bench to release any air bubbles. The cuvette was inserted into an Eppendorf Eporator and electroporated at 1.6 kV. Immediately after electroporation, 250 µL of S.O.C. medium (Invitrogen) was used to resuspend the cells inside the cuvette. The resuspension was transferred to a microcentrifuge tube and incubated at 30°C for 90 minutes. After incubation, 10 µL was taken out and plated onto a pre-warmed LB plate containing

kanamycin. The remaining volume was plated to a separate LB plate also containing kanamycin. The plates were left to incubate overnight at 30℃ and 220 RPM. The following day a colony was picked and placed in an overnight culture so that a glycerol stock could be made for storage at -80℃.

Testing the pJH0204 + P_{infC}-mCardinal Flag Plasmid

To confirm the plasmid has been integrated into the genome of *P. putida* a culture PCR was performed. A culture containing the plasmid construct was grown overnight at 30℃ and 220 RPM and the next morning a 1:100 dilution of the culture was made. 1 µL of culture was suspended in 99 µL of dH_2O and heated between 90℃ and 100℃ for 5 minutes to lyse the cells. 1 µL of the cell lysis was then used as the template for a PCR using DreamTaq polymerase and a primer set to amplify the synthetic mCardinal sequence in the integrated plasmid. Two controls were included in this PCR, the negative control was a culture of WT KT2440, and the positive control was a culture of pJH0204 + *taclac*-mCardinal, a construct created by Dr. Andres Felipe Carrillo proven to have powerful constitutive mCardinal production. Both control samples were prepared in the same way as the sample containing the *infC* promoter. The PCR products were run on an agarose gel to check for amplification. In addition to the PCR, a plate reader assay was also completed. The overnight cultures of the three constructs were diluted down to an OD_{600} of 0.05 on an Eppendorf BioPhotometer in LB, M9 supplemented with 15 mM succinate, and SESOM. 200 µL of each strain/media combination was plated in triplicate on Costar 96 well flat bottom plate and covered with an optically clear cover and sealed with parafilm to prevent evaporation. The plate was placed in the BioTek Synergy H1 plate reader and the instrument was set to incubate the plate at 30℃ with shaking while taking a reading for OD_{600}nm and a reading for fluorescence set at an excitation wavelength of 604 nm and an emission wavelength of 659 nm every hour for 24 hours. Data from the plate reader was then exported to Excel for fluorescence per OD calculations. This plate reader assay was repeated but with the constructs in *E. coli* and the plate reader set to 37℃ taking readings for 18 hours to test for fluorescence in *E. coli* in addition to *P. putida*. The final test to diagnose any issues with the pJH0204 + P_{infC}-mCardinal Flag plasmid was an RT-PCR which will check if mCardinal was being transcribed. Overnight cultures of the three constructs in *P. putida* were back diluted into LB and grown until an OD_{600}nm of 0.5 was reached. Samples

were taken, frozen at -80°C, RNA was extracted and DNase treated, and cDNA was synthesized using the same procedure as previously. A PCR was performed with DreamTaq polymerase and promoters that will amplify mCardinal. The PCR product was again run on an agarose gel to check for amplification.

3.3 Results

Preparation for Genetic Expression Measurements with qPCR

Primer sets and qPCR amplification conditions were first optimized before testing the genetic expression levels of *Pseudomonas putida* genes under stressed conditions. This includes ensuring primer sets amplify their target gene, calculating primer efficiency, and determining what cDNA concentration to use. RNA from *P. Putida* KT2440 grown under ideal conditions was extracted and treated with DNase using an RNA extraction kit with a modified procedure (Hay et al., 2017). Upon imaging the RNA on an agarose gel, only two bands should be seen which represent the 16S and 23S ribosomal RNA. The presence of a larger band would suggest gDNA contamination (Figure 3-2). Next, cDNA was synthesized from the RNA, including controls where no reverse transcriptase was added. A non-quantitative RT-PCR was performed with the cDNA and primers for target genes to ensure that the primers produced a single amplicon of the appropriate size (Figure 3-5). If no band was observed from the reaction containing cDNA template, a new primer set was ordered, and the PCR test was repeated This step also served to confirm that there is no gDNA contamination in the cDNA, as no amplification was observed in the no RT controls.

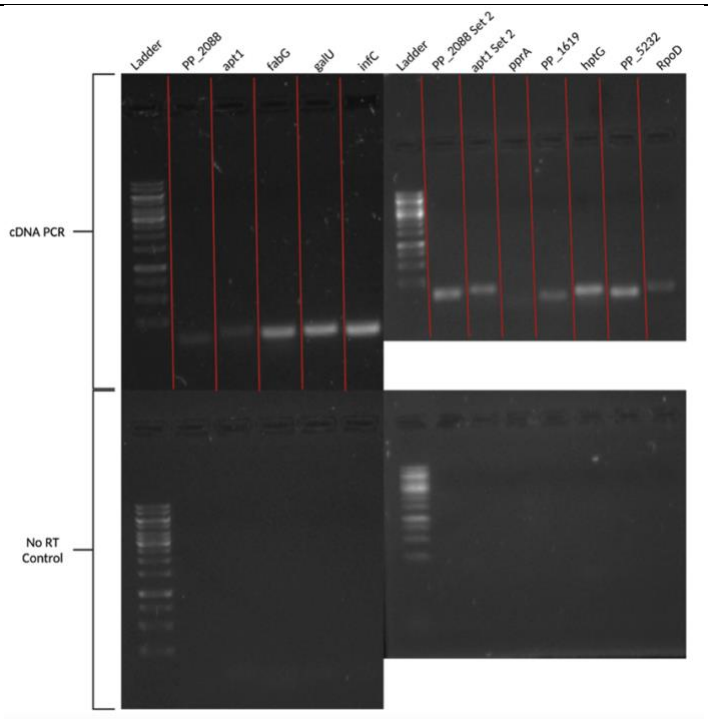


Figure 3-5, RT-PCR of KT2440 with qPCR primer sets: Amplification of target gene cDNA by standard RT-PCR. The lack of amplicons observed in the no RT control confirm the lack of gDNA contamination. Primer sets with no amplicon were discarded and a new primer set for the gene was ordered that did produce an amplicon.

Having confirmed that cDNA samples are free of gDNA and that the primers amplify their target, primer efficiency can be calculated. Dilutions of cDNA were combined with each primer set and SYBR Green qPCR master mix. Primer efficiency can be calculated by graphing the generated C_T values against logarithm of cDNA concentration and taking the slope of the line generated. Ideally, primer set efficiencies would lay between 90% and 100% efficiency. Table 3-3 summarizes the calculated efficiency of the primer sets. Since some of the calculated efficiencies obtained lay outside the optimal range, that decreased efficiency needs to considered when analyzing qPCR results in later sections. Given more time it would have been beneficial to test more primer sets for each gene to find primer sets that all functioned at a similar high efficiency.

Table 3-3, qPCR primer efficiency

Gene	cDNA Concentration	Log_{10}(cDNA) Concentration	Average C_T	Slope	Efficiency
PP_2088	100	2	24.4937344	-3.5291	0.92

	50	1.69897	25.6003396		
	20	1.30103	26.2618574		
	5	0.698970004	29.2209955		
ATP1	100	2	22.9851364	-3.9742	0.78
	50	1.69897	23.9678121		
	20	1.30103	25.4786884		
	5	0.698970004	28.1234575		
infC	100	2	19.2203232	-3.7525	0.85
	50	1.69897	20.0265382		
	20	1.30103	21.707563		
	5	0.698970004	24.0033695		
thrS	100	2	22.8395479	-4.082	0.76
	50	1.69897	24.0388764		
	20	1.30103	25.5546978		
	5	0.698970004	28.162784		
PP_1619	100	2	26.2329386	-4.0518	0.77
	50	1.69897	26.4395975		
	20	1.30103	28.5410685		
	5	0.698970004	31.2109242		
PP_5232	100	2	23.6714642	-3.1323	1.08
	50	1.69897	24.1454554		
	20	1.30103	25.5536998		
	5	0.698970004	27.6218864		
RpoD	100	2	24.618359	-3.7583	0.85
	50	1.69897	25.3378281		
	20	1.30103	26.6610512		
	5	0.698970004	29.4519559		

Summary of genes and their primer set's calculated efficiency determined by qPCR. Average C_T values are the average of three technical replicates.

Genetic Expression of KT2440 WT Under Stressed Conditions

Temperature Stress

Determining genetic expression of target genes under temperature stress was necessary to validate bioinformatic findings preceding this project (Harding et al., 2021). Wild type *P. putida* KT2440 were exposed to cold temperature of 10°C following the procedure performed by Frank

et al., 2011. Samples for RNA extraction were taken before and after cold exposure to measure the difference in expression between the two conditions. RNA was extracted, treated with DNase, and cDNA was synthesized as before. After qPCR, raw C_T values were analyzed using the $2^{-\Delta\Delta C_T}$ method (Livak & Schmittgen, 2001). As Kyle Harding's MQP had shown, these target genes were expected to have very small fold changes when comparing stressed to non-stressed conditions. In these results a small fold change, relative to the *RpoD* control, would mean a value approaching 1. It was also expected that PP_1619 would see a relative fold change greater than 1 as it is a gene thought to be upregulated in a cold environment (D'Arrigo et al., 2016). Alternatively, PP_5232 was expected to have a fold change less than 1 because it should be downregulated in a cold environment (D'Arrigo et al., 2016). The results from the temperature stress qPCRs, displayed in Figure 3-6, show that all the genes have a relative fold change much less than 1, indicating downregulation in the cold condition.

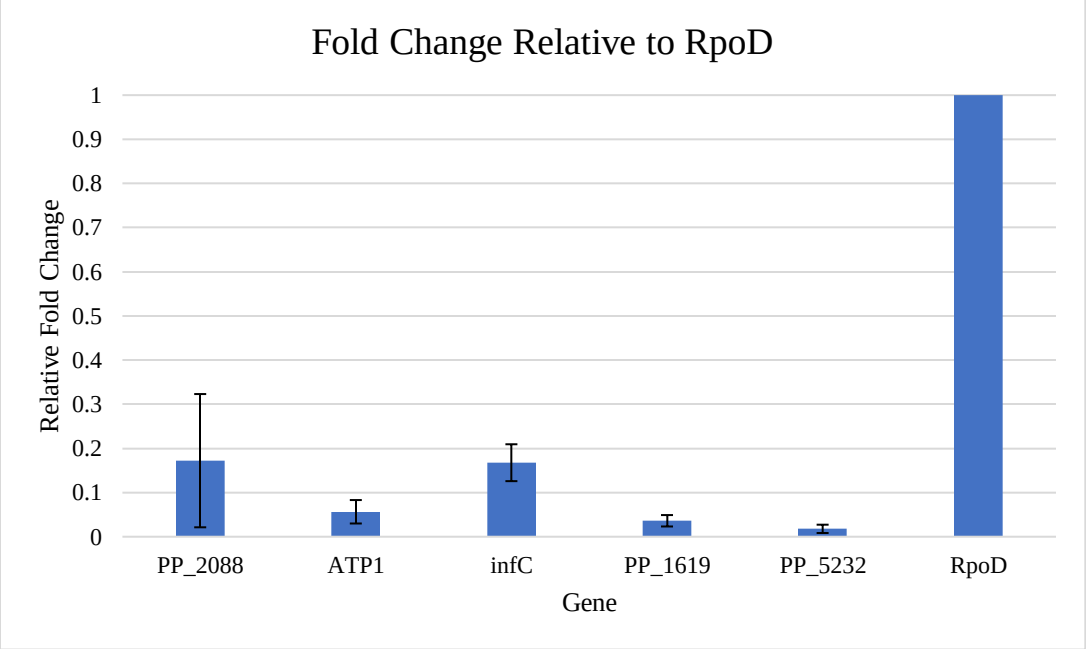


Figure 3-6, temperature stress qPCR results, RpoD control: Fold change, relative to RpoD, of genes of interest averaged across three trials of temperature stress. Error bars are +/- S.E.M. with n=3.

When looking at the raw C_T values, included in the Appendix A, a possible explanation appears. *RpoD* did not have consistent expression across the temperature stress, so when all the

other C_T values are normalized to *RpoD*'s variable C_T values, they become skewed. Although *RpoD* has been used as a qPCR control in *P. putida* (Q. Chang et al., 2009) it has not previously been used as a control for temperature stresses. It may be that the expression of *RpoD* is not consistent enough to be used as a control when considering temperature as a stress. In addition to performing $2^{-\Delta\Delta C_T}$ calculations with *RpoD* as the control, calculations using *infC* as the control were also performed (Figure 3-7). *InfC* had the smallest ΔC_T values across the raw C_T data (Appendix A), making it a potential candidate to normalize the other genes against. When analyzing the data again using *infC* as the control, the data shifts slightly more to what is expected although all the genes of interest still have a relative fold change less than one. After performing a one-way ANOVA test for the temperature stress data where *RpoD* was the control no significance was observed between the non-control genes (p=0.416). When changing the control gene to *infC* there is a significant difference when comparing non-control PP_2088 (p=0.003), *ATP1* (p=0.002), PP_1619 (0.002), and PP_5232 (p=0.001) to *RpoD* in the post hoc. The significance values from each experiment my be skewed because of the relatively high error and the differing primer efficiencies.

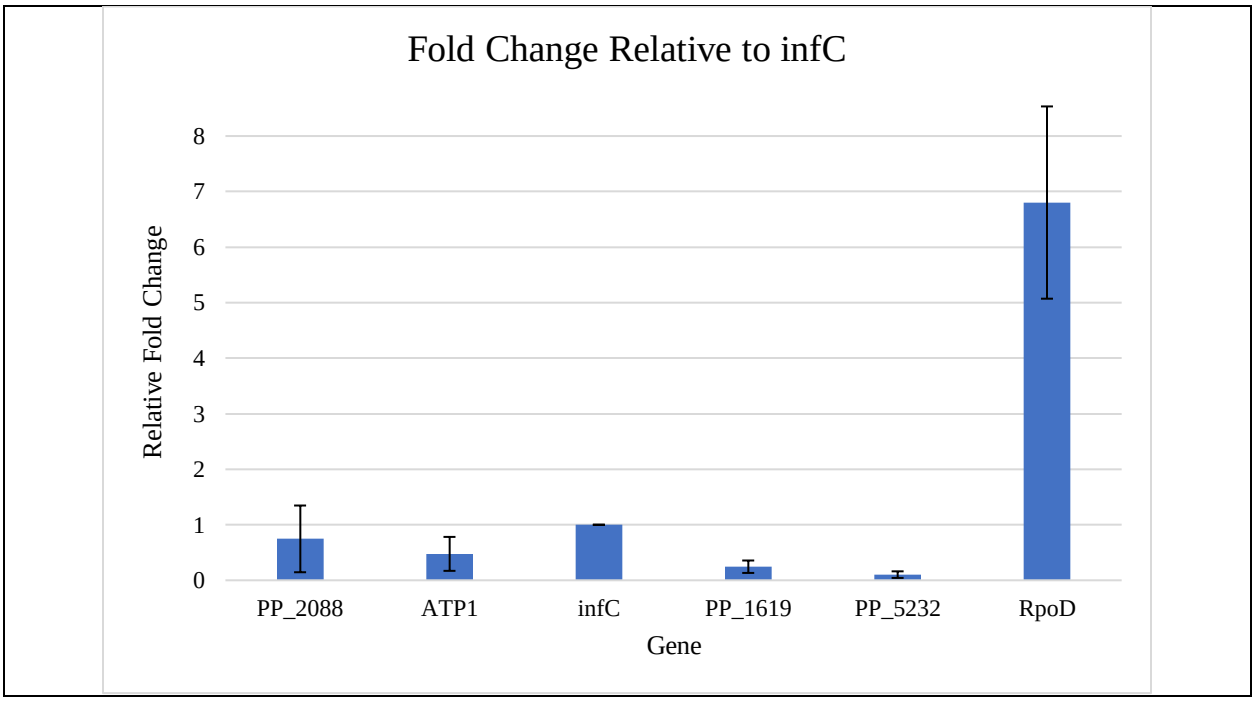


Figure 3-7, temperature stress qPCR results, infC control: Fold change, relative to infC, of genes of interest averaged across three trials of temperature stress. Error bars are +/- S.E.M. with n=3.

Soil Extract Stress:

Soil extract was not among the conditions identified in which the target genes have high expression and low fold change. However, it was hypothesized that growth in soil extract would be a stressful condition, and therefore genes such as *infC* that remain stable under other stress conditions may also remain stable during growth in soil. In addition to temperature stress a series of qPCR experiments were performed using growth in SESOM (solubilized extract of soil organic matter) (Vilain et al., 2006) after 2 hours and after 24 hours. Unlike the temperature stress, *RpoD* expression remained consistent with SESOM as the stress. Similar to the results seen in the temperature stress, the genes of interest were being expressed in smaller amounts than expected (Figure 3-8). This is even more apparent after 24 hours in SESOM, where every gene's relative fold change drops greatly. These findings may echo other findings that genes identified through bioinformatic analysis of FPKM values may not necessarily match results obtained by measuring expression with qPCR, a challenge that is discussed further in the Discussion Chapter.

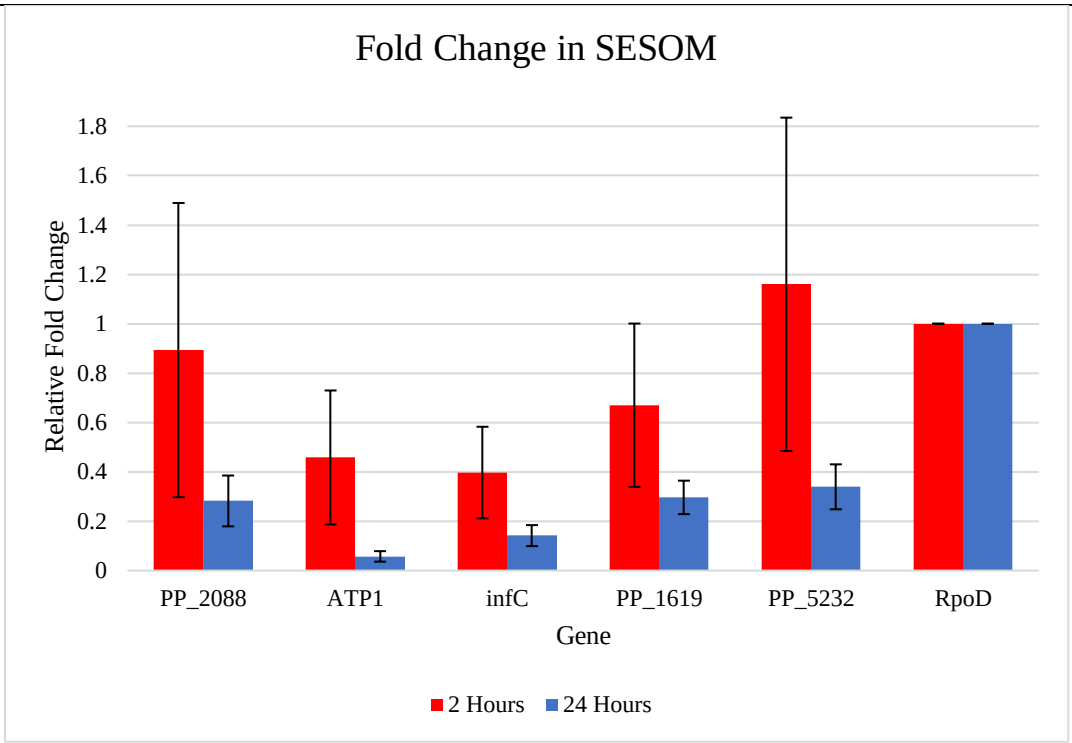


Figure 3-8, SESOM stress qPCR results: Fold change, relative to RpoD, of genes of interest averaged across three trials of growth in SESOM. Error bars are +/- S.E.M. with n=3

Identification and Testing of a Possible Promoter

Identifying a promoter for infC

The next goal was to create a fluorescent reporter construct that behaved as predicted by the qPCR experiments, under a variety of stress conditions. The hypothesis was that promoter activity accounted for the observed changes in gene expression, and therefore applying the promoter to a reporter gene would create a genetic circuit with predictable function. Therefore, to further characterize expression, the suspected promoter region of *infC* was attached to a fluorescent (mCardinal) reporter. *InfC* was chosen because it had the most consistent C_T values under temperature stress and thus best validated the initial bioinformatic analysis (Table 3-1). Between *infC* and the region suspected of being its promoter is another gene called *thrS*. Many genes are bi- or multi-cistronic in bacteria. To test whether *infC* and *thrS* are on the same mRNA strand, and therefore could be controlled by the same promoter, an RT-PCR analysis was performed with primers spanning the *thrS* and *infC* open reading frames (Figure 3-3).

Figure 3-9, thrS-infC RT-PCR: Amplicons produced by an RT-PCR, lanes two and three are testing to make sure that the qPCR primer sets ordered for thrS amplify as intended. Lane four is the amplicon produced verifying that thrS and infC are on the same mRNA strand.

The results of Figure3-9 confirm that *thrS* and *infC* were on the same mRNA strand. Then qPCR was performed using the same cDNA stocks as the SESOM stress experiments to confirm that the expression levels of *thrS* are similar to *infC*, as would be expected if they are on the same mRNA strand. *ThrS*'s primer efficiency is included in the Table 3-2. In both the 2 hour ($p=0.761$) and the 24 hour ($p=0.949$) there was no significant difference in the expression values of *infC* and *thrS*, supporting the idea that these genes are on the same mRNA. Inquiry into possible explanations for the large error in the results is included in the Discussion Chapter.

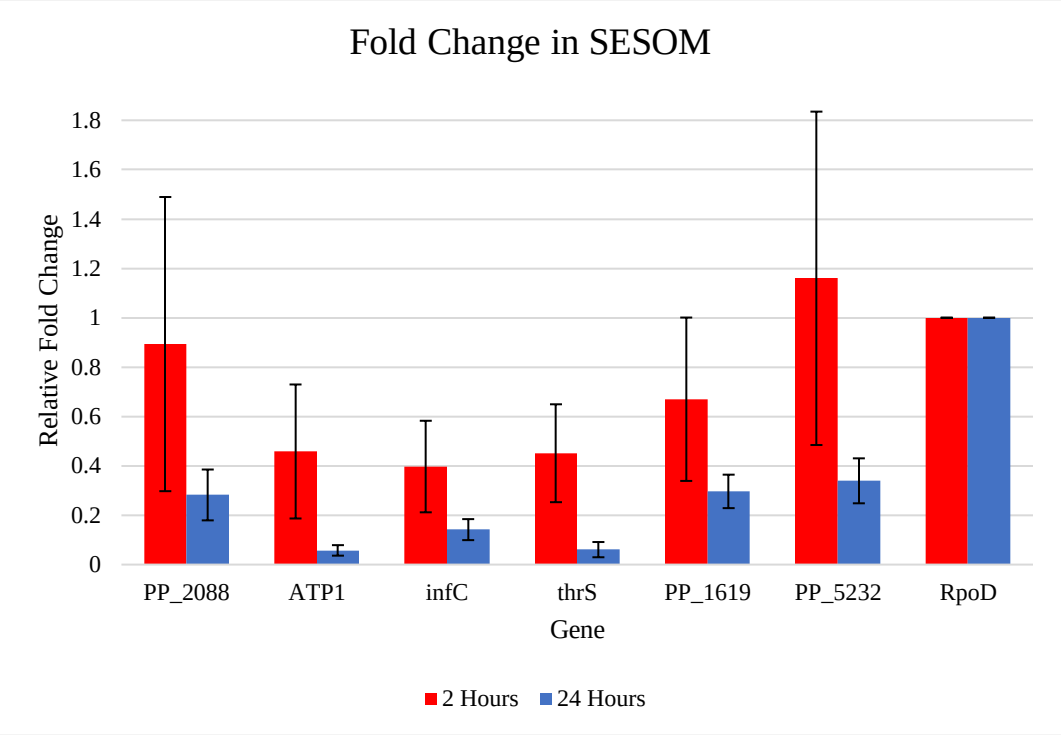


Figure 3-10, SESOM stress qPCR results, including thrS: Fold change, relative to RpoD, of genes of interest averaged across three trials of growth in SESOM including expression levels of thrS. Error bars are +/- S.E.M. with n=3

Building and testing a reporter construct using the promoter region of infC

To test if the suspected promoter region for *infC* would be a useful promoter for future genetic engineering applications, it was attached to a reporter for analysis. The suspected promoter region was cloned into the existing pJH0204 plasmid with a mCardinal as the reporter. After transformation, 5 colonies were picked for verification with restriction digests and sequencing. All the colonies picked and miniprepped had the correct restriction digest pattern, as seen in Figure 3-11 and one of the plasmids with the correct sequence was used for integration into *P. putida* for testing.

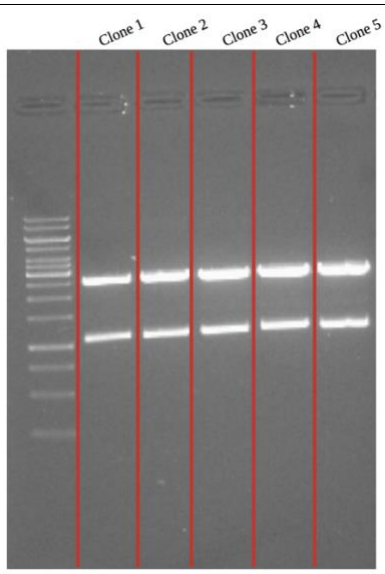


Figure 3-11, digested pJH0204 + PinfC-mCardinal plasmid: Restriction pattern of clones digested with BamHI and XhoI. Clone 1's plasmid had the correct sequence and was eventually integrated into P. putida.

To ensure that the plasmid was correctly integrated into the genome of *P. putida* a culture PCR was performed with primers to amplify the sequence of mCardinal which is not present in the wild type parent strain. Also included in this PCR was a positive control construct known to have strong mCardinal expression, and negative control of the wild type parent strain KT2440. The results of running the PCR product on an agarose gel confirm that the plasmid containing the *infC* promoter and the mCardinal reporter was integrated correctly (Figure 3-12).

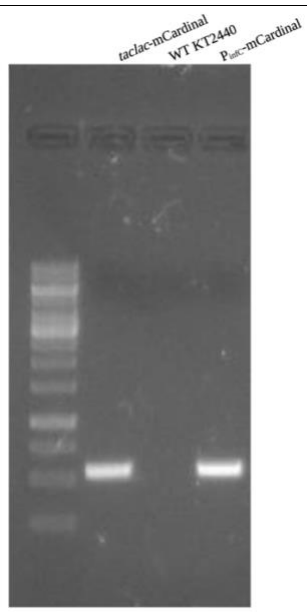


Figure 3-12, P. putida *plasmid integration PCR: Culture PCR ensuring the integration of the P_{infC}-mCardinal plasmid into the P. putida genome.*

Growth curves of the new pJH0204 + P_{infC}-mCardinal Flag plasmid were created alongside growth curves of the *taclac*-mCardinal positive control and the WT KT2440 negative control to test if fluorescent protein was being produced. Overnight cultures were back diluted into LB, M9, and SESOM to an OD_{600} of 0.05 and were then plated in a 96-well plate. The plate was inserted into the plate reader programmed to shake at 30°C and take an absorbance and fluorescence reading every hour. Over 24 hours no fluorescence was observed in the strain with P_{infC}-mCardinal integrated in any of the medias tested (Figure 3-13 A, B, and C). The fluorescence of the WT and P_{infC}-mCardinal were indistinguishable, indicating that either the region originally suspected of being the promoter is not the correct promoter region, or that the promoter does not work well in genetic engineering applications. To diagnose the problem a similar growth curve was generated but with the constructs in *E. coli* to see if the construct worked in a different organism (Figure 3-14 A, B, and C). Again however, no expression of mCardinal was observed from the P_{infC}-mCardinal reporter.

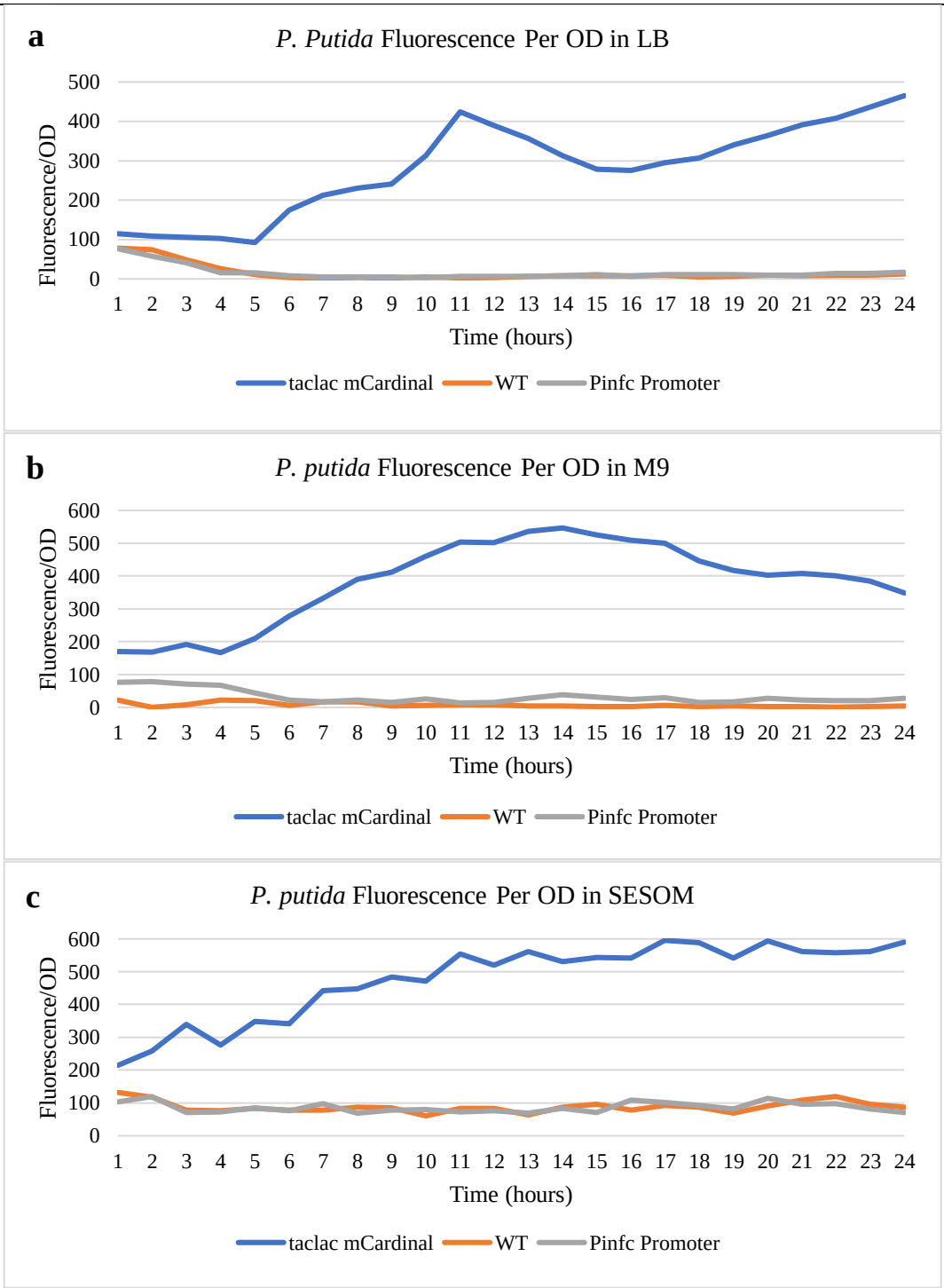


Figure 3-13, P. putida *growth curve: Fluorescence over 24 hours of pJH0204 + P_{infC}-mCardinal, taclac-mCardinal, and WT KT2440 in P. putida in three different media types: LB (a), M9 (b), and SESOM (c). Each OD_{600} and fluorescence reading is an average of three technical replicates.*

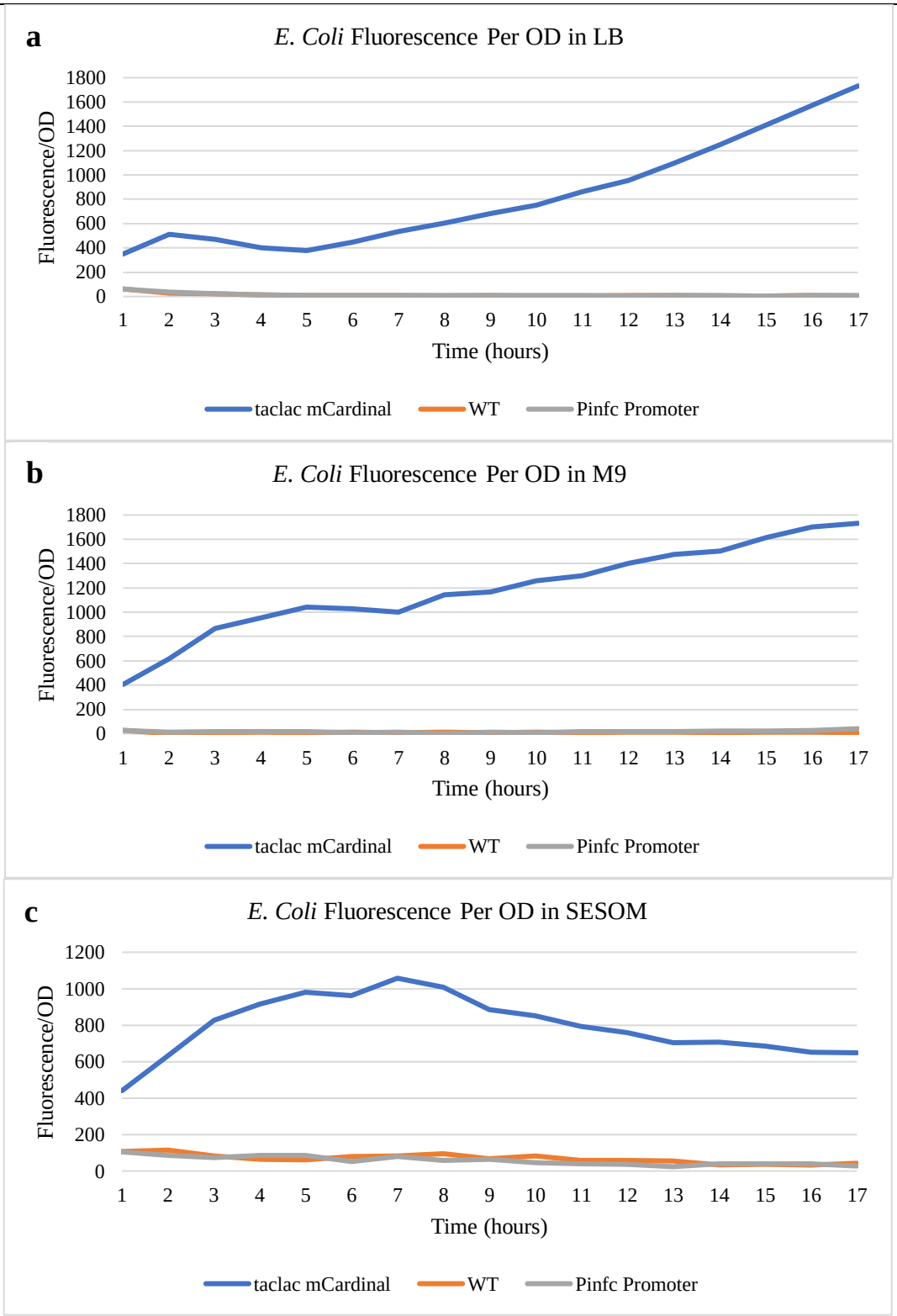


Figure 3-14, E. coli growth curve: Fluorescence over 24 hours of pJH0204 + P_{infC}-mCardinal, taclac-mCardinal, and WT KT2440 in E. coli in three different media types: LB (a), M9 (b),

and SESOM (c). Each OD_{600} and fluorescence reading is an average of three technical replicates.

In addition to the *E. coli* growth curves, an RT-PCR was completed to determine if mCardinal was being transcribed to RNA but just not translated to protein inside of *P. putida*. Overnight cultures of P_{infC}-mCardinal, taclac-mCardinal, and WT KT2440 were back diluted and allowed to grow to an OD_{600} of 0.5. Samples were taken and their RNA was extracted to make cDNA as before. This RT-PCR, Figure 3-15, showed that there is some transcription of P_{infC}-mCardinal mRNA being produced, but not nearly as much as the constitutive mCardinal producer.

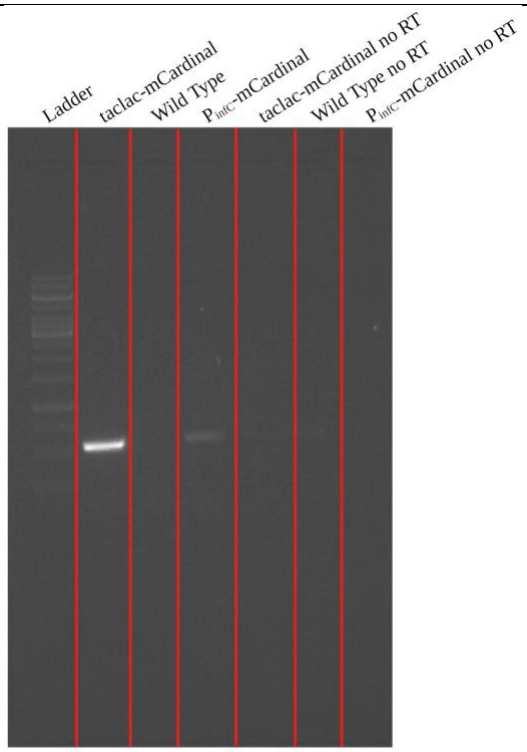


Figure 3-15, transcription RT-PCR: RT-PCR of pJH0204 + P_{infC}-mCardinal, taclac-mCardinal, and WT KT2440, including their respective no RT controls.

It was concluded that the P_{infC} promoter that was identified upstream of the *thrS* gene was a generally weak promoter. Therefore, additional research was performed in the literature to identify possible explanations for the behavior of this promoter. The homologous region of the *E. coli* genome, which has been much more thoroughly studied than *P. putida*, was aligned to the *P. putida thrS-infC* genomic locus (Wertheimer et al., 1988). Figure 3-16 shows the area just

upstream of *infC* in *P. putida* aligned with area defined in *E. coli* to have a strong promoter for *infC*. This strong promoter, known as P_{12}, lies within the coding region of *thrS* (Wertheimer et al., 1988). The -10 and -35 consensus boxes for σ^{70} found in the P_{12} promoter are conserved across the two organisms (Figure 3-16, line 2) It is likely therefore that the analogous strong promoter for *infC* in *P. putida* also resides in that area, though this has not been validated in the literature. Future directions of the project utilizing this promoter region are discussed in the Discussion Chapter.

Figure 3-16, analogous infC alignment: Alignment of the area upstream of infC in E. coli *and* P. putida*. A longer alignment including all of infC is included in Appendix C.*

3.4 Discussion

The overall goal of this part of the project was to use RNA-seq data validated with qPCR to find a promoter which could serve as a reliable constitutive promoter for genetic engineering applications. Ultimately, data obtained through qPCR did not match the expected bioinformatic findings of consistent genetic expression across environmental stresses. Also, the region suspected of being a strong promoter did not drive expression of a mCardinal in a reporter assay. There are a few possible explanations as to why the qPCR data did not match the transcriptomics. The selected genes may be part of the 20% of genes that don't behave as expected and have a low fold change (Coenye, 2021). There may be other issues in the RNA-seq data, such as those outlined in Bruning et al., 2015 that, when analyzed, generated results that were not representative of actual genetic expression. There are also confounding variables in the qPCR experiments that could be resolved such as further optimizing primer efficiencies so that a more direct comparison of the expression values can be made.

The promoter region selected for use in a reporter assay did not function as intended and more investigation needs to be done into the genetic sequences surrounding *thrS* and *infC* in order to understand why. While the location and function of these genes in *P. putida* may be known, there is an analogous site in *E. coli* which has been better studied. Genes encoding for *infC* and *thrS* are conserved across *E. coli* and *P. putida*, making the coding area in *E. coli* a good place to start the search for possible promoter regions (Keseler et al., 2021). The analogous genetic area in *E. coli* is shown in Figure 3-17. Due to the locations of *thrS* and *infC* it was assumed that the transcription of both genes would be controlled by a sequence just upstream of *thrS*, as thought to be the case in *P. putida* as well (Mayaux et al., 1983). Upon further investigation it was determined that there are three promoters inside the coding region of *thrS* that may be able to control the expression of *infC* (Plumbridge et al., 1980). The first promoter is in fact the promoter for *thrS* that is located 170 nucleotides upstream from its translation initiation start site (Wertheimer et al., 1988). This promoter is responsible for co-transcription of *thrS* and *infC*, validating the findings that these genes can be on the same mRNA strand. However, this promoter is not the most efficient promoter for the production of *infC* (Wertheimer et al., 1988). There are two other promoters which lay inside of the coding region for *thrS*, one of which, called P_{12}, seems to the most efficient promoter for driving the expression of *infC* (Wertheimer et al., 1988). This promoter is located 178 to 267 base pairs upstream from the start codon of *infC* and it has a -35 box of 5'-TTGAGA-3' and a -10 box of 5'-TAAAAT-3' (Pramanik et al., 1986; Wertheimer et al., 1988). The canonical

-35 and -10 boxes for σ^{70} are 5'-TTGACA-3' and 5'-TATAAT-3' respectively, confirming the region's candidacy as a promoter (He et al., 2018). The sequence surrounding this area is shown in Figure 3-16 and 3-17.

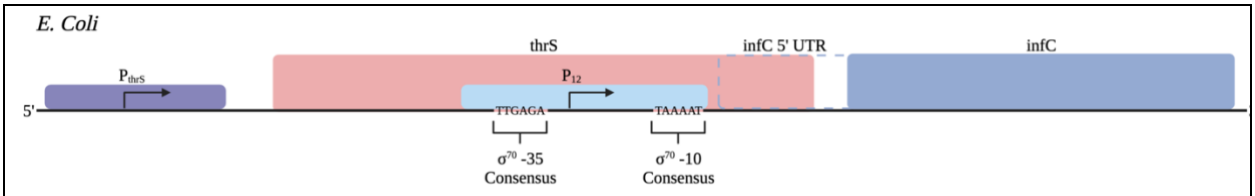


Figure 3-17, infC and thrS in E. coli: Annotated area surrounding the infC gene in E. coli.

It is also important to note that this P_{12} promoter allows expression of *infC* to be fivefold higher than that of *thrS* in steady state conditions, possibly explaining why even though *thrS* and *infC* can be transcribed on the same mRNA strand, *thrS* was not identified in the original bioinformatic gene identification preceding this work (Wertheimer et al., 1988). It also explains the low levels of *infC* seen in the cDNA in the final RT-PCR of this work. Expression levels in *E. coli* for *infC* were also quantified under each of the three *infC* promoters. Importantly, the transcript levels originating from the P_{12} promoter remained constant under stress where transcript levels under the other two promoters were variable under stress (Giuliodori et al., 2007) This finding further proves that the site analogous to P_{12} in *P. Putida* is likely the best promoter candidate to try.

The next step to creating a successful construct with a promoter for *infC* is to employ the findings from the analogous region in *E. coli*. *InfC* and *thrS*, along with the genes downstream from them, rpmI and rplT, are thought to be part of a multi gene operon in *P. putida* (D'Arrigo et al., 2016). However, the same was thought to be true in *E. coli* before other promoters in the coding regions of these genes was discovered. It is entirely possible that the strongest promoter for *infC* has not been characterized but exists inside of the coding region of *thrS* in *P. putida* just like that of *E. coli*. One approach to try to identify the promoter would be to separate the coding region of *thrS* into chunks and to test each one of those chunks as a possible promoter with a reporter. Constructs that see expression would be indicative of containing the region of *thrS* that is the promoter for *infC* and could in turn be split into smaller sequences to get a more exact promoter location. This process would be time consuming, and a more targeted approach can be taken that looks at the untranslated region (UTR) of *infC*. In *E. coli*, the promoter for *infC* lays in a certain area just outside the 5' UTR of the gene, so a similar area could be taken from *P. putida*. The UTR

of *infC* in *P. putida* is 142 base pairs in length and stretches into the coding region of *thrS* just as it does in *E. coli* (D'Arrigo et al., 2016), indicating that *P putida*'s *infC* may also be regulated from inside the coding region of *thrS*. The sequence just upstream of the UTR can then be taken and cloned into a reporter system to test for promoter strength. Since these genes are structured similarly in *E. coli* and in *P. putida* it is likely that this region is the strongest promoter for *infC*, and therefore is a good candidate to use in genetic engineering applications.

Chapter 4, Conclusions

This work aimed to characterize and improve existing butanol biosensors along with use validated transcriptomic data to determine to identify and test a possible promoter that could be used in future genetic engineering applications. After testing existing butanol biosensors, it was determined that they lack clear and reproducible induction patterns. To improve the sensors and reduce observed leakiness a few changes were made to the genetic circuit. The reporter was changed from GFP to mCardinal to reduce background fluorescence, a cloning site for a promoter library was included to test different promoters, and an epitope tag was added to the BmoR protein in order better understand the behavior of the activator protein. After more test inductions degradation of BmoR was observed in the new construct. Given the inconclusive data of prior published butanol biosensors combined with BmoR's apparent instability *E. coil* and inability to regulate the P_{BMO} promoter no further action was taken to develop the biosensor. There were inconsistencies when validating transcriptomic data with qPCR. The genes tested all had expression levels much less than was expected from the bioinformatic analysis. In addition, the region thought to be the promoter of the *infC* gene did not produce fluorescence when integrated into a reporter assay. Although, after comparing the genetic locus surrounding *infC* in *P. putida* to the much more characterized analogous site in *E. coli*, a better promoter candidate was discovered based on the endogenous *E. coli* P_{12} promoter.

Despite the shortcomings of this work there is still an important broader impact to consider. The first broad lesson to take away is that to create a fully functional microbial biosensor, all the parts need to be validated, more than just by sequence, to ensure proper functionality. Often, when fluorescence is observed from a system it is assumed that all the parts, including associated proteins, are functioning as intended. As seen in the published butanol biosensors, this is not always the case as the fluorescence observed was likely due to leakiness while this work showed that BmoR was not stable and not able to induce reporter expression. The second broad lesson is to use bioinformatic meta-data analysis as a tool, it must also be validated. The disconnect between large scale transcriptome analysis across multiple studies and benchtop qPCR validation became apparent in this work. Transcriptome meta-analyses still remain a very helpful tool to identify novel systems and to compare one data set to a larger set, but results obtained should be verified before they are implemented into future steps in a study.

References

Abbas-Aghababazadeh, F., Li, Q., & Fridley, B. L. (2018). Comparison of normalization approaches for gene expression studies completed with high-throughput sequencing. *PLOS ONE*, *13*(10), e0206312. https://doi.org/10.1371/journal.pone.0206312

Beal, J., Overney, C., Adler, A., Yaman, F., Tiberio, L., & Samineni, M. (2019). TASBE Flow Analytics: A Package for Calibrated Flow Cytometry Analysis. *ACS Synthetic Biology*, *8*(7), 1524–1529. https://doi.org/10.1021/acssynbio.8b00533

Benner, S. A., & Sismour, A. M. (2005). Synthetic biology. *Nature Reviews Genetics 2005 6:7*, *6*(7), 533–543. https://doi.org/10.1038/nrg1637

Bloch, K. D. (1992). Digestion of <scp>DNA</scp> with Restriction Endonucleases. *Current Protocols in Immunology*, *2*(1). https://doi.org/10.1002/0471142735.im1008s02

Blount, Z. D. (2015). The unexhausted potential of E. coli. *ELife*, *4*. https://doi.org/10.7554/eLife.05826

Bokinsky, G., Peralta-Yahya, P. P., George, A., Holmes, B. M., Steen, E. J., Dietrich, J., Soon Lee, T., Tullman-Ercek, D., Voigt, C. A., Simmons, B. A., & Keasling, J. D. (2011). Synthesis of three advanced biofuels from ionic liquid-pretreated switchgrass using engineered *Escherichia coli*. *Proceedings of the National Academy of Sciences*, *108*(50), 19949–19954. https://doi.org/10.1073/pnas.1106958108

Browning, D. F., & Busby, S. J. W. (2016). Local and global regulation of transcription initiation in bacteria. *Nature Reviews Microbiology*, *14*(10), 638–650. https://doi.org/10.1038/nrmicro.2016.103

Bruning, O., Rauwerda, H., Dekker, R. J., de Leeuw, W. C., Wackers, P. F. K., Ensink, W. A., Jonker, M. J., & Breit, T. M. (2015). Valuable lessons-learned in transcriptomics experimentation. *Transcription*, *6*(3), 51–55. https://doi.org/10.1080/21541264.2015.1064195

Caldas, J., & Vinga, S. (2014). Global Meta-Analysis of Transcriptomics Studies. *PLoS ONE*, *9*(2), e89318. https://doi.org/10.1371/journal.pone.0089318

Cameron, D. E., Bashor, C. J., & Collins, J. J. (2014). A brief history of synthetic biology. *Nature Reviews Microbiology 2014 12:5*, *12*(5), 381–390. https://doi.org/10.1038/nrmicro3239

Chang, M. C. Y., & Zhao, H. (2015). Editorial overview: Opportunities and challenges in synthetic biology. *Current Opinion in Chemical Biology*, *28*, v–vi. https://doi.org/10.1016/j.cbpa.2015.08.001

Chang, Q., Amemiya, T., Liu, J., Xu, X., Rajendran, N., & Itoh, K. (2009). Identification and validation of suitable reference genes for quantitative expression of xylA and xylE genes in Pseudomonas putida mt-2. *Journal of Bioscience and Bioengineering*, *107*(2), 210–214. https://doi.org/10.1016/j.jbiosc.2008.09.017

Chavez, M., Ho, J., & Tan, C. (2017). Reproducibility of High-Throughput Plate-Reader Experiments in Synthetic Biology. *ACS Synthetic Biology*, *6*(2), 375–380. https://doi.org/10.1021/acssynbio.6b00198

Chevez-Guardado, R., & Peña-Castillo, L. (2021). Promotech: a general tool for bacterial promoter recognition. *Genome Biology*, *22*(1), 318. https://doi.org/10.1186/s13059-021-02514-9

Clarke, E. J., & Voigt, C. A. (2011). Characterization of combinatorial patterns generated by multiple two-component sensors in E. coli that respond to many stimuli. *Biotechnology and Bioengineering*, *108*(3), 666–675. https://doi.org/10.1002/BIT.22966

Coenye, T. (2021). Do results obtained with RNA-sequencing require independent verification? *Biofilm*, *3*, 100043. https://doi.org/10.1016/j.bioflm.2021.100043

Cuenca, M. del S., Roca, A., Molina-Santiago, C., Duque, E., Armengaud, J., Gómez-Garcia, M. R., & Ramos, J. L. (2016). Understanding butanol tolerance and assimilation in Pseudomonas putidaBIRD-1: An integrated omics approach. *Microbial Biotechnology*, *9*(1), 100–115. https://doi.org/10.1111/1751-7915.12328

Daegelen, P., Studier, F. W., Lenski, R. E., Cure, S., & Kim, J. F. (2009). Tracing Ancestors and Relatives of Escherichia coli B, and the Derivation of B Strains REL606 and BL21(DE3). *Journal of Molecular Biology*, *394*(4), 634–643. https://doi.org/10.1016/j.jmb.2009.09.022

D'Arrigo, I., Bojanovič, K., Yang, X., Holm Rau, M., & Long, K. S. (2016). Genome-wide mapping of transcription start sites yields novel insights into the primary transcriptome of *Pseudomonas putida*. *Environmental Microbiology*, *18*(10), 3466–3481. https://doi.org/10.1111/1462-2920.13326

D'Arrigo, I., Cardoso, J. G. R., Rennig, M., Sonnenschein, N., Herrgård, M. J., & Long, K. S. (2019). Analysis of Pseudomonas putida growth on non-trivial carbon sources using transcriptomics and genome-scale modelling. *Environmental Microbiology Reports*, *11*(2), 87–97. https://doi.org/10.1111/1758-2229.12704

Davis, M. C., Kesthely, C. A., Franklin, E. A., & MacLellan, S. R. (2017). The essential activities of the bacterial sigma factor. *Canadian Journal of Microbiology*, *63*(2), 89–99. https://doi.org/10.1139/cjm-2016-0576

de Las Heras, A., & de Lorenzo, V. (2011). Cooperative amino acid changes shift the response of the σ54-dependent regulator XylR from natural m-xylene towards xenobiotic 2,4-dinitrotoluene. *Molecular Microbiology*, *79*(5), 1248–1259. https://doi.org/10.1111/j.1365-2958.2010.07518.x

Dietrich, J. A., Shis, D. L., Alikhani, A., & Keasling, J. D. (2013). Transcription factor-based screens and synthetic selections for microbial small-molecule biosynthesis. *ACS Synthetic Biology*, *2*(1), 47–58. https://doi.org/10.1021/sb300091d

Elmore, J. R., Furches, A., Wolff, G. N., Gorday, K., & Guss, A. M. (2017). Development of a high efficiency integration system and promoter library for rapid modification of Pseudomonas putida KT2440. *Metabolic Engineering Communications*, *5*, 1–8. https://doi.org/10.1016/j.meteno.2017.04.001

Escherich, T. (n.d.). The intestinal bacteria of the neonate and breast-fed infant. 1884. *Reviews of Infectious Diseases*, *10*(6), 1220–1225. https://doi.org/10.1093/clinids/10.6.1220

Fernandez-López, R., Ruiz, R., de la Cruz, F., & Moncalián, G. (2015). Transcription factor-based biosensors enlightened by the analyte. *Frontiers in Microbiology*, *6*. https://doi.org/10.3389/fmicb.2015.00648

Follonier, S., Escapa, I. F., Fonseca, P. M., Henes, B., Panke, S., Zinn, M., & Prieto, M. A. (2013). New insights on the reorganization of gene transcription in Pseudomonas putida KT2440 at elevated pressure. *Microbial Cell Factories*, *12*(1), 30. https://doi.org/10.1186/1475-2859-12-30

Frank, S., Schmidt, F., Klockgether, J., Davenport, C. F., Gesell Salazar, M., Völker, U., & Tümmler, B. (2011). Functional genomics of the initial phase of cold adaptation of Pseudomonas putida KT2440. In *FEMS Microbiology Letters* (Vol. 318, Issue 1, pp. 47–54). Oxford Academic. https://doi.org/10.1111/j.1574-6968.2011.02237.x

Fu, J., Sharma, P., Spicer, V., Krokhin, O. v., Zhang, X., Fristensky, B., Wilkins, J. A., Cicek, N., Sparling, R., & Levin, D. B. (2015). Effects of impurities in biodiesel-derived glycerol on growth and expression of heavy metal ion homeostasis genes and gene products in Pseudomonas putida LS46. *Applied Microbiology and Biotechnology*, *99*(13), 5583–5592. https://doi.org/10.1007/s00253-015-6685-z

Garmendia, J., de Las Heras, A., Galvão, T. C., & de Lorenzo, V. (2008). Tracing explosives in soil with transcriptional regulators of Pseudomonas putida evolved for responding to nitrotoluenes. *Microbial Biotechnology*, *1*(3), 236–246. https://doi.org/10.1111/j.1751-7915.2008.00027.x

Garrido-Cardenas, J., Garcia-Maroto, F., Alvarez-Bermejo, J., & Manzano-Agugliaro, F. (2017). DNA Sequencing Sensors: An Overview. *Sensors*, *17*(3), 588. https://doi.org/10.3390/s17030588

Giuliodori, A. M., Brandi, A., Giangrossi, M., Gualerzi, C. O., & Pon, C. L. (2007). Cold-stress-induced de novo expression of *infC* and role of IF3 in cold-shock translational bias. *RNA*, *13*(8), 1355–1365. https://doi.org/10.1261/rna.455607

Goers, L., Kylilis, N., Tomazou, M., Yan Wen, K., Freemont, P., & Polizzi, K. (2013). Engineering Microbial Biosensors. *Methods in Microbiology*, *40*, 119–156. https://doi.org/10.1016/B978-0-12-417029-2.00005-4

Gray, D. (1997). Overview of Protein Expression by Mammalian Cells. *Current Protocols in Protein Science*, *10*(1). https://doi.org/10.1002/0471140864.ps0509s10

Gu, P., Liu, L., Ma, Q., Dong, Z., Wang, Q., Xu, J., Huang, Z., & Li, Q. (2021a). Metabolic engineering of Escherichia coli for the production of isobutanol: a review. *World Journal of Microbiology and Biotechnology 2021 37:10*, *37*(10), 1–9. https://doi.org/10.1007/S11274-021-03140-0

Gu, P., Liu, L., Ma, Q., Dong, Z., Wang, Q., Xu, J., Huang, Z., & Li, Q. (2021b). Metabolic engineering of Escherichia coli for the production of isobutanol: a review. *World Journal of Microbiology and Biotechnology*, *37*(10), 1–9. https://doi.org/10.1007/S11274-021-03140-0/TABLES/1

Harding, K., Farny, N., & Young, E. M. (2021). *Transcriptomic Analysis of Pseudomonas putida for Implementation of Soil Microbial Biosensor A Major Qualifying Project*.

Hausjell, J., Weissensteiner, J., Molitor, C., Halbwirth, H., & Spadiut, O. (2018). E. coli HMS174(DE3) is a sustainable alternative to BL21(DE3). *Microbial Cell Factories*, *17*(1), 169. https://doi.org/10.1186/s12934-018-1016-6

Hay, M., Yu, M. L., & Ma, Y. (2017). *RNA extraction of Escherichia coli grown in Lysogeny Broth for use in RT-qPCR | jemi.microbiology.ubc.ca*. https://ujemi.microbiology.ubc.ca/node/124

He, W., Jia, C., Duan, Y., & Zou, Q. (2018). 70ProPred: a predictor for discovering sigma70 promoters based on combining multiple features. *BMC Systems Biology*, *12*(S4), 44. https://doi.org/10.1186/s12918-018-0570-1

Holmsgaard, P. N., Norman, A., Hede, S. C., Poulsen, P. H. B., Al-Soud, W. A., Hansen, L. H., & Sørensen, S. J. (2011). Bias in bacterial diversity as a result of Nycodenz extraction from bulk soil. *Soil Biology and Biochemistry*, *43*(10), 2152–2159. https://doi.org/10.1016/j.soilbio.2011.06.019

Huynh, K., & Partch, C. L. (2015). Analysis of Protein Stability and Ligand Interactions by Thermal Shift Assay. *Current Protocols in Protein Science*, *79*(1). https://doi.org/10.1002/0471140864.ps2809s79

Jörnvall, H. (1994). The alcohol dehydrogenase system. In *Toward a Molecular Basis of Alcohol Use and Abuse* (pp. 221–229). Birkhäuser Basel. https://doi.org/10.1007/978-3-0348-7330-7_22

Keseler, I. M., Gama-Castro, S., Mackie, A., Billington, R., Bonavides-Martínez, C., Caspi, R., Kothari, A., Krummenacker, M., Midford, P. E., Muñiz-Rascado, L., Ong, W. K., Paley, S., Santos-Zavaleta, A., Subhraveti, P., Tierrafría, V. H., Wolfe, A. J., Collado-Vides, J., Paulsen, I. T., & Karp, P. D. (2021). The EcoCyc Database in 2021. *Frontiers in Microbiology*, *12*. https://doi.org/10.3389/fmicb.2021.711077

Khalili, B., Weihe, C., Kimball, S., Schmidt, K. T., & Martiny, J. B. H. (2019). Optimization of a Method To Quantify Soil Bacterial Abundance by Flow Cytometry. *MSphere*, *4*(5). https://doi.org/10.1128/msphere.00435-19

Kim, M., Lim, J. W., Kim, H. J., Lee, S. K., Lee, S. J., & Kim, T. (2015). Chemostat-like microfluidic platform for highly sensitive detection of heavy metal ions using microbial biosensors. *Biosensors and Bioelectronics*, *65*, 257–264. https://doi.org/10.1016/j.bios.2014.10.028

Kivisaar, M. (2020). Narrative of a versatile and adept species Pseudomonas putida. *Journal of Medical Microbiology*, *69*(3), 324–338. https://doi.org/10.1099/jmm.0.001137

Kukurba, K. R., & Montgomery, S. B. (2015). RNA Sequencing and Analysis. *Cold Spring Harbor Protocols*, *2015*(11), pdb.top084970. https://doi.org/10.1101/pdb.top084970

Lim, J. W., Ha, D., Lee, J., Lee, S. K., & Kim, T. (2015). Review of Micro/Nanotechnologies for Microbial Biosensors. *Frontiers in Bioengineering and Biotechnology*, *3*. https://doi.org/10.3389/fbioe.2015.00061

Lindahl, V., & Bakken, L. R. (2006). Evaluation of methods for extraction of bacteria from soil. *FEMS Microbiology Ecology*, *16*(2), 135–142. https://doi.org/10.1111/j.1574-6941.1995.tb00277.x

Liu, J., Li, J. Q., Feng, L., Cao, H., & Cui, Z. (2010). An improved method for extracting bacteria from soil for high molecular weight DNA recovery and BAC library construction. *Journal of Microbiology*, *48*(6), 728–733. https://doi.org/10.1007/s12275-010-0139-1

Livak, K. J., & Schmittgen, T. D. (2001). Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2−ΔΔCT Method. *Methods*, *25*(4), 402–408. https://doi.org/10.1006/METH.2001.1262

Marbach, A., & Bettenbrock, K. (2012). lac operon induction in Escherichia coli: Systematic comparison of IPTG and TMG induction and influence of the transacetylase LacA. *Journal of Biotechnology*, *157*(1), 82–88. https://doi.org/10.1016/j.jbiotec.2011.10.009

Martin-Pascual, M., Batianis, C., Bruinsma, L., Asin-Garcia, E., Garcia-Morales, L., Weusthuis, R. A., van Kranenburg, R., & Martins dos Santos, V. A. P. (2021). A navigation guide of synthetic biology tools for Pseudomonas putida. *Biotechnology Advances*, *49*, 107732. https://doi.org/10.1016/j.biotechadv.2021.107732

Mayaux, J. F., Fayat, G., Fromant, M., Springer, M., Grunberg-Manago, M., & Blanquet, S. (1983). Structural and transcriptional evidence for related thrS and infC expression. *Proceedings of the National Academy of Sciences*, *80*(20), 6152–6156. https://doi.org/10.1073/pnas.80.20.6152

McKinnon, K. M. (2018). Flow Cytometry: An Overview. *Current Protocols in Immunology*, *120*(1). https://doi.org/10.1002/cpim.40

Meyer, A. J., Segall-Shapiro, T. H., Glassey, E., Zhang, J., & Voigt, C. A. (2018). Escherichia coli "Marionette" strains with 12 highly optimized small-molecule sensors. *Nature Chemical Biology 2018 15:2*, *15*(2), 196–204. https://doi.org/10.1038/s41589-018-0168-3

MONOD, J., & JACOB, F. (1961). Teleonomic mechanisms in cellular metabolism, growth, and differentiation. *Cold Spring Harbor Symposia on Quantitative Biology*, *26*, 389–401. https://doi.org/10.1101/sqb.1961.026.01.048

Mozejko-Ciesielska, J., Pokoj, T., & Ciesielski, S. (2018). Transcriptome remodeling of *Pseudomonas putida* KT2440 during mcl-PHAs synthesis: effect of different carbon sources and response to nitrogen stress. *Journal of Industrial Microbiology and Biotechnology*, *45*(6), 433–446. https://doi.org/10.1007/s10295-018-2042-4

Nakamura, H., Shimomura-Shimizu, M., & Karube, I. (n.d.). Development of Microbial Sensors and Their Application. In *Biosensing for the 21st Century* (pp. 351–394). Springer Berlin Heidelberg. https://doi.org/10.1007/10_2007_085

Nakamura, H., Shimomura-Shimizu, M., & Karube, I. (2007). Development of Microbial Sensors and Their Application. In *Biosensing for the 21st Century* (Vol. 109, pp. 351–394). Springer Berlin Heidelberg. https://doi.org/10.1007/10_2007_085

Namdev, P., Dar, H. Y., Srivastava, R. K., Mondal, R., & Anupam, R. (2019). Induction of T7 Promoter at Higher Temperatures May Be Counterproductive. *Indian Journal of Clinical Biochemistry*, *34*(3), 357–360. https://doi.org/10.1007/s12291-019-0813-y

Nielsen, D. R., Leonard, E., Yoon, S. H., Tseng, H. C., Yuan, C., & Prather, K. L. J. (2009). Engineering alternative butanol production platforms in heterologous bacteria. *Metabolic Engineering*, *11*(4–5), 262–273. https://doi.org/10.1016/j.ymben.2009.05.003

Niesen, F. H., Berglund, H., & Vedadi, M. (2007). The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nature Protocols*, *2*(9), 2212–2221. https://doi.org/10.1038/nprot.2007.321

Part:BBa R0051 - parts.igem.org. (n.d.). Retrieved September 21, 2021, from http://parts.igem.org/Part:BBa_R0051

Peng, J., Miao, L., Chen, X., & Liu, P. (2018). Comparative transcriptome analysis of Pseudomonas putida KT2440 revealed its response mechanisms to elevated levels of zinc stress. *Frontiers in Microbiology*, *9*(JUL), 1669. https://doi.org/10.3389/fmicb.2018.01669

Plumbridge, J. A., Springer, M., Graffe, M., Goursot, R., & Grunberg-Manago, M. (1980). Physical localisation and cloning of the structural gene for E. coli initiation factor IF3 from a group of genes concerned with translation. *Gene*, *11*(1–2), 33–42. https://doi.org/10.1016/0378-1119(80)90084-0

Pramanik, A., Wertheimer, S. J., Schwartz, J. J., & Schwartz, I. (1986). Expression of Escherichia coli infC: identification of a promoter in an upstream thrS coding sequence. *Journal of Bacteriology*, *168*(2), 746–751. https://doi.org/10.1128/jb.168.2.746-751.1986

Pseudomonas butanovora BmoR (bmoR), IstA (istA), and IstB (istB) genes - Nucleotide - NCBI. (n.d.). Retrieved June 5, 2021, from https://www.ncbi.nlm.nih.gov/nuccore/AY093933.3/

Ragoonanan, V., & Aksan, A. (2007). Protein Stabilization. *Transfusion Medicine and Hemotherapy*, *34*(4), 246–252. https://doi.org/10.1159/000104678

Renneberg, R., Pfeiffer, D., Lisdat, F., Wilson, G., Wollenberger, U., Ligler, F., & Turner, A. P. F. (2007). Frieder Scheller and the Short History of Biosensors. In *Biosensing for the 21st Century* (pp. 1–18). Springer Berlin Heidelberg. https://doi.org/10.1007/10_2007_086

Renneberg, R., Pfeiffer, D., Lisdat, F., Wilson, G., Wollenberger, U., Ligler, F., & Turner, A. P. F. (2008). Frieder Scheller and the Short History of Biosensors. In *Biosensing for the 21st Century* (Vol. 109, pp. 1–18). Springer Berlin Heidelberg. https://doi.org/10.1007/10_2007_086

Rocha, D. J. P. G., Castro, T. L. P., Aguiar, E. R. G. R., & Pacheco, L. G. C. (2020). *Gene Expression Analysis in Bacteria by RT-qPCR* (pp. 119–137). https://doi.org/10.1007/978-1-4939-9833-3_10

Ryan, B., & Henehan, G. (2001). *Current Protocols in Protein Science* (J. E. Coligan, B. M. Dunn, D. W. Speicher, & P. T. Wingfield, Eds.). John Wiley & Sons, Inc. https://doi.org/10.1002/0471140864

Shemer, B., Shpigel, E., Glozman, A., Yagur-Kroll, S., Kabessa, Y., Agranat, A. J., & Belkin, S. (2020). Genome-wide gene-deletion screening identifies mutations that significantly enhance explosives vapor detection by a microbial sensor. *New Biotechnology*, *59*, 65–73. https://doi.org/10.1016/j.nbt.2020.06.002

Shemer, B., Yagur-Kroll, S., Hazan, C., Belkin, S., & Kelly, R. M. (2018). Aerobic Transformation of 2,4-Dinitrotoluene by Escherichia coli and Its Implications for the Detection of Trace Explosives. *Aem.Asm.Org 1 Applied and Environmental Microbiology*, *84*, 1729–1746. https://doi.org/10.1128/AEM

Shi, S., Choi, Y. W., Zhao, H., Tan, M. H., & Ang, E. L. (2017). Discovery and engineering of a 1-butanol biosensor in Saccharomyces cerevisiae. *Bioresource Technology*, *245*, 1343–1351. https://doi.org/10.1016/j.biortech.2017.06.114

Sørensen, H., & Mortensen, K. (2005). Soluble expression of recombinant proteins in the cytoplasm of Escherichia coli. *Microbial Cell Factories*, *4*(1), 1. https://doi.org/10.1186/1475-2859-4-1

Swain, P. S., Elowitz, M. B., & Siggia, E. D. (2002). Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proceedings of the National Academy of Sciences*, *99*(20), 12795–12800. https://doi.org/10.1073/pnas.162041399

T. Das, A., Tenenbaum, L., & Berkhout, B. (2016). Tet-On Systems For Doxycycline-inducible Gene Expression. *Current Gene Therapy*, *16*(3), 156–167. https://doi.org/10.2174/1566523216666160524144041

Taylor, S. C., & Posch, A. (2014). The Design of a Quantitative Western Blot Experiment. *BioMed Research International*, *2014*, 1–8. https://doi.org/10.1155/2014/361590

Thévenot, D. R., Toth, K., Durst, R. A., & Wilson, G. S. (2001). Electrochemical biosensors: recommended definitions and classification1International Union of Pure and Applied Chemistry:

Physical Chemistry Division, Commission I.7 (Biophysical Chemistry); Analytical Chemistry Division, Commission V.5 (Electroanalytical Chemistry).1. *Biosensors and Bioelectronics*, *16*(1–2), 121–131. https://doi.org/10.1016/S0956-5663(01)00115-4

Tu, Q., Yin, J., Fu, J., Herrmann, J., Li, Y., Yin, Y., Stewart, A. F., Müller, R., & Zhang, Y. (2016). Room temperature electrocompetent bacterial cells improve DNA transformation and recombineering efficiency. *Scientific Reports*, *6*(1), 1–8. https://doi.org/10.1038/srep24648

Vilain, S., Luo, Y., Hildreth, M. B., & Brözel, V. S. (2006). Analysis of the Life Cycle of the Soil Saprophyte *Bacillus cereus* in Liquid Soil Extract and in Soil. *Applied and Environmental Microbiology*, *72*(7), 4970–4977. https://doi.org/10.1128/AEM.03076-05

Wan, X., Volpetti, F., Petrova, E., French, C., Maerkl, S. J., & Wang, B. (2019). Cascaded amplifying circuits enable ultrasensitive cellular sensors for toxic metals. *Nature Chemical Biology*, *15*(5), 540–548. https://doi.org/10.1038/s41589-019-0244-3

Wertheimer, S. J., Klotsky, R.-A., & Schwartz, I. (1988). Transcriptional patterns for the thrS-infC-rplT operon of Escherichia coli. *Gene*, *63*(2), 309–320. https://doi.org/10.1016/0378-1119(88)90534-3

Winsor, G. L., Griffiths, E. J., Lo, R., Dhillon, B. K., Shay, J. A., & Brinkman, F. S. L. (2016). Enhanced annotations and features for comparing thousands of Pseudomonasgenomes in the Pseudomonas genome database. *Nucleic Acids Research*, *44*(D1), D646–D653. https://doi.org/10.1093/nar/gkv1227

Xiao, Y., Zhu, W., Liu, H., Nie, H., Chen, W., & Huang, Q. (2018). FinR regulates expression of nicC and nicX operons, involved in nicotinic acid degradation in Pseudomonas putida KT2440. *Applied and Environmental Microbiology*, *84*(20). https://doi.org/10.1128/AEM.01210-18

Yan, Q., & Fong, S. S. (2017). Challenges and advances for genetic engineering of non-model bacteria and uses in consolidated bioprocessing. *Frontiers in Microbiology*, *8*(OCT), 2060. https://doi.org/10.3389/FMICB.2017.02060/BIBTEX

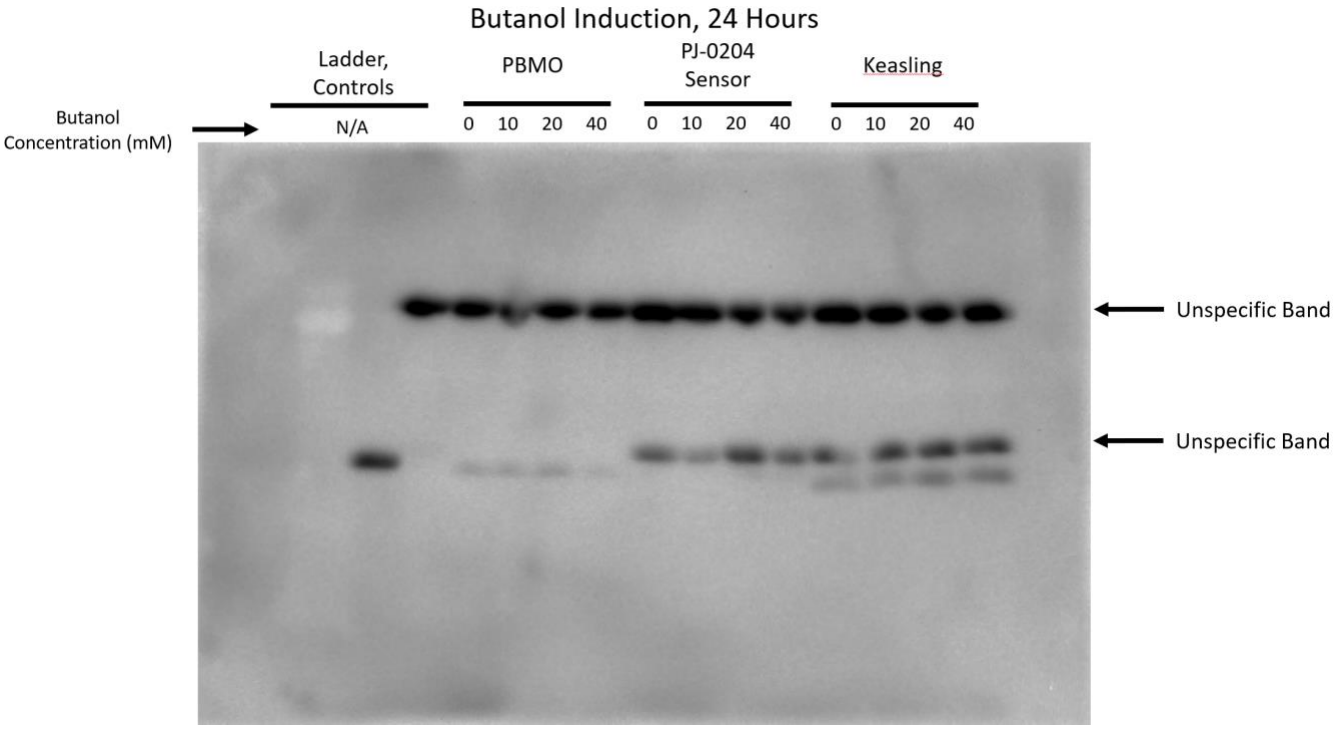
Yu, H., Chen, Z., Wang, N., Yu, S., Yan, Y., & Huo, Y. X. (2019). Engineering transcription factor BmoR for screening butanol overproducers. *Metabolic Engineering*, *56*, 28–38. https://doi.org/10.1016/j.ymben.2019.08.015

Yu, H., Wang, N., Huo, W., Zhang, Y., Zhang, W., Yang, Y., Chen, Z., & Huo, Y. X. (2019). Establishment of BmoR-based biosensor to screen isobutanol overproducer. *Microbial Cell Factories*, *18*(1), 1–11. https://doi.org/10.1186/s12934-019-1084-2

Appendix A – Raw Data

RawDataxlsx

Appendix B – Butanol Induction Western Blot



Appendix C – Sequences Used

Improved Butanol Biosensor Sequence.pdf

infC Genetic Locus.pdf